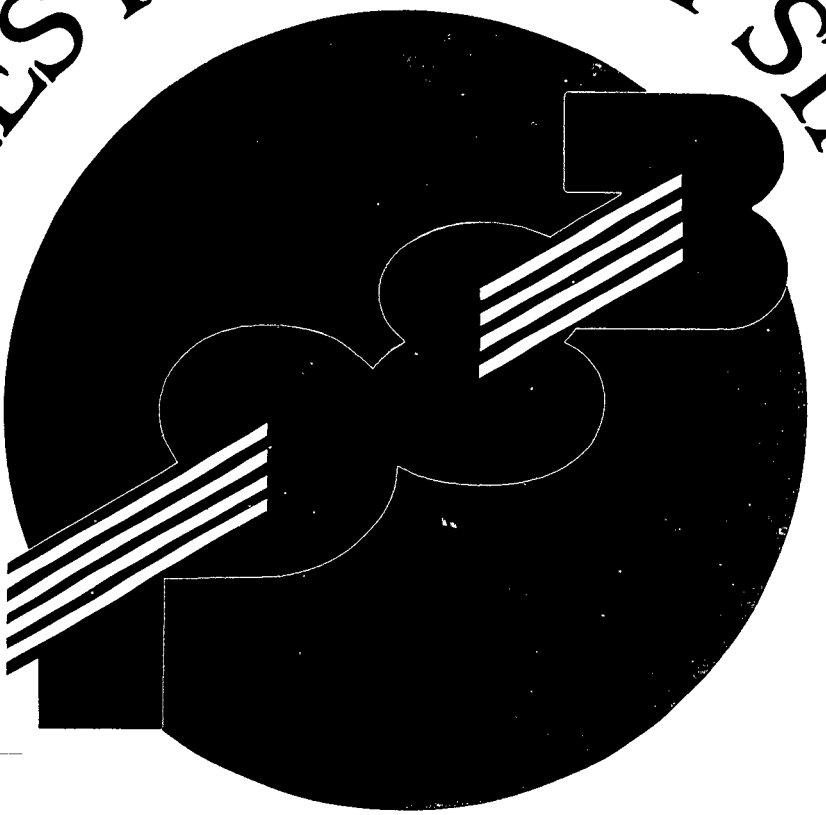
PRIORITIES IN HEALTH STATISTICS

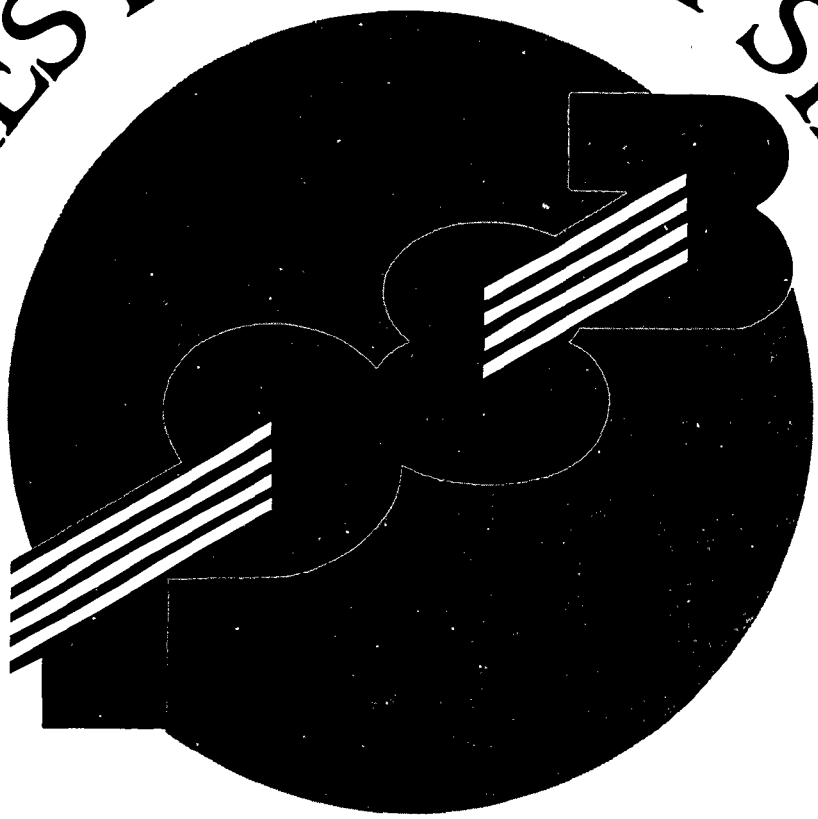


PROCEEDINGS OF THE 19TH NATIONAL MEETING OF THE PUBLIC HEALTH CONFERENCE ON RECORDS AND STATISTICS AUGUST 1983

U.S. Department of Health and Human Services • Public Health Service • National Center for Health Statistics

2/2/84

PRIORITIES IN HEALTH STATISTICS



PROCEEDINGS OF THE 19TH NATIONAL MEETING OF THE PUBLIC HEALTH CONFERENCE ON RECORDS AND STATISTICS AUGUST 1983

DHHS Publication No. (PHS) 81-1214

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
National Center for Health Statistics

Hyattsville, Maryland
December 1983

CONTENTS

CONCURRENT SESSIONS

CONCURRENT SESSIONS

CONCURRENT SESSIONS

STANDBY PRESENTATIONS

APPENDIX

FOREWORD

The Public Health Conference on Records and Statistics (PHCRS) is convened biennially by the National Center for Health Statistics to consider the latest developments in public health reporting systems, vital statistics registration, and health surveys. Starting as a relatively small group of State registrars and health statisticians in 1949, participation in the Conference has grown each year to include other health professionals such as epidemiologists, health planners, and economists. This year, for the 19th National Meeting of the PHCRS, there were 32 sessions, and over 1,000 registrants.

Information and ideas on a wide variety of health statistics programs, techniques of data collection and analysis, latest computer applications, and related matters are exchanged through formal presentations, exhibits, and audience discussion. This year, for the first time, a call for papers was issued and the response was gratifying. Unfortunately, so many excellent papers were submitted that not all could be placed on the program. Summaries of those selected and presented are included in these Proceedings.

The theme for the 19th National Meeting of the PHCRS was "Priorities in Health Statistics." The program addressed three areas having primary need for health information and statistics -- health promotion and disease prevention, health care costs, and monitoring and evaluating health care programs. Within these sets, concurrent sessions focused on major areas of interest relating to systems design and program administration, methodology, and data use and analysis. Special sessions were presented by the National Institute of Mental Health and the Health Care Financing Administration.

Judging from the many comments we have received, the 1983 PHCRS was indeed a great success. The participants felt that the program was timely, relevant, and informative, offering them new information which they can put to practical use in their work. In addition, both the speakers and the topics presented were highly praised.

Our thanks to all who participated in the 1983 Conference, whether by planning the sessions, preparing papers, or contributing to the discussions. Each one helped to make the Conference the informative, well-structured, and beneficial interchange that it was. It is our hope in publishing these Proceedings that the information and ideas exchanged at this Conference will be a valuable contribution toward meeting the rapidly changing data needs of the 1980s.

Manning Feinleib, M.D., Dr.P.H.
Director
National Center for Health Statistics

First Plenary Session

CALL TO ORDER AND OPENING REMARKS

Manning Feinleib, M.D., Dr.P.H., Director, National Center for Health Statistics

Good morning, and welcome to the 19th national meeting of the Public Health Conference on Records and Statistics. For those of you who have not had a chance to look at your programs, I am Manning Feinleib, Director of the National Center for Health Statistics, which is sponsoring this meeting. It is a great pleasure to have you here.

I would like to begin this morning by introducing a few of our special visitors. I hope that they will stand as I call their names.

Dr. Baruch Modan, Director-General of the Ministry of Health of Israel.

Mr. John Coombs, Director of the Health Division of Statistics Canada in Ottawa, Canada.

Four former directors of the National Center for Health Statistics are with us: Forrest Linder, Ted Woolsey, Ed Perrin, and Dorothy Rice.

Thank you. We also are delighted to have a number of other visitors from Canada and the World Health Organization. The members of the National Committee on Vital and Health Statistics are in attendance. So are the deans or their representatives from the schools of public health in the United States. We have several other distinguished guests who will be addressing us this morning, and I will introduce them individually a little later. It is good to have all of you here.

These meetings go back to 1942, when a group of State registrars met in St. Louis, Missouri, to discuss ways to improve the vital registration system. When the title Public Health Conference on Records and Statistics was first used in 1958, the National Health Survey was still young and the National Center for Health Statistics did not exist at all. While all of the States had vital registration offices, none had a State center for health statistics. Today, a majority of the States have such centers. The Public Health Conference on Records and Statistics provides a forum for representatives from Federal, State, and local health agencies, as well as universities and professional associations, to share their knowledge and experience. This diverse gathering lends the Conference a rich variety of perspectives on current issues concerning health information systems in the United States. These Conferences have been held on a biennial basis with but a few interruptions. I hope we will be able to resume the biennial schedule henceforth.

Before turning to a discussion of the content of our Conference during the next three days, I would first like to bring you up to date on some of the activities at the National Center.

As many of you may know, the Center conducts about a dozen surveys on an annual or periodic basis. The National Vital Statistics Cooperative Program, the National Health Interview Survey, and the National Hospital Discharge Survey are conducted annually on a continuing basis. The other surveys which span a variety of data collection systems are now being conducted according to a Periodicity Plan established two years ago. Altogether, our data collection activities cover the spectrum of health and illness from birth to death. We are covering every facet of the health spectrum. The birth and death registration systems provide total population coverage. The other surveys are based on stratified, random samples carefully designed and selected by our research and methodology group to provide representative samples of the United States. I will not go over each of these systems but you will be hearing a lot about them during the next three days of the Conference.

Despite various constraints in terms of budget and personnel, the Center is in fairly good shape. All of the surveys that are currently planned for Fiscal Years 1983 and 1984 are expected to be conducted on schedule. However, we are continuing to explore changes that we can make to increase our capacity and capability for producing needed health data. One of the most important directions that we are exploring is more effective integration of our data systems through utilization of a common sample for the population-based surveys. This promises to be a cost-efficient change that will also enhance the analytical potential of the separate surveys. We are also entering into formal agreements with other Public Health Service agencies that will enable us to collect information of particular importance to them in coordination with our ongoing surveys.

Many of you may know that I have been involved for many years with several major longitudinal studies at the National Heart, Lung, and Blood Institute including the Framingham Study. Drawing on this experience I am urging the various surveys at the Center to consider the possibility of linkage of the base information of the national surveys with followup procedures so as to establish longitudinal information which could be extremely useful in formulating health policy. At a minimum, insofar as our confidentiality statutes permit and insofar as it does not hamper recruitment for the various surveys, we expect to link the surveys with the National Death Index which has proven to be an extremely useful research tool.

One initiative that we are pursuing with great vigor can be summed up with the word "automation." I am convinced that the technological changes in the computer field will have great impact on our own operations and on that of other statistical agencies in solving the universal problem of the time lag between collection and release of data. When this is coupled with expected improvements in the accuracy of tabulations, and improvement in the publication process, I am sure that the initial heavy investment will be well worth it in the long run. Related to the area of computer automation we are also conducting feasibility studies in related areas, such as random digit telephone dialing, to reduce the cost of various surveys without impairing the validity of the data collected. If these feasibility studies indicate that there will be no loss of validity of the data, we are hopeful that these techniques can also be used by other agencies at the State and local level to meet their growing needs for small-area data.

After a period of inactivity, the National Committee on Vital and Health Statistics was reconstituted earlier this year to serve as an advisory group to the Secretary for the Department of Health and Human Services and the Assistant Secretary for Health. The full Committee has already met three times this year and just yesterday heard reports from three newly established Subcommittees which dealt with three important areas of current interest. One is dealing with problems involved with disease classification and automated coding of medical diagnoses, including the beginning steps for the next revision of the International Classification of Diseases, and a consideration of the Diagnostic Related Groups to be used for Medicare reimbursement effective October 1. A second Subcommittee is pursuing vigorously the issues related to the establishment of uniform minimum health data sets which HHS would like to institute in its various data collection systems; and finally a Subcommittee on the vital statistics cooperative program is guiding us in NCHS in establishing more effective and collegial relations with the State vital statistics systems.

Probably one of the most important pieces of evidence of the vigor of the National Center for Health Statistics and health statistics activities in the country is your attendance at this meeting. You have come to this meeting in record numbers--more than 900 people were preregistered and we are expecting over 1,000 to attend in total. You have all come to this meeting to share your knowledge and experience and your perceptions of the priorities and systems necessary to meet health statistics requirements for the 1980s.

Some of you are producers of health information. Some of you represent the wide community of data users--universities, health care providers, health planners, businesses, and government agencies. Whatever our backgrounds we are united in a common concern for the range and quality of information available for assessing health in this country and the steps we can take in our own jobs to assure the availability of timely, complete, and accurate data to meet the various needs of our programs.

At a time of tight budgets it is more essential than ever that all those involved in collecting, analyzing, and using health data work together and establish close communication. We need cross-fertilization, cooperation, and sharing if we are to make the best use of our available resources. That is the important purpose of this meeting. Today all of us who produce statistical data have to be keenly aware of the relative priorities and demands that are placed upon us. This requirement has given us the theme for this meeting--Priorities in Health Statistics--which will form the basis of our discussions and exchanges for the next three days.

You should be aware that nearly all of the papers to be presented in the concurrent sessions of this Conference are contributed papers. Last winter, in preparation for this meeting, NCHS sent out a call for papers. The response was an outpouring--we received more than 300 abstracts. The staff of the Center reviewed all the abstracts and made difficult choices in order to obtain the approximately 80 papers that will be presented at the Conference. We would have liked very much to use all of the papers that were submitted, but we have only three days. The response that we had, however, I take as a measure of the enthusiasm and vigor of the health information enterprises in this country. I think you will find some breakdown of the contributors to be interesting. Of the 80 papers to be presented, about one-fourth of them were prepared by personnel of State and local agencies, primarily health agencies. Another quarter of the contributors are associated with universities, medical schools and various survey research centers. Another quarter of the contributors are personnel with various Federal agencies. And the final group is made up of a variety of representatives from voluntary organizations, businesses, and hospitals.

Let us now turn to a brief overview of the organization of the meeting during the next three days. Each day of the Conference will highlight a different area of high priority for health data collection and information needs.

Both the plenary session and all of the concurrent sessions on a given day will be devoted to one priority area. Today, the priority area is data needs for health promotion and disease prevention. Health promotion and disease prevention is one of the highest priorities of the Department of Health and Human Services. Working with a coalition of many organizations, some 200 specific objectives have been set as national goals for achievement by the year 1990. Federal, State, and local governments, employers, unions, schools, voluntary agencies, and many others are involved in education and informational programs designed to help people achieve the behavioral changes needed for healthful living styles, and thereby to reduce morbidity and mortality for a wide sector of diseases. The data needs to track and evaluate our progress in achieving these goals will be the basis for discussions and presentations this morning and this afternoon.

For Tuesday, the priority area to be discussed is data on health care costs. This priority reflects the national concern about the steep and sustained rise in the costs of medical care. This fall we will see the implementation of various measures intended to contain the rise in costs of Medicare and Medicaid. This will add a new challenge for all those involved in developing and analyzing data in the health costs area.

For Wednesday, the priority area to be considered is the monitoring and evaluation of health service programs. This area encompasses data needs for administration and accountability and is increasingly important in guiding us in establishing priorities under tight budgets.

There is one other organizational aspect during the next three days which you should note. Papers in the concurrent sessions are grouped according to three tracks: statistical program administration, methodology, and data use and analysis. You may follow one track throughout a day or over the entire three days or you can mix and match as you please.

The program also includes three special sessions. Late Tuesday afternoon, we will hold the traditional afternoon session with the staff of the National Center for Health Statistics. This is an informal session at which we will attempt to answer any and all questions you may have about our various programs. On Wednesday, two of our fellow agencies from HHS will hold special sessions: staff of the National Institute of

Mental Health will present a session in the morning on statistical perspectives on the U.S. mental health service delivery system; in the afternoon, the staff of the Health Care Financing Administration will hold a session on sources and uses of data for evaluating Medicare and Medicaid reimbursement, eligibility, and coverage demonstrations and experiments. These sessions are designed to meet some of the specific needs many of you have expressed. All Conference participants are invited to attend and participate in these sessions.

This Conference, like all the others held over the past 40 years, blends talent, knowledge, and experience of people in many different jobs in many different parts of the country. It offers one of the few opportunities that we have to get together with many of our colleagues in health statistics and to learn from their experiences. I hope that each of you will benefit from the sessions of the next few days, and that each of you will also contribute to the discussions. The priorities that we must address are a tough challenge; by learning from each other we will better be able to meet them.

PRIORITIES IN PUBLIC HEALTH DATA

Edward N. Brandt, Jr., M.D., Assistant Secretary for Health
U.S. Department of Health and Human Services

I'm pleased to be your keynote speaker this morning. I feel quite honored, in fact, since after me comes a full program of excellent people dealing with every important issue in the world of public health records and statistics. This year's program is clearly as important to the field of public health as programs in the past have been. And maybe more so.

At this point, I would like to congratulate the Conference Committee for putting together a program that is serious, substantive, and stimulating. The papers--and the people giving them--are sure to make a solid contribution this week to the future course of health statistics. I'm impressed.

As a biostatistician myself, I read a program like this and I wonder why I strayed so far from my computer. Where did I go wrong? There's so much good work to be done and so many good people with whom to share that work.

For this morning, I've been asked to say just a few words about the "Priorities" in health data...what they seem to be and what the role of government might be in responding to those authorities...and then I want to ask something of you. But more of that in a moment.

First, the priorities.

Of a number of priorities that could be discussed, I want to bring to your attention this morning three priorities in particular. I really can't put them in any rational descending order of importance, since, in my judgment, each one is as important as the other. But let me discuss the turnaround priority first, because it's a relatively simple one to state. Also, I'm sure that everybody in this room has "turnaround" somewhere on his or her personal list of data priorities.

Before the advent of the computer, it was common for the public health worker or the services manager to be working with published data that was 7 to 10 years old. It wasn't the best situation, but then we didn't seem to need data that was much more timely.

The computer is changing all that. With its help, we have reduced lag time to the current 30 to 36 months. But that's still not good enough. A base year of 1980 is simply inadequate for public health planning in 1983 for the years 1985 through 1988. When you recall the many important medical and health care changes as well as the profound economic and social changes that have occurred in the society since 1980, you have to say that 1980 is just not good enough. For today and for the future, it is an atypical year for planning.

But timeliness is not an end in itself. Timeliness of data provides the means by which we gain better understanding of the health needs of today's citizens and we achieve greater insight into the effectiveness--or the deficiencies--of our system of health care delivery.

Timeliness is also a key to helping us achieve one of the objectives of any good data system: comparability among as many data sets as possible. In public health, we still have a distance to travel before reaching that objective of general and broad comparability.

The time lag may be no more than a function of our culture...that is, our human work systems are still not organized to take full advantage of the electronic systems humming all around us. Or we may not be as agressive as we ought to be in pushing the computer to carry more of the processing load and, by so doing, liberating people to exercise their brainpower with more effect...in a shorter space of time...at the most critical points.

Is progress possible on this priority? I believe it is and in the forefront will be the National Center for Health Statistics.

Staff members of the NCHS know a great deal about the issue of turnaround time. For many years they have taken the brunt of the criticism, whether they deserved it or not. But I am pleased to recognize the efforts they are making--and will continue to make--to further modernize many of the Center's activities in order to put the turnaround issue behind them. I know this is a high priority issue with Dr. Feinleib, also, and he has my complete support.

Let me just mention two areas where we hope to see some progress fairly soon.

In the area of vital statistics, our problem has been that, at the Federal level, we can go no faster than the slowest State. However, that argument, if it had any validity, was only valid for the past, not the present and certainly not for the future. The pace of collecting vital statistics has been accelerating at every level of government. Hence, our explanation for yesterday's long lag time will not be good tomorrow. In fact, I would say that if we do not focus more attention on the problem, it might well be the Federal level that becomes the slowest level of all.

The record until now certainly points in that direction. Over the past few years, automation was introduced into the National Center on a piecemeal basis. And, in many instances, it was just as well that the NCHS did not plunge into the purchase of an expensive, agencywide, but quickly obsolescent system. But that was yesterday. Today, the state of the art is such that we are able to--and we must--speed up the degree of automation throughout the Center and go at the problem with an overall, cohesive plan.

One of the first areas to benefit would be vital statistics. The lag time for publication of vital statistics has been about 3 years. This year we are doing a little better, closer to a gap of about 30 months. However, by the end of 1984 we intend to have reduced the lag time on vital statistics down to 12 to 15 months. That is, final data for calendar year 1983 will be available in January or February 1985. That's our goal and our timetable for reaching it. Dr. Feinleib and I agree that it's a realistic goal and we are working to achieve it.

A related effort focuses on that important work..."available." We may have improved the collecting and manipulating of data, but what's the use of that, if it all bogs down in the publications process?

We hope to eliminate that problem in 1984, when our "front-end computer publication capability" will be fully operational. This capability will allow us to generate final computer tabulations that are camera-ready for the printer. As many of you know, particularly those of you who shepherd publications through the Government Printing Office, the ability to generate camera-ready computer runs will cut many months off the timetable of available data. For some publications the timetable may be shortened by as much as a year.

These efforts in vital statistics and in data publication should do much in the years to come to lower our anxiety levels about turnaround time at the National Center for Health Statistics. It will still be a priority, but I believe it will generate more data and less worry.

The second priority I want to share this morning concerns a need to develop at the State, county, and municipal levels the capacity of gathering and using "small-area data." The term "small-area" is meant to cover much more than vital statistics. Nearly every jurisdiction already does pretty well on those or is well on the way to improvement. The "small-area data" I have in mind would give a more three-dimensional picture of the delivery and use of specific health services.

Such data would yield the following kinds of information:

- The status of prenatal care, the priority medical needs in perinatal and maternal care, and predictive models of delivery and birth outcomes in a given geographical area...
- Also available would be hard indicators of the prevalence of communicable diseases among specific age, sex, occupational, and zip code groups...
- "Small-area data" can substantially illuminate the impact on local residents of certain discreet environmental events, such as severe heat or cold, or the improvement or the deterioration in the quality of local drinking water, or the dumping nearby and the subsequent clean-up of toxic wastes...
- And such data also reveal, tally by tally, the emergence of long-term challenges for local health service, such as an expanding population of elderly residents who have comparatively weak health profiles or a rising curve of alcoholism among school-age children.

These kinds of data are available on a national level but in varying depth. And, to recall the first priority I mentioned a moment ago, the collections are done by different agencies whose goals, objectives, and timetables are not always consistent, one with another. Let's take drunk driving as just one example.

Each year an estimated 25,000 Americans are killed in alcohol-related highway accidents. Another 700,000 are injured. It is the leading cause of death among young people 16 to 24 years of age. Drunk driving is one of this country's leading epidemics affecting the public health. But where do these statistics come from? How good are they? And can they be any better?

The Department of Transportation is one main data source. The DOT's "fatal accident reporting system" aggregates State data giving the age and sex of the decedents, the causes of accidents to the extent they're known, the types of vehicles involved, and so on. The role of alcohol is deduced from the reports of blood-level tests of the drivers involved in the accidents. Most States routinely do the BAC testing.

This past January, DOT published a report titled "Alcohol in Fatal Accidents." The data for the DOT study were drawn from highway fatality reports out of 29 States. This DOT report has been helpful in clarifying the magnitude of the problem of drunk driving in the United States. It's very useful for policymakers and decisionmakers at the national level.

As good as the report is, however, it is nevertheless limited by the uneven nature of reporting at the State level. The aggregated FARS data give an indication of the problem...they do not provide the sought-after final statistical definition.

Unfortunately, we do not have any compatible, complementary set of data elsewhere to use as confirmation. The NCHS mortality data give accidental deaths involving motor vehicles, but these may be off-the-road accidents, they include such specialty vehicles as earth movers, and the data may or may not include the blood alcohol concentration levels in the victims.

Also, the FARS reports have a 30-day limit. That is, the system will carry all fatalities that occur within 30 days of the accident. But the NCHS system, based upon death certificates, has no such limitation. In addition, and probably of most significance, is the fact that the NCHS data are about 3 years behind FARS. So for quite different reasons, neither FARS nor NCHS can give us high quality information on this important public health problem...the impact of drunk driving upon the Nation's health status.

At the Federal level, we can move ahead with a number of improvements, but the entire system ultimately rests on the depth and the integrity of the data systems at all levels of government. In the example of drunk driving, DOT cannot produce a high quality FARS report nor can NCHS produce an indepth mortality report unless we receive uniformly high quality reports from State and local safety and public health authorities. But those kinds of reports will only be written when each State and locality sets as its own goal the development of quality "small-area data."

One more point. Until 1981, the Federal Government played the major role in health service delivery. At the time President Reagan took the oath of office, the Public Health Service alone was operating 69 grant-in-aid service delivery programs. Despite the unevenness and inadequacy of our data at the Federal level, we nevertheless had to make some statistical assumptions to support all our program and budgetary planning.

These "macro"-statistics may have been useful on the theoretical or intellectual level, but they were difficult--often impossible--to relate to the local, individual, human level...the level at which services were needed and used.

It was a priority of this Administration to change that. In August 1981, the Congress agreed to and authorized the block grant program. As a result, during the past 2 years, the center of gravity in health service delivery has shifted from Washington out to the States and territories.

In a sense, the States now have the same problem the Federal Government had. That is, they need quality data to provide the underpinning of all their program planning and budgeting. But unlike the Federal Government, the States and communities can have direct and continuous access to the primary source of all public health data: the people themselves. In addition, they have this access within the manageable confines of small areas.

"Small-area data" are more than vital statistics...that point is critical to understand. Because the universe may be small does not mean that the data should lack scope or depth. To return again to my example of drunk driving, a truly comprehensive look at the problem would require statistics from FARS, from death certificates, from emergency services personnel, from the schools and colleges in the area, and possibly from the automobile dealers, too. With these kinds of data in hand, it might be possible for us to come much closer to a true assessment of the risk posed by drunk drivers and a clearer sense of what needs to be done about it.

Is this challenge regarding "small-area data" purely theoretical, or is it something we can move on today? In point of fact, we not only can move on it today, but a good start has already been made by experts drawn from all levels of government. I am referring to the "model standards for community preventive health services."

This exemplary effort was begun almost 7 years ago. Then, after more than 3 years of planning and development, the "model standards" were published, a truly "collaborative project" that brought together the United States Conference of Local Health Officers, the National Association of County Health Officials, the Association of State and Territorial Health Officials, the American Public Health Association, and the U.S. Public Health Service--principally the Centers for Disease Control.

The "model standards" were offered to State and local health officials as a companion to Healthy People, which was published at the same time, the fall of 1979. Healthy People, you may remember, brought together in one document the rationale for the health promotion and disease prevention strategy adopted by the Public Health Service.

It was thought that, with the help of the proposed model standards, health officials could begin shaping their service delivery programs and budgets according to a set of specific, desirable, and attainable public health goals for their individual States and localities. Those goals, in turn, would have been based upon the ones published in the parent document, Healthy People.

These standards are at once both the stimulus and the beneficiary of improved systems for collecting and analyzing small-area data. Responsible health agencies have seen the value of developing data systems with significantly more community breadth and depth than usual...systems that can deliver more specific and more localized epidemiological data.

I understand that this has been the experience in the cities of Seattle and Birmingham and in the States of Tennessee, Utah, and Pennsylvania, where the objectives of the Surgeon General's report have been adopted and the community model standards are being applied to measure program performance, effective cost levels, and health status outcomes.

With four years of experience behind us, CDC will reconvene the collaborative work group--representatives of the four professional organizations and the Public Health Service--in order to review the standards, revise them where needed, and see what else should be done to make them more acceptable and useful to local and State health officials. The new group will come together next month and will be working on the project for about a year. I feel confident that their efforts will contribute significantly to our collective ability to develop and use quality small-area data.

And that brings me to the third priority I want to discuss with you this morning.

Many years ago, when life was much simpler--and our methods of data collection were simpler, too--health officials could be satisfied with a few measurements related to the health status of the population they served. Much was suspected but little was known about the relationships between housing and health...employment and health...diet and health...or recreational behavior and health.

- Those days are well behind us. Our experiences with lead-based paint and urea-formaldehyde foam insulation, for example, have taught us a great deal about the impact of housing upon health.
- The widely discussed diseases of "black lung" and asbestosis are just two of a large number of job-related health risks of special concern to public health personnel, employee associations, management, and all levels of government.
- And anyone engaged in providing emergency services can tell you what happens in a community following the sudden popularity of motorbikes and dirt bikes, sailboards and surfboards, raquetball, hang-gliding, and rafting.

The life of our society is organic. You can't totally isolate one aspect of society's experience from all the others. There are too many important intersections with other aspects of social experience. We must therefore recognize and accept the interrelated, organic nature of our social experience and devise our data-gathering systems to reflect it.

That's certainly not the easiest method to choose. An interrelated data system will no doubt multiply by several-fold most of the problems I've mentioned thus far:

- Better data technology must be aplied to more--and more different--sets of data in order to achieve quicker turnaround across the board.

- o The negotiation process for achieving comparability among dissimilar data sets becomes much more complex.
- o And the issues of confidentiality and the protection of proprietary data become more difficult to resolve.

If I were asked to identify a few of the toughest challenges to the field of public health records and statistics, I would definitely include among them the development of an organic system of data that presents health information within a context of the total life of the population.

This is not a brand new idea. Granted, we have very few examples to show off, but we do have some. And one is right in our own Department of Health and Human Services.

At the request of the Secretary, we have organized a Departmentwide "Health Information Policy Council." Included in the charter of this Council is the charge to plan "for the development of comprehensive, Departmentwide health information systems which meet policy and program needs." This means expanding the diet of the HHS statistics professional to absorb data from the Social Security System, the child protection services, the Administration on Aging, Medicare and Medicaid, and others.

Meanwhile, the Public Health Service itself is taking an inventory of the many health-related data systems now functioning in the private sector. Most of these serve particular professional interests or are specific to certain medical or social health problems. I would hope that such an inventory will provide us with the clues for ways to integrate the data from different sources into a broadened, more dimensional picture not only of health itself but of health in the context of our country's total national life.

The next few years should be very exciting ones for those of us involved in developing better ways to collect and use records and data in health and medicine. The only problem I see is not one of resources or abilities. It is the age-old problem of purpose.

Why bother? Why improve our data technologies? We solve any of the problems that are on the agenda of this Conference? For an answer, I will recall the words of Thomas J. Watson, Jr., written back in 1960, when he was directing the fortunes of IBM. Tom Watson's concern was "technological change" and he raised the same question: Why bother with it? His answer was simple, direct, and worthy of repetition:

> "Its total purpose (is) the benefit of mankind. We cannot afford to let the interest and excitement of the process distract us from its main purpose, the improvement of man."

In our field of records and statistics, we are easily distracted by the bells and whistles of the process. But we must shake ourselves awake and return to the real purpose of our work. And that should be, as Tom Watson advised, "the improvement of man."
Thank you.

TRANSLATING OBJECTIVES INTO REALITY: DATA NEEDS TO MEET THE NATIONAL OBJECTIVES

Hugh H. Tilson, Burroughs Wellcome Co.

Perhaps the most exciting part of this symposium is its point of departure - having the courage to continue to dream in the face of some pretty tough fiscal realities facing public health for the 1980s and beyond. But without the vision - even when times are tough - things will never get better; and without looking ahead, the scale-down priorities and activities - if scaling down in fact be necessary - will never be based upon "where we are going" rather than the opposite - making budget cuts on the basis of tradition and "where we have been." But even then, translating these lofty and worthwhile national objectives into an action plan for the nation must involve every one of the building blocks of our nation's public health system - the over 3,000 local health departments, health districts, and other health jurisdictions in the country who, working together, represent the hope for achieving these objectives or, perhaps said equally well, without the involvement of which the objectives are only hollow rhetoric. Yet translating these objectives into a realistic work plan for an individual community requires far more work than has been done as of yet. Therefore it is especially desirable that we take a moment to take a look at the one major effort currently on the table which may shed some light on the work which needs to be done in the next five years in order to get this critically important translational job done in a way that will let us know once it happens that it has happened and where we stand so that we may then plan where to go from there.

The work which I have the privilege of discussing today has taken over 5 years to bring from a twinkle in the eye of several of us to the current state of readiness upon which I am happy to report today. Specifically, I refer to an effort which started many years before but was finally brought to a critical mass by the Director of the Centers for Disease Control in 1976. A coalition, constituted by the leadership of the National Association of County Health Officials, U.S. Conference of City Health Officers, Association of State and Territorial Health Officers, The American Public Health Association, and the Public Health Service, was possible because of a shared sentiment that someone needed to get on with translating a related dream into an important reality. The dream was that it would be possible to develop standards by which every community could know what to expect of itself in the way of community preventive health services without sacrificing the wonderful flexibility which characterizes the heterogeneity of public health in America as we know it today. This effort was quickly recognized by Congress (I guess it goes without saying that that was not entirely accidental) in the language for the Joint Conference Committee reconciliation of the Health Services Extension Act of 1977:

> "The Secretary of Health, Education, and Welfare, in cooperation with appropriate professional entities and individuals, shall within two years of the date of the enactment of this section (1) establish model standards with respect to preventive health services in communities, and (2) report such standards to the Congress. Such standards shall be developed to identify populations in need of preventive or protective health services and to maintain community-oriented preventive health programs."

We could surely spend hours around these tables haggling over one standard or another for virtually every program that anybody here has ever studied or even read about. That is one of the reasons that we never got on as a nation to developing standards. Every discipline, every jurisdiction, every funding program, every separate category has had not just one but dozens of special ways of describing what it does. The result has been model proposal after model proposal for model data system after model data system which has left us all not modeled, but muddled, and in the process has blocked our progress to what, after all, should have been a relatively easy consensus exercise.

The working group assembled in 1976 at the CDC under Congressional mandate worked for a succeeding two years to bring these concepts to that level of consensus. Let me hasten to point out that it was a modest effort - a fledgling first step, if you will - at model standards building. But, indeed, it was a first step. I hope that you will agree that it was in the right direction. Basically, the document embodies three critical concepts.

First, the group realized that it was in error in thinking that it knew what a standard was. My guess is that there are many of you who think you know what standards are. I propose to disabuse you of that thought in just a moment. Second, the group had to come to grips with just who the standards really were for - a local health department, a constituency, a program, or a community. More of that in just a minute, too. And, finally, the standards working group had to develop a strategy by which whatever standards were established could be planned, measured, assessed, and replanned, lest this set of standards be yet another of the myriad documents which those of us with the best of intentions have developed over the years only to have them sit on the dusty shelves as monuments to our thoughts, rather than our deeds.

First then, we wrestled with what is a standard? Webster's definition of a standard is (paraphrased) a description of the way things should be. We determined that every community needed an inventory of kinds of services which had the right to expect of itself and proceeded

to array a list of 26 general areas (Figure 1). Digging deeper we concluded that such a statement - to be understood and meaningful - needed an inventory of measures by which the community could know exactly what it was within these categories that it might expect of itself. The work group committed to building into each standard at least one statement of a health outcome . . . that is to say a health status measure for the community that would help to bridge the usual service delivery and mechanism orthodoxy with the purpose of undertaking the activity. Thus, you will hear about outcome "standards" and process "standards" as statements of relevant conceptual-level objectives within each standard, as illustrated by a representative page from the Standards (Figure 2).

Let me take a moment to digress into one area which will be of particular interest to many of the participants in this audience - the question of errors of the first and second kind. There are several standards not in the collection which many of us at the time, and certainly many more subsequently, felt should have been included. Several were not because we just couldn't get around to them. Several others were not for political or strategic reasons. Notable in the latter category is a standard on mental health services. No one would deny that those are central for community level programming and that good mental health programming involves a heavy emphasis upon prevention. Thus, by anybody's definition they would qualify as community preventive health services and in order to help to achieve national objectives, establishing standards for those services certainly seems appropriate. However, because of political separatism of some mental health professionals from the rest of public health, including an organizational and bureaucratic separatism at federal, state, and local levels, and because mental health spokespersons made it clear at the time that they were interested in developing their own model standards for community preventive mental health services, our work group deferred to the latter effort. In retrospect, I think this was an error. Independent of the quality and quantity of any mental health "standards" which existed before and have been developed since, the establishment of mental health as a legitimate, important member in the family of community preventive health services seems worthwhile and the establishment of standards in this simple and congruent format seems desirable.

This seems especially appropriate in light of the second contribution - would it be self-aggradizing to call it "major contribution?" - of the standards development effort. In trying to conceive of the most useful way to state community level preventive expectations for the purposes of getting on with the national agenda, the work group determined that its very mandate gave it an insight into what may have been the compelling reason for the lack of substantial standards for community preventive health services up to then. Specifically, public health people tend to think in terms of the bricks and mortar of the health department. Yet all studies of local public health departments, including an extremely noteworthy one which was seeing the light of day just at the time of the standards effort under the direction of Dr. Arden Miller at the University of North Carolina, looking at exemplary local health departments, have concluded if nothing else at least one profound truth about local public health: every local health department is quite unique based upon the unique situation in the community whose health it is trying to preserve, promote, and protect. And in this uniqueness, then, comes the key to the second concept - the "AGPALL" (a governmental presence at the local level). The notion is that it is the community which is the frame of reference and not the department of public health. Whether the health department does or doesn't provide one service or another needs to be driven by the presence in that community of alternative methods of achieving the same objectives and the relative importance of those objectives compared to other objectives which the department or some other agency could also be pursuing. Relating this back to comments concerning a standard for mental health, what is apparent is that every community needs a program in the prevention of psychiatric problems. Whether this program is operated all or in part by a separate department of mental health, whether public health nurses have something to do with promoting the mental health, or whether the health department becomes the primary provider of services depends upon the constellation of activities, the perspectives, and the priorities in the local community. But utterly unacceptable would be failure to have considered this as an area for legitimate community pursuit. Thus, it seems appropriate that a standard for community preventive health services in this and many areas would be developed. And, thus, those who know full well that the local health department in Merced, California, is too small to be able to afford its own separate air quality control program and who, therefore, might have been skeptical about the inclusion of an air quality controls standard in the inventory of standards necessary for every community will be surprised to know that Merced met this standard in field tests. This translation in Merced of this standard had to do with the presence within the environmental health program of the local health department of a liaison officer, part of whose duties were to relate to the State Department of Air Quality Control and register air pollution alerts with them. In the process, however, the local health department acquitted itself of the responsibility to be sure that the community was served by a community-wide air quality control protection effort. Indeed, many standards were met with a minimum of actual activity by the local health department and the AGPALL worked by involving other responsible public and private actors in the achievement of the community's objectives.

If I've given you an example of a standard which probably should have been included, it would be only fair to give you an example of one which not everybody was crazy about including. The standard on primary care was controversial to say the least. Indeed the work group was split right down the middle and in its first vote on which standards to be developed defeated the

motion to include a standard on primary care. Here the sentiment was that primary care was the province of private medicine, that inclusion might be inflammatory, and that some of the poor, small health departments simply couldn't afford to be involved in primary care medicine. As such democratic processes go, when the minority (those of us interested in primary care) lost the first vote, the biggest of us (since he is not in the audience, I won't name him) metaphorically stood on the table, pounded his shoe and said you people haven't been out there if you think that primary care isn't the concern of the local health authority (or should I say AGPALL). And so it went. Eventually a standard on primary care was developed and, here again, the forcing agrument was the AGPALL one - namely that while no health department could be expected to provide all primary care for all people in the community (although I guess if John Sbabaro from Denver were here he might say yes, but we tried it in the halcyon days of that wonderful system, and I might say something about Project Health in Portland and its rather ambitious objectives), the converse is nevertheless equally unacceptable. The truth is that the absence of a good primary medical care system (and for that matter through that system access to a backup, secondary, and tertiary care system, generally located elsewhere except in the biggest of the communities) a community could hardly consider itself to be adequately served for preventive health services. And in those instances in which the private medical marketplace is inadequate to generate the primary health care system with full accessibility, availability, and acceptability for all persons, the governmental presence at the local level in instance after instance has determined that it is, in fact, necessary to put some of its scarce preventive resources into addressing this high priority area. This may mean an investment as modest as participating in a local health planning agency or as ambitious as a government financed, third-party reimbursement system to supplement Medicaid and build the capacity for community health centers. And these, then, are depicted by the various levels of specificity in the standard for primary care, the first page of which is reproduced as Figure 3.

The third unique feature of this standards effort is the negotiation concept. Like the other two, it attempts not to rediscover the wheel, but rather th redefine it. Every community, by this notion, once it establishes its own goals and objectives and translates them into outcome and process statements, and once it establishes a focal point for health - the AGPALL - needs as the third element a balance wheel, a process by which it can help itself to hold itself accountable. The role for the state government in health is to negotiate with the governmental presence at the local level on a periodic (generally annual or biannual) basis the appropriate targets for that community. An outside or neutral negotiation process assures that the community will not only look objectively internally at its own competing priorities and its interim achievement of objectives (and those of us who have done expedient planning at the local level recognize the need for that kind of internal accountability holding!) but also that the community's level of commitment to establishing and accomplishing health objectives is appropriate given the overall commitments of other communities within and ultimately aggregated across the entire state. This, then, relates back to the comments of Dr. Brandt. In order for us to achieve our health promotion and disease prevention objectives as a nation by 1990, each state needs to make its commitment to those objectives. Several states, including California and Tennessee, have already made that commitment through fairly lengthy and scholarly documents. Part of that commitment must, of course, be a commitment for implementation in every community across the state as appropriate. Whether that commitment comes from the community to the state or the state to the community depends upon the level of prior commitment, level of communication, and point in the planning cycle. What is clear is that if every community has a standard and objectives in the area of each health promotion objective, then the state should be quite aware of what it is reasonable, given the current commitment, to expect as productivity in the aggregate from all communities in the state. And if this falls short of what the state expects "from itself," then a process of negotiation by which to advocate for additional output as appropriate needs to be initiated. I can already see the raised eyebrows around the room, especially among those who have gone out telling local communities that they needed to do more with less in this era of budget cuts. It's not popular for a state agency to go into a local community and tell them how remiss they are being not doing their "fair share of the state's work." My response as a local health officer to that kind of pep talk always use to be "if you've got the money, I've got the time." What the negotiation process built into this set of model community preventive health standards is meant to achieve, therefore, is not to wring water from a stone, but rather to allow all actors in the health enterprise a better "handle" on community level priorities and trade-offs; or conversely, the commitment to these objectives will not be made unless either new resources may be brought to bear or, lacking these, old resources traded off.

But of course negotiation is appropriate for decreases as well as increases in the level of one or, in fact, all objectives. If as the reality of block grants in the absence of carryover categorical funds and the presence of a tight, though recovering, economy forces the issue of reduction in funding level for government programs in general, including public health programs (however meritorious those might be of an exemption from such cuts), then again a process of negotiation which allows the community to see its contribution in perspective with contributions and needs statewide is eminently appropriate.

Of course when the document was written these were more than mere concepts, because bits and snatches of flexible outcome oriented standards, examples of the AGPALL, and plenty of precedents for state/community negotiation were

already in place. Nevertheless, putting them together into a single document and arraying a new set of standards upon this framework was new. It required field testing, first as ideas through the relative experts and opinion leaders nationwide, and then as tools through appropriate state and local agencies. These were both done before actual publication of the document and the concepts and contents were found adequate to the task. Subsequent to publication, the Centers for Disease Control determined to do intensive field feasibility testing under contract. Contracts were awarded to the states of California and Maine during 1980 and 1982. The results of this field test have been published by Phil Weiler et al. in The American Journal of Public Health (November, 1982) and submitted as Final Reports of the Centers for Disease Control last year. They provide fascinating reading and rather than attempting to summarize them here, let me rather commend them to you. I would, however, like to touch on three valuable lessons learned from the California experience and one from the Maine experience, the second first.

After contracting to field test the standards in the state of Maine, the state came to the realization that given its current state/local political relationships, the lack of a strong local health department system in Maine, and problems with internal accountability, the AGPALL concept central to the negotiation process was simply not at a level where the conditions of the contract could be met. This is a critical lesson and one which must not be finessed with a simple "Oh, yes, of course but that's only in those states that don't have local health departments and/or strong regional state operated health departments." The problem is that as long as there are any states which qualify in this category, it is not the feasibility of standards or the appropriateness of the AGPALL concept which ought to be worrisome, it is rather the more fundamental implementation feasibility for programs of national scope, such as our Objectives for the Nation or any national contribution to Health for All by the Year 2000. It was a lesson that we learned during the gearing up for the swine flu immunization program - an effort which has been undervalued for the important lessons learned. One of the important lessons from the swine flu preparation was that we, as a nation, could not count on a significant coordinating force for public health in every community in our nation and that we could not afford not to have such a force.

From the California experience came three very different conclusions.

First, it is clear that negotiation is not only feasible but desirable. If there is no other product to getting state level people together with local level people it is the realization that we are colleagues not adversaries in public health striving toward shared goals, call it Hawthorne effect or, you'll pardon my borrowing from my current role as a pharmaceutical epidemiologist, placebo effect. It is nevertheless true that the sharing of efforts developed a spirit of cooperation not previously described. Of course, there were points of friction and breakdowns in communication and even negotiations, points of disagreement, and inevitably, costs. And so the conclusion is good government isn't easy and isn't free - but it's probably worth it!

Second, we learned something about the AGPALL - most particularly that even though the local health department might not be specifically responsible for the area under negotiation, it knew plenty about the area and/or was plenty interested and able to gain information about the area from its community.

And, third, we learned something about the data base to support the establishment and evaluation of standards; namely that much to everyone's surprise, there was a great deal of data around at the local health department level and while it wasn't always exactly on the money and/or not every data point that might be envisioned in the standards is available in every community in which the standards have been tested, much less those less fortunately, more remote, less progressively, or whatever, nevertheless the negotiation of a congruent, consonant objective-oriented data base against which to evaluate and document our progress is much more than the twinkle in the eye of the director of a state center for health statistics.

Based on this, California has had the temerity now to articulate objectives for 1990 based upon the nation's objectives for health promotion (need I point out that the document was generated by a state center for health statistics or comment Kay Moser and Sheila Dumbauld for their outstanding work in this translation of national priorities?) and the process by which these are linked to individual community standards (outcome and process objectives) already underway.

My colleagues from North Carolina will, I know, be disappointed if I don't at least wave a hand at the outstanding work which they have done. I see Charlie Rothwell, founder and current Director of the North Carolina State Center for Health Statistics, in the audience. In North Carolina, standards are linked to outcomes and processes; they are negotiated annually (actually on a 3-year recurring cycle); and the State Center for Health Statistics generates regular updates in preparation for an annual negotiated block grant contract between state and local health department of a list of key health indicators which is a major data subset of the data points required by the standards. The issue of standards in North Carolina continues to be a matter of hot debate - and plenty of matters remain to be resolved. And North Carolina is only one example. Perhaps there are those here from one of the several other states currently trying out the concepts - Utah, Oregon, Ohio, Illinois, Maryland, among others.

I would be remiss not to point out that the Model Standards document also envisions a federal role. As you have heard and will no doubt continue to hear, our nation is experiencing a "new federalism" - a rediscovery that states have

mentalities which merit respect and support and that the federal government does not need to do and say everything which, in health, in many people's eyes, is primarily a state responsibility and privilege. In the process, however, there are many roles which must be occupied by the federal government which must not be lost in the presence of a creative transfer of funding and oversight back to the states. Specifically, the federal government must help our nation to see national dreams. Hence, the objectives for the 1990s and the commitment to the American leadership in Health for All by the Year 2000 are national priorities appropriately supported by the governmental presence at the national level, the AGPAFL, if you will. Second, someone needs to help the state negotiator and arbitrator to hold the local agency accountable and to understand the vision and commitment in state role itself. This is then a convener and opinion leader role - one which must not be lost at the national level. And, oh yes, there is always the residual guarantor problem. Specifically if there are jobs to be done and no resources at the state and local level to do them, it is neither inappropriate nor undesirable for the federal government to show the leadership by putting its money where its mouth is and help to foot the bill for urgent national priorities. Block grants are wonderful. They just need to be funded and accounted for! Standards provide a logical framework for this agenda.

Where do we go from here? First, we need to know a lot more than we currently do about problems and progress in the development and use of standards by various states around the nation. You should know that the Health Administration Section of the American Public Health Association has put together a nationwide survey of states which will be ongoing and which will evaluate the states' progresses toward model standards. In a related effort, the Centers for Disease Control has developed, with full collaboration from state and local health officials and the APHA, a national survey to be conducted in the fall (pending OMB approval) designed to tap the experience and attitude of local health departments with regard to standards as they are currently used nationwide.

I am excited to be able to commend Dr. Brandt on his vision in supporting Bill Foege's proposal to reconvene the working group which developed the original model standards to review, 5 years later, the standards. While the scope of work of the contract, just announced as having been awarded to the American Public Health Association, is complex, there are three basic thrusts of the work to be undertaken this year: first the surveys regarding current status of standards are to be reviewed and evaluated to learn as much as we can from existing field experience and the standards, themselves, are to be evaluated both for clarity and feasibility; second the need for new standards is to be assessed and if any needed (I hereby nominate mental health in case anybody wondered), such a standard is to be developed. And Finally, the need and suggested methods for harnessing the national wisdom regarding targets, goals, and objectives (including those which we've just heard Dr. Brandt talk about) and tying them more directly to standards (e.g., using a specific referenced linkage strategy, standard-by-standard, data-point-by-data-point) is to be assessed. Of course any effort as ambitious and central to what public health is all about as this one will require the understanding and involvement of everyone in this room. I have some ideas about how you might be involved; and knowing many of you, I would guess that you have dozens for every one that I have. I'd enjoy discussing these with you either in the formal discussion or over the next three days.

Let me thank you, Mr. Chair, for the privilege of participating in this morning's program. I deeply believe that Model Standards for Community Preventive Health Services represent not a plaything for under-occupied bureaucrats but rather a major conceptual advance which may just help us to get where public health intuitively has always known it needs to go. Indeed, it may be the indispensable tool for the achievement of the objectives laid out by Dr. Brandt.

Thank you.

Figure 1.

PROGRAM AREAS FOR MODEL STANDARDS

Preamble
How To Use Standards
Administration and Supporting Services
Air Quality
Chronic Disease Control
Communicable Disease Control
  Immunization
  Sexually Transmitted Diseases
  Tuberculosis
Dental Health
Emergency Medical Services
Family Planning
Food Protection
Genetic Disease Control
Health Education
Home Health Services
Housing Services
Injury Control
Institutional Services
Maternal and Child Health
Noise Control
Nutritional Services
Occupational Health
Primary Care
Public Health Laboratory
Radiological Health
Safe Drinking Water
Sanitation
  Child Care Facilities
  Governmental and Nongovernmental Public Buildings
  Mobile Home Parks
  Recreational Areas
  Schools
School Health
Solid Waste Management
Surveillance/Epidemiology
Vector and Animal Control
Wastewater Management

Figure 2.

AREA: Communicable Disease Control (Immunization)

GOAL: There will be no cases of officially designated vaccine-preventable diseases in the community.

NOTES:

(1) The need for a system of surveillance, epidemiologic assessment, and containment of outbreaks of vaccine-preventable diseases is addressed by the general surveillance/epidemiology standard.

(2) This standard is intended to cover all immunizing agents (e.g., toxoids, not just vaccines).

FOCUS	OBJECTIVES	INDICATORS	POPULATION IN NEED
	OUTCOME		
	0-1. By 19____the incidence of__*__will not exceed_____, or the absence of disease will be maintained. * Insert name of each officially designated vaccine-preventable disease.	(a) Incidence (b) Secondary spread	The Community
	PROCESS		
Childhood Immunization Deterioration	P-1. By 19____and in each succeeding year, at least 90 percent of the 2-year-old population will have completed primary immunization for the officially designated vaccine-preventable diseases. CROSS-REFERENCE: MCH	Percent of 2-year-old population which has completed primary immunization	2-year olds
	P-2. By 19____ and for each succeeding year, all school enterers will have complied with one of the following alternatives: (a) 100 percent of primary and appropriate booster immunizations complete (b) A remedial course to bring immunizations up to 100 percent has been initiated and certified by an appropriate provider (c) Exemption for medical or religious reasons from immunization requirements. CROSS-REFERENCE: MCH SCHOOL HEALTH	Percent of school enterers complying with alternatives	School enterers

Figure 3.

AREA: Primary Care

GOAL: Residents of the community will have primary health care services to promote their achieving and maintaining optimal health status.

FOCUS	OBJECTIVES	INDICATORS	POPULATION IN NEED
	PROCESS		
Availability and Accessibility of Services	P-1. By 19____the official health agency or other appropriate governmental agency will, in the absence of the provision of minimum primary health care services in the community, provide such services directly or through purchase agreements (including arrangements for specialist and hospitalization referrals); in addition, this agency will supplement existing services where they are inadequate.	Availability and accessibility of ambulatory acute and chronic primary care services	The community
Promotion of Services	P-2. By 19____the official health agency or other appropriate governmental agency will promote the utilization of primary care services, including dental services, by the community. CROSS-REFERENCE: HEALTH EDUCATION	Existence of ongoing promotional effort	
Personal Preventive Services	P-3. By 19____the official health agency or other appropriate governmental agency will assure the availability and promote the utilization of personal preventive health services (e.g., childhood immunizations, family planning) as part of an integrated system of primary care services. CROSS-REFERENCE: CHRONIC DISEASE CONTROL COMMUNICABLE DISEASE CONTROL DENTAL HEALTH FAMILY PLANNING GENETIC DISEASE CONTROL HEALTH EDUCATION MCH NUTRITIONAL SERVICES	(a) Percent of residents for whom basic immunization series is completed (b) Percent of pregnant females who receive pre-natal care beginning in the first trimester (c) Percent of babies who have received basic infant care series of exams (d) Number of screening exams performed by categories such as: (1) Vision screening (2) Dental screening (3) Hearing screening (4) Lead poisoning (5) Hyptertension (6) Anemia (7) Etc. (e) Percent of high-risk screened individuals referred to diagnostic follow-up with documented follow-up examination (f) Number of prophylactic visits to dental hygenists	

PERSPECTIVES ON HEALTH PROMOTION

Lester Breslow, University of California at Los Angeles

Only in very recent years have we been able to devote any significant attention to health promotion.

Until almost the present time the overwhelming health problem of mankind has been to avoid premature death and the ravages of communicable disease. Focus on those aspects of health must still prevail in developing nations of the world. Malaria and schistosomiasis continue to shrivel the lives of millions around the globe. Even in the United States at the beginning of this century high infant mortality, pneumonia and influenza, tuberculosis and the diarrheal diseases dominated the health scene. Certain segments of the population in our country continue in the 1980s to suffer from excessive rates of many conditions long known to be preventable.

In the latter part of this twentieth century, however, the United States and other industrially developed nations are encountering a new kind of health problem. We no longer must contend so exclusively with the threats to health that have required action throughout history. People in several countries are living generally into the eighth and ninth decades of life, largely free of disease during most of that time.

It has therefore become possible to think about promoting health, not merely avoiding disease and premature death. The World Health Organization stimulated thinking about this matter with the definition of health it adopted in the late 1940s: "physical, mental and social well-being, not merely the absence of disease and infirmity." That concept has enabled us to consider a new health agenda. The latter certainly includes some holdover items from the previous agenda; we haven't completed some health tasks that have confronted mankind for centuries.

We can and should, however, turn to the current challenge: health promotion. It encompasses all measures that enhance the possibility of a full life, both in extent and quality. This must be essentially what the WHO had in mind.

A conference on health statistics is necessarily concerned with measurement. Thus a primary issue here is how to quantify health, as conceived by WHO and meaning a full life both in extent and quality. Some would assert that we can measure the extent of life by its longevity, but that it is impossible to deal quantitatively with the essence of life: physical, mental and social well-being. Others of us are seeking a scientific, quantifiable approach to well-being, and believe that some progress is being made. To begin, all people have some degree of health. They fall somewhere between the high end and the low end of physical, mental and social well-being. Crude attempts have been undertaken with some success to determine that "somewhere" on the health scale. While far from being completely satisfactory, advances toward the measurement of health as well-being are being made.

If health includes but means more than avoiding disease and premature death, then health promotion includes but embraces more than measures to prevent specific diseases and death from them. With the curtailment of disease and the extension of life, health promotion increasingly takes the form of steps to maintain and expand life's function and enjoyment generally, and to build reserves against the forces that detract from health. For example, appropriate exercise and good nutrition may both curtail the risk of coronary heart disease and enhance physical fitness.

Strategy for Health Promotion

A comprehensive strategy for health promotion entails three main lines of endeavor: medical, environmental and behavioral. These constitute the ways in which progress against disease has been achieved, and they are the means of promoting health.

Considering how to deal with almost any health problem leads one to realize that it may be approached through medicine, the environment and behavior. For example, to curtail infant deaths good prenatal and pediatric care are necessary; together with home hygiene, including reduction of exposure to toxic agents; and parent education. To prevent loss of teeth and maintain oral health, caries are repaired and calculus removed; adequate fluoridation of water established; and people encouraged to brush their teeth as well as follow a prudent diet. For high blood pressure control, the detection and vigorous medical treatment of the condition can be effective; along with making available foods having less fat and salt; and heightening public awareness of the health significance of being overweight. All of these measures are well known to you. They are mentioned only to emphasize the triumvirate strategy--medical, environmental and behavioral--that we have followed in disease prevention, in keeping away from the negative end of the spectrum.

The same strategy seems appropriate for moving toward the positive end of the health spectrum, i.e. health promotion. To achieve greater physical, mental and social well-being, medical, environmental and behavioral measures are all in order. These may be employed to strengthen people's capacity for enjoying a full life as well as avoiding disease.

Important issues surround each of these modalities for health promotion in the United States.

As a nation we have made progress, but clearly not enough, toward assuring equitable access to medical services. Passage of the Medicare-Medicaid legislation in 1965 extended medical services to

the elderly and the poor of our country who before that time frequently had extreme difficulty in obtaining medical care needed for health. While those programs have alleviated the situation somewhat, the elderly and the poor as well as other segments of the population too often still encounter overwhelming problems in obtaining hospital and physician services that most Americans take for granted. The disparity in access to medical care is a continuing blot on the American social scene. It sets us apart from the other industrialized and most of the developing nations of the world. Recently our situation in this matter has been aggravated by growing emphasis on the economic gains rather than the health gains to be achieved by providing medical care.

That emphasis is closely associated with a second major issue, one receiving great public attention in respect to medical care: rapidly rising costs. The latter, of course, reflect a great many influences. A considerable part of the increase, however, particularly for hospital services, derives from the construction, equipping and staffing of unnecessary facilities; the legal wrangling about the dollars involved, especially in services that injure people; and the income sought from investing in hospitals, as well as from loaning funds for expansion of hospital plants not needed for health purposes. While the trend toward using the medical care system for economic benefit does not explain all the cost increase, it does account for a significant part. Also, the growing commercialization of medicine often conflicts with the originally intended benefit of the system, i.e. health benefit. Arnold Relman, Editor of the distinguished New England Journal of Medicine, has recently been pointing out that danger.

A third, and related matter, is that of medical care quality. Major advances in techniques for investigating that matter have permitted the delineation of questions about the quality of medical service in America. Too much of it is not up to a reasonable standard. That statement applies to virtually every aspect of medical care: what happens in physicians' offices, at the operating table, the x-ray machines and in the laboratory. In medical care the bottom line should be health, not dollars.

Thus, to assure appropriate health benefit from medical services it will be necessary to deal further with the issues of equitable access, health vs. economic interest, and quality.

In the case of environmental measures for the protection and enhancement of health we have also made considerable progress but still face difficult problems. Environmental health issues are similar to those in the medical field.

While disparities among Americans in housing, workplaces and other aspects of the environment have been curtailed, and we are generally much better situated than our grandparents were, our nation still tolerates too much inadequate housing, too many hazardous workplaces, too much air pollution, and too many accidents in transport. Perhaps saying that as a nation we tolerate these situations may be regarded as too strong. The fact is, however, that resistance to necessary steps for health in important circles of our society continues to delay needed improvements.

That resistance derives mainly from efforts to protect economic interests against moves to "clean-up" the environment as a health measure. This may be seen particularly in matters of regulation. Every nation in the world faces the issue of how to deal with health threats from the environment that are created in the course of economic development. Seeking the proper balance is one of the most important political problems of our time, worldwide.

Assessing environmental health risk is, in many respects, comparable to assessing the quality of medical care. In both instances the ultimate question is, what does it mean for health?

Thus, in the environmental as well as in the medical approach to health promotion, attention must be given to issues of equity, balancing health and economic interests, and quality.

In the third modality of health promotion, the behavioral, we encounter these same three issues. In the behavioral approach to health, however, another problem arises. Being called upon to spend money on medical or environmental measures for health protection provokes some to emphasize personal responsibility for behavior affecting health over social responsibility for medical and environmental measures to enhance health. That tendency in turn provokes those committed to medical and environmental approaches to express the view, "Don't blame the victim." The behavioral approach to health, it is alleged by some, merely detracts from the necessary emphasis on the other two modalities.

Concern about personal versus social responsibility for health has arisen before in the history of public health. For example, in his 1941 Preface to Communicable Disease Control, Gaylord Anderson noted, "While it is true that the community is merely the sum of its individual members, nevertheless the problems of protection are not simply the mass application of personal prophylaxis. The community presents a complex mixture of social, political and economic influences that may either facilitate or impede the spread of disease. These same influences affect the control measures that may be developed." That statement applies with equal force to non-communicable disease control. The fact that disease agents are now often chemical products of industrialized society rather than biologic microorganisms, and that the term "spread" refers to a social rather than a biologic force does not change the fundamental point that Gaylord Anderson was making.

Personal behavior does affect health; that behavior is, in turn, substantially affected by the circumstances of life. A person's health-related actions do not occur in a vacuum; they occur in and mainly reflect the social milieu.

Consider how two different men would respond to the availability of a package of cigarettes and a bottle of whiskey: (1) a 19-year old resident of the South Bronx, high school dropout, alienated from his home, with no job and no prospects for a job or his own family, whose friends smoke all the cigarettes and drink all the whiskey they can obtain; (2) a 35-year old resident of Westchester County, college graduate, living with his family and good prospects of grandchildren, with a good job, whose friends do not smoke cigarettes and drink whiskey in moderation. Who would expect men in such different circumstances to behave the same way? We cannot properly "blame the victim," nor can we ignore the social responsibility for his conditions of life.

Neither can we ignore the fact that smoking cigarettes and drinking alcohol to excess are personal actions subject to influence. As professionals in the health field we are obligated to encourage health-enhancing behavior. Fulfilling that obligation includes striving for social conditions that will foster such behavior in all persons.

There is no contradiction between personal and social responsibility for health. They are intertwined.

Health professionals should make clear the interrelationship between the two kinds of responsibility for health, and particularly how they pertain to a comprehensive strategy for health promotion, a strategy that embraces medical, environmental and behavioral modalities.

The Measurement of Health Promotion

Full understanding of where we stand in health promotion requires both ascertaining the extent of health that a person or a population has achieved, and determining how successfully we are applying the three modalities of health promotion.

Measurement of health per se in the past has taken the form almost exclusively of quantifying its absence: death, disease and disability. Now, consistent with an expanding focus toward the positive end of the health spectrum, efforts are underway to measure well-being (fitness, wellness). Its parameters include anatomical, physiological, chemical, bacterial, immunological and genetic. Examples of how these may be measured are indicated below:

Anatomical

Optimum weight/height ratio
Normal epithelial tissue throughout body

Physiological

Blood pressure, approximately 120/80
No electrocardiographic abnormalities

Chemical

Blood cholesterol level, about 200
Substantial glucose tolerance

Bacterial

Freedom from bacteruria
Absence of tuberculosis infection

Immunological

Immunity to current strains of influenza
Immunity to poliomyelitis

Genetic

Absence of trisomy 21 in fetus
No Tay-Sachs affected fetuses

You will recognize these markers as being what are now often called risk factors. The latter term connotes their predictive nature for disease and premature mortality. For example, obesity is a risk factor for heart disease and early death, and therefore to be avoided.

As we turn to the positive side of the health spectrum, however, it is possible to use the same items (but generally in the opposite direction) as predictors, and therefore to be sought, for health. It is important to emphasize generally in the opposite direction. The extreme opposite of obesity, for example, would be the equally unhealthy extreme loss of adipose and other tissue. An optimum weight/height ratio is obviously desirable.

Conceptually, this notion of optimum can be extended through the several categories of health predictors listed, and more commonly known now as risk factors. Thus, not only is there an optimum weight/height ratio for health but also an optimum blood pressure, an optimum cholesterol level, an optimum glucose tolerance curve. Rather than limiting our attention in such measurements to the diagnosis of hypertension, hypercholesterolemia, or diabetes--or risk of these conditions--we can now begin to consider the optimum level for health. Quantifying such items would be the corollary of measuring health, not just death, disease, and disability. A step in that direction, for example, would be to report from surveys not just the proportion with systolic blood pressure over 160 mm Hg. or over 140 mm Hg., but the proportion less than 110, 110-120, 120-130, 130-140, 140-150, and so on. Such reporting would indicate a health focus and, over the years, the attainment of health, not merely the extent of what we call hypertention.

Determining how well we are doing in health promotion also entails measuring the extent of progress in using the three components of health strategy: medical, environmental and behavioral. Thus we need to monitor such items as extent of immunization among children, and length of time since last mammogram among women over 50 years of age; amount of asbestiform fibers in the air, and toxic chemicals in streams and ground water; how many and which people still smoke cigarettes, or use alcohol to excess.

The Public Health Service, in Health Promotion-Disease Prevention: Objectives for the Nation, has already assembled available data on such matters and set objectives that we ought to reach. That would be a good start. It would be highly desirable to extend the range of items and the geographic locales of reporting them as the basic means of measuring health promotion in our country.

This is an exciting time to be in the field of health records and statistics. It does not only provide the opportunity to improve ways of collecting and reporting information long, and still, needed as a basis for disease prevention. Now it calls for creating the means for tracking progress in the next aim: health promotion.

TRACKING THE NATIONAL PREVENTION OBJECTIVES

J.M. McGinnis
Joel Kavet
Office of Disease Prevention
and Health Promotion

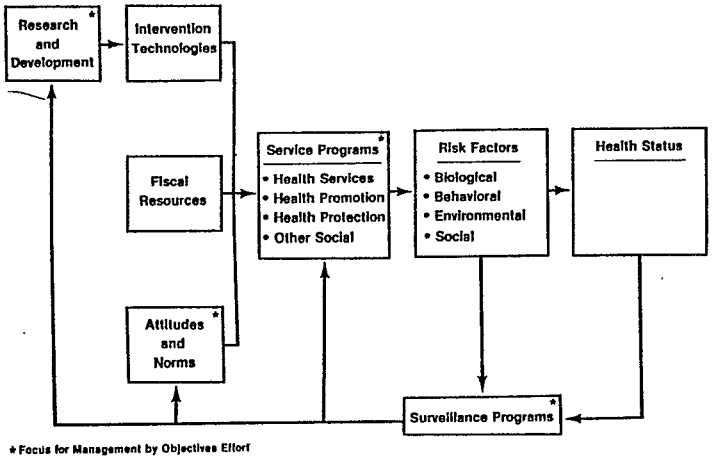
This paper speaks to the issue of the national perspective on health promotion and disease prevention as represented by our efforts to attain certain national and measurable objectives by 1990. The paper is divided into two portions: a) a presentation on the background and context for the general approach to disease prevention and health promotion objectives; and b) a more detailed assessment of the data needs for those objectives specifically targeted to health promotion.

In 1979 a Public Health Service-sponsored work shop marked the beginning of a noteworthy endeavor that was to result in the publication of a report entitled Promoting Health/Preventing Disease: Objectives for the Nation.[1] Published in the Fall of 1980, that report represents the labors of several hundred dedicated men and women from across the spectrum of public health and social service professions and it embodies a set of measurable objectives designed to guide the national effort in health promotion and disease prevention through the 1980's.

The motivating context for the objectives effort is noted in the chart in the schematic on Figure 1, presenting a summary of the various factors influencing health outcomes.[2] Indicated in Figure 1 is the fact that morbidity and mortality are influenced by the interplay of various biological, behavioral, environmental and social risk factors. These risk factors can in turn be influenced through the delivery of a variety of treatment, health promotion, health protection and other social services. The ability of a society to deliver those services is dependent upon the availability of appropriate intervention technology, the fiscal resources necessary to pay for the services, and societal willingness to commit the resources to the services. Research efforts are important to developing effective intervention technologies. At the heart of the system is the need for appropriate surveillance systems which can gather information about health status and risk factor prevalence and feed that information back to affect the nature of service programs, societal attitudes and norms, and research and development activities.

Figure 1

Factors Determining Health Status



Several loci on this chart are susceptible to the management-by-objectives approach, including the activities undertaken within service programs, the societal attitudes and norms, the research and development exercises, and the surveillance activities. As we move further into our discussion of the objectives themselves, the role of the various stages in this analytic construct will become apparent.

The objectives developed are designed to foster the achievement of the five broad goals outlined in the 1979 report Healthy People: The Surgeon General's Report on Health Promotion and Disease Prevention.[3] These goals, listed in Figure 2, present the Nation's aspirations for health improvement for people at five major life stages: infants, children, adolescents and young adults, adults and older adults.

Figure 2

Health Status Goals

Goal One:	**Healthy Infants (below age1)**
	Subgoal: To reduce the incidence of low birth weight infants
	Subgoal: To reduce the incidence of birth defects
Goal Two:	**Healthy Children (age 1—14)**
	Subgoal: To enhance childhood growth and development
	Subgoal: To reduce childhood accidents and injury
Goal Three:	**Healthy Adolescents/Young Adults (age 15—24)**
	Subgoal: To reduce death and disability from motor vehicle accidents
	Subgoal: To reduce misuse of alcohol and drugs
Goal Four:	**Healthy Adults (age 25—64)**
	Subgoal: To reduce heart attacks and strokes
	Subgoal: To reduce the incidence of cancer
Goal Five:	**Healthy Older Adults (age 65 and above)**
	Subgoal: To increase the proportion of older people who can function independently
	Subgoal: To reduce premature death and disability from influenza and pneumonia

In addition to the presentation of these five broad measurable goals, Healthy People contains a discussion of fifteen priority areas which were necessary to address in order to achieve the goals. Those areas, presented in the three groupings of health promotion, health protection

and preventive health services, are noted in Figure 3.

Figure 3

Health Strategy Targets

Health Promotion for Population Groups

- Smoking cessation
- Alcohol and drug abuse reduction
- Improved nutrition
- Exercise and fitness
- Stress control

Preventive Health Services for Individuals

- Family Planning
- Pregnancy and infant care
- Immunizations
- Sexually transmissible diseases services
- High blood pressure control

Health Protection for Population Groups

- Toxic agent control
- Occupational safety and health
- Accidental injury control
- Community water supply fluoridation
- Infectious agent control

These fifteen areas have served as the focus for the development of the measurable objectives for prevention. In all, 223 specific objectives have been developed across the 15 areas with five kinds of objectives, noted in Figure 4, developed for each of the 15 areas.

Figure 4

NUMBER OF OBJECTIVES BY CATEGORY

—Improved health status	59*
—Reduced risk factors	47**
—Increased public/professional awareness	38*
—Improved services/protection	51
—Improved surveillance/evaluation systems	32
	227

*1 duplicate
**2 duplicates

The course of the discussion which follows sketches the form and the thrust of the approach to implementing the Federal segment of the initiative in health promotion and disease prevention. Special emphasis is given to the challenge posed by the need to be able to document where we stand at a point in time, and how we are progressing over time in relation to the measurable targets embedded in the objectives.

It is important to emphasize at the outset that, by themselves, Federal initiatives in health promotion and disease prevention, no matter how vigorously implemented, cannot assure attainment of the goals and objectives established for 1990. Our discussion focuses on pursuit of the objectives from a Federal perspective, but the ultimate success of such endeavors will depend in large measure on whether, and the extent to which, the Federal commitment is matched by likeminded support and similarly directed efforts at other levels of government, in the private sector, and among the citizens of our Nation. In a very real sense, then the agenda embodied among the Objectives for the Nation is indeed national in its scope and in its aspirations.

The Federal Role in Health Promotion and Disease Prevention: Pursuit of the Objectives for the Nation

Shortly after their publication, the Public Health Service instituted a number of measures to focus attention on the objectives and promote their integration into the programs of the Department of Health and Human Services, as well as elsewhere in the Federal Government. Early on, an agency of the Public Health Service, designated by the Assistant Secretary for Health, was assigned principal responsibility for coordinating Public Health Service and related Federal activities in each of the fifteen priority areas. The assignments are noted on Figure 5. They were made on the basis of programmatic or statutory responsibilities, experience and expertise.

Figure 5

Lead HHS Agencies for Objectives

Category	HHS Agency/Office
Preventive Services	
High Blood Pressure Control	National Institutes of Health
Family Planning	Office of Population Affairs
Pregnancy and Infant Health	Health Resources and Services Administration
Immunizations	Centers for Disease Control
Sexually Transmitted Diseases	Centers for Disease Control
Health Protection	
Toxic Agent Control	Senior Advisor for Environmental Health
Occupational Safety and Health	Centers for Disease Control
Accident Prevention and Injury Control	Centers for Disease Control
Fluoridation and Dental Health	Centers for Disease Control
Surveillance and Control of Infectious Diseases	Centers for Disease Control
Health Promotion	
Smoking and Health	Office on Smoking and Health
Misuse of Alcohol and Drugs	Alcohol, Drug Abuse, and Mental Health Administration
Nutrition	Food and Drug Administration
Physical Fitness and Exercise	President's Council on Physical Fitness and Sports
Control of Stress and Violent Behavior	Alcohol, Drug Abuse, and Mental Health Administration

The lead agencies for each of the fifteen priority areas were charged with the responsibility for formulating a coordinated set of plans which outline the array of programs and activities the Federal sector has under way or under consideration that might contribute toward achievement of the objectives. Taken together, these implementation plans provide a rather detailed inventory of the Federal commitment to health promotion and disease prevention.

The complete set of implementation plans has been published as a supplement to the September/October 1983 issue of Public Health Reports.[4] Our purpose in offering these plans for the broadest dissemination is to share with all

concerned the intentions of the Public Health Service, as well as offer a model for replication with whatever modifications others may deem suitable and appropriate elsewhere in government or in the private sector. Careful examination of the implementation plans will reveal, in each instance, those activities the Federal Government has under way or under consideration in a number of categories: education and information measures; grants to the States and service delivery measures; technical assistance and cooperative measures; economic and other incentive measures; and research and surveillance measures.

The implementation plans represent a clear declaration of intent, but their utility as a program management tool is limited in the absence of the means to monitor progress toward meeting the targets specified in the objectives. As a consequence, a number of steps have been taken to assure and enhance the capacity of the Public Health Service to keep abreast of the nature and rate of progress toward the objectives.

Notable among the oversight mechanisms are the progress reviews conducted to keep the Assistant Secretary for Health and his principal associates posted on the status of activities related to the objectives. Each month, a progress review is conducted on one of the fifteen priority areas. The sessions afford the lead agencies and their collaborators an opportunity to report on recent developments in the field and present information which reflects progress toward attainment of the objectives. The sessions also provide a forum for discussion of problems encountered along the way and how they have been or might be overcome. The proceedings of each session are summarized by the lead agency and submitted for publication in Public Health Reports.

However important and informative they may be, the progress review sessions alone are not best suited to the task of keeping track, on an ongoing basis, of where we stand with respect to each of the objectives. Recognition of the importance of being able to monitor the status of activities associated with each of the objectives and progress toward those objectives had an essential influence on their formulation. To the extent possible and practicable, the objectives were articulated in such a way as to embody a quantifiable target and they included, where they were available, baseline data intended to serve as a measure of then-current status and point of departure for the initiatives of the 1980's.

The progress reviews and other oversight activities of the lead agencies and their collaborators will eventually be augmented by the resources of an Information Tracking System which takes fullest advantage of the quantitative emphasis adopted in expressing the objectives. Relying largely on existing data sources, the tracking system will provide a centralized resource capable of compiling and managing a body of data and related information which reflects the status and progress of efforts to move toward the targets embodied in the measurable objectives. While full implementation of the tracking system is still 8-10 months in the offing, a preliminary glimpse of the kinds of information it will handle will be available later this year when the second triennial Prevention Profile is published as part of Health: United States, 1983.

Monitoring Progress: An Examination of the Data Available for the Objectives in Health Promotion

The capacity of the progress review sessions, the tracking system, or the Prevention Profile to reflect adequately the degree of progress toward the objectives is necessarily limited by the existence of suitable data and information. Not surprisingly, there are objectives for which data adequate to permit specification of a baseline or measurement of progress do not exist. In the case of others, experience may have shown existing baseline data to be erroneously or poorly specified. One approach to these problems, worthy of particular note, is the development of a survey instrument on health promotion and disease prevention which is to be fielded as a supplement to the FY 1985 edition of the National Center for Health Statistics' highly regarded Health Interview Survey. This survey supplement will provide much needed data on more than 30 different objectives across the spectrum of priority areas, thereby enhancing our capacity to monitor and assess progress toward the objective targets.

From time to time, ad hoc surveys on a particular priority area may also be undertaken. Thus it is, for example, that a survey scheduled to be conducted shortly will provide a body of data that will permit measurements to be made in conjunction with a number of the objectives associated with the control of stress and violent behavior. The so-called "stress survey" will be conducted by a private contractor using an instrument developed with the advice and guidance of a panel of experts from the agencies of the Public Health Service.

The capacity to allocate and apply resources in a responsible and responsive manner depends heavily on the availability of data by which to gauge the effects of our efforts. Oversight of many activities which contribute to realization of the objectives is facilitated by the presence, in many instances, of ongoing data and information collection mechanisms. In general, this is more likely to be the case for objectives falling into the priority areas that have been assigned to the categories labeled Preventive Health Services or Health Protection. Of particular concern to us in this instance, however, is the capacity to track progress toward the objectives in the areas which come under the rubric Health Promotion: Smoking Control; Alcohol and Drug Misuse Prevention; Improved Nutrition; Physical Fitness and Exercise; and Control of Stress and Violent Behavior. These are of notable interest inasmuch as they involve activities designed to influence or alter the behaviors or attitudes of individuals. It is an area where changes in awareness as well as changes in overt behaviors are especially important. And, it is an area where there have not always been data collection mechanisms in place. In many cases, measurements of the type necessary to track the objectives have been hard to come by or non-existent. With attention to these matters we hope the problem will, in time, abate.

To illustrate some of the problems and plans related to the health promotion objectives, let's assess how the data availability situation now stands and suggest how it might change as a consequence of the two survey activities alluded to earlier. We should also note that the two surveys were designed with an eye toward acquiring

and only two of those have data subsequent to the baseline measurement. Twelve of the fifteen--four-fifths of the total--in this category are without any baseline measurement.

The importance to this effort of the kind of data gathering activities discussed earlier is reflected in Table 5. As a result of timely consideration of data needs and careful collaboration, the two surveys identified above were designed in such a way as to make it possible

TABLE 5: Priority Health Promotion Objectives by Priority Area and Category of Objectives Scheduled to be Addressed in the HIS Supplement and the Stress Survey

	Category of Objective					
Priority Area	Improved Health Status	Reduced Risk Factors	Improved Awareness	Improved Services/ Protection	Improved Surveillance/ Evaluation	Total
Smoking Control*...	0	1	4	0	0	5
Alcohol and Drug Misuse Prevention*.	0	0	2	0	0	2
Improved Nutrition*.	0	2	2	2	0	6
Physical Fitness and Exercise*.....	0	2	1	0	0	3
Control of Stress and Violent Behavior #...	0	0	2	0	2	4
Total.....	0	5	11	2	2	20

* - Data to be collected through the HIS Supplement
# - Data to be collected through the HIS Supplement and the Stress Survey

to yield some form of measurement or assessment of progress for 20 of the health promotion objectives. In light of some of the concerns cited earlier, it is worth noting that these two surveys will acquire data for 11 "improved awareness" objectives. For more than half of the affected objectives, the surveys will acquire baseline data not heretofore available. For six others, there will be follow-up data, with four of the objectives being measured on such a basis for the first time.

TABLE 6: Status of Data Availability for Priority Objectives, by Priority Area, Projected Post HIS Supplement and Stress Survey

		Data Available			
Priority Area	No. of Objectives	Baseline	Follow-up	Not Applicable	Unavailable
Smoking Control.	10	7	2	3	0
Alcohol and Drug Misuse.....	14	13	10	1	0
Improved Nutrition.....	15	12	1	1	2
Physical Fitness and Exercise.....	11	5	2	4	2
Stress and Violent Behavior..	10	5	2	3	2
Total.....	60	42	17	12	6
Total from Table 3.....	60	30	13	12	18

Table 6 shows the anticipated status of data availability if the two surveys proceed as planned and acquire the data we anticipate they will. While the table reflects a measure of progress, it also leaves little doubt that there is much to be done before we can be comfortable with our capacity to track the health promotion objectives. We will have reduced by a substantial margin the number of measurable objectives for which there are no baseline data, but fewer than half will have any follow-up data from which it might be possible to get even a hint of a trend. And while the number of objectives for which we have no data will have been reduced by two-thirds, there is still the remaining third--six objectives--with which we must be especially concerned.

Generally speaking, the objectives that will not benefit from the data collection activities of the two surveys cited in this discussion are those which require measurement reflecting performance or awareness on the part of other than household survey respondents (e.g., providers or practitioners). These will necessarily have to be addressed in other data gathering activities. And even though the Health Interview Survey supplement and the stress survey will improve our position with respect to data availability, it is important to point out that these are not routine, ongoing data collection activities. There remains a continuing need for follow-up data and not too much time will be able to pass before the need for timely data will again call attention to the full set of health promotion objectives.

The Relation Between Federal Efforts and Those Conducted at the State and Local Levels

At the beginning of this discussion, we noted that the health promotion and disease prevention goals and objectives for the Nation consitute an agenda which merits nationwide and not just Federal attention. Collaborative and complementary activities at the subnational levels are vital to our collective aspirations and efforts to improve still further the health of the American people. At the same time, however, it is important to stress that the agenda suggested by the Objectives for the Nation is neither comprehensive or exhaustive, nor was it intended to be.

Others at this conference have called attention to the Model Standards for Community Preventive Health Services developed as a collaborative undertaking by organizations representing Federal, State and local governments, as well as the voluntary sector. Careful examination of the Model Standards reveals more clearly the true scope of the challenge we face in health promotion and disease prevention. Due to the fact that the Model Standards were issued before work was complete on the objectives, there are differences in some of the approaches. But these will be resolved with next year's revision of the Model Standards and the most striking feature is the reinforcing nature of the two exercises. The co-existence and complementary nature of the Objectives for the Nation and the Model Standards can only heighten our awareness of the essential fact that the prospects for success depend on far more than the energies, commitment, or resources of any single player.

It is worth noting that data collection activities at the State level may also contribute to our ability to comprehend more fully the significance of national data and enhance our capacity to monitor and explain what progress is recorded toward the objective targets. Several states have undertaken projects to adapt the objectives to their own conditions and priorities. Additionally, the Centers for Disease Control has initiated cooperative efforts with a number of states to elicit state-based data on several key health-related behaviors.

data through which progress toward the objectives could be monitored.

There are 223 discrete objectives and 78 of these appear in the five priorities classified under the Health Promotion heading. Table 1 shows the objectives as they are distributed among the five priority areas.

TABLE 1: Health Promotion/Disease Prevention Objectives by Priority Area and Priority Status

Priority Area	Total	Priority: High	Medium	Other
Smoking Control	17	10	0	7
Alcohol and Drug Misuse Prevention	19	9	5	5
Improved Nutrition	17	11	4	2
Physical Fitness and Exercise	11	11	0	0
Control of Stress and Violent Behavior	14	6	4	4
Total	78	47	13	18

Sixty of the 78 have been assigned a medium or high priority by the agencies of the Public Health Service and, as a consequence, are the subjects of implementation plans. The 18 objectives classified as "Other" are not considered further in this discussion. That omission should not be taken as any reflection on the overall importance or worthiness of those objectives. As we noted earlier, pursuit of the objectives is more than a Federal undertaking. The designation of priority alluded to here reflects only the judgment of the Federal agencies as to where and how they believe their limited resources can be best and most appropriately applied.

More than half of the measurable health promotion objectives are directed at reducing risk factors (16) or increasing public or professional awareness (15). Their share of total pool is, in one sense, larger than it appears at first glance. Virtually all of the objectives assigned to the column in Table 2 headed "Improved Surveillance/Evaluation" are directed at systemic

TABLE 2: Health Promotion Objectives by Priority Area and Category of Objective *

Priority Area	Category of Objective: Improved Health Status	Reduced Risk Factors	Improved Awareness	Improved Services/ Protection	Improved Surveillance/ Evaluation	Total
Smoking Control	0	2	4	2	2	10
Alcohol and Drug Misuse Prevention	4	5	3	1	1	14
Improved Nutrition	2	5	3	4	1	15
Physical Fitness and Exercise	0	4	2	1	4	11
Control of Stress and Violent Behavior	3	0	3	1	3	10
Total	9	16	15	9	11	60

* Includes only those objectives assigned high or medium priority.

issues. Aside from such notations as "present," "absent," "under development," and the like, progress toward these objectives is not measurable in terms comparable to those used to measure the bulk of those in the other categories (where it is possible in many instances, for example to use a variety of rates).

Table 3 portrays the current status of data availability for the Health Promotion objectives. Some form of baseline data exist for 30 of the 60

TABLE 3: Current Status of Data Availability for Priority Objectives, by Priority Area

Priority Area	No. of Objectives	Data Available: Baseline	Follow-up	Not Applicable*	Unavailable#
Smoking Control	10	4	1	3	3
Alcohol and Drug Misuse Prevention	14	12	9	1	1
Improved Nutrition	15	7	1	1	7
Physical Fitness and Exercise	11	4	0	4	3
Control of Stress and Violent Behavior	10	3	2	3	4
Total	60	30	13	12	18

* — Systemic type of objectives for which measurement data are not likely to be available in a form comparable to the objectives in other categories.

\# — Objectives for which we currently have neither baseline nor follow-up data.

priority health promotion objectives, but for only 13 of those 30 are there any data available for one or more intervals subsequent to the baseline period. For only those 13, therefore, is it possible to make any judgment, however tentative, as to whether any progress is being made toward the objective targets. Of particular concern is the fact that there are currently 18 objectives for which no baseline data are available at all. In a sense, then, these represent items for which objective measurement has yet to begin.

We noted earlier the particular importance assigned to efforts to improve awareness in conjunction with pursuit of the health promotion objectives. Fifteen of the objectives--a quarter of the total in health promotion--fall into the category dealing with improving awareness. At

TABLE 4: Current Status of Data Availability for Priority Objectives, by Category of Objectives (Improved Public/Professional Awareness) and by Priority Area

Priority Area	No. of Objectives	Data Available: Baseline	Follow-up	Not Applicable	Unavailable
Smoking Control.	4	1	1	0	3
Alcohol and Drug Misuse	3	2	1	0	1
Improved Nutrition	3	0	0	0	3
Physical Fitness and Exercise	2	0	0	0	2
Stress and Violent Behavior	3	0	0	0	3
Total	15	3	2	0	12

the present time, as Table 4 demonstrates, there are available baseline data for only 3 of those 15

CDC has proposed institutionalizing these surveys as surveillance mechanisms to be conducted by the states. At the present time they are thinking of sponsoring the activity through cooperative agreements under which the states would conduct the surveys with the assistance and benefit of CDC training, core questionnaire development and data processing. CDC will explore the prospects of conducting the survey on a monthly basis, year-round, with a minimum number of interviews per month being set at 50. Steps will also be taken in processing and analyzing the data to weigh the results for seasonal considerations, and other important variables. The national 1990 objectives provide an important basis for the interview questions.

Looking Forward

The acquisition and analysis of data to measure progress toward the health promotion and disease prevention goals and objectives can also be expected to call attention to the substance of the objectives themselves. In addition to changes growing out of advances in the science base, measurement or tracking data will begin to suggest which of the objectives might be in need of reconsideration, modification or reformulation. In some cases, circumstances may point to the need for adding new objectives or deleting existing ones. The agencies of the Public Health Service have already begun in a number of instances to consider such eventualities as the first round of progress reviews comes to a close, and they have begun in a number of the priority areas to formulate recommendations regarding the future form and content of the objectives.

As we look forward to the mid-point of the decade, we also look forward to a full mid-course review of the objectives and the implementation plans. Tentatively scheduled for early in Fiscal Year 1985, the review will allow us, with help from a broad constituency, to fine-tune the objectives in order that they may appropriately guide Public Health Service health promotion and disease prevention activities in the second half of the 1980s.

While our discussion has focused on the Federal approach to the objectives, they are, as noted earlier, the elemental blocks of a national--and not just a Federal--initiative in health promotion and disease prevention. Recognizing this, the Public Health Service is making a concerted effort to promote more broadly careful consideration of the objectives at the State and local levels of government as well as by organizations in the private sector. A meeting convened at the Centers for Disease Control last fall was devoted to exploring the prospects for, and issues associated with, application of the objectives at the State and local levels. Earlier comments suggest that there has been some gratifying progress in this area. We are hopeful this trend will continue and we look forward to the spread of similar activities into the private sector. Indeed, a conference to foster such developments will be held in Fiscal Year 1984. By actively seeking to expand the number of participants in activities directed toward attainment of the objectives, we can only enhance our prospects for success.

References

1. U.S. Department of Health and Human Services, Promoting Health/Preventing Disease: Objectives for the Nation, DHHS (PHS) Publication No. 80-3316999, Government Printing Office, Washington, D.C. 1980.

2. J.M. McGinnis, "Setting Nationwide Objectives in Disease Prevention and Health Promotion: The United States Experience," in W. Holland, R. Detels and E.G. Knox, eds., Oxford Textbook of Public Health (in press).

3. U.S. Department of Health, Education, and Welfare, Healthy People: The Surgeon General's Report on Health Promotion and Disease Prevention, DHEW (PHS) Publication No. 79-55071, Government Printing Office, Washington, D.C. 1979.

4. U.S. Department of Health and Human Services, "Public Health Service Implementation Plans for Attaining the Objectives for the Nation," in Public Health Reports, Volume 98 No. 5 Supplement, September-October 1983

Health Promotion: The State Perspective

James O. Mason, Utah Department of Health

I am pleased to be invited to speak at this conference on priorities in health statistics to provide a state's perspective on health promotion. Too often the important role of statistics in public health initiatives, such as health promotion, is not fully appreciated. Health promotion is a relatively new direction for public health; and for us to determine where we are going and measure our progress, we must be able to tell where we are and where we have been.

Too many times over eager researchers use statistics like a drunk uses lamp posts, for support rather than illumination. In health promotion, we need statistics to illuminate our progress in making a difference in the lives of individuals. With all our service responsibilities and research on the effects of health promotion programs, we cannot forget basic statistical work to keep us aware of where we are.

Jonathan Fielding notes that any success in health promotion/disease prevention may be measured by how well the programs: 1) avoid premature death; 2) reduce avoidable morbidity; and, 3) minimize disability that interferes with usual functioning.[1]

These outcomes can be measured. The importance of looking at outcome data rather than process measurements is paramount. We not only need the data which has traditionally been available, but also new data bases including morbidity and small area data. Our ability to plan and evaluate progress is severly hampered by the absence of this information.

Questions are being raised about costs and benefits. This will grow as the private sector adopts health promotion techniques. Some improperly or inadequately evaluated confidence may be lost. We must evaluate. We must look at cost benefits. We must assess relative risks and the value of incremental gains, and have the date to manage risks. The credibility of the progress in health promotion may be undermined without the data needed for these critical analyses.

Until recently, infectious diseases were responsible for the vast majority of illnesses and deaths in this country. The 1850 census on mortality in the United States shows that approximately three-fifths of all deaths were caused by infectious diseases.[2]

These epidemics were brought under control through major public initiatives including improved water purification and wastewater treatment, general sanitation, better housing and improved nutrition, as well as immunization against vaccine preventable diseases. Surveillance, epidemiology, microbiology and immunology were the basic sciences of this revolution which added so much to the well being of this and other nations.

This magnificent accomplishment is often referred to as the "First Public Health Revolution." It is an ongoing effort involving public and private agencies to keep these diseases in check. With the exception of smallpox, all of the past causes of death and disability would again give rise to epidemics if the means of control were even temporarily relaxed.

We are now engaged in what many call the "Second Public Health Revolution." The battle here is being fought on a broad field encompassing the individual's total personal universe: lifestyle, personal habits, environment, workplace and home. The killers and cripplers today are the chronic and genetic diseases and the results of violence which take an unnecessary toll in lives, productivity, and resources. Table 1 shows Major Causes of Death by Age in Utah during 1978-1980.[3] Note the major killers by age group and how many are lifestyle related and can best be influenced through changes largely under individual control.

The importance of an individual's lifestyle and behavior as related to the relative risk of death from today's leading killers is shown in Table 2.[4]

Lifestyle contributes to 44 percent of these deaths. The other components of the Health Field Concept, health care organization, environment and human biology, trail in significance. The Health Field Concept was developed in 1975 by Marc Lalond, then national Minister of Health and Welfare in Canada.[5]

The effect of health promotion/infectious disease prevention is encouraging. We see progress in declining deaths on our roads and highways, fewer childhood poisonings, decreasing dental caries, control of some congenital metabolic disorders, advances in cervical and breast cancer detection and early diagnosis of cardiovascular disease and chronic lung disease. These are examples of successful and practical health promotion initiatives.

However, despite these impressive early successes, advances in medical science and technology and the billions of dollars being spent on medical care services, our nation's health indices are still far from where they should and could be.

The time has come for much greater emphasis of our national and state health priorities. The states applaud the efforts on the federal level in setting specific goals and measurement criteria to improve the health of Americans during the next decade. The publications Healthy People and Objectives For The Nation[6] bear the message that we can and must do more to keep our people vigorous, strong and healthy.

Many states are preparing their own plans of action with goals and objectives to improve the health of their citizens.

Historically, public health has made a significant contribution in improving health and in lengthening life. Nevertheless, there still are significant health problems that should be of concern to policy makers. In 1980 in Utah there were 8,103 deaths that resulted in 173,625 early years of life lost and 71,305 working years of life lost. This resulted in a direct economic loss in terms of earning power of $642.3 million in 1980, not to mention the health care costs involved with caring for these diseases.[7] This is a substantial loss, since much of it could have been prevented.

In our policy document, we have identified six major health status problems in Utah which should be of major concern to policy makers:

--Cancer.
--Congenital Anomalies.
--Dental Conditions.
--Heart Diseases and Stroke.
--Motor Vehicle Accidents.
--Problems of Early Infancy.

We have established a specific goal for each of these six problems. Table 3 shows the goal we have set for cancer.[8]

Similar goals have been established for the other priority areas. We have identified specific risk factors which contribute to the health priority problems. Table 4 outlines the risk factors in order of their priority.[9] For cancer, smoking, environmental exposure, genetic pre-disposition and alcohol abuse are the significant association risk factors.

The most important risk factors for Utahns in the six health problems are summarized on Table 5.[10]

It does not surprise you that smoking is clearly the single most significant controllable risk factor in terms of Utah's priority health problems. It is a major contributor to heart disease, stroke, cancer and problems of early infancy. It has also been found to compound the effects of other risk factors: smoking, in combination with certain environmental exposure, has been found to increase the possibility of cancer 10 times. Smoking accounts for 350,000 premature deaths - unnecessary deaths - annually in America. Economically this means 77 million excess work loss days per year and 150 million excess sick bed days per year.[11]

With the goals established and major contributing health risks identified, our next step was formulating a plan of action. This is being done within the context of five major objectives.[12]

These are shown in Table 6:

1. Get Utahns to Assume more Individual Responsibility for Their Own Health.
2. Identify and Appropriately Refer Utahns with Genetic Predispositions.
3. Assure all Prospective Mothers Receive Adequate Pre-Natal Care.
4. Increase the Number of Utahns Consuming Appropriate Levels of Flouride.
5. Reduce Utahns Exposure to Substances Harmful to Health.

Utah has a significant health status head start on the rest of the nation. Utah residents, for the most part, have traditionally accepted and appreciated the importance of individual responsibility in staying well. This is illustrated by comparing Utah's age adjusted death rates to national figures.

As an example of the health status of Utahns, these tables show Utah and United States death rates for heart disease and cancer during the past twenty years.[14] (Tables 7 & 8)

These differences, diminishing for heart disease and increasing for cancer, can be attributed, to a large degree, to a healthy lifestyle and reduced self-imposed risks including smoking and alcohol consumption. Utah's highly favorable mortality rates may not be achievable on a national scale for a number of reasons. The rates, do, however, serve as a target for what is possible in any state or community. They are largely related to health promotion and disease prevention rather than benefits brought on by medical technology.

There is an inverse relationship between the health of Utahns and utilization of medical care and hospital beds. Prudent lifestyle reduces the need and therefore use of the medical care system and its attendant costs. Were the nation to achieve Utah's mortality rates for leading causes of death by appropriate changes in lifestyle, over 284,000 lives could be saved annually.

If national hospital utilization as respresented in age adjusted annual patient days per 1000 could be reduced from 1214 days to Utah's 707 days, over $17 billion could be saved annually. (1981 data).[15]

The key to this type of progress is motivating Americans to a personal awareness and responsibility for their own health and wellness. We must build a national consciousness for wellness for all Americans.

Up until now, too many of the behavior changes we have promoted have involved the better educated, mostly white, upper and middle class segments of the U.S. population. These are the people who join health and fitness clubs, play tennis, golf, ski, bicycle -- all activities that generally take some sort of capital investment. All these activities are wonderful for those able to pursue them, but, there are many who do not fit into this convenient, and affluent, fitness pattern.

Health promotion and risk reduction activities must reach far beyond these people if we ever hope to reach a true level of health consciousness in this country. Unless we are able to reach all segments of the population, we will never meet the goals we have set for a national consciousness for wellness in America. Health promotion and risk reduction must reach into each home, apartment or condominium regardless of race, age or financial status. It must reach into our poorer neighborhoods where death rates are absolutely disgraceful. It must reach into the educational system, the inner cities, and our suburban and rural areas where accidents, violence, stress and a vast array of social factors take a deadly toll in homicides, suicides and mental illness.

During the past thirty years, the United States has achieved more in the realm of civil rights than at any time during the country's history. The benefits of American citizenship in civil rights, social justice, and social programs finally reached great segments of our population. Of course, we are by no means finished with this endeavor, nor should we ever be finished with it.

But now is the time to expand this awareness of social concerns to health concerns. Now is the time for the Second Revolution in Public Health - health promotion, risk reduction and disease prevention - to reach all segments of our population. We need a health promotion program to match our progress in civil rights.

Dr. William Foege, Director of the Centers for Disease Control posed the question, "How can we measure and compare civilization?"[16] His conclusion was that the true measure of a civilization is founded on how people within it treat each other.

The active application and teaching of health promotion/disease prevention principles are important aspects of loving, caring and sharing. This process extends beyond the scope of aerobics, accident prevention and nutrition, to the full spectrum of threats to our health and wellbeing, including what has been labeled "The Last Epidemic" or the threat of nuclear war.

Healthy People and Objectives for the Nation have become national banners. Now is the time for each state to move ahead on those goals and objectives. We need not wait for additonal proof. We must move ahead.

This will require the best efforts and commitment of each state. It will require the sharing of expertise across professional lines as well as national state and community boundaries. It will also extend beyond the traditional health sector. All public health professionals have a role in this effort. The data to identify needs, establish priorities and measure accomplishments and outcomes is an essential, integral part of this process. May our united accomplishments reflect well on our contemporary civilization.

Table 1

Major Causes of Death by Age
Utah
1978-1980

Age Group	Major Causes of Death in Priority Order	Number of Deaths	Age Specific Rates of Mortality Per 100,000
Under Age 1	Problems of Early Infancy	790	623.5
	Congenital Anomalies	336	265.2
	Non-Motor Vehicle Accidents	28	22.1
Age 1-14	Non-Motor Vehicle Accidents	147	11.7
	Motor Vehicle Accidents	126	10.0
	Congenital Anomalies	65	5.2
Age 15-24	Motor Vehicle Accidents	386	43.7
	Non-Motor Vehicle Accidents	128	14.5
	Suicide	122	13.8
Age 25-44	Motor Vehicle Accidents	285	24.8
	Cancer	225	19.6
	Suicide	212	18.4
Age 45-64	Cancer	1,221	192.1
	Heart Disease	1,688	265.6
	Motor Vehicle Accidents	171	26.9
Over Age 65	Heart Disease	6,453	1,996.9
	Cancer	2,364	721.5
	Pneumonia and Influenza	614	187.4

Table 2

An Epidemiological Model for Health Policy Analysis

Leading Causes of Death	Proportional Allocation of Risk Factors for Each Cause of Death to Each Health Field[3]				
	Relative Weight	Medical Care Establishment	Lifestyle	Environment	Human Biology
Cancer	16.4%	10	37	24	29
Heart Disease	16.0	12	52	9	27
M. V. Accident	16.1	12	69	18	1
Certain Diseases of Early Infancy	14.1	27	30	15	28
Birth Defects	10.4	6	9	6	79
Other Accidents	9.9	14	51	31	4
Suicide	6.5	3	60	35	2
Stroke	2.7	7	50	22	21
Alcoholism	2.7	3	70	9	18
Influenza Pneumonia	2.4	18	23	20	39
Diabetes	1.9	6	26	0	68
Bronchitis, Emphysema, and Asthma	0.9	13	40	24	24
Total[2]	100.0%	12.4	44.1	17.8	25.3

1. Percent of years of life lost before age 65 from the cause of death in relation to the total years of lifes lost from the 12 causes of death listed.

2. Weighted average based on the relative weight of each cause of death.

3. G. E. Alan Dever, "An Epidemiological Model for Health Policy Analysis," Social Indicators Research 2 (1976) pp. 453-666

Table 4

Major Risk Factors Associated with Utah's Primary Health Problems

Health Problem	Risk Factors in Order of Priority
Cancer	Smoking Environmental Exposure Genetic Pre-Disposition Alcohol Abuse
Congentital Anomalies	Genetic Pre-Disposition Alcohol Abuse Drug Abuse Improper Nutritional Habits Age of Mother Smoking
Dental Disease	Improper Fluoride Intake Improper Nutritional Habits Poor Dental Hygiene
Heart Disease and Stroke	Smoking Hypertension Genetic Pre-Disposition Improper Nutritional Habits Excessive Stress Poor Physical Conditioning
Motor Vehicle Accidents	Alcohol Abuse Vehicle Safety Restraint Usage Excessive Speed Drug Abuse Driver Age Availability of Emergency Medical Services
Problems of Early Infancy	Inadequate Pre-natal Care Age of Mother Alcohol Abuse Improper Nutritional Habits Smoking Drug Abuse

Table 3

G O A L

To reduce cancer mortality for Utah residents below the 1980 rate of 471.4 working years of life lost per 100,000 population by 15 percent to 406.7 in 1990 and by 30 percent to 330.0 in the year 2000.

Assuming the 1980 ratio of deaths to working years of life list, this will result in a savings of 1,383 future working years for about 282 people who would have died in 1990 and a savings of 3,216 working years for about 656 people who would have died in the year 2000.

As indicated in the Table, inadequacies or problems within the Health Care Organization such as unavailable or inaccessible service accounted for only 12.4% of the early years of life lost while Lifestyle accounted for 44.2%, Environment 17.8%, and Human Biology 25.3% of early years of life lost.

Table 5

MAJOR RISK FACTORS

SMOKING

ALCOHOL ABUSE

DRUG ABUSE

IMPROPER NUTRITIONAL HABITS

EXCESSIVE STRESS

POOR PHYSICAL CONDITIONING

LACK OF SAFETY RESTRAINT USAGE

EXCESSIVE SPEED

GENETIC PREDISPOSITION TO DISEASE

INADEQUATE PRE-NATAL CARE

IMPROPER FLUORIDE INTAKE

ENVIRONMENTAL AND OCCUPATIONAL EXPOSURE

Table 6

OBJECTIVE 1.	GET UTAHNS TO ASSUME MORE INDIVIDUAL RESPONSIBILITY FOR THEIR OWN HEALTH.
OBJECTIVE 2.	IDENTIFY AND APPROPRIATELY REFER UTAHNS WITH GENETIC PREDISPOSITIONS.
OBJECTIVE 3.	ASSURE ALL PROSPECTIVE MOTHERS RECEIVE ADEQUATE PRE-NATAL CARE.
OBJECTIVE 4.	INCREASE THE NUMBER OF UTAHNS CONSUMING APPROPRIATE LEVELS OF FLUORIDE.
OBJECTIVE 5.	REDUCE UTAHNS EXPOSURE TO SUBSTANCES HARMFUL TO HEALTH.

Table 7

DISEASES OF THE HEART,
AGE-ADJUSTED MORTALITY RATES/100,000 POPULATION
UTAH AND UNITED STATES, 1960, 1970, 1980

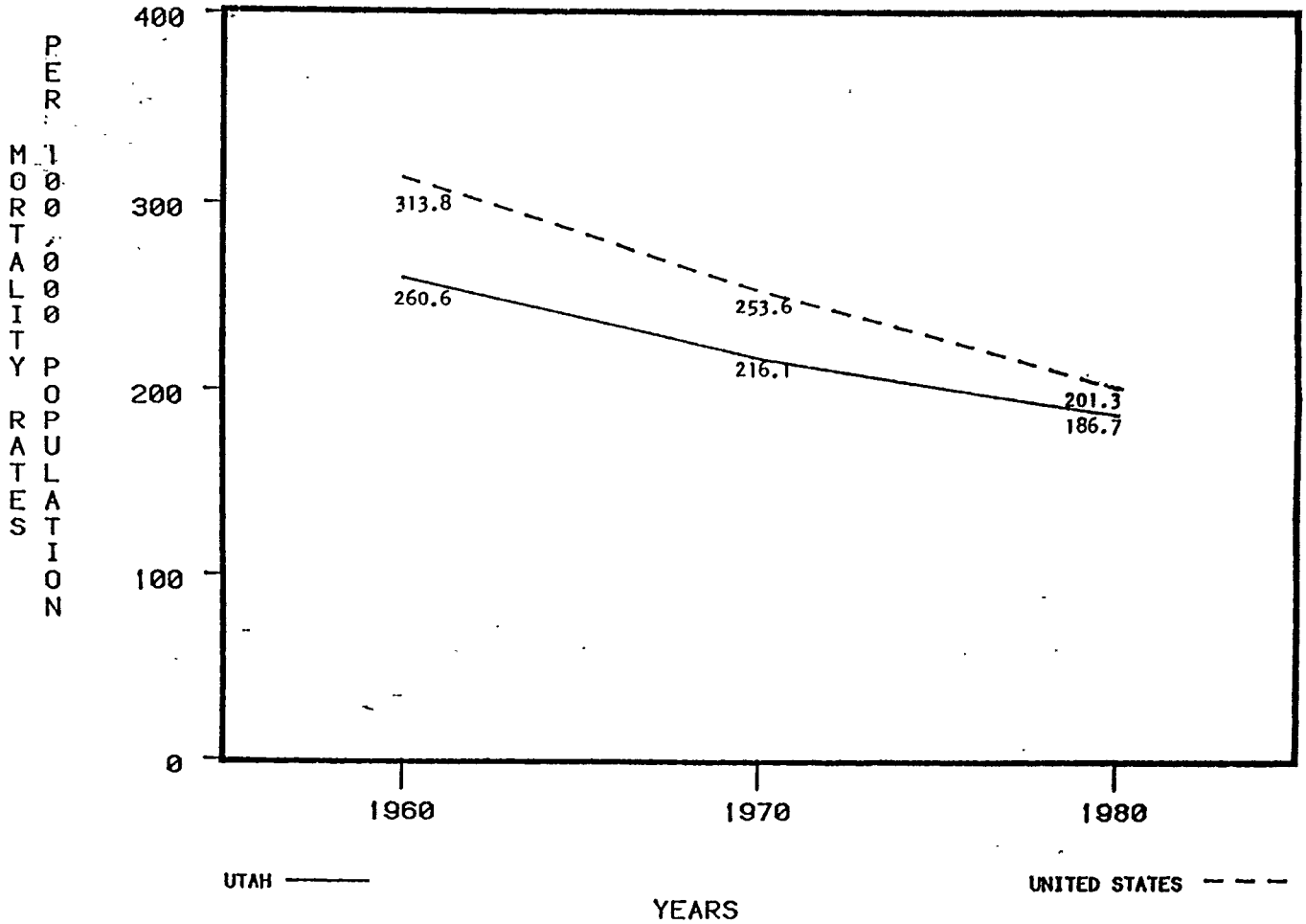
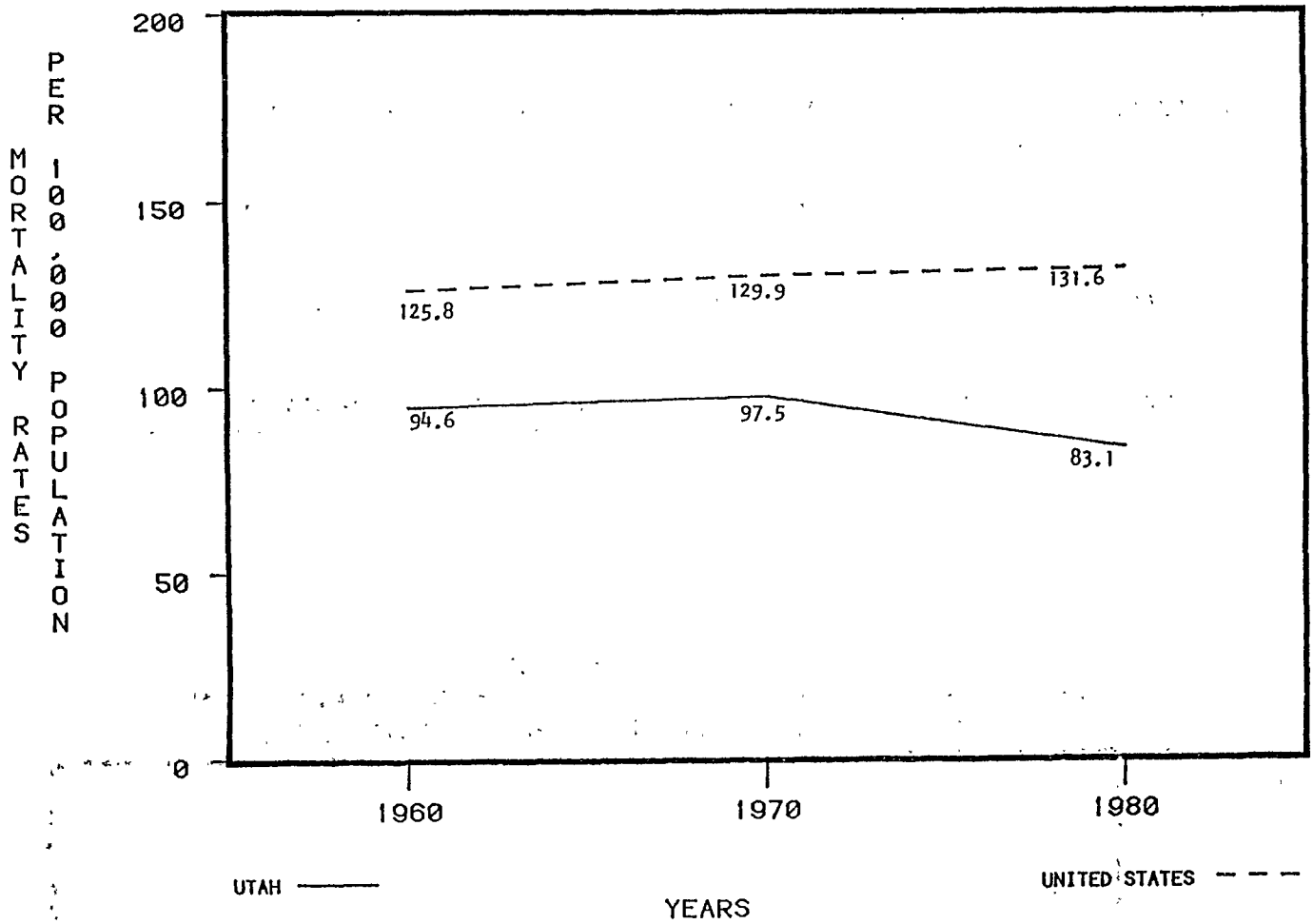


Table 8

MALIGNANT NEOPLASMS, AGE-ADJUSTED MORTALITY RATES/100,000 POPULATION: UTAH AND UNITED STATES, 1960, 1970, 1980



References

1. Fielding, Jonathan E., "Success of Prevention," Millbank Memorial Fund Quarterly, Health and Society, November 3, 1978, Vol. 56
2. Utah Department of Health, A Prescription for Health Status in Utah, May 1982 (Draft), pg 4
3. Ibid, pg 23
4. Dever, Alan, "An Epidemiology Model for Health Policy Analysis," Social Indicators Research 2 (1972) pp 453-666
5. Marc Lalonde, A New Perspective on the Health of Canadians, 1975
6. Healthy People, The Surgeon General's Report on Health Promotion and Disease Prevention, U.S. Department of Health, Education, and Welfare, 1979. Objectives for the Nation, Promoting Health -- Preventing Disease, U.S. Department of Health and Human Services, Fall 1980
7. Dever, Alan, "An Epidemiology Model for Health Policy Analysis," Social Indicators Research 2 (1972) pp 453-666
8. Utah Department of Health, A Prescription for Health Status in Utah, May 1982 (Draft), pg 31
9. Ibid, pg 41
10. Ibid, pg 42
11. Cloud, Daniel T., M.D., "Affirmative Leadership," American Medical Association Interim Meeting, December 6, 1981, Las Vegas, Nevada
12. Utah Department of Health, A Prescription for Health Status in Utah, May 1982 (Draft) pg 47-55
13. Monthly Vital Statistics Report, Vol. 30, No. 13, Department of Health and Human Services, Pub. No. (PHS) 83-1120, December 1982
14. Ibid
15. Utah Department of Health, A Statement of Health Policy: A Prescription for Health Care Costs in Utah, February 1983
16. Foege, William H., M.D., "Prevention and World Health: The Next Two Decades," American College of Preventive Medicine, November 4, 1979, New York City

DEVELOPMENT OF A COMPREHENSIVE CITY HEALTH POLICY AND GUIDE TO RESOURCE ALLOCATION DECISIONS

Gillian Marsden, Seattle-King County Department of Public Health

I am here today to present information on what one city, Seattle, has done to focus local resources on health promotion. To do this I will firstly describe the methodology and process we used to develop a health policy, secondly, show how we have used that policy, and thirdly, I will tell you something of the activities emanating from the policy. I will also provide recommendations for those of you who might be contemplating producing a health policy for your local areas. Before I embark on a discussion of the methodology, I would like to give you a brief overview of what the Seattle Health Policy is and why we wrote it.

Essentially the Seattle Health Policy is an analysis of the health status of Seattle residents and the delineation of a set of health status goals, by age group, for Seattle for 1990. Our Health Policy also includes strategies for achieving those goals and a set of resource allocation criteria. We developed a health policy because city government was struggling with what it wanted the Health Department to do. The mission of a public health department is often extremely unclear to elected officials. Other city departments tend to have self evident missions -- for example they produce water or maintain roads. However, given that the provision of health services is largely a responsibility of the private sector, the role of the public health department is often obscure at best. Therefore, the Health Policy provides the City Council with goals and guidelines for allocating funds to health activities. Furthermore, it enables the Council to understand and determine how health department services should relate to those of other community health resources.

Lest you think the production of the health policy was entirely occasioned by external forces, let me hasten to add that the health department also had a very vital interest in developing a clear policy. We wanted to be sure that our limited resources were focused on those areas where public health services could make a difference to the health status of our community in the 1980's. We wanted our services to be understood both by elected officials and by the community at large. We wanted to forge a more effective partnership between public and private resources. We wanted measurable goals that could be related to nationally accepted objectives.

Now that I have given you a brief overview of what the Seattle Health Policy is and why it was produced, I would like to turn to the methodology used. I should preface my discussion by mentioning that when we set out to write the Seattle Health Policy, we were not allocated any resources above and beyond our existing staff. We had no money for primary data collection. Indeed we had to beg and borrow to finance the extensive copying and printing bills that we ran up during the course of writing the policy. Since we had no money for primary data collection, we reviewed all existing local, state and federal health policy documents. We found that the 1979 Surgeon General's Report on Health Promotion, "Healthy People", was by far the most relevant and exciting of these documents. We felt strongly that the Surgeon General's Report was an important document which already addressed, on a national level, many of the issues about which we were concerned. We also felt that for all its value, the Surgeon General's Report would remain an academic exercise unless its principles were applied by local government working in concert with the private sector. If the Surgeon General's recommendations were to be effective, we believed that they must be implemented at the grass roots level. Therefore we decided to offer the Mayor and the City Council a report to complement "Healthy People" and provide a guide to local action.

The Seattle Health Policy used the same methodology as was used in "Healthy People". We first examined health status by age group for Seattle, and then established overall age group goals. Adjustments were made to fit our local situation. For example, Seattle's children have historically enjoyed better health than the national average, so the recommended goal for Seattle's children is to reduce deaths to fewer than 29 per 100,000 by 1990 rather than the 34 per 100,000 specified as the national goal. We then examined the health status of each neighborhood. As with any city there is considerable variation from the average when specific neighborhoods and minority groups are analyzed. Through our analysis, we identified those neighborhoods and groups which faced major health problems. Based on this work we recommended policies and strategies that the city could adopt to raise the health status of specific neighborhoods, of minority groups and of the city as a whole.

Our next step was to assess programs currently operated or funded by the city in the light of the suggested policies. A survey of nine other cities was conducted to compare the range of public health programs offered in Seattle with services in cities of similar size. Finally, general criteria were developed to guide funding decisions in the immediate future.

While my description of the methodology may have made producing the Health Policy sound relatively simple, it was in fact no easy undertaking. As I am sure you recognize, we had to deal with substantial gaps in the data. When we started work on the Health Policy much of the 1980 census data were not available. Therefore, we had to use population projections. We also had to make assumptions regarding vital data and income because we had no income data on birth and death certificates. We were severely hampered by the fragmented morbidity data that exist at the local level. At least in Seattle, there is no comprehensive compilation of morbidity data. There were tantalizing pieces of information that were compiled by the School District for the city as a whole but these data were not available by neighborhood. Since much of our analysis was on a neighborhood basis, this lack of data was quite frustrating. Where local morbidity data were not available, our only recourse was to use national

survey statistics and apply them to the local level. As a result, we were forced into too heavy a reliance on mortality data. This caused a particular problem in understating such health issues as the need for dental care. We tried to counteract that reliance by numerous caveats and by using what piecemeal morbidity data were accessible.

Turning from the methodology and its limitations, I would like to talk briefly about the process we used to prepare the health policy. Process is as important as methodology if the goal is to produce a policy which is accepted and used by the community. The Seattle Division chose to use a process that combined a small inter-departmental task force with extensive review by professional and community groups and by academic experts. The 12 member task force was composed of health department staff, representatives from the Budget Office, City Planning Office, the Mayor's Office and the Department of Health Services at the University of Washington. Task Force members not only made the overall decisions on format and direction but also did all the staff work for the report. To broaden involvement we compiled a mailing list of all the relevant professional groups, community groups, academicians and individuals that we could think of. The mailing list was continually expanded. As people called and said they wanted to be a part of developing the policy, we welcomed them, put them on the mailing list and sent them policy drafts. We also held community meetings, presented information at meetings of professional societies and conducted periodic briefings for City Council members.

Having listened to my description of the process, you may be wondering, well, did it work? Overall I would say "Yes", but we certainly did not come up with a health policy that was universally accepted. However, people had an opportunity to understand our methodology, give us their comments during the production phase and see, at least, some of their concerns addressed. Hence, I believe that the process built the basis for adoption of the policy by the City Council and for its use in the larger community during the 1980's.

This then brings me to some of the critical decisions that we made regarding the approach and the consequences of those decisions. Firstly, while adoption of the "Healthy People" methodology had the advantage of producing measurable local goals that could be compared to similar national goals, it also had the disadvantage of producing a set of yardsticks that are very hard to use on a year-to-year basis. It is difficult to track progress at the local level due to lack of age, race and sex denominator data. There is very little morbidity data available on a city-wide and neighborhood basis. Further, health services have a limited ability to affect health status, as measured by vital data. The economy and unemployment are likely to have more profound effects on health status than is the provision of specific public health services. Many causes of death (accident, suicide, homicide and alcoholism) are related as much to social conditions as to health services. The Health Policy did not contain a comprehensive analysis of the social factors affecting risk and health status. It is our intent to deal with all of the above issues by monitoring progress and conducting a review and update of Seattle's health status in 1985. At that time the Health Policy goals and strategies will be adjusted. We are also attempting to develop improved morbidity data but we have few resources to devote to this effort.

A second critical decision regarding methodology was our attempt to examine health status by neighborhood and to propose targeting of services on a geographic basis. As you can imagine that approach was well received in certain neighborhoods but thoroughly opposed in others. In Seattle low income and minority people are concentrated in certain areas of the city. Typically, the health status of those neighborhoods is lower than the city-wide average. We believed that targeting certain services to "high risk" neighborhoods would be more likely to help achieve our health status goals than scattering resources more widely. While we tempered our targeting recommendations with overall considerations of accessibility for high risk people who are geographically dispersed, we did not articulate clear guidelines for handling this issue.

Having listened to something of the methodology and process, you may be interested to know how the policy has been used. It has been used in decisions on allocation of funds, development of new services and refocusing existing services toward prevention and health promotion. On completion of the policy and after extensive review, the City Council adopted a resolution which embodied the major features of the document. Subsequently, the City Council has used the policy with almost religious fervor in their annual review of the Seattle Division budget. We are now attuned to expect a battery of questions as to how this service or that position is related to the achievement of the health policy goals. It is, of course, no easy task to respond to these questions. Explaining how tinkering with the microcosm affects the macrocosm, with no hard data on cause or effect, never ceases to challenge the creative mind. However, I should not mislead you into thinking the Council is unaware of how to use the policy. In the three years that have followed the adoption of the health policy we have had to deal with consistent funding cuts for health services. Both the Department and the City Council have found the policy useful in identifying top priorities for service maintenance. A second way in which the policy has been used is to help with decisions on pass-through funds. The Seattle Division acts as a pass-through agency for block grant and other funds. We have to make recommendations to the City Council on allocation of pass-through funds. It has proved useful to have the Health Policy as a common basis for both the Health Department and the City Council to make those resource allocation decisions.

A third way in which the Health Policy has been used is to focus our interest and resources on new services that are likely to effect change and are geared to the objectives laid out in Healthy People. To give you just a few examples, since putting together the health policy the department has developed a program for dealing with car safety for children. We have an Infant Car Seat Loan Program for low income families

that is coupled with comprehensive education on use of car seats. The education and loan program is provided to clients of our maternity and child health services throughout the Health Department and community clinics. We also came to recognize that the Health Department has much latent power for advocacy. Consequently we have mobilized the City and County government lobbyists to work for a child passenger restraint bill. Our efforts, along with those of other groups, culminated this year when the State Legislature finally adopted requirements for the use of passenger restraint systems for children less than 5 years old. Similarly, the Seattle Division has worked with the Police Department and the City Council in promoting handgun legislation and with the State Legislature in advocating for restricted availability of handguns.

Writing the Health Policy and analyzing our services made us recognize that we had almost no services directed toward working adults. Furthermore occupationally related health issues were key factors affecting the health status of adults. Subsequent to the adoption of the health policy we have developed a number of occupational health services that we now provide to the local community. Similarly, our services for the elderly were very limited and not well focused on prevention of disability. We have now reorganized services for the elderly to link health promotion with the provision of primary care. We have also set up education groups for caregivers for the elderly. The groups are led by Public Health Nurses and help those caring for elderly relatives to understand the aging process, the local resources available for the elderly and how to provide maximum health and mobility for their relatives.

Just as Healthy People has been followed by the delineation of specific objectives to help achieve the goals, the Seattle Health Policy has been followed by internal and external planning and evaluation. Within the Health Department we have developed planning and evaluation guidelines that we use for assessing all Seattle Division services. The guidelines require staff to identify specific objectives for each service and relate those objectives to the health problems of high risk neighborhoods or groups and to the achievement of the Health Policy goals. We have also recognized the inadequacies of our internal data collection systems and embarked on a computerized registration and encounter system that will vastly increase our planning and evaluation capabilities.

Looking beyond the Department we are negotiating with the State for a consolidated contracting system. The new contracting system would allow allocation of funds in relation to state and local outcome goals and service objectives, thus putting internal and external planning and reporting requirements on the same basis.

I would like to close with a few thoughts for others who may be contemplating producing a health policy. In 1981 if you had asked me for my recommendation regarding producing a health policy, I would have said "Don't". It was a great deal of work. We had no funds. We had to deal with a lot of community interaction, some of it hostile and some of it negative. However, in the two years that have elapsed since the adoption of the Health Policy by the City Council, I have come to view the availability of the document in a much more favorable light. From the things that I have already said you can see that the policy has proved useful to the Health Department. Thus I have the following suggestions for those considering developing a health policy. Firstly, involve as many segments of your local health system as possible in order to improve the comprehensiveness of the policy and its acceptance by the community. Secondly, use "soft data", for example community need surveys, to reduce the dependence on mortality figures. Thirdly, use geographic boundaries that are consistent with the "real world" neighborhoods and which will be conducive to data gathering and updates for example, census tracts, zip codes and those planning areas used by other branches of local government. Fourthly, if you choose to use geographic targeting to identify highest risk neighborhoods, then clearly set out guidelines under which modifications of that approach should be used. Finally, I would suggest that you read our health policy. You will readily see, both from what I have said today and from reading the policy, the limitations of our approach. It certainly is not an all encompassing document. It has a lot of shortcomings. However, it has proved to be one of the most useful planning tools that we have developed.

Copies of the Seattle Health Policy may be obtained for $7.50 from the Seattle Division, Seattle-King County Department of Public Health, 1500 Public Safety Building, Seattle, Washington, 98104.

Health Promotion in the Work Place

Session A

USING DEATH CERTIFICATE DATA TO CONTRIBUTE TO A STATEWIDE OCCUPATIONAL HEALTH SURVEILLANCE SYSTEM: THE RHODE ISLAND EXPERIENCE

David M. Gute, Massachusetts Department of Public Health
Bruce C. Kelley, Rhode Island Department of Health

INTRODUCTION

There is increasing interest in documenting the effects of occupational and environmental hazards in the expression of disease. Both the U.S. National Committee on Vital and Health Statistics (1) and the National Center for Health Statistics (2) have expressed interest in the adaptation of morbidity and mortality data systems to contribute information on occupational hazards.

In keeping with this interest in occupational health, the National Center for Health Statistics is encouraging states to code the usual occupation and industry of the decedent as obtained on death records. The initiation of a surveillance cooperative agreement program between states (SCANS) to participate in occupational health surveillance activities by the National Institute for Occupational Safety and Health (NIOSH) has also led to an increase in the number of states which routinely code the occupation and industry as obtained on death certificates. (3) These data will enable investigators to estimate the relative magnitude of state-specific occupational health problems as well as to study differential mortality by occupation and industry as pioneered by Guralnick (4) and Milham (5) in this country.

This paper will summarize work undertaken from 1978 to 1982 in collaboration between the Surveillance Branch of NIOSH and the Rhode Island Department of Health (RIDH). This work represents an attempt to fashion a statewide occupational health surveillance system. A model occupational health surveillance system should enable researchers to identify potential health hazards in the workplace and attempt to isolate those industries and occupations which place workers at excess risk. A surveillance system should also have the ability to detect trends and monitor these changes to assess if they are etiologically meaningful. An important component of any surveillance system is the dissemination of information so that it may be used to affect intervention activities and the allocation of public health resources. Ideally a surveillance system should make use of existing data sets thus keeping costs low. Emphasis should also be placed on the use of data sets which are widely available so that a surveillance system would be transferable from one geographic area to another.

Given these attributes of an occupational health surveillance system the first requirement in Rhode Island was to choose appropriate data sets to use. There are a wide variety of data sources which could be adapted to contribute to an improved understanding of occupational morbidity and mortality. For the purpose of this paper emphasis will be placed on the death certificate as a source of data with the wide availability of the death certificate being the most important factor contributing to this decision.

The first use of death certificate data involved the calculation of standardized mortality ratios (SMR) and proportionate mortality ratios (PMR) for white Rhode Island resident decedents $\geq$ sixteen years of age at the time of death. Non-whites were coded but excluded from the analysis as their small representation in Rhode Island's population, 5.6 percent in 1980, precluded analysis(6). In addition, decedents identified as being institutionalized, students, or members of the armed forces were excluded from the analysis.

The occupational and industrial coding system was the same as the system used by the U.S. Bureau of the Census in the 1970 Census (7). Coding progressed at the fullest level of specificity under this system. SMRs were performed for the 16-64 population for the years 1968-1972 using the Census year of 1970 as the mid-point. PMRs were calculated for the following age intervals: 16-34, 35-44, 45-54, 55-64, 65-74, 75-84, and $\geq$85. The expected values for the PMR analysis were obtained from the mortality experience of all decedents included in the study.

The PMR analysis was tabulated for males on thirty occupational groups. For females, the PMR analysis was conducted on twenty-one occupational groups. For both males and females, PMR anlysis was conducted for twenty-two industrial groups. The results of the PMR analysis are reported by sex, age, occupational group, industrial group, and cause of death. The cause of death categories number seventeen. PMRs are displayed for the 16-64 and the $\geq$65 age categories.

The method to assess the significance of the mortality ratios is taken from Bailer and Ederer.(8) This method expresses 95 percent and 99 percent significance factors for an observed value of a Poisson variable to its expectation. This approach is not strictly correct in the sense that an a priori test is being made in an a posteriori situation. However, it does provide a basis for identifying associations between employment and cause of death which would benefit from scrutiny. Because the main emphasis of this analysis is to generate hypotheses for further study, the use of this test is felt to be appropriate.

PROPORTIONATE MORTALITY RATIO CONCLUSIONS

The interpretation of mortality analyses such as the PMR is complicated by certain difficulties. These difficulties have been previously summarized and need not be reviewed except to note that certain key elements (occupational and industrial data and cause of death) can be missing or hampered by uneven validity and reliability for any analysis using

the death certificate as a primary source of information. In interpreting PMR analyses, every effort must be made to keep these caveats in mind. Balanced against these caveats is the desire to use the results of the PMR analyses as one part of an occupational health surveillance system to attempt to define hypotheses for occupational disease research and identify hazards amenable to control technologies. In setting such agendas attention must be paid to the following factors: the magnitude of the association of the occupation with a given disease, the consistency of the association with previous studies, the biological plausibility of the association, the presence or absence of a dose-response relationship, and the presence of suspected etiologic agents in the occupations or industries at excess risk.

Having considered the magnitude and quality of the evidence supporting the association and by assessing the overall public health significance of the association, policymakers can begin to translate the expression of risk generated by occupational health surveillance systems into meaningful action. This action can be translated into either a more accurate characterization of the excess risk borne by workers in certain occupations or into initiatives which move beyond analytic responses and consider control techniques which seek to minimize the risk to workers.

Based upon the results of the mortality analyses carried out in Rhode Island the associations listed in Table 1 represent those most in need of further research or control. Only findings for males are summarized in Table 1. Because the Rhode Island analysis was the first United States population-based PMR study to include women it was thought to be premature to summarize these findings in the manner of Table 1.

Table 1 synthesizes the findings from four different analyses of the Rhode Island data set. They are as follows:

1. PMR analysis 1968-1972 (16-64, $\geq$65)
2. SMR analysis 1968-1972 (16-64)
3. PMR analysis 1973-1978 (16-64, $\geq$65)
4. PMR analysis 1968-1978 (16-64, $\geq$65)

The four occupational associations and the two industrial associations in Table 1 were each found to be significantly elevated ($p<.05$) in at the least 3 of the analyses. They also were generally consistent with previously published findings of similar studies. The finding of an increased colon cancer experience among the Professional and Technical Worker occupational category also satisfied these criteria but this relationship was judged to be confounded by a social class gradient stronger than any occupational risk factor. In addition, causes of death associated with amorphous occupation and industrial categories (e.g., Never Worked) were dropped from this summary.

The associations reported in Table 1 could partially result from the influence of personal risk factors. Given the constellation of conditions, the differential patterns of smoking by occupational group are perhaps the most salient to address.

It is evident from inspecting the smoking rates of the occupational groups listed in Table 1 that two of the groups, Transport Operatives and Construction Craftsmen, have been documented as exhibiting excess prevalence of cigarette smoking.(9) Although no attempt was made in the Rhode Island analysis to adjust for smoking, attempts in the Third National Cancer Survey and Roswell Park studies to do similar adjustments generally left the association between occupation and lung cancer unaffected.(10)

It is also important to note that in the case of Transport Operatives, Construction Craftsmen and the Construction industry important job-related exposures (diesel exhaust, asbestos) must also be evaluated in the expression of disease. The possibility of these exposures acting in concert (either in an additive or multiplicative fashion) with personal risk factors such as smoking must be monitored.

From what is already known about the influence of asbestos in the expression of lung cancer and mesothelioma the further intensification of control measures or substitution of less hazardous substances is warranted. Given the lack of firm evidence with regard to the relationship between diesel exhaust and other petroleum products and lung cancer the emphasis for further activity should be placed on the sharpening of epidemiologic analyses to better understand this association.

The two findings of increased acute myocardial infarction among Managers and Administrators and Policemen would also benefit from further study. In the former example rigorous adjustment for social class could well account for the increased experience of this population. In the case of Policemen the association between this form of employment could well be tied to the personal risk factors which recruits bring with them rather than explicit occupational risk factors. Further study is necessary in pursuing both associations.

The other conclusion which can be gleaned from the PMR analyses is the high probability of positive benefit which could be attained in the aggressive mounting of health promotion programs at the worksite. The importance of personal risk factors in association with occupational factors in contributing to increased risk of mortality can be clearly seen. Although the discrete influence of these personal risk factors is difficult to estimate, attempts at minimizing their impact can only be beneficial.

SENTINEL HEALTH EVENTS (OCCUPATIONAL)

A further use of available data sources has attracted increasing interest and will represent the next section of this paper. This use entails the application of the "sentinel health event" concept as developed by Rutstein (11) to health conditions suspected to be of occupational origin.

To further refine the use of this approach Rutstein and the Working Group on Man Made Diseases in collaboration with NIOSH,

Surveillance Branch undertook the development of a list of Sentinel Health Events - Occupational Disease (SHE/O)(12). This effort generated a group of 50 ICDA-9 rubrics thought to be indicative of possible occupational influence based upon review of published epidemiologic literature and clinical judgment. The RIHD and the Rhode Island Health Services Research Inc. produced data from the Rhode Island vital records system and hospital discharge abstracts system to conduct the initial testing of the SHE/O list. Specifically, deaths and hospital discharges experienced by Rhode Island residents during the years 1974 through 1978 were tabulated by selected SHE/O diagnoses, by age, and by sex. These are presented as Tables 2 and 3.

It should be noted that these are preliminary tabulations which have been produced for the purpose of discovering problems associated with the use of existing data systems for occupational disease surveillance. With this caveat in mind, however, the data in Table 2 are of some interest. Table 2 presents the number of SHE/O diagnoses, other diagnoses, and total deaths and hospitalizations in Rhode Island between 1974 and 1978. Although SHE/O diagnoses represent only 7.5 percent of all deaths and 1.6 percent of all hospitalizations during the period, the direct and indirect costs to society of these 3,400 deaths and 11,450 hospitalizations are large. Table 3 presents this experience in detail for selected SHE/O diagnoses (involving both deaths and hospital episodes) for 1974-1978.

LIMITATIONS OF PRELIMINARY SHE/O DATA

The use of retrospective data to test a list which was developed and specified in the context of ICDA-9 presented problems. First, due to code conversion problems some of the SHE/O rubrics included in the list of Sentinel Health Events - Occupational Disease could not be included in Table 2. In converting death records from ICDA-9 to ICDA-8, some diagnostic rubrics were either broader in ICDA-8 or did not exist in ICDA-8.

Second, these data represent SHE/O diagnoses, not the joint product of diagnosis and occupational exposure. Neither the death data nor the hospital discharge data were screened for occupational/industrial experience to select on decedents or patients having both a SHE/O diagnosis and experience in one of the listed occupations or industries. The hospital discharge data do not include information on occupational or industrial experience. Thus, such screening is not possible. Therefore, the events in Tables 2 and 3 are a superset of SHE/O cases in that they also include people with occupational/industrial experiences other than those hypothesized to be involved in the etiology of true SHE/O cases. To estimate the proportion of SHE/O diagnoses that may be SHE/O cases, death certificates were manually screened to select decedents reported as having both a SHE/O diagnosis as the underlying cause of death and experience in one of the listed occupations and/or industries. This screening was only undertaken for selected SHE/O diagnoses. For several of the SHE/O diagnoses, none of the decedents were reported as having worked in one of the listed occupations and/or industries. With regard to the other selected SHE/O diagnoses, only two to six percent of the decedents were reported as having worked in one of the listed occupations and/or industries. It should be noted that the proportions of SHE/O diagnoses that are cases are probably underestimated because this represents a testing of an initial draft of the SHE/O list and many of the heavy manufacturing industries of interest are not represented in the Rhode Island economy.

EVALUATING THE RHODE ISLAND OCCUPATIONAL HEALTH SURVEILLANCE SYSTEM

Though the positive effects of the SCANS projects to data are evident, weaknesses in the disease surveillance protocols delineated above must be resolved. Further, to implement the NIOSH surveillance strategy, weaknesses in the overall surveillance system must be addressed. These weaknesses of the system include the following:

1. The present surveillance system is overly dependent on mortality data. Mortality data suffer from a variety of deficiencies. The major concerns include the number of events available for study, and the limited possibilities of discovering associations in the face of long latency periods compounded by the inability to intervene after the fact of death.
2. The PMR and SMR analyses employed in population-based mortality analyses can only be effectively used in the generation of epidemiologic hypotheses. They cannot be appropriately used in the rigorous testing of existing hypotheses.
3. The use of hospital discharge data in the SHE/O analysis addresses some of the limitations of mortality data, but is compromised by the fact that it only addresses the portion of morbidity that is treated in hospitals. Further, at this time hospital discharge abstracts do not contain information about patients' work experiences.
4. The SHE/O approach functions by selection of cases from a predetermined list and is therefore a static system. It will not be especially useful in uncovering new relationships between work and health.
5. A main weakness of the existing occupational disease surveillance system is that it has not been linked to a subsequent analytic system which would seek additional epidemiologic data and examine possible associations between work and health. Presently, associations found to exist by population-based analyses can only be followed up on an ad hoc basis.
6. In addition to further analytic investigation, the capacity is inadequate for investigative follow-up of cases of likely occpuational origin. Not only SHE/O cases, but also PMR and SMR associations confirmed by previous analytic investigations should be reviewed by groups of experts. Such review groups should select

cases or associations for further investigation which have a high probability of being of occupational origin. Multi-disciplinary teams (i.e., epidemiologists, practitioners of occupational medicine, industrial hygienists, etc.) could then be deployed to investigate the selected cases.

Given these caveats, the existence of the SCANS project has been a positive influence on improving the prominence of occupational health as a major health problem in Rhode Island. The launching of a similar surveillance effort in other states would likely produce advantages similar to those observed in Rhode Island. These advantages begin with the accumulation of a state-specific body of data which is useful for planning purposes and for the generation of epidemiologic hypotheses linking work and health. Additionally, surveillance efforts at the state level allow for the investigation of possibly unique industries which occur only in some states. Although these industries on the national level may not be sizeable, they may be quite large in specific states. The other main advantage offered by the SCANS approach is that it offers a highly valuable return of useful data for relatively modest expenditure of personnel time and funds. Such efficiencies are best attained through the adaptation of existing data sources for occupational disease surveillance.

Ultimately, the worth of any surveillance system can only be determined through the use of the data. The lack of an effective intervention capacity can reduce a surveillance system to a sterile and hollow exercise. Such development takes time and must be done in a competent and effective manner. Occupational health resources should ideally be allocated on a sound epidemiologic data base. Lacking a data base, such allocation decisions are impossible. The SCANS approach in its most basic sense attempts to provide the necessary data to foster this type of decision-making at the state level.

REFERENCES

1. Statistics needed for determining the effects of the environment on health. Vital and Health Statistics, Series 4, No. 20, DHEW Publication No. (HRA) 77-1457, July 1977.
2. National Center for Health Statistics. Environmental health - A plan for collecting and coordinating statistical and epidemiologic data. DHEW Publication No. (PHS) 80-1248, 1980.
3. Kaminski R, Brockert J, Sestito J, et al: Occupational information on death certificates: A survey of state practices. Am J Public Health 71: 525-526, 1981.
4. Guralnick L: Mortality by occupation and industry among men 20-64 years of age: United States, 1950. Vital Statistics, Special Reports. Vol. 53 No 2, 1962.
5. Milham S: Occupational mortality in Washington state, 1950-1971. Washington, D.C.: GPO, DHEW Publication No. (NIOSH) 76-175-A, B, C, 3 Vols., 1976.
6. U.S. Department of Commerce, Bureau of the Census, General Population Characteristics, 1980 Census of Population. Part 41 Rhode Island PC80.
7. United States Bureau of the Census, 1970 Census of Population. Alphabetical Index of Industries and Occupations. U.S. Government Printing Office, Washington, D.C.
8. Bailer JC, and Ederer F: Significance factors for the ratio of a poisson variable to its expectation. Biometrics 20: 639-634, 1964.
9. Sterling TD, Weinkam JJ: Smoking characteristics by type of employment. JOM 18:743-754, 1976.
10. Dubrow R, Wegman D: Synthesis of the results of occupational disease surveillance studies: Setting priorities for occupational cancer research and control. In press.
11. Rutstein D, Berenberg W, Chalmers T, et al: Measuring the quality of medical care: a clinical method. New England Journal of Medicine 294:582-588, 1976.
12. Rutstein DD, Mullan R, Frazier TM, et al: Sentinel Health Events (Occupational): A basis for physician recognition and public health surveillance. Am J Public Health 73: 1054-1062, 1983.

Supported in part by the National Institute for Occupational Safety and Health.

TABLE 1
Priorities for Control and Further Study
by Occupation and Industry and Cause of Death*

Male Occupation Title Cause (ICD code)	16-64 Significant Findings	≥65 Significant Findings
Managers and Administrators		
Acute Myocardial Infarction (410)	3	2
Construction Craftsmen		
Trachea, Bronchus and Lung Cancer (162)	0	3
Transport Operatives		
Trachea, Bronchus and Lung Cancer (162)	3	1
Policemen		
Acute Myocardial Infarction (410)	3	0
Industry Title Cause (ICD code)		
Construction		
Trachea, Bronchus and Lung Cancer (162)	2	2
Accidents (800-949)	3	0

* Found to be significant ($p<.05$) in at least 3 of the following assessments of Rhode Island mortality data. 1. SMR ages 16-64 (1968-1972)[1], 2. PMR ages 16-64 and ≥65 (1968-1972)[2], 3. PMR ages 16-64 and ≥65 (1973-1978) and 4. PMR ages 16-64 and ≥65 (1968-1978).[3]

Sources: 1,2 Technical Report No. 23 The Association of Occupation and Industry with Mortality in Rhode Island (1968-1972). 3, 4 Available from authors.

TABLE 2

Total SHE/O Diagnosis and Non-SHE/O Deaths and Hospital Episodes of Care: Rhode Island, 1974-1978

	Deaths					
	1974	1975	1976	1977	1978	Total
SHE/O	604	657	708	731	729	3,429
Other	8,551	8,380	8,607	8,475	8,374	42,387
Total	9,155	9,037	9,315	9,206	9,103	45,816
% SHE/O Diagnosis	6.6	7.3	7.6	7.9	8.0	7.5

	Hospital Episodes					
	1974	1975	1976	1977	1978	Total
SHE/O	2,129	2,106	2,095	2,467	2,657	11,454
Other	136,633	137,397	136,286	137,981	135,482	683,779
Total	138,762	139,503	138,381	140,448	138,139	695,233
% SHE/O Diagnosis	1.5	1.5	1.5	1.8	1.9	1.6

TABLE 3

Deaths and Hospital Episodes of Care Attributed to Selected Sentinel Health Events of Occupational Origin: Rhode Island 1974-1978

ICD-9	Condition	DEATHS						Hospital Episodes					
		1974	1975	1976	1977	1978	Total	1974	1975	1976	1977	1978	Total
011,502	Pulmonary Tuberculosis, Silico-Tuberculosis	9	8	7	4	5	33	58	45	47	54	40	244
020	Plague	0	0	0	0	0	0	0	0	0	0	1	1
021	Tularemia	0	0	0	0	0	0	0	0	0	0	0	0
022	Anthrax	0	0	0	0	0	0	0	0	0	0	0	0
023	Brucellosis	0	0	0	0	0	0	1	0	1	0	0	2
037	Tetanus	0	0	0	0	0	0	0	0	0	1	1	2
056	Rubella	0	0	0	0	0	0	21	17	8	7	7	60
070.0-.1	Hepatitis A	3	4	4	3	2	16	114	75	50	45	31	315
070.2-.3	Hepatitis B							36	37	30	47	29	179
071	Rabies	0	0	0	0	0	0	0	0	0	0	0	0
073	Ornithosis	0	0	0	0	0	0	0	0	0	0	1	1
160.0	MN Nasal Cavities	0	3	1	1	2	7	11	12	10	11	12	56
161	MN Larynx	18	23	21	16	18	96	136	88	84	140	109	557
162	MN Trachea, Bronchus, Lung	376	417	460	484	485	2222	687	776	749	904	989	4150
158,163	MN Peritoneum, Pleura	6	4	10	10	8	38	17	16	24	22	21	100
170	MN Bone	13	6	7	19	9	54	36	32	36	32	18	154
187.7	MN Scrotum	1	0	0	0	0	1	1	2	2	1	1	7
188	MN Bladder	61	54	71	56	74	316	437	423	499	543	575	2477
189	MN Kidney & Other Urinary Organs	41	33	33	42	42	191	113	104	109	127	135	588
204	Lymphoid Leukemia, Acute	9	7	8	8	2	34	58	43	53	52	39	245
205	Myeloid Leukemia, Acute	15	20	24	24	18	101	72	60	53	60	88	333
207.0	Erythroleukemia	1	3	1	0	6	11	2	4	7	6	19	38
283.1	Hemolytic Anemia Non-autoimmune	0	2	0	0	0	2	2	1	0	5	1	9
284.8	Aplastic Anemia	4	9	6	2	4	25	8	10	7	14	20	59
288.0	Agranulocytosis or Neutropenia	1	1	1	2	0	5	CCP	CCP	CCP	CCP	CCP	CCP
366.4	Cataract	NA	NA	NA	NA	NA	NA	6	2	0	7	9	24
443.0	Raynaud's Syndrome (Secondary)	NA	NA	NA	NA	NA	NA	16	12	10	9	13	60
500	Coal Worker's Pneumoconiosis	1	0	0	0	0	1	4	1	3	2	0	10
501	Asbestosis	1	0	0	1	1	3	2	0	0	0	0	2
502	Silicosis Talcosis	0	3	2	1	1	7	1	3	5	1	3	13
503	Berylliosis	0	0	0	0	0	0	0	0	0	0	1	1
504	Byssinosis	0	0	0	0	0	0	0	0	0	0	0	0
570, 573.3	Toxic Hepatitis	5	0	0	2	2	9	19	24	19	12	12	86
584,585	Acute or Chronic Renal Failure	28	48	43	45	35	199	181	239	216	279	392	1307
606	Infertility Male	NA	NA	NA	NA	NA	NA	15	11	10	10	9	55
692	Contact and Allergic Dermatitis	NA	NA	NA	NA	NA	NA	75	69	63	76	81	364
	TOTAL	604*	657*	708*	731*	3429*	2106*	2129	2106	2095	2467	2657	11454

NA: Not Applicable CCP: Code Conversion Problems

* Death totals includes 58 deaths due to accidental causes of possible occupational etiology.

THE PENSION COSTS OF WORKSITE HEALTH PROMOTION PROGRAMS

Halley S. Faust, HealthCare of the Bluegrass

Multiple advocates of worksite health promotion programs claim reduced absenteeism, improved health insurance costs, reduced Workers' Compensation claims, reduced disability insurance costs, and increased productivity and life expectancy for employees.[1, 2] Up to 1983, there have been very few empirical studies to support these claims for the worksite setting. In fact, no studies are available determining alteration of health insurance costs to the employer through alteration of risk factors for future disease. There have been few studies which have supported the longitudinal relationship between risk factors and future increased or decreased health insurance utilization. Increased life expectancy claims have been made by projecting alterations in health risk effects on longevity. This is usually done through a health risk appraisal technique. Yet no experimental studies have shown an actual (vs. projected) increase in life expectancy due to alteration in risk factors because of employee health promotion programs.

This paper reports on some theoretical and empirical findings of the effects of the Blue Cross and Blue Shield of Michigan Go To Health (GTH) project on life expectancy as projected by health risk appraisal and subsequent increased pension liability costs.[3]

The effect of increased pension liability costs was addressed theoretically in a paper by Gori and Richter, who found that gradual elimination of minimum preventable portions of major causes of death would increase the number of Social Security beneficiaries by 9.23% by the year 2000, requiring an increased tax rate of 13.55% to support the increased benefits accruing to these numbers of increased beneficiaries.[4] The immediate effects of increased life expectancy would be an increased productivity with gains in the gross national product and government revenues, but with a gradually increased retired population drawing pension fund and Social Security benefits requiring increased taxes. Presumably, a shift in health care costs from shorter-term illnesses to degenerative, longer-term illnesses would occur with a probable hiatus of increased cost for a short period of time. There would be alterations in supply and demand ultimately in the health care sector with a need for more health care workers, more intensive hospital costs due to degenerative diseases, and higher social costs of financing health care.

There are multiple factors which affect pension costs. Some of these factors include eligibility requirements, retirement age, benefits at retirement, death (survivor) benefits, and vesting privileges. By altering the assumptions under one or more of these factors, estimates of the ultimate costs of increased pension liabilities can be made.

METHODS

The Blue Cross and Blue Shield of Michigan Go To Health project was a two-year, quasi-randomized, controlled trial of health promotion at the worksite. Four groups were used: Group D, our strict control group, received no intervention or data gathering except what was available from employer company records and a health attitude and knowledge questionnaire distributed at baseline, at the end of year one, and at the end of year two. Group C received similar treatment as Group D, except they also were provided an health risk appraisal (HRA) (General Health, Inc., Washington, DC) with feedback only in the form of the health risk appraisal printout, a forty-page multi-color, personalized, computerized feedback brochure detailing the effects of risk factors on morbidity and mortality for selected diseases. The third group, Group B, received everything similar to Group C but they were also offered screening at the worksite for blood pressure, weight, height, cholesterol, HDL and blood sugar. They were then provided a feedback session in group format to discuss the results of the health risk appraisal and screening, and methods they could use to interact with community resources to help reduce risks. The fourth group (Group A) was provided the same information as Group B plus individualized counseling on existing risk factors and how they can be modified. Members at high risk were offered group risk reduction programs at the worksite held in a combination of lunchtime and company-sponsored time programs.[3]

The BCBSM pension plan allowed eligibility at age 25 following one full year of employment. "Early" retirement age was 63, and "normal" retirement age was 65. While there were several formulae for benefits available at retirement, the two most frequently used formulae were: a) 2% of "average monthly earnings" multipled times the total years of credited service (up to 30 years) less one and two-thirds percent of the "estimated Social Security benefit" multiplied times the number of years of credited service (up to 30 years); or b) $11 times the number of years of credited service (up to 30 years). "Early" retirement reduced the benefits by .3 percent for each full month that "early" retirement preceded the "normal" retirement date. Depending upon the retiring employee's preference, one of five formulae could be used to determine the payout of benefits at retirement, all related to death (survivor) benefits. Ninety percent of employees used a level income throughout the lifetime of the retiree and the beneficiary. Vesting privileges required ten full years with the company and attaining the age of 55 years.

The origin of the results of the data shown (methods of calculation, statistical data analysis, estimated life expectancy, estimate of probability of spouse survival, estimate of

probability that employee is vested, estimate of earnings at retirement, and estimate of number of years of service at retirement) can be found in the final report of the Go To Health project, available from Blue Cross and Blue Shield of Michigan.[3] The method used to determine the long-term and net discounted value of the pension liability program was as follows: the estimated monthly pension benefit was multiplied by the number of months of increased life expectancy, and then multiplied times the total number of people in the experimental groups to arrive at a net benefit payout if all persons were vested and not survived by their spouse. This amount was then multiplied times the probability of being vested and the probability of not being survived by a spouse to produce a net pension liability of the total employed population likely to reach retirement and be vested. This result was then divided by the number of years over which the extra payments would occur (fifty) to arrive at a yearly total pension cost resulting from increased life expectancy. The final amount was then discounted at 3% in 1979 dollars beginning thirteen years from 1979 and extending sixty-three years from 1979 (the minimum remaining life expectancy was 13 years and the maximum was 63 years).

Decision-makers in industry look at pension liabilities related to the other costs of a health promotion worksite program. Thus, the figures that follow are integrated with the net gains or losses due to changes in absenteeism, disability, productivity, and the net out-of-pocket cost of the Go To Health program. Details of the findings in absenteeism, disability, productivity and out-of-pocket costs can be found in the GTH final report.[3] Absenteeism showed a reduction of approximately 8.11 hours per year per employee for a net savings per year of $86.94. The net present value of absenteeism benefits over the expected tenure of the employee with the company was $556.79. The disability cost savings per person (reduced long-term absenteeism) was 9.15 hours per employee. The net present value of benefits over the tenure of the average employee was $494.08. There is no net change in health insurance utilization over two years. Productivity gains in the first year of the study (1980) showed a 10% increased productivity for Group A as compared with Groups B or C; this was statistically significant at the .05 level. By the second year this productivity gain, while the magnitude remained the same, lost statistical significance. The total cost of the project excluding research and evaluation costs, but including capitalizing the start-up costs, was $107,561 (1979 dollars discounted).

RESULTS

The empirical findings of the Go To Health study compared Group A with Group C (the health risk appraisal/life expectancy control group). There was an increase in life expectancy of 1.02 years. The estimated average monthly earnings at retirement in 1979 dollars was $1,725 per month. The estimated average total numbers of years of service at retirement was 21.07; the probability of being vested was 0.983. The probability of not being survived by a spouse was .6187.

A summary of the measurable benefits and costs for Group A is shown in Table I.

Table I

Summary of Measurable Benefits and Costs For Group A, by Average Wage Values

BENEFITS	Group A (Average)
Aggregate Absent Hours	$ 190,979
Disability Hours	169,471
Productivity	*
Life Expectancy	(539,010)
TOTAL BENEFITS	(178,560)
COSTS	(107,561)
NET BENEFITS	(286,121)
Number in Group	343
NET BENEFITS PER PERSON	$(834)

Numbers in parentheses are negative.

*No net statistically significant marginal benefit could be defined for this category for this group.

The net positive benefits were derived from aggregate absent hours and long-term disability hours ($360,450). This was considerably offset by the total life expectancy cost of $539,010 and the cost of the project of $107,561. There is a negative net benefit of $834 per person, a value which would have been totally reversed to a positive $252,889 or $737 per person if there had been no net increase in life expectancy.

Because the life expectancy figure used does not assume any changes in the pension factors related to the increased life expectancy, a sensitivity analysis was done. I hypothesized that an increase in life expectancy may: 1) increase employee time with the company which would have a net effect of increasing savings in absenteeism and disability, increasing contributions to the pension plan and increasing the retirement age yielding increased average monthly earnings and increased years of service at 10%, and increasing the net benefit from a negative $834 per person to a negative $1,182 per person; 2) increase the probability that the employee would be vested, altering the current assumption of .983 probability of vestment to .99, increasing the net negative benefit per person by $10; 3) increase the probability that the employee would not be survived by the spouse. An increase in 10% of survivability compared to the spouse would increase the net liability by $109 per person.

There has been a suggestion that a short-term health promotion program may not confer long-term benefits; the benefits may decay at

some fixed rate or at some variable rate. Assuming a fixed rate of decay would simply require us to increase our discount rate to 3% plus whatever decay rate seems appropriate. Thus a 3% per year decay rate would have an effect of discounting our figures by 6%. Given a decay of absenteeism, disability hours, and life expectancy of 3% per year, our net benefit per person dropped exactly $800 to a negative $43 per person.

Finally, how would a productivity increase offset the pension liability? Productivity appears to be the most powerful factor which could be increased in these calculations because benefits accrue every minute of every working hour. Instead of assuming a 10% productivity gain as was found in the first year and was sustained in the second year (but losing statistical significance) I assume a 1% net increase in productivity for the tenure of the employee with the company (an estimated 7.65 years following the end of the study). The net present value of this productivity savings would be $440,906. Table II shows the effect on the cost benefit analysis.

Table II

Adding Productivity to the Summary of Measurable Benefits and Cost for Group A Using Average Wage Values

BENEFITS	Group A (Average)
Aggregate Absent Hours	$ 190,979
Disability Hours	169,471
Productivity	440,906
Life Expectancy	(539,010)
TOTAL BENEFITS	262,346
COSTS	(107,561)
NET BENEFITS	154,785
Number in Group	343
NET BENEFITS PER PERSON	$ 451

The net change in net benefits per person because of productivity is an increase of $1,285 (from negative $834 to postiive $451).

DISCUSSION

The Go To Health project shows the importance of estimating the long-term pension liabilities related to worksite health promotion programs. A simple increase in life expectancy of 1.36 percent caused an estimated ultimate liability in the GTH promotion program of $834 per participant assuming no decay of the effects of the two-year project. If the project were to continue for the tenure of the employee with the company, the ultimate costs of the project would continue to grow. If this continuation of the project was necessary to prevent further decay of benefits, then the net negative benefit to the company would have been even larger. If there were a decay in benefits over some period of time so that by the time the employee reached spouse survivability, or increases in life expectancy were lost, then there would be a positive net benefit to the company of $737 per employee.

The BCBSM pension program is a defined benefit program. If it were a defined contribution program, then the increased life expectancy would have no net effect on pension liability because only a pre-defined amount of money could be paid out to the retiree or his/her beneficiary.

Finally, if productivity increases accrue (in the BCBSM project a longer follow-up period and/or larger productivity measurement group could settle this question) there would be a significant impact on the net cost benefit analysis which would easily help to offset a defined benefit pension program net liability due to increased life expectancy.

Acknowledgements

George Ferrell, Ken Warner, Doug Vilnius assisted in the original GTH benefit-cost analysis. Monica Kroft typed the manuscript.

References

1. Berry, C.A. An Approach to Good Health for Employees and Reduced Health Care Costs for Industry. Washington, D.C.: Health Insurance Association of America, 1981.

2. Kristein, M.M. Economic Issues in Prevention. Preventive Medicine 1977; 6:252-264.

3. Blue Cross and Blue Shield of Michigan. Go To Health Project: The Final Report as Submitted to the W.K. Kellogg Foundation. Detroit, MI, 1983.

4. Gori, G.B., Richter, B.J. Macroeconomics of Disease Prevention in the United States. Science 1978; 200:1124-1130.

Measurement and Analysis Issues in Health Promotion

Session B

THE SURVEY AS A SOURCE OF HEALTH PROMOTION DATA

Thomas Stephens, Canada Fitness Survey

Introduction

This presentation is probably more aptly titled "The Household Survey of Fitness and Physical Activity as a Source of Health Promotion Data." It goes without saying that surveys, in the generic sense, have been indispensable in providing useful data for planning and evaluating health promotion efforts. This presentation, therefore, is more modest in its ambitions: in the next 20 minutes, I will review some of the findings of eight population surveys on fitness and exercise, point out some implications for health promotion, and offer some guidance to those of you planning or even considering a survey on these topics.

Between 1972 and the present, fitness and exercise have figured as topics in countless surveys; eight of these will be considered today. They were chosen as being national in coverage (four each in the US and Canada) and because the resulting data are (or will be) available in machine-readable form. Table 1 identifies the surveys, their sponsors, dates and methods of data collection, sample size and age coverage.

A Brief Review of Findings

We now have amassed enough data that answers to four key questions should be possible:

1. How physically active is the population in its leisure time?
2. Who is active? Who is not?
3. How have activity levels changed over the years? Are they increasing, as observation suggests?
4. What are the attitudes, motives and perceptions associated with physical activity?

Activity Levels

Just as defining the number of smokers or seat-belt users is fundamental to planning and evaluating anti-smoking or buckle-up campaigns, so it is essential to establish how active the population is. Ignoring for the moment the fact that our eight surveys are spread over 10 years, we find that anywhere from 15% - 65% of the population is classified "active", depending on the definition used. On the face of it, this is not very helpful. However, when the more stringent definitions are set aside, there is a good deal of consensus that 50% - 60% of the population is physically active at least occasionally. This figure is halved when regular weekly frequency is required (see Table 2).

The implication of these numbers, however rough they may be, is that a large majority of the population is probably not yet active enough to achieve any health benefit. A continued promotion of fitness, and of regular activity, seems warranted.

Participant Profiles

Considering all the variations on the theme of "active" which appear in Table 2, it is perhaps surprising to find a good deal of consistency in the profile of the active American or Canadian. The typical active person is: young (1,2,3,4,5,8), with a higher income (1,2,4,5,8), educated (1,2,8), a professional or manager (2,3,8), and residing in the Midwest or West (2,4,8), or the suburbs (4,5). None of the surveys contradicted this profile, although some did not report some of these characteristics. Gender, incidentally, does not appear to distinguish the active from the inactive. The Canada Fitness Survey, one of the more recent of the surveys, reports that males and females are equally likely to be classified active (57% and 55%, respectively).

The implication for health promotion planning is that the older population (50 and over), blue collar and lower-income groups are not participating in the fitness movement. For reasons of health benefits alone, it is important that these groups be reached and it will evidently require a special effort.

Trends Over Time

Quite apart from different definitions of participation, it is next-to-impossible to establish good trend data with most of these surveys, given their variations in sample composition and data collection methods. This makes the few genuine trend studies all the more valuable. The PARTICIPaction polls, for example, report an increase of 12% in the active population, based on identical techniques with equivalent samples over three years, while the two NSPHPC waves, separated by only one year, uncovered virtually no change. The only other surveys capable of comparison are the 1976 Fitness and Amateur Sport survey and the 1981 Canada Fitness Survey, which employed similar (but not identical) methods and samples. This comparison shows an increase of 18% in sport participation and no change in exercise activities.

These meagre sources, supplemented by other data, e.g., on equipment sales (9), indicate an increase of potentially sizeable proportions in the active population. The implication of this rolling fitness band-wagon for health promotion is as follows: efforts to promote fitness activities will probably appear successful, and conclusive evaluation of these efforts will definitely require the use of control groups.

Attitudes and Perceptions

Documenting the level of participation in physical activities and profiling the active person are the essential building blocks of a useful data

base on exercise and fitness. Necessary as they are, these topics fall short of providing health promoters with the insights to develop effective programs. Nevertheless, probing the motivations, attitudes, perceptions and beliefs underlying active vs. sedentary use of leisure time is rare. The exceptions are instructive. For example, the President's Council survey reported that 57% of their sample believed they were getting enough exercise, and that this belief was more likely to be held by the older respondents, who were, in fact, less likely to be active. This same survey revealed that the US Government was the most frequently cited source of fitness information, while the General Mills study found the government to be the least often mentioned source of health information. There may be a message here for the professional pamphleteers.

In a similar vein, two surveys (2,8) found that doctor's orders were the least likely reason for being active, but also the only reason to increase in importance with age (8), while Perrier found doctor's orders were most likely to be effective with inactive people. There are clear implications here for increasing the role of physicians in the promotion of healthy lifestyles. Finally, the relatively common (2,4,5,8) finding that perceived lack of time due to work is the main obstacle to activity has implications for employers who could alleviate this with flexible work scheduling and on-site fitness facilities.

Future Considerations

With this very brief sampling of results in mind, let's turn to the prospect of future surveys on physical activity, of which there are certain to be many, as the fitness boom itself moves into high gear.

In this regard, there are two issues to be considered:

1. What important questions can be answered with existing data files?
2. What important changes to past methodological practice are needed?

Secondary Analysis

Compared to all the questions which a new survey can answer, and all the exciting prospects presented by collecting new data, no question is as important and no prospect is as challenging as the following: Is a new survey really needed? Can existing data answer my questions at least satisfactorily, if not perfectly?

Unquestionably, there is a large number of important practical issues that can be fruitfully examined with existing data. Indeed, if the frustrations of working with other people's data can be tolerated, there are some real benefits to be expected from secondary analysis, in addition to economy and immediate availability of data. Foremost is the robustness of conclusions which do emerge, as the analysis of one issue in several data sets is akin to a multi-trait/multi-method approach. For example, if higher SES is found to be associated with higher activity levels in six different surveys over a ten-year period in two countries with various data collection methods, it is reasonable to conclude that this is a genuine relationship.

Some other outstanding questions which would benefit from this approach are the following:

1. Are women as active as men when intensity of activity is taken into account?
2. Does work activity compensate for inactive leisure time for the blue-collar worker?
3. What underlies the east-west differences in activity level?
4. What role do parents' activity patterns play in the formation of youthful habits of leisure-time use?
5. Do physical activity, smoking, alcohol use, dietary habits and other health behaviors form discernible patterns?

All of these questions can be examined with at least two data sets and many with three or more. There are even four surveys out of the eight reviewed here (2,3,5,8) which collected data from several members of the same household, thus allowing for examination of intergenerational influences.

Two important questions can be answered only by secondary analysis, requiring, by their nature, two or more existing data sets:

1. What have been the real changes in population activity levels over the last decade?
2. How has the profile of the active American or Canadian changed during this period?

The only important class of questions which cannot be satisfactorily answered by secondary analysis of these survey data sets involves the causal links between and among activity, fitness, health status and the use of health care services. Only the Canada Health Survey measured health status in a comprehensive manner, along with activity and (cardiovascular) fitness, and these data are strictly cross-sectional.

But even this class of question can be fruitfully, if not conclusively, examined by linkage of data sets. For example, linking the records of either the Canada Health Survey or the Canada Fitness Survey to the death index of Statistics Canada or to provincial health insurance files would provide a prospective design for studying the relationship of activity and other health-related behaviors to morbidity and mortality (10).

A Definitional Problem

Finding answers to these questions by means of secondary analysis will be hampered by the lack of a consistent definition of "participant". As noted in the review of findings, great variations in this definition have resulted in estimates of the active population which vary by a factor of four or more. Nevertheless, secondary analysis can produce some uniformity of definition by creating an index of activity (11) based on type, frequency, intensity and duration of participation. This approach can be applied to five of the eight surveys (1,2,3,4,8).

In addition to its relevance to secondary analysis, no issue is an important for the design of future surveys on exercise and fitness as the definition of activity. Several points should be observed in the design of new surveys:

1. Identify for the respondent what "activity" means. Does this include or exclude work, housework, school activities, farm chores? (The CFS covered all of these but identified them separately so we could focus on leisure-time activity.)

2. Do not constrain the choice of activity to be reported without careful consideration of previous surveys. (The CFS used a flashcard with 120 activities listed and, as a consequence, found that gardening ranks fifth and popular dance eleventh of all leisure-time physical activities. Many earlier surveys overlooked these pursuits.)

3. The average duration and frequency of each activity done during the reporting period should be recorded, and the intensity determined if space permits. (The CFS measured all three characteristics for activities done regularly or recently and omitted intensity for those activities which were more infrequent.)

4. If seasonality is not a concern, use a short recall period (e.g., two weeks). Otherwise, a longer period is essential even if some detail is lost. (The CFS went as far back as one year, but collected less detail than for activities which were done weekly and/or within the previous month.)

All of these recommendations are directed at surveys intended to probe activity in some depth. As a guideline, this level of detail on physical activity is equivalent to determining frequency and amount of tobacco or alcohol use.

Generalized recommendations for polls and surveys where space is at a premium are more difficult to provide; these will depend on the sponsor's purposes. If the intent is to acquire a general indication of leisure-time physical activity, a two-part question should suffice:

1. "On average during your leisure time, how often do you engage in any physical activity (e.g. brisk walking, bicycling, swimming, jogging, vigorous gardening or others)?" (daily/4-6 times per week/2-3 per week/once per week/2-3 times per month/monthly/less often)

2. "On average, how much total time per week do you spend actually doing these activities?" (___hrs. ___min.)

In both their long and short versions, these questions will yield an index of activity by multiplying total time by a constant representing the energy cost of the activities (an average value for the short version). Active/moderate/sedentary or other labels can be arbitrarily attached to scale values according to analysis requirements. Even more simply, levels of activity can be defined on the basis of hours per week, by treating energy cost as a constant (which of course it is not).

Even this simplified approach can yield results. One CFS definition of "active", as shown in Table 2, used total time per week with an indicator of year-round consistency. "Active" individuals had participated in physical activities a minimum of three hours per week for at least nine months. The "sedentary" had participated for fewer than three hours weekly and for fewer than nine months. Using this definition, the active were distinguished from the sedentary in the expected direction on several dependent variables: five fitness measures, three health status measures, three health behaviors, the importance of regular activity and other health habits, readiness to become more active, and perceptions of obstacles to activity.

Conclusion

What can we conclude about the (household fitness and exercise) survey as a source of health promotion data? Experience to date suggests that some useful information has been accumulated concerning participation levels, profiles and changes over time. Observation suggests that there will be many more future surveys on this topic, due more to interest in the fitness phenomenon than to any particular need for more data. If new surveys are to be launched, however, it should be only after attempts to answer the outstanding questions via secondary analysis and a careful consideration of the definition of active participation in physical activity.

REFERENCES

Principal Sources

1a. National Adult Physical Fitness Survey. Newsletter, The President's Council on Physical Fitness and Sports, May 1973 (Special Edition).

1b. President's Council on Physical Fitness and Sports, National Adult Physical Fitness Survey. Physical Fitness Research Digest, 1974, 4 (2), 1-27.

2a. Statistics Canada. Culture Statistics/ Recreational Activities 1976. Ottawa: Minister of Supply and Services, 1980. (Catalogue No. 87-501, occasional).

2b. Statistics Canada. Participation in Downhill and Cross-country Skiing, October 1976. Ottawa: Minister of Supply and Services, July 1979.

2c. Statistics Canada. Participation in Ice Skating and Ice Hockey, October 1976. Ottawa: Minister of Supply and Services, October 1979.

2d. Statistics Canada. Participation in Jogging or Running, and Bicycling, October 1976. Minister of Supply and Services, April 1980.

2e. Statistics Canada. Participation in Tennis and Golf, October 1976. Minister of Supply and Services, July 1981.

3. Statistics Canada/Health and Welfare Canada. The Health of Canadians: Report of the Canada Health Survey (ch. 3). Ottawa: Minister of Supply and Services, 1981. (Statistics Canada Catalogue No. 82-538E).

4. The Perrier Study: Fitness in America. New York: Perrier-Great Waters of France, Inc., 1979.

5. Family Health in an Era of Stress. Minneapolis: General Mills, Inc., 1979.

6a. U.S. Department of Health and Human Services, National Center for Health Statistics. Highlights from Wave I of the National Survey of Personal Health Practices and Consequences: United States, 1979. Vital and Health Statistics, Series 15, No. 1, June 1981.

6b. U.S. Department of Health and Human Services, National Center for Health Statistics. Basic Data from Wave I of the National Survey of Personal Health Practices and Consequences: United States, 1979. Vital and Health Statistics, Series 15, No. 2, August 1981.

6c. Revelyn K. Eisenstadt and Charlotte A. Schoenborn. Basic Data from Wave II of the National Survey of Personal Health Practices and Consequences: United States, 1980. National Centre for Health Statistics Working Paper Series No. 13, October 1982.

7. Behaviour and Attitude toward Physical Activity among Canadians. Results of a National Survey Conducted for PARTICIPaction by Contemporary Research Centre Ltd., May 1982 (mimeo). PARTICIPaction, 80 Richmond Street West, Suite 805, Toronto, Canada M5H 2A4.

8a. Fitness and Lifestyle in Canada. Ottawa: Canada Fitness Survey, May 1983.

8b. Fitness and Aging. Ottawa: Canada Fitness Survey, November 1982.

8c. A User's Guide to CFS Findings. Ottawa: Canada Fitness Survey, January 1983.

Other references

9. Statistics Canada. Recreation equipment ownership 1974-1982. Travel-log, 1982, 1 (4), Catalogue No. 87-003.

10. F. White. The Canada Fitness Survey: Implications for health research and public health practice. Canadian Journal of Public Health, 1983, 74 (2), 91-95.

11. A.D. Furrie and T. Stephens. Energy Expenditure Patterns in the Canadian Population. Presentation to the annual meeting of the Society for Prospective Medicine, Québec, Québec, October 21, 1982.

TABLE 1. EIGHT SURVEYS OF EXERCISE AND FITNESS

	Survey/date age coverage	Sample size/ methods	Data collection	Remarks
1.	National Adult Physical Fitness Survey/President's Council on Physical Fitness and Sports (1972)	3875 age 22+	one interview per household	really an activity survey, despite the name
2.	Survey of Fitness, Physical Recreation and Sport/ Fitness and Amateur Sport Canada (1976)	70,000+ age 14+	self-completed questionnaire	also an activity survey, modelled in part on the President's Council survey (i.e. same flaws in distinguishing 'sport' and 'exercise')
3.	Canada Health Survey/Health and Welfare Canada, Statistics Canada (1978)	13,507 households 26,388 questionnaires, age 15+ approx. 6000 fitness tests, age 15-64	interview, self-completed questionnaire, physiological tests, blood analyses	included an inventory of recreational activities and physical household chores, and a step-test of oxygen uptake
4.	Fitness in America Perrier/1978	1510 age 18+	personal interview, telephone interview of runners	a very comprehensive look at activities and related motivation
5.	American Family Health Report/ General Mills (1978)	1254 families 2181 interviews age 12+	interview	section on exercise as part of an inventory of lifestyle behaviors
6.	National Survey of Personal Health Practices and Consequences/ NCHS (1979,1980)	Wave I-3025 Wave II-2436 age 20-64	telephone interview	minimal activity questions, but one of only 2 surveys to be repeated
7.	PARTICIPaction polls (1979,1982)	1982-2000 age 15+	personal interview	little detail beyond activity levels, but also a repetition yielding rare trend data
8.	Canada Fitness Survey/Fitness and Amateur Sport Canada (1981)	11,884 households 21,568 questionnaires age 10+ 15,519 fitness tests age 7-69	self-completed questionnaire,fitness tests, anthropometry	the largest sample ever to have both fitness and activity measured, includes other lifestyle behaviors, designed for repetition every 5 years

TABLE 2. ACTIVE PARTICIPATION DEFINED

Survey/date	Key definition of "active"	% reported "active"
President's Council/1972	now doing at least one of 6 listed activities	55%
Fitness and Amateur Sport/ 1976	(a) any exercise activities in last month	59%
	(b) any sport activities in last 12 months	50%
Canada Health Survey/1978	score 3000+ on index incorporating frequency, intensity, duration	36%
Perrier/1978	(a) participated on a regular basis any time during the year	59%
	(b) "high active" based on energy expenditure index	15%
General Mills/ 1978	planned physical exercise at least several times per week	36%
NSPHPC/ 1979, 1980	often take long walks (highest of 7 listed activities)	37% (1979) 37% (1980)
PARTICIPaction/ 1979,1982	physically active a minimum of 2 or 3 sessions per week	25% (1979) 37% (1982)
Canada Fitness Survey/1981	(a) any exercise activities in last month	58%
	(b) any sport activities in last 12 months	68%
	(c) participated on average 3 hours/week for 9 months of last 12	56%
	(d) energy expenditure index	(forthcoming)

TELEPHONE INTERVIEWING AS A TECHNIQUE FOR COLLECTING BEHAVIORAL RISK FACTOR DATA THE TEXAS DEPARTMENT OF HEALTH EXPERIENCE

Harrold P. Patterson, Ph.D. , Texas Department of Health

Background

During the summer of 1982 the Texas Department of Health conducted a Behavioral Risk Factor telephone survey to obtain information on the lifestyles of Texas residents. This survey was noteworthy because it is the first time the relatively new social science survey technique of telephone interviewing had been used by the agency. It also represented a departure from the manner in which the Department usually obtains information about the health of people in the state.

This event reflects, in many aspects, the change in relative importance of chronic versus infectious diseases processes to public health. As the major causes of death and disability have changed since 1900, many public health agencies have been left with mechanisms for obtaining and maintaining data on the reportable, infectious diseases, yet with little or nothing for the chronic diseases. As the impact of chronic diseases increases, planning and implementing ameleortive measures becomes more problematical without adequate data.

The Texas Behavioral Risk Factor Survey was planned and developed during May and June of 1982, with data collection occurring during a four-week period in July and August. Data processing and report writing occurred during September-December. A total of 1,840 Texas residents aged 18 and over were randomly selected for a 10-15 minute telephone interview. Eight topics were addressed in the survey. These were:

1. Seat belt usage
2. Hypertension
3. Physical activity
4. Weight
5. Cholesterol
6. Stress
7. Cigarette smoking .. and
8. Alcohol use and abuse

The survey was a part of the Centers for Disease Control's effort to establish a nationwide base of risk behavior data.

Although the results of the survey are valuable and quite interesting, this is not the focus of this paper. There are reports available which describe the survey's major findings and its methodology and administration. Instead, what I would like to discuss today is the experience we had doing this survey. Since this was a new experience for the TDH, I thought that perhaps others might be interested in how we went about doing the survey, what difficulties and successes we had, and how such a survey was conducted by a large, bureaucratic organization.

The Decision to Survey

The Texas Department of Health currently employs over 4000 people statewide and administers a 200 million dollar plus budget. Like many other state health agencies, it originated from efforts to control epidemic diseases which were wide-spread in a largely unsettled, frontier territory. Since the beginning of the twentieth century its main interests have been the prevention of outbreaks of epidemic diseases and the maintenance of vital records. Over the years, however, many other tasks have been added to its basic purpose. Such things as monitoring industrial radiation; hazardous waste surveillance; control of health hazards in the workplace; shellfish sanitation; food and drug inspection; licensing, regulation, and inspection of health facilities; nutrition services; and emergency medical services accrediation have all been added to its activities.

It has become increasingly apparent to the Department that many of its surveillance systems no longer provide information about what currently affects the health of most of the residents of the state. The systems which feed in data about "reportable" diseases and conditions such as veneral disease, tuberculosis, the vaccine-preventable diseases (diptheria, pertussis, tetanus, poliomyelitis, measles, rubella, and mumps) and others are not much help when most of the people are dying of heart diseases, malignant neoplasms, cerebro-vascular degeneration, and some form of accident or violence. Some progress has been made by starting a cancer registry program and cooperative ventures with foundations and voluntary organizations. But there is a long way to go before anyone has definitive information about current health problems. This situation has not, of course, gone unnoticed. Attempts are made to deal with it by making synthetic estimates or by extrapolating information from other areas. But this "gap" in information remains.

Given this situation, the Department was ripe for an overture from the Centers of Disease Control, DHHS, in Atlanta, Georgia, to consider conducting a Behavioral Risk Factor Survey. A meeting was held in April of 1982 to discuss the feasibility of the TDH undertaking such a survey. The meeting was attended by representatives of the CDC, Center for Health Promotion and Education, and by TDH personnel representing the chronic disease, health planning, vital statistics and data processing programs. This meeting resulted in a decision to actively pursue the setting up of a mechanism to organize and conduct such a survey. A proposal outlining why it was needed, what would be done, and what would be achieved was prepared for the TDH decision heirarchy. Considerable support was obtained from the Commissioner of Health and the Deputy Commissioner, the highest levels of the decision heirarchy. Without this support it is doubtful that the survey would have even gotten off the ground. Their support also had the effect of cutting through myriads of red tape and other bureaucratic obstacles. It is not irrelevant to point out that the Commissioner and the Deputy Commissioner were

among those who were most acutely aware of the information gap mentioned earlier.

A number of technical or procedural difficulties had to be overcome to initiate the survey. The first was availability of funds. The CDC's method for conducting the survey involved the use of mainly in-house personnel, equipment, and resources. One direct expense, however, involved the hiring of part-time people to actually do the telephone interviewing. Once a source of funds was found for this, the preparations began in earnest. Another problem had to do with the technique of telephone interviewing itself. Although this technique is relatively established as a data collection method in the social sciences, it was viewed as somewhat suspect by many TDH people. It was seen as something political pollsters use to obtain rather sketchy attitudinal data. Also, another large state agency in Texas had conducted a needs survey using the personal interview technique and involving very large expenditures, and there was concern over whether or not the TDH could afford to get into information gathering of this type. However, after checking with CDC expert consultants and with TDH staff experienced in this area, these fears were pretty much dispelled.

Another issue which developed, but was not really a difficulty, was that once other state agencies discovered the TDH was planning a survey, they wanted to participate too. Several offered to provide funds if we would consider adding their questions to the survey. These interests had the effect of further convincing the Department that such a survey was very much needed. As it turned out, the decision was made to conduct a limited survey and not to add additional questions. It was felt that the experience of one survey was needed before more complex work could be attempted. And, by adopting the CDC questionnaire unchanged, we would be able to compare Texas data with other states.

Organizing the Survey

One of the first steps taken to conduct the Risk Factor Survey was to draw together from various parts of the Department individuals who were interested, who had experience with surveys, and who could contribute time to the project. As it turned out, there were several people who were interested, but only a few with surveying backgrounds. Eventually, two primary groups emerged during the development stage: a technical group and an administrative group.

The technical group had those people with experience in survey work. Their functions involved sample design, questionnaire development, computer programming and operations, interviewer screening, organizing and supervising the data collection effort, overseeing data processing, and other such technical detail. The administrative group was composed of those not experienced in survey work but with considerable expertise in the interworkings of the Department. Their functions were to obtain the necessary clearances, arrange for telephone lines to be available, find the necessary space to work, take care of paperwork, make financial arrangements, develop work schedules, and other administrative tasks. Both groups interacted with CDC depending upon the nature of the communication.

The ability of the two groups to achieve a completed survey depended greatly on their efforts to cooperate. A great deal of effort went into the problem of making the two groups function in concert. Occasionally, disparities would arise, but most of these were resolved.

Conducting the Survey

Since the survey was conducted in a public agency and not in a research organization, certain adjustments had to be made. Originally, we had planned to work on the survey in the evening, letting the staff draw compensatory time and not relieving them of their regular work tasks. It did not actually work out that way. The planning and development of the survey necessarily had to be done during normal working hours because most of the decision-makers were available only then. Even during the data collection phase, where the interviewing was scheduled from 6 to 9 in the evening, we found that it was necessary to have several people do survey work during 8 to 5 in order to prepare for the coming evening's work and to provide continuity for the study. We ended up with several people putting in 10 to 12 hour days, plus weekends, so that their regular work and the survey work could both be done. This arrangement is probably inevitable when the survey is taken on as an additional task to one's regular job.

Conclusions

After going through this survey and after about a year's reflection, the following conclusions are offered:

1. Most of the costs for a survey of this type in a public agency will be indirect rather than direct. Direct outlays at the TDH were approximately $7200. This was for interviewer pay, telephones, keypunching services, and printing. However, about 1500 professional man-hours were expended for the survey. This translates into about $25,000 to $30,000. With 1,840 completed interviews, the per interview cost comes to about $17-$20. This compares to $50 to $60 for face-to-face interviews.

2. Valid data can be obtained by a public agency by using the survey method. When the data obtained by the Texas survey are compared with U.S. census data, the results are quite favorable. In five age categories (18-34, 35-44, 45-54, 55-64, and 65+) the sample differs from the census by less than two percent in any one category. Likewise, the differences among racial/ethnic groups is not more than 3% in any one group. The only substantial difference is that the survey population is somewhat more educated than the general population.

3. Useful information can be obtained from a Risk Factor Survey. Since one question asked about the respondent's propensity to drink and drive, we were able to contribute to the state legislature debate about this important issue. Another question enabled us to estimate

hypertension prevalence. This was important for input into the TDH application for a federal hypertension grant. Most importantly, the overall results of the survey are being used extensively by the Department in its health promotion activities.

4. A good, reliable survey can be achieved in a public agency under the following conditions: first...support at the highest level is obtained, second...personnel with survey experience are available, and third...a definite commitment is made by the agency.

With these conclusions in mind I would offer three suggestions to any public agency contemplating a similar survey: The first suggestion is - It is very important to find someone with previous survey experience. The attributes of this person or persons are (1) previous hands-on expereince with doing scientific sample surveys, (2) experience in sample design and execution, (3) strong statistical skills, and (4) leadership abilities. This person must be able to work full-time during the planning and data collection phases of the study. My second suggestion is - The survey work should be directed by the most experienced technical person even if funds are provided for the survey from other units within the agency. And the third suggestion is - Staff chosen to work on the survey should be relieved of their normal work tasks. A separate group should be established to do the survey.

In conclusion, I would encourage other agencies and organizations to undertake telephone surveys themselves when information is needed on lifestyles or other health attributes of the population. This is a viable method producing quality data. I believe its usefulness and its use will continue to become more important in the future.

Nutrition and Health Promotion

Session C

DATA FOR MONITORING DIET, NUTRITION, AND CANCER

Mary Grace Kovar, National Center for Health Statistics

In June 1982 the National Academy of Sciences, National Research Council, Assembly of Life Sciences issued the report Diet, Nutrition, and Cancer followed a year later by a second report Diet, Nutrition, and Cancer: Directions for Research (1,2).

After assessing the research, the Committee on Diet, Nutrition, and Cancer concluded that "the differences in the rates at which various concerns occur in different human populations are often correlated with differences in diet. The liklihood that some of the correlations reflect causality is strengthened by laboratory evidence that similar dietary patterns and components of food also affect the incidence of certain cancers in animals." The Committee found the combined epidemiological and laboratory evidence strong enough to recommend dietary guidelines, which are shown in figure 1.

Figure 1

1. Reduce fat intake from its present level (approximately 40%) to 30% of total calories in the diet.
2. Include fruits, vegetables, and whole grain cereal products in the daily diet.
3. Minimize the consumption of foods preserved by salt-curing (including salt-pickling) or smoking.
4. Minimize contamination of foods with carcinogens from any source.
5. Minimize or remove mutagens from foods.
6. If they consume alcoholic beverages, do so in moderation.

A Public Health Service Task Force was convened to evaluate the report and make recommendations. The Task Force stated that "A clear need exists for a system to accurately track individual nutritional intake and status and to relate them to the presence and subsequent development of disease." The report contained strong recommendations about the need for monitoring with large enough samples to study populations subgroups and detect differences, frequent enough data collection to measure change, standardized measures, and simultaneous measurements of as many risk factors as feasible. It also recommended that these objectives be accomplished by augmenting and strengthening the National Nutrition Monitoring System. A summary of its recommendations is shown as figure 2.

Figure 2

1. The population base of the monitoring system must be large enough to detect small differences in disease.
2. The system must be large enough to incorporate sufficient numbers of people in subcategories of interest; these categories include ethnic groups, racial groups, age or regional groups - subgroups that may have unique dietary patterns or unique metabolic charateristics.
3. Major hypothesized relationships should be monitored. Dietary intake and nutritional status must be monitored simultaneously and the monitoring system should include information on as many factors suspected of increasing cancer risk as possible.
4. Monitoring must be frequent enough to measure change, and to raise warning of adverse practices in time to prevent or alleviate the potential damage caused by their continuation.
5. The system must be flexible enough to incorporate the results of research as they become available.
6. The system should be capable of establishing an interface with cancer surveillance systems.
7. Methods of measuring dietary intake and nutritional status must be improved, and the methods should be standardized so that results from one study can be compared with those from another.

The purpose of this paper is to evaluate whether we have the data to assess dietary intake and nutritional states and their association with specific forms of cancer as recommended by the PHS Task Force and whether we have the data to monitor progress towards the guidelines proposed by the NAS Committee on Diet, Nutrition, and Cancer. In order to do so, it is necessary to have some understanding of the National Nutrition Monitoring System.

The National Nutrition Monitoring System is an array of data collection activities that include per capita availability of food, food composition, food consumption studies and nutritional status surveys, and special surveys.

Several of these activities provide information useful to understanding the relationship between diet and cancer. For example, the Department of Agriculture has collected information on the per capita availability of food since 1909. These estimates are based on quantities of food flowing through the food distribution system. They can be used to evaluate qualitatively changes that have occurred in the American diet since the beginning of the century. Estimates of per capita availability of food are frequently the only data on food consumption available from other countries. Therefore, the U.S. data provide the only basis for most international comparisons. However, they provide no information about individuals.

The two major surveys in the National Nutritional Monitoring System that do provide information about individuals are:

- The Nationwide Food Consumption Survey (NFCS) of the U.S. Department of Agriculture.
- The National Health and Nutrition Examination Survey (NHANES) of the Department of Health and Human Services.

The Surveys

The household portion of the NFCS has been conducted six times beginning in 1936-37. Over the years the survey design has changed and evolved; the data are not strictly comparable over time. Nevertheless, it is the best historical record on diet in the United States that we have (3). However, this portion of the NFCS is not the source of data on individual consumption.

Information on dietary intake of individuals was obtained in addition to the information on food used in the household, from the last two surveys only - in the spring only of the 1965-66 survey and in all four seasons of the 1977-78 survey (4). The sample consisted of members of the households participating in the NFCS household phase. In the 1977-78 survey, all household members were eligible during the first 3-month period (April-June 1979). In the remaining 3 quarters, all household members under age 19 and half those age 19 and older were eligible (people in 1-person households were included regardless of age). Information was obtained for approximately 31,000 individuals in 15,000 households in 114 locations in the 48 coterminous states. Weighting for non-response provided the basis for the 37,785 individuals in published reports.

Information on dietary and nutritional intake was obtained from 3 consecutive days of dietary reporting consisting of a 1-day interviewer-assisted recall and a 2-day self-administered record. The questionnaires were designed to collect information on quantity, form, and source of specific foods, food supplements, and self-reported height, weight, and health status. A leaflet, a set of stainless-steel measuring cups and spoons, a plastic ruler, and a card of equivalents was provided to each household to help estimate quantities (4).

The NHANES has also been conducted twice - in 1971-74 and 1976-80. It developed from the National Health Examination Surveys of the 1960's and reflects the medical orientation of those surveys; medical histories, physical examinations, and laboratory assessments, including x-rays and other physical measurements, are major components of the survey.

The sample in 1976-80 consisted of people 6 months through 74 years of age with oversampling of children under age 5, women of child-bearing age, and people 65-74 years of age (5). Information was obtained for approximately 21,000 individuals in 64 locations. Weighting for non-response and for the probability of selection and post-stratification provide the basis for the civilian non-institutional population estimates in published reports.

All sample people were interviewed first in the household and then invited into the mobile examination centers that were moved to each site. At the center there was a dietary interview consisting of a 24-hour recall and a food frequency questionnaire. Food models were used to help estimate amounts. Height and weight were measured and blood samples were taken for later analysis. The need to move the mobile examination centers meant that interviewing was conducted throughout the year, but only once at each site.

In both surveys, people age 12 and over responded for themselves. A parent or other adult in the household responded for children under the age of 12.

The Data

The surveys as they have been conducted are of limited usefulness for evaluating hypotheses about relationships between diet and cancer.

The samples of approximately 21 thousand and 31 thousand are not large enough to detect small differences, and both surveys have smaller samples at older than at younger ages. Even though cancer is the second leading cause of death, the associations being sought are with specific sites of cancer and the annual incidence rates of specific cancers are low. Moreover, incidence is higher at older ages where sample sizes are relatively small in both surveys (5). Nor are the samples large enough for analyses and comparisons of population subgroups. The Hispanic HANES of 1982-84 will provide information about that subpopulation, but others are not currently covered.

The surveys do not include data on all major hypothesized relationships. Some data on alcohol consumption are collected in both surveys, but only NHANES includes questions on smoking, occupation, and pesticide exposure, and the measurement of carboxyhemoglobin and pesticides.

The surveys have not been conducted frequently enough to monitor critical changes in dietary practice. Currently, they are conducted only every decade although there are plans for changing the frequency of both surveys. Measures for new hypotheses and new measurement techniques cannot be incorporated, nor can sudden or critical changes be monitored.

Neither survey currently has a longitudinal component, nor is either linked with cancer surveillance systems. Beginning with the Hispanic HANES, the HANES will collect the information to link the people in the sample with the National Death Index. USDA is exploring this possibility for the NFCS.

The methods of data collection and the nutrient data banks are similar but are not strictly the same. They are, however, closer to one another than to many of the local epidemiological studies. A great deal of work on standardization remains to be done.

Despite their limitations, we do have the two surveys, conducted at approximately the same time, designed to measure the food consumption of individuals in the American population. We can look at data from them to see whether we at least have a baseline for monitoring the guidelines in Diet, Nutrition, and Cancer that are shown in figure 1.

Neither survey has the ability to estimate the proportion of Americans who consistently obtain 30 percent or less of their total calories from fat. According to NHANES II, 23 percent of the civilian noninstitutional population ages 6 months - 74 years had fat consumption that low in a 24-hour period. According to the NFCS, only 6 percent of the population had average fat consumption that low over a 3-day period. The true proportion on any given day is probably 10-20 percent, but the proporation of the population that maintains fat consumption over a long period of time is unknown.

Acquiring that knowledge will require either longitudinal studies or relying on respondents' ability to report accurately their food consumption over a period of years instead of days. Continuing longitudinal surveys have many advantages, including linking food consumption with later development of disease, but they are expensive. If people can recall their food consumption over a long period accurately, that is the more cost-effective method, although it has the limitation of obtaining information only for survivors. The National Cancer Institute is funding several

research studies to investigate the accuracy of long-term recall.

Neither survey obtains the information about what proportion of the American population consumes fruits, vegetables, and whole grain cereal products daily. According to NHANES II, 45 percent of the population consumed both fruits and vegetables (including juices) in a 24-hour period. Questions needed to ascertain who consumes all three each day were not asked. If they had been, the answers might not be reliable because of uncertainty about what constitutes whole grain cereals.

The questionnaires were not designed to obtain information about how the foods were preserved. Such questions may be appropriate only in countries where foods are preserved at home. In a country where relatively little food is grown and processed at home, household respondents may not know the method of preservation. Package labeling (old-fashioned, charcoal flavored) may be more misleading than revealing.

Two of the guidelines, minimizing contamination of foods with carcinogins and minimizing mutagens, cannot be monitored through food consumption surveys. It is possible to monitor which foods people eat of course. If science progresses to the point where it is possible to know which foods are likely to contain carcinogins or mutagens, we can know the quantity of those foods being eaten and by whom. However, the surveys are not designed to follow foods grown in specific areas through the national food processing and distribution system, so knowledge about who ate foods contaminated by, say, a localized application of pesticides, cannot be obtained from such surveys.

Data on alcohol consumption are obtained in both surveys, but the questions on past surveys have not been designed to define light, moderate, and heavy drinkers. A battery of questions was designed for the National Health Interview Survey in cooperation with the National Institute of Alcohol Abuse. They are currently being used for that survey and for the Hispanic HANES. If the data are valid and reliable, the questions will also be used on future national HANES.

Summary

There have only been four national surveys of the dietary intake of individuals in the United States. They have contributed a great deal to our understanding of the nutritional status of the American population and have taken us a long way from relying on the per capita availability of food.

We are now placing new demands on the surveys - demands for monitoring nutritional status of subgroups of the population, demands for data to provide the basis for food fortification programs, and demands for elucidating relationships between diet, nutrition, and the subsequent development of disease. The surveys were not designed to meet all of these demands and, as can be seen from this evaluation, they don't. They can, however, be strengthened and redesigned with larger samples of the total population and of specific subpopulations of special interest, with longitudinal components, and with linkages to cancer and other chronic disease monitoring systems. It can be done, but it will take research in measuring food components and in methods of data collection and analysis, an improved standardized, and constantly current nutrient data bank, and a major committment from all interested parties to ensure that it is done.

References

1. Diet, Nutrition, and Cancer: Committee on Diet, Nutrition, and Cancer, Assembly of Life Sciences, National Research Council, National Academy Press, Washington, D.C. 1982.

2. Diet, Nutrition, and Cancer: Directions for Research, Committee on Diet, Nutrition, and Cancer, Assembly of Life Sciences, National Research Council, National Academy Press, Washington, D.C. 1983.

3. Rizek, R. L. The Nationwide Food Consumption Survey. Presented at the Workshop on Evaluation of Food Consumption Surveys, National Academy of Sciences, June 6, 1983.

4. Woteki, C. E. The Individual Intake Phase of the Nationwide Food Consumption Survey. Presented at the Workshop on Evaluation of Food Consumption Surveys, National Academy of Sciences, June 6, 1983.

5. National Institutes of Health, Surveillance Epidemiology End Results: Incidence and Mortality Data: 1983-77, Public Health Service, NIH, Pub. No. 81-2330, U.S. DHHS, Washington. U.S. Government Printing Office, June 1981.

6. Pao, E. M. and S. J. Mickle. Problem Nutrients in the United States. Food Technology 35 (9): 58-79, September 1981.

7. National Center for Health Statistics, M.D. Carroll, S. Abraham, and C. M. Dresser: Dietary Intake Source Data: United States, 1976-80. Vital and Health Statistics. Series 11 No. 231. DHHS Pub. No. (PHS) 83-1681. Public Health Service, Washington. U.S. Government Printing Office, March 1983.

IRON-DEFICIENCY ANEMIA IN THE UNITED STATES: A REINTERPRETATION OF HEMATOLOGIC DATA FROM THE TEN-STATE AND HANES I SURVEYS

Jack Hegenauer, Tracy Dale, and Paul Saltman
University of California San Diego

Iron-deficiency anemia remains a widespread nutritional problem in developed nations in spite of the fact that its cause--inadequate dietary iron intake--is generally recognized. Measures to eradicate iron deficiency in the U. S. have been stalemated by lack of agreement over the best means of fortifying American diets, as well as over issues related to the epidemiology of iron deficiency: criteria for diagnosis, prevalence among susceptible subpopulations, and interpretation of racial differences in hematologic data.

Our experience with the efficacy of iron fortification in the Mexican school lunch program (1) has convinced us that epidemiologic criteria of anemia seriously underestimate the incidence of iron deficiency among school-aged children and adolescents, and that iron-deficiency "anemia" is best defined as the ability to increase hemoglobin concentration (Hb) or hematocrit (Hct) in response to iron supplementation. Our studies suggest that epidemiologic criteria alone are inappropriate for determining the "normality" of a population if it is malnourished with respect to selected nutrients, i.e., if it is iron-deficient to an unrecognized degree. Such biological criteria for iron deficiency have long been applied in Scandinavia to define normal (i.e., optimal) hematologic values (2,3,4).

An estimate of the incidence of anemia or iron deficiency in any age, sex, or race subpopulation is impossible without appropriate criteria for abnormality. There is significant disagreement over the anemic classifications proposed by different agencies (WHO, CDC, NCHS). Normative values based on HANES I data have been calculated by Dallman (5) and Garn (6), but these may be flawed by the assumption that certain population samples represent a well-nourished universe of optimal Hb values.

In this report, we will first bring a semi-quantitative analysis of population modes to bear on the question of the black-white Hb difference (7,8). This is necessary because the frequency distributions for Ten-State and HANES I Hb data are decidedly non-normal, and formal statistical treatment obscures the many subtle similarities between black and white populations. The use of normative criteria for anemia, rather than optimal values for iron-supplemented populations like those studied in Norway and Sweden (2,3,4) leads to the inescapable conclusion that 97.5% of American men and women are not anemic. Attempts to find a more satisfying definition of anemia have relied on multivariate analysis using risk factors of iron deficiency. We will present evidence that the most commonly used risk factor--transferrin saturation (TS)--is statistically unreliable for this purpose. We will discuss our preliminary attempts to develop a more discriminating risk factor to identify anemic individuals.

METHODS

Hb, Hct, serum iron (SI), and TS data were selected by age, sex, race, and absence of hemoglobinopathies from survey datatapes. Data from the Ten-State Nutrition Survey were kindly provided by Drs. James Goldsby and Ellen Borland of the Centers for Disease Control on datatapes TAUNOR and TAUSOU. Data from the HANES I were generated by the NCHS as datatape HEHANESI. DU480010, available from the National Technical Information Service. The analyses, interpretations, and conclusions contained in this report are solely those of the authors, not of the NCHS.

RESULTS

Normality of Hb distributions

The frequency distributions of Hb deviate strongly from normality for every subpopulation examined. For example, Figure 1 shows the distribution of Hb (in 0.1-g/dL increments) for black Americans, aged 18-45 years, sampled in the Ten-State and HANES I surveys. The distributions show a curious "sawtooth" pattern that is not expected from large population samples unless certain Hb values were reported with greater favoritism or were "forbidden" by the data collection process. A detailed analysis of this phenomenon is in preparation. We are convinced that this problem is caused by roundoff error (loss of significant digits) when laboratories convert photometric measurements made with older instruments into Hb concentration using a standard formula. Hb is then "quantized" into discrete values separated by 0.3-0.5 g/dL. All age, sex, and race subpopulations sampled in the Ten-State and HANES I surveys show this clustering, which causes an unfortunate loss of significant information regarding the shape and normality of the population distribution.

Distribution means versus modes

The mean Hb, Hct, and MCHC (mean cell hemoglobin concentration) of blacks is lower than that of whites at every age and socioeconomic level (8,9). Simple simulations verify that the lower means are not merely due to the dilution of a normal population by anemic individuals, because elimination of up to 20% of the lowest values has little effect on the mean. Although blacks have lower means, many black populations in HANES I show modal, or most frequently observed, values identical to the modes of their white counter-

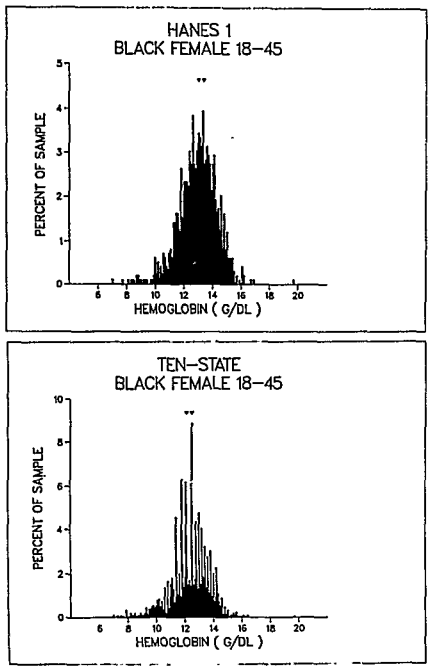


FIGURE 1. Frequency distributions of Hb (in 0.1-g/dL increments) for black women sampled in the HANES I (N=401) and Ten-State (N=703) surveys. Lack of normality and clustering of Hb values is discussed in text.

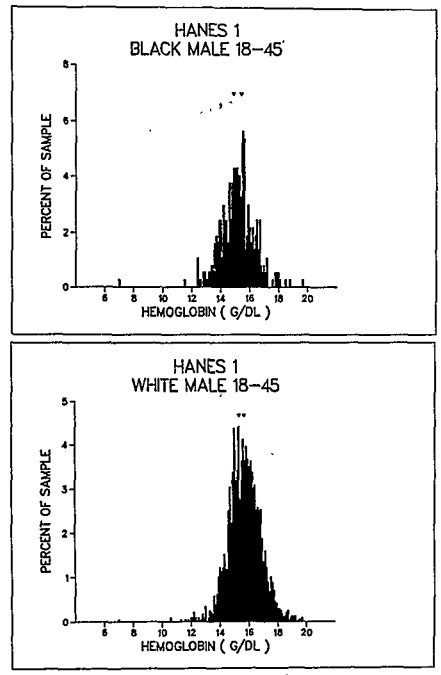


FIGURE 2. Frequency distributions of Hb for black (N=371) and white (N=2402) men sampled in HANES I. Blacks have lower mean Hb (filled arrows) than whites, but both have identical modal values (open arrows).

parts. The modes appear prominently in the Hb distributions of adult men and women in spite of the ragged shape of the distribution (Figures 2 and 3). This information will often be lost from Hb frequency distributions using intervals greater than 0.1 g/dL, so that such distributions then appear approximately normal. Black and white adolescents and adults of both sexes in HANES I show virtually identical modal values for Hb, Hct, and MCHC (Table 1). For some other subpopulations the modal values for blacks are generally higher than the black means but are lower than comparable white modes (Table 1). Since the greatest number of adult blacks and whites sampled have identical values for Hb, Hct, and MCHC, it is difficult to argue persuasively for the hypothesis (9) that blacks have a lower genetic "setpoint" for the synthesis of Hb and red cells. The possibility has not yet been excluded that blacks are in fact more anemic than whites because of their cultural, educational, and nutritional habits, even though they can "afford" to be normal.

Estimates of prevalence using risk factors

Transferrin saturation. Preliminary analysis of HANES I hematologic data has suggested that iron deficiency as measured by low TS levels (less than 16%) is more prevalent than low Hb (10). This analysis concluded, however, that iron-deficiency anemia, measured by the combination of low Hb and low TS, is not widespread in the U. S. and appears to be a problem only among very young children. We feel there is a substantial margin for error in estimating prevalence by

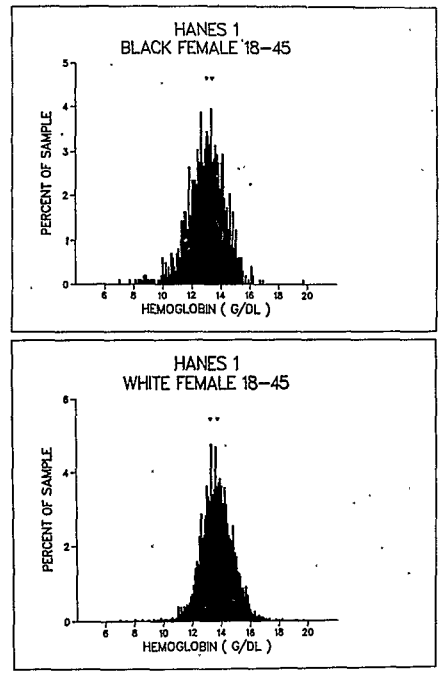


FIGURE 3. Frequency distributions of Hb for black (N=984) and white (N=4693) women sampled in HANES I. Blacks have lower mean Hb (filled arrows) than whites, but both have identical modal values (open arrows).

TABLE 1. Means and modes of Hb, Hct, and MCHC for white and black age and sex subpopulations surveyed in HANES I.

Age	Sex	Hemoglobin (g/dL)				Hematocrit (% PCV)				MCHC (g/mL)			
		White		Black		White		Black		White		Black	
		Mean	Mode	Mean	Mode	Mean	Mode	Mean	Mode	Mean	Mode	Mean	Mode
1- 3	M + F	12.4	13.1	11.7	12.3	36.8	38	36.3	36	.335	.342	.322	.325
5-11	M + F	13.2	13.1	12.5	12.5	38.9	39	37.8	37	.339	.335	.332	.337
13-17	M	14.9	14.8	14.0	14.0	43.7	43	42.0	43	.341	.334	.334	.334
	F	13.7	14.1	12.9	13.2	40.5	39	38.9	39	.337	.335	.332	.335
18-45	M	15.7	15.31	15.2	15.5	45.9	45	45.3	45	.342	.331	.334	.331
	F	13.7	13.3	13.0	13.3	40.7	40	39.5	40	.335	.330	.328	.328
55-74	M	15.4	15.0	14.4	15.0	45.4	45	43.5	43	.338	.334	.330	.322
	F	14.0	14.0	13.1	13.0	45.4	45	43.5	43	.335	.330	.327	.327

the simultaneous occurrence of two conditions (low Hb = anemia and TS less than 16% = iron deficiency), since appropriate criteria for defining these conditions are debatable. Some would argue that anemia without iron deficiency, in the absence of other nutritional pathology, is a contradiction in terms. For example, our studies in Mexico (1) have shown that children even without signs of anemia or low TS may respond to efficacious iron supplementation. In fact, the individual correlation between Hb and TS or SI is extremely poor for every adult subpopulation we have examined in HANES I (Figures 4 and 5). The very low correlation coefficients for white men and women (Table 2) and the absence in the scatterplot (Figure 4) of significant features at TS less than 16% lends no support to the use of SI or TS as measures of iron status. Without better evidence of strong statistical interrelationship with Hb--the primary measure of iron status--it is unwise to use TS as a risk factor for iron deficiency.

MCHC. The close homeostatic coupling between synthesis of Hb and maturation of red cells is shown in the strong correlation between Hb and Hct for all populations examined (Figure 6). Small differences in the regression relationship differentiate men from women, but blacks and whites of the same sex are virtually indistinguishable (Table 2). There is no evidence from this relationship that blacks are a unique subpopulation, only that Hct is lower because of the tendency toward lower Hb. The MCHC (Hb divided by Hct, or the fraction of the red cell occupied by Hb) has been used to diagnose iron deficiency because it has been observed that a low MCHC can be raised by iron supplementation (2,3,4). Low MCHC has been taken to mean that the red cell's "reserve" capacity for Hb is limited only by iron availability. Optimal MCHC, however, is not a constant in spite of the close linear relationship between Hb and Hct (Figure 6). If we know the equation for the regression line of Hct on Hb for a population, we can compute the theoretical MCHC corresponding to a given Hb. The relationship is, in fact, curvilinear (Figure 7), so we must be cautious about the interpretation of low MCHC, because women will have significantly lower MCHC than men within the normal range of Hb for each. Correlation plots appear to fall on the continuum of the theoretical curve, and mixtures of subpopulations (e.g., blacks and whites) can be distinguished from each other only by a lower mean Hb, not by a significantly different relationship between Hb and MCHC (Figure 8). It can be shown that the linear regression of 1/MCHC on 1/Hb gives a slope and intercept that are virtually identical to the intercept and slope, respectively, of the linear regression of Hct on Hb discussed earlier. Thus, MCHC or its transformations give no new mathematical information that is not inherent in the original Hb or Hct data. We can thus easily distinguish men and women by the regression relationship (Table 2), but blacks simply look like whites with a lower Hb. Work in progress will attempt to use the bivariate normal distribution of Hct and Hb to define an acceptable range of hematologic values for normal populations.

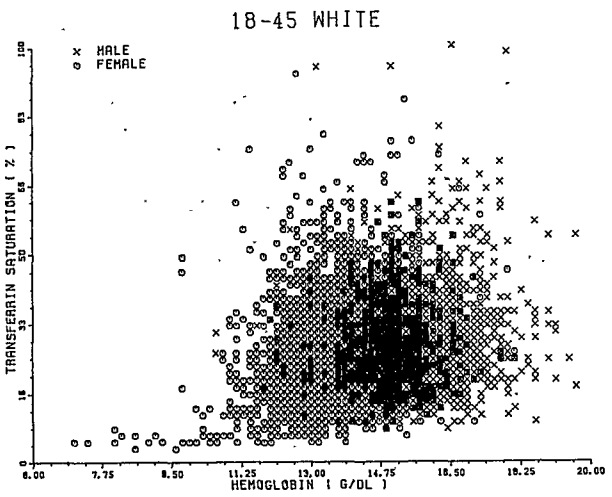


FIGURE 4. Correlation plot of Hb and TS for white men (N=2402) and women (N=4693) sampled in HANES I.

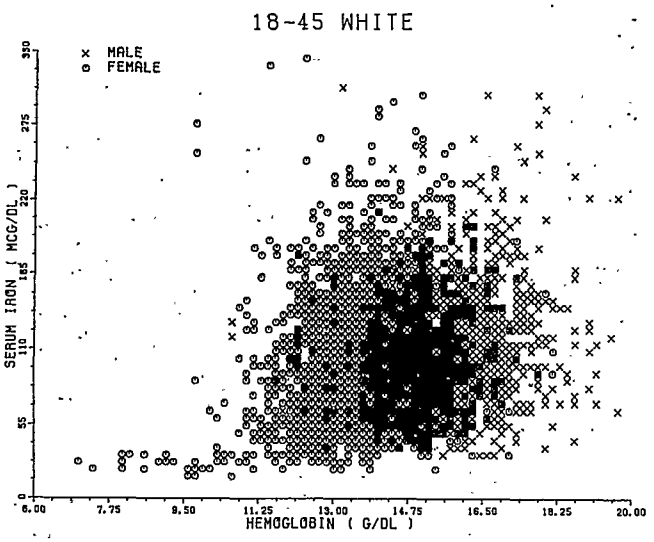


FIGURE 5. Correlation plot of Hb and SI for white men (N=2402) and women (N=4693) sampled in HANES I.

TABLE 2. Correlation coefficients and regression statistics for relationship of Hb with other measures of iron status and "risk factors" of iron deficiency.

Hb : Hct	Black Female	Hct	= 10.33 + 2.25 . Hb	r = 0.85
	Black Male	Hct	= 12.38 + 2.17 . Hb	r = 0.80
	White Female	Hct	= 9.90 + 2.26 . Hb	r = 0.85
	White Male	Hct	= 13.13 + 2.09 . Hb	r = 0.79
Hb : TS	White Female	TS	= -4.19 + 2.26 . Hb	r = 0.22
	White Male	TS	= 14.72 + 1.02 . Hb	r = 0.10
Hb : Serum Iron	White Female	SI	= -4.14 + 7.72 . Hb	r = 0.21
	White Male	SI	= 27.40 + 5.25 . Hb	r = 0.15
1/Hb : 1/MCHC	Black Female	1/MCHC	= 2.24 + 10.38 . (1/Hb)	r = 0.53
	Black Male	1/MCHC	= 2.07 + 13.90 . (1/Hb)	r = 0.46
	White Female	1/MCHC	= 2.22 + 10.35 . (1/Hb)	r = 0.50
	White Male	1/MCHC	= 2.01 + 14.32 . (1/Hb)	r = 0.48

CONCLUSIONS

In the absence of suitable criteria for anemia, it is not yet possible to measure the true prevalence of iron deficiency in America. We can approach the question of the relative amount of iron deficiency among blacks, however, by comparing them to whites as a reference population. When we carefully scrutinize the modal values of distributions of hematologic values of adolescent and adult black and white populations, we find unmistakable evidence for central tendency toward the same values of Hb, Hct, and MCHC. Correlation analysis of Hb and Hct shows clearly that blacks and whites are not separate subpopulations in the same sense that males and females show distinct differences in regression statistics.

The hypothesis that blacks have lower mean Hb than whites because of genetic predisposition should be shelved until we have excluded the possibility that blacks are not in fact more iron deficient than whites (11). Unfortunately, traditional multivariate analysis of anemia using transferrin saturation as a risk factor gives a statistically improbable result that will be of little value in diagnosing "true" iron deficiency. The bivariate distribution of Hb and MCHC (or Hct) gives us a far more statistically reliable tool for determining the normalcy of a population. Future work should be focussed on exploiting valid risk factors and on applying measures of optimal iron status to the HANES data.

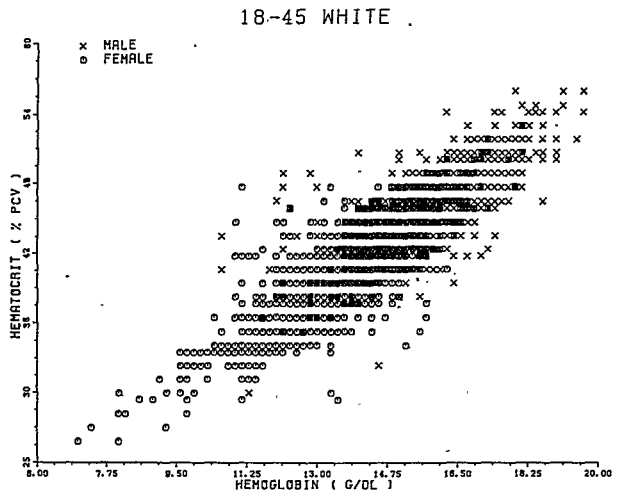


FIGURE 6. Correlation plot of Hb and Hct for white men (N=2394) and women (N=4679) sampled in HANES I.

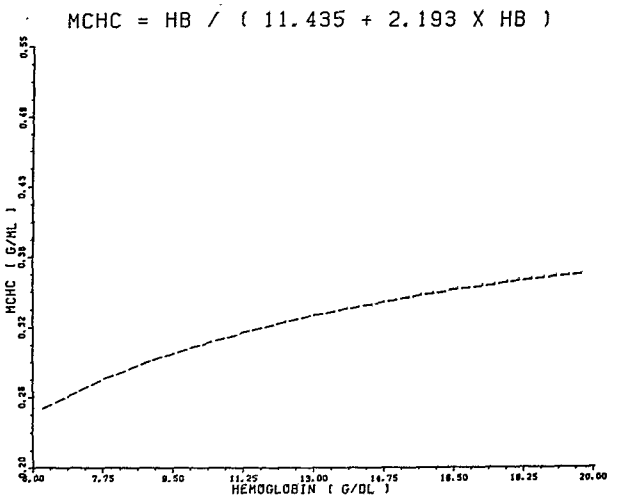


FIGURE 7. Theoretical relationship between Hb and MCHC calculated for arbitrary coefficients. Equation uses a value for Hct predicted by the regression relationship between Hb and Hct (see TABLE 2).

ACKNOWLEDGEMENTS

This work was supported by the Weingart Foundation and by USPHS research grant AM-12386.

REFERENCES

1. Rivera, R., R. Ruiz, J. Hegenauer, P. Saltman, and R. Green. 1982. Bioavailability of iron- and copper-supplemented milk for Mexican school children. Am. J. Clin. Nutr. 36: 1162-1169.

2. Larsen, O. 1966. Studies on hemoglobin values in Norway. VI. Hemoglobin concentration, hematocrit and MCHC in 19 year-old men. Acta Med. Scand. 180: 621-630.

3. Natvig, H., and O. D. Vellar. 1967. Studies on hemoglobin values in Norway. VIII. Hemoglobin, hematocrit and MCHC values in adult men and women. Acta Med. Scand. 182: 193-205.

4. Garby, L., L. Irnell, and I. Werner. 1969. Iron deficiency in women of fertile age in a Swedish community. Acta Med. Scand. 185: 113-117.

5. Dallman, P. R., and M. A. Siimes. 1979. Percentile curves for hemoglobin and red cell volume in infancy and childhood. J. Ped. 94: 26-31.

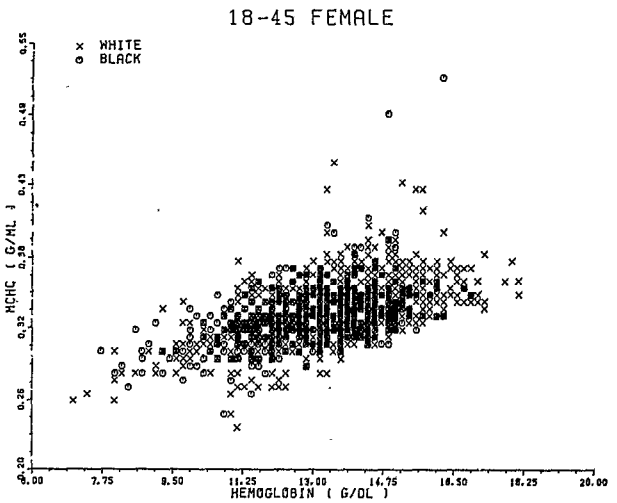


FIGURE 8. Correlation plot of Hb and MCHC for white (N=4679) and black (N=2394) women sampled in HANES I. Note that scatterplot is clustered around the theoretical line shown in FIGURE 7.

6. Garn, S. M., A. S. Ryan, S. Abraham, and G. Owen. 1981. Suggested sex and age appropriate values for "low" and "deficient" hemoglobin levels. Am. J. Clin. Nutr. 34: 1648-1651.

7. Dallman, P. R., G. D. Barr, C. M. Allen, and H. R. Shinefield. 1978. Hemoglobin concentration in white, black, and Oriental children: is there a need for separate criteria in screening for anemia? Am. J. Clin. Nutr. 31: 377-380.

8. Garn, S. M., A. S. Ryan, G. M. Owen, and S. Abraham. 1981. Income matched black-white hemoglobin differences after correction for low transferrin saturations. Am. J. Clin. Nutr. 34: 1645-1647.

9. Meyers, L. D., J.-P. Habicht, and C. L. Johnson. 1979. Components of the difference in hemoglobin concentrations in blood between black and white women in the United States. Am. J. Epidemiol. 109: 539-549.

10. National Center for Health Statistics. 1982. Diet and iron status, a study of relationships: United States, 1971-74. Data from the National Health Survey. Series 11, No. 229. U. S. Department of Health and Human Services Publication No. (PHS) 83-1679.

11. Jackson, R. T., H. E. Sauberlich, J. H. Skala, M. J. Kretsch, and R. A. Nelson. 1983. Comparison of hemoglobin values in black and white male U. S. military personnel. J. Nutr. 113: 165-171.

Health Promotion Programs: Evaluation and Cost Benefit

Session D

A Cost-Benefit Analysis Using Comparative Projections of Outcomes With Natural Controls for the American Health Foundation Promotion System at Blue Cross/Blue Shield of Indiana

K. Per Larson, Marvin Kristein, Ph.D., Consultants to the American Health Foundation

Introduction

Cost pressures are mounting in disease treatment, pressures that are prodding companies to act. Companies have been swiftly proceeding with cost-containment actions, doing what they know best: applying management techniques to create efficiencies. However, cost-containment is often short-term and small-scale; it takes the existing system as a given. As a result there is widespread employer interest, with both large and small companies, in cost-prevention -- and health promotion.

Moreover, based on the Framingham logistic, it would appear that in a typical company of 1000 people with utilization costs of $1,000,000, it would be possible to impact nearly 50% of those costs traceable to lifestyle risk factors (Table 1).[1] Together, cost-containment and cost-prevention could constitute an overall health cost management program.

However, health promotion requires skills and knowledge outside the experience of most managers. They seem highly reluctant to try an unfamiliar approach in an area hitherto consigned to medical institutions. These managers say they want added assurances that the methods of health promotion work -- to a degree of certainty far beyond what is normally required in more customary management techniques. The key to providing some degree of certainty is evaluation.

Yet no long-term, independently evaluated studies have been published for health promotion programs at the worksite. This type of work is currently going on principally at Control Data, Johnson & Johnson, and Kimberly-Clark. This presentation reports on the long-term study just completed at Blue Cross/Blue Shield of Indiana (BBI).

Background

BBI implemented a risk factor-based intervention model health promotion system seven years ago, in 1976. The model was created by the American Health Foundation, which trained BBI's staff to implement it. Prior to this, the staff was BBI's health service. Basically they traded in their beds, aspirins and band-aids for risk factor questionnaires, mini-screenings, and three classroom-based interventions for nutrition, weight control, and smoking.

The system ran unevaluated until BBI applied to the Kellogg Foundation in 1982 for an evaluation grant. K. Larson provided a cost-benefit model for analyzing four years of data for three groups: those exposed to corporate culture changes in health; those additionally exposed to mini-screenings and health risk questionnaires; and those targeted for risk factor intervention who actually participated in an intervention program. The evaluation of these three groups is in the process of completion.

This presentation reports on comparative projections of outcomes[2] between the employees at BBI and three natural controls: people in the state of Indiana; a bank with a similar number of people doing similar work three blocks down the street in Indianapolis; and Blue Cross/Blue Shield of Wisconsin (which intends to implement the program in the near-future).

There is some basis for the comparability of these natural controls: BBI's employees are predominantly urban female whose bias in medical care utilization, if any, would be toward high utilization compared to the balanced urban/rural, male/female population of the state; the bank's demographically similar employees were not affected by any form of health promotion activity until the administration of questionnaires and mini-screenings after the control period; the Wisconsin group is also Midwestern and performs work similar to BBI.

Results

Participation rates. Over 95% of all employees have been exposed to the program. Top management support, employee communications, noontime lectures and community events were generated by the health promotion service. Over 85% of all employees took a mini-screening. Of these, 4 out of 10 signed up for an intervention program. Half of these completed the program. By any measure, participation and exposure rates were high.

In fact, the program as implemented by BBI's staff had the characteristics of Health Promotion That Works cited by Jonathon Fielding:

- Top management support
- Strong identity, high visibility
- Employee involvement and enough participants
- Modified work environment ("they mean it!...")
- Assessment of health risk
- Programs based on assessments
- Peer group support
- A "Do as I Do" staff
- Evaluation built-in

Such total programs are said to get the multiplier effects.

Overall utilization. Table 2 presents the actual utilization experience of BBI for 1980-82 compared to actual trend figures for similar types of health insurance coverage for the state of Indiana as a whole for the same period. In this 3-year period, BBI saved $1,050,000 in utilization costs compared to what was experienced by the general population of the state. The appended graphs give more details.

The costs of running the health promotion service at BBI for the same period were stabilized at approximately $71,000 per year. This includes full-time salaries, materials, and overhead. The program is deliverable at a cost of $33.32 per year (or $2.86 per person per month).

Comparing costs against benefits would give BBI a stabilized return in the 4th-6th year of operation of the program of 5-to-1. Costs in the first years either increased or showed wide variation.

An analysis of program costs by intervention subcategories described above is currently being completed for comparison with the utilization rates of these three subgroups.

Absenteeism.[3] Lower absenteeism is reported to be an early response to health promotion. Based on the NCHS 1974 Health Interview Survey, this excess absenteeism was estimated at 2 days per smoker. More detailed analyses in progress indicate 50-75% less absenteeism for program participants in general, plus fewer outpatient claims and up to 3 times less total utilization costs. Inpatient care, once a person is admitted, appears to be the same for all employees.

Alcoholism and drug counseling. A subprogram in drug and alcohol counseling reports costs of $19,500 as against savings in worktime and out-of-pocket costs of $71,600. Subjective ratings cite 70% of the program participants as making "good progress." The program costs $0.66 per month per employee (or $7.91 per year) to operate for the entire company.[4]

Percentage increases in utilization. BBI's medical care utilization costs grew 8% in the 3-year period compared to 17-20% trend factors experienced for various insurance coverage components for the state of Indiana for each year of the same period.

During this same period, the bank's utilization costs grew 59.6% and the Wisconsin group experienced an increase of 81.4%.

There appear to have been no substantial shifts in employee numbers or demographics during this period in any of the natural controls. However, these percentage comparisons are merely gross indicators of comparison.

Further Analyses

This preliminary analysis emphasizes gross outcome measurements. It would appear from the detail in the data and from the high exposure and participation rates that a factor best termed a "corporate culture" or a "systems effect" had a significant role in these outcomes, at least for this employer group. Since the original evaluation design focused only on explicit, "official" program components (such as the risk factor questionnaires, mini-screenings, and classroom interventions), the data are being reanalyzed to isolate this factor. Moreover, due to the nature of BBI's business, this database will be the subject of numerous further analyses.

At this time, however, it would appear that the whole is greater than the sum of the parts and that corporate culture bears further investigation. This will be the subject of a further paper.

References

[1]Kristein MM: How much is chronic disease costing the typical American company? Mimeographed, American Health Foundation, New York, 1980.

[2]Warner KE and Murt HA: Premature deaths avoided by the antismoking campaign. Am J Public Health 1983; 73:672-77.

[3]American Cancer Society: Proceedings, Natl. Conference on Smoking and Health, New York, Nov. 18-29, 1981. Publ.No.82-IM-0245.

American Telephone & Telegraph Co. Managing Health Promotion in the Workplace (RS Parkinson et al, Eds). Palo Alto, CA: Mayfield, 1982.

Bjurstron LA & Alexiov NG: A program of health disease intervention for public employees. J Occup Med 1977; 20:521.

Kristein MM: How much is chronic disease costing the typical American company? Mimeographed, American Health Foundation New York, 1980.

Wood L: Statement on estimated company savings associated with smoking cessation efforts. Available from NY Telephone Co, New York.

[4]This compares well with the literature on alcohol abuse treatment. See annual reports of NIAAA and 1970 Comptroller General of the US. Substantial cost savings from establishment of Alcoholism Programs for Federal Civilian Employees. Report to Subcommittee on Alcoholism of Committee on Labor, US Senate.

Table 1
Company XYZ: 1000 Employees Screened

Number At Risk	Risk Factor	Per Person Direct Cost	Totals
50	Smoking/HBP/HCH	$2760	$138,000
25	HBP/HCH	1500	37,500
75	Alcohol	1325	99,375
50	Smoking/HBP	820	41,000
125	Smoking/HCH	630	78,750
50	HBP	250	12,500
100	Smoking	160	16,000
100	HCH/CVD	50	5,000
300	HCH/CA	31	9,300
			$437,425

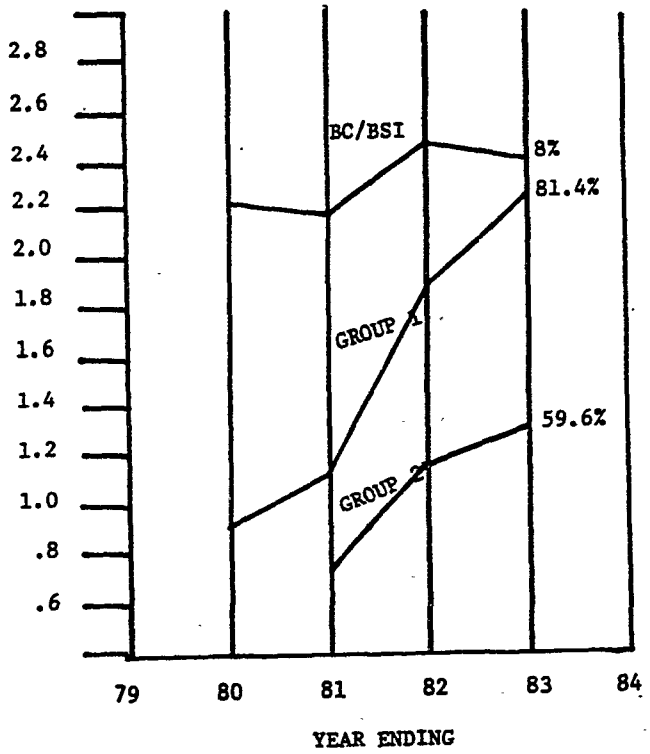
HCH=High Cholesterol (CVD=Cardiovascular disease; CA=Cancer)
HBP=High Blood Pressure
All figures are 1976 dollars

Table 2
Utilization Data

	1980	1981	1982	Total Savings
Trend	$2,777,000	$2,794,000	$3,162,000	$8,733,000
Actual	2,388,000	2,687,000	2,608,000	7,683,000
Difference	389,000	107,000	554,000	1,050,000

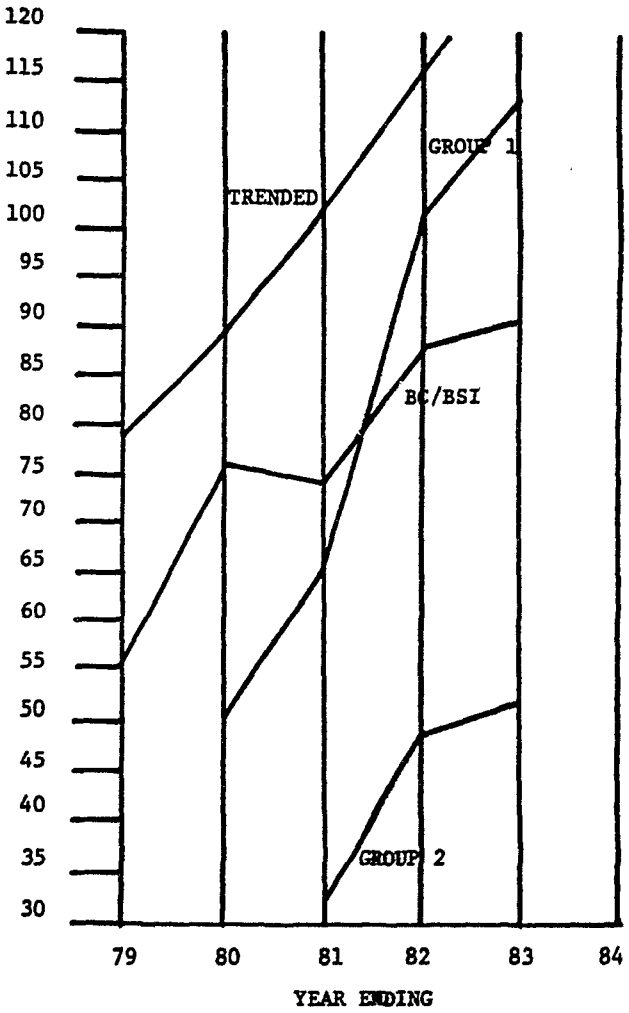
PERCENTAGE INCREASE IN

TOTAL HEALTH CARE COSTS



AVERAGE COST OF HEALTH CARE

PER CONTRACT PER MONTH (EXCLUDES DENTAL)



EVALUATING A HEALTH PROMOTION PROGRAM BY EXAMINING HEALTH CARE CLAIMS

Barbara E. Merrill
Control Data Corporation

Health promotion programs at the worksite are becoming popular, with companies hoping that the programs will subsequently help contain and/or reduce health care costs. Although some data does exist on the cost effectiveness of controlling certain health risk factors,[1] the cost effectiveness of a comprehensive health promotion program that includes multiple risks such as smoking, weight, hypertension, fitness, stress, and nutrition, has not been established.

For health promotion programs at the worksite to remain a viable option for controlling rising health care costs, their cost effectiveness needs to be established. Before industrial leaders decide to invest in health promotion, they are demanding to know what return on investment they can expect.

In 1980, Control Data Corporation developed its own health promotion program called STAYWELL. The key question that Control Data's top management wants answered is whether STAYWELL contains health care costs over time. To evaluate this issue most effectively, Control Data is monitoring patterns and trends in our health care claims costs.

This paper will describe:

- The evaluation of the STAYWELL program through the merging of health care claims data with health risk factor data.
- How health care claims data are collected.
- The advantages and potential biases in using health care claims data to evaluate STAYWELL.
- The analysis plan for this data set.
- Some preliminary findings of the evaluation using this data set.

The Evaluation of the STAYWELL Program

The STAYWELL program is a voluntary health promotion program offered free as a benefit to full-time Control Data employees and their spouses. The STAYWELL program includes three phases: the promotion phase, the educational phase, and the support/follow-up phase.

The promotion phase consists of an orientation session for all eligible employees, the completion of the Health Risk Profile (Control Data's health hazard appraisal) and health screening, and a group interpretation session that examines the results of the Health Risk Profile.

The educational phase consists of lifestyle change courses in smoking cessation, weight control, stress management, nutrition, and fitness. Each of the courses is available in three different media: traditional instructor-led courses, PLATO courses (Control Data's educational computer-based instruction), and self-study courses.

The support phase consists of employee-led special topic groups such as walking groups, weight loss groups, and stress reduction exercise groups. Also, employees can retake the Health Risk Profile at specified intervals to determine their progress in reducing health risks.

The STAYWELL program implementation was begun in the spring of 1980. The implementation has occurred on a phased-in basis. Currently over 25,000 employees at 17 major Control Data work sites throughout the country are eligible for the program. This represents over 50 percent of the total national Control Data employee base. Of those employees eligible, about 70 percent voluntarily enroll in the STAYWELL program, and of those enrolled, about 85 percent complete one or more Health Risk Profiles.

The evaluation effort also began in 1980. The STAYWELL program evaluation effort has centered around three key objectives; 1) to monitor STAYWELL program participation and reactions to program activities; 2) to determine the impact of the STAYWELL program in changing individual health behavior and worksite culture; 3) to assess consequences of individual and worksite changes on employee health and performance and on company health care costs. Monitoring the company's health care claims falls under the third objective.

The uniqueness of the STAYWELL program evaluation is its ability to merge STAYWELL program data on individual participation in specific activities and risk factor change with the health care claims data. In addition, because the claims data are coded for both type of medical procedure and diagnostic categories, this effort has the capability

of tracking changes in specific disease categories as well as types of procedures over time. With this unique data base, the STAYWELL program evaluation will examine key health promotion evaluation questions such as:

- Do individuals with lifestyle health risk factors (smoking, hypertension, lack of exercise, overweight), have higher health costs than individuals without these risk factors?
- What kinds of health care costs are higher for those with lifestyle risk factors? In which diagnostic groups do differences occur?
- Do health care costs at STAYWELL sites decline over a period of time (five years), relative to Control Data sites without STAYWELL?
- Do individuals who change their lifestyle risk factors (e.g., quit smoking), have proportionately lower health costs over time compared to those who do not change (e.g., still smoke)?

How Health Care Claims Data are Collected

Control Data offers two major health insurance benefit options to its employees. The first option is Control Data's self-insured, self-administered health insurance plan which had deductibles of under $100 for the 1980-1982 period. The second option is local Health Maintenance Organizations (HMOs), for locations that have HMOs available.

Over two-thirds of Control Data employees are insured through the first option, Control Data's own health insurance program. Only those employees who are covered under Control Data's health insurance are included in the health care claims data base. No health care cost data are available for those employees who are insured through HMOs. Employees are allowed to switch from one plan to the other only during an annual open enrollment period.

Employees insured by Control Data's health insurance submit claims whenever they occur. The claims adminsitration process determines whether the claim is covered under the insurance plan, whether the deductible has been met, and whether the claim is reasonable and customary for the particular procedure code and geographic location.

Each claim submitted by employees is coded and keyed into the computer file. A claim record is created for each procedure on the claim with the following information:

- Procedure code.
- Primary and secondary diagnosis code.
- Location of service.
- Service date.
- Charges billed.
- Amount paid.
- Provider code.
- Hospital admission/discharge dates (if appropriate).

The coders utilize ICD-9-CM diagnostic codes, UB-16 hospital procedures codes, CPT-4 physician procedure codes, ADA dental procedure codes, and special internal codes for key procedures not coded elsewhere. The coders are well trained and quality control is maintained by having the coding supervisors check daily random samples to confirm appropriateness of coding. In addition, there is also an automated claim editing capability that checks on obvious coding errors, (e.g., sex by certain procedure codes). An independent audit indicated that coding was significantly more accurate than the industry standard of 95 percent.

About 50 percent of claims submitted come directly from providers with diagnostic and procedure codes already completed. No independent code checking is done on these diagnostic and procedure codes submitted by the providers, except for the automated claim editing.

In the third quarter of each year a computer file is created containing all the employee claims for the past year. This file is then merged with the other STAYWELL evaluation files in an integrated data base management system.

The STAYWELL evaluation data base contains four major record types on individuals enrolled in the STAYWELL program. These record types include:

1. Demographics of the individual, including whether the employee is enrolled in an HMO or Control Data's health insurance.
2. Participation in specific STAYWELL activities.
3. Health Risk Profile results (multiple years starting in 1980).
4. Health claims data (multiple years starting in 1980).

Because employee I.D. number is the key used in all four data records, combined

records can be created by merging data across record types. However, because of the sensitive nature of this data, employee I.D. numbers are scrambled, to ensure confidentiality of an individual's information. In addition, individual names are not part of any of the record types in the STAYWELL data base, and only aggregate results are reported by the evaluation.

Advantages and Potential Biases of the Claims Data

Like most other data sets, there are advantages and potential biases in the data being input into the STAYWELL data base. The STAYWELL data base has four major advantages over most health promotion evaluation data sets. First, individual employee health risks (from the multiple year Health Risk Profiles) and claims data are merged. This merging allows examination of the correlations between health risk factors and costs.

Second, the records are kept over time. This allows study of trends and changes in utilization and costs to emerge over time. In addition, changes in utilization and costs can be statistically analyzed to determine "real" changes versus "random" changes.

Third, this data set collects information pertinent to whether the STAYWELL health promotion program impacts health risks and health care costs. By including information on initial health risks, changes in health risks, participation in STAYWELL activities, and health care costs, this data set can trace individuals through the STAYWELL program and correlate the various participation levels with health risk factor changes and health care cost differences over time. Finally, this data set allows the examination of health risks as related to specific diagnostic categories.

The advantages of the data set allow for a wide range of analysis, examining many questions concerning the effects of risk reduction on subsequent health outcomes, the effects of the STAYWELL program on health care costs, and the relationship between health care utilization and health risks. But despite these tremendous advantages there are several sources of potential bias within the data set. Some of the potential biases can be statistically controlled once they are studied and understood. Other sources of bias must be acknowledged and may in the future lead to further research, information gathering and study by Control Data or other interested parties.

Five major sources of potential bias within this data set have been identified.[2] First, health risk data is available only for those who choose to enroll the in the STAYWELL program and complete the Health Risk Profile. These volunteers may differ in health status and health risk factors from those who do not volunteer. In addition, some of the information in the Health Risk Profile, such as smoking status and fitness level, are self-reported.

A second potential bias is that those individuals who are covered by Health Maintenance Organizations are not included in this data set. To date, individual Health Maintenance Organizations have been unable or unwilling to supply utilization information on their patients. These HMO users may have different health risks and/or utilization patterns, or they may be affected differently by the STAYWELL program than those employees covered by Control Data's health insurance. In addition, the demographic profile of HMO users may be different, thus skewing some of the demographic utilization and cost patterns.

Third, some coding may not be consistent across providers. Since 50 percent of the diagnostic and procedure codes are completed by the providers, this could be a serious problem in some coding areas. Specifically, some codes may be underutilized by providers. For example, psychiatric or alcoholism-related codes may be underutilized. This may be due either to societal conventions or to perceived lack of extensive insurance coverage in these areas. Whatever the reason, the pattern of code underutilization or overutilization by the providers can only be estimated. Additionally, since standard code definitions are still imprecise, the same disease could be coded differently by various providers who differ slightly in their interpretations of the codes.

The fourth potential bias is that employees who do not satisfy the deductible may not submit any claims, therefore under-representing health care costs across the company. Finally, the fifth potential bias is that changes in the geographic distribution of the workforce at Control Data could affect claims costs, since health costs differ remarkably across regions.

Analysis Plan for this Data Set

In general, epidemiological studies have concentrated on finding risk indicators after a specific health outcome has occurred. Outcomes studied include death, stroke, specific cancers,

myocardial infarction, etc. The STAYWELL program evaluation data set, however, will examine the effects of risk reduction on subsequent health costs and outcomes. The analysis of this data set will occur in three steps.

Stage 1 will consist of an examination of health costs and utilization patterns of employees with good health habits compared to those with poor habits. Specifically, smokers versus non-smokers, sedentary employees versus exercisers, hypertensives versus employees with normal blood pressure, and employees who are overweight versus those not overweight. In addition, demographic differences, such as age, sex, education, job group, and geographical location, in health care costs and utilization will be examined.

Health risks and demographic variables will be used first as independent variables. Any significant differences found will be controlled in future analysis, so that real changes over time will not be masked.

Several considerations will need to be taken into account when the three-year trend analysis of the demographic differences is done. First, during the three years, changes in the price of medical procedures will occur. This price inflation may not be constant or consistent across all medical procedures. In addition, these price changes may be different in the various geographical regions where Control Data has its facilities. Therefore, the locations with the largest employee concentrations will be identified and the prices for major procedures at each location will be tracked in order to control adequately for price inflation over time.

Another consideration in tracking trends is that providers may change some of their coding preferences. For example, certain mental illness codes may not be as taboo in 1982 as they were in 1980. Therefore, some changes may be due not to changes in incidence, but rather to changes in coding preferences.

A third consideration to be examined in analyzing the three-year trend data is controlling for changes in the distribution of overall employee demographics and geographical locations. Correction factors will need to be established in these areas where warranted.

A final consideration in analyzing this data is that medical technology may change over time. Changes in medical technology may impact health costs as well as coding preferences. Any major change in costs will need to be examined for this possible alternative explanation.

The second stage of the analysis of this data set will examine changes in health care costs of participants in STAYWELL activities compared to the costs of employees not active in STAYWELL activities, controlling for demographic variables. This stage will probably need three to five years of claims data in order for established changes in cost to emerge.

The third stage of the STAYWELL evaluation analysis will be the examination of claims for employees at Control Data facilities where STAYWELL has been offered regardless of how much employees have participated in STAYWELL activities, compared to Control Data facilities where STAYWELL has yet to be introduced. We estimate that five years of claims data will be necessary to see cost changes emerge in this stage.

Preliminary Findings from the Health Claims Data Set

Currently the analysis of the health care claims data set is in the middle of Stage 1. Charts 1 and 2 show some preliminary findings for the examination of health care costs of employees with low health risk compared to those with high health risk. Four health risk factors were examined:

- Smokers versus non-smokers.
- Hypertensive versus non-hypertensives.
- Sedentary employees versus exercisers.
- Overweight employees versus those not overweight.

The total dollars paid by Control Data and the average days in the hospital in 1980 of all employees who had completed a Health Risk Profile were calculated. F-ratios were calculated on each health habit for both the total dollars paid and the average number of days in the hospital.

Smokers' health care costs were found to be significantly higher on the average than non-smokers health care costs. In addition, smokers spent more days in the hospital, on the average, than non-smokers.

Those employees with blood pressure greater than or equal to 160/95 had, on the average, more than twice the health care costs as those employees with blood pressure below this level. There was no difference, however, in the average

number of days spent in the hospital.

Employees who exercised cost significantly less than employees who got no exercise on a regular basis. No significant difference was found in the average number of days spent in the hospital between the two groups, however. No significant differences were found in comparing employees who were overweight with employees not overweight.

Conclusion

The preliminary findings from the STAYWELL evaluation data set suggest that major cost differences exist within the employee population between employees with low health risks and employees with higher health risks. The next group of analyses will determine how demographic variables impact these cost differences. Future analysis will then concentrate on whether the STAYWELL program has an impact on health care costs and utilization.

The size and complexity of the STAYWELL evaluation data base makes this analysis challenging. However, even our results to date suggest that knowledge concerning relationships between lifestyle and health care costs can be greatly enhanced by an effort of this scope.

Footnotes

[1]Fielding, J. E.: Effectivness of Employee Health Improvement Programs. Journal of Occupational Medicine, 24: 907-915, 1982.

[2]Dr. J. E. Fielding helped articulate potential biases and potential pitfalls in logitudinal data analysis under a private consulting agreement with Control Data.

CHART 1

BASELINE HEALTH CARE CLAIMS BY HEALTH RISK

	AVERAGE TOTAL $ PAID	NO. AVERAGE DAYS IN HOSPITAL
RISK:		
SMOKING	$390.87	0.60
CURRENT SMOKERS AND THOSE WHO QUIT LESS THAN FIVE YEARS AGO N = 2,376		
NEVER SMOKED OR QUIT MORE THAN FIVE YEARS AGO N = 3,193	$313.27	0.28
SIGNIFICANT LEVEL	P<.03	P<.05
HYPERTENSION		
GREATER THAN OR EQUAL TO 160/95 N = 300	$692.95	0.53
LESS THAN 160/95 N = 5,269	$325.65	0.41
SIGNIFICANT LEVEL	P<.02	P<.51

NOTE: POPULATION IS NON-HMO EMPLOYEES WHO TOOK HRP IN 1980

CHART 2

BASELINE HEALTH CARE CLAIMS BY HEALTH RISKS - CONT'D

	AVERAGE TOTAL $ PAID	NO. AVERAGE DAYS IN HOSPITAL
EXERCISE		
SEDENTARY N = 1,219	$436.92	0.57
SOME OR VIGOROUS EXERCISE HABITS N = 4,350	$321.01	0.37
SIGNIFICANT LEVEL	P<.01	P<.19
OVERWEIGHT		
GREATER THAN 20% OVERWEIGHT N = 1,637	$362.42	0.61
20% OR LESS OVERWEIGHT N = 3,932	$339.71	0.33
SIGNIFICANT LEVEL	P<.55	P<.22

NOTE: POPULATION IS NON-HMO EMPLOYEES WHO TOOK HRP IN 1980

MEASURING AND VALUING THE ECONOMIC BENEFITS OF DIABETES CONTROL

Timothy M. Smeeding, University of Utah
LaVonne A. Booton, Western Illinois University

The use of cost benefit analysis in health care program evaluation is still in its infancy, however, rational resource allocation in health care demands effective evaluation of health care interventions designed to reduce disease costs. The long term nature and physiological behavior of chronic disease have implications for resource utilization that differ from other diseases and health care problems. Diabetes Mellitus is an excellent example of an economically costly disease, but one in which intervention strategies which improve disease management can be effective in reducing both direct and indirect costs associated with its prevalence. The purpose of this paper is to lay out a framework for estimating the economic costs of diabetes and for measuring the net potential dollar benefits from interventions designed to reduce these costs.

I. Cost-Benefit Analysis in Health Care. Cost-benefit analysis has been used effectively in a variety of choice situations involving allocation of limited resources to alternative uses. A major impediment to its widespread incorporation in the allocation of health care expenditures is that its use forces the decision maker into the sensitive area of placing dollar valuations on human life. In order to compare programs and set priorities across programs with noncomparable outcomes, and in order to distinguish between differing values of lives which could be saved, cost-benefit analysis is a must. Society can no longer afford to provide unlimited amounts of health care based on "needs" alone. Estimates of the value of lives saved are an integral part of the necessary rationing process, and certain elements of this valuation process are fairly well defined and generally accepted by most analysts. The Cooper-Rice approach, which has been widely accepted as the dominant economic framework for determining the value of life, advocates the use of cost-benefit analysis in choosing among investments in health care programs (Rice & Cooper 1967; Cooper & Rice 1976). Their framework considers direct outlays for medical care and indirect losses of earnings attributed to disease. Indirect costs are estimated using a human capital approach whereby the value of life is determined by earnings capability, so when productivity is reduced by morbidity or eliminated by mortality, the value of lost subsequent earnings is translated into the value foregone of the person. In considerations of policy decisions, this approach has negative implications for many groups within our society, especially those who are unable to generate earnings because of the debilitating effects of disease. Although these limitations are restrictive, the Cooper-Rice approach offers a quantitative method of evaluating the indirect costs of output lost from disease, and provides a structure to build upon for achieving greater efficiency in cost-benefit analysis of health care programs. This paper moves this process forward by incorporating and quantifying more of the costs associated with a disease, and by disaggregating these costs according to the party which bears the cost.

The benefits to be realized from reducing the negative effects of an illness are broader in scope than conservation of medical resources and maintenance of earned income and family life style, personal and emotional problems, financial stress, increased anxiety and resentment, as well as the more general limitations on opportunities to be full productive members of society. Although these costs vary considerably among patients and their families, there is growing recognition of the need for their inclusion to prevent serious underestimation of the costs of disease.

Evaluating Diabetes Intervention Efforts. The evaluation of intervention strategies designed to minimize the adverse health effects of diabetes is an area where the application of cost benefit analysis can be extremely useful to decision makers. Most economic evaluation in this area has been concerned with the benefit, usually in terms of reduced hospitalization only, of education strategies designed to instruct diabetics to self-monitor and self-treat their illness. Because these studies usually concentrate on adult-onset, insulin using diabetics, they deal with the class of tertiary preventive health care strategies wherein the objective is not to prevent or forestall the onset of a disease, but to minimize its adverse health effects once experienced. Because of the nature of the disease, this type of diabetes intervention has much to offer in terms of outcomes which are amenable to cost-benefit analysis in general and the cost accounting framework suggested here in particular.

Foremost is the relatively short time needed to observe differences in hospitalization for pre and post education interventions (Neresean & Zaremba 1982). It has been estimated that between 30 and 50 percent of admissions, readmissions and extended hospital stays for diabetes are related to knowledge or practice of self-management skills, and that 19 percent of all hospital admissions for diabetes are preventable when proper education is available and practical application of that education is followed through with effective self-management (Geller & Butler 1982). There appears to be a high potential payoff in reducing direct hospitalization costs via self-help education programs, however, additional indirect benefits in terms of fewer lost work days, as well as personal benefits in the form of increased quality of life, might also be realized. Due to measurement difficulties, these benefits are rarely, if ever, included in evaluation studies in this area. The resultant understatement of the value of benefits has serious implications for resource allocation. Even though market prices are not evident for many of the costs that can be reduced or eliminated with educational interventions, these can be estimated with more

sophisticated techniques that use opportunity cost and shadow pricing concepts. Economists have applied these tools successfully in other areas of research and policy, and the magnitude and growth of health care expenditures provide more than ample motivation for their use in this area. Another problem in evaluation attempts is that the net benefits from intervention programs have not been presented in terms of who might gain from such strategies. In addition, the long term benefits of diabetes self-control, which may be extensive, are not covered in these studies.

In summary, it appears that education programs for tertiary stage diabetes may be a cost beneficial strategy. However, the entire breadth and depth of benefits have not yet been laid out in a standardized format which is accessible and useful for specific groups who might be interested in such outcomes. As Most, Sinnock and Alogna (1982) clearly point out, there is a need for greater uniformity in experimental design, intervention design, classification procedures and identification of objectives and outcomes in diabetes control interventions. Consistency among health program evaluations is necessary if evaluation is going to be used as an effective tool for resource allocation. While cost-benefit analysis can provide such consistency, its value and efficiency are limited if some of the costs are not included and if the recipients of potential benefits are not designated.

II. A Framework for Measuring the Costs of Diabetes. The costs associated with diabetes are used to illustrate the framework developed for measuring the costs of disease, however, it should be realized that this framework can be easily adapted to other chronic diseases. Figure I presents a data matrix in which rows of figures classify the effects of diabetes into three separate groupings: direct (diabetes related) medical costs, indirect costs and personal costs, and their sum. The aggregate cost to society is shown in the first column. The next three columns of the table exhaustively distribute these costs according to the parties affected by each type of cost: private third party insurers, the taxpaying public, and affected individuals and their families. Thus, this matrix presents not only the aggregate costs of diabetes, but also their distribution according to both types of effects and affected parties. It is best to explain this framework by considering the perspective of each of the parties interested in the costs of diabetes and the benefits from its diminution.

Third Parties. One of the goals expounded by proponents of preventive health care is the hope that third party payers would be willing to pay for preventive services which would produce a substantial long-run reduction in their costs for hospitalizations and related services. Insurance companies have been wary of such programs, but in the case of tertiary interventions, where the results can be gauged over a relatively short time span, there is a much greater chance that third parties will be willing to pay for preventive efforts. Many of the growing number of diabetes education intervention experiments have been designed exactly for this purpose. Because third party payers are generally interested in their savings alone, only direct medical costs related to diabetes are relevant to their decision to cover outpatient education programs.

It is relatively straightforward to utilize this one cell of the matrix to evaluate an outpatient education program for a third party provider. If the intervention study is designed on a before and after basis, baseline third party costs for diabetes and diabetes related claims can be measured over a given time period. Following the education intervention and an appropriate waiting period (1 to 2 years), third party diabetes related costs, including the direct costs of the education intervention, are counted again. If the study is set up on a before and after basis, proponents of the education strategy hope to show third party payers that the reduction in their share of direct medical costs for diabetes treatment exceeds their charges for outpatient education. If the experiment is set up on a scientifically preferred, randomly selected control group-treatment group basis, the difference in cost experiences between the control and treatment groups is compared to the third party costs of the intervention.

Taxpayers. A second relevant perspective is that of the "taxpaying public", in the case of diabetes, this sector is responsible for some portion of both the direct and indirect costs. To the extent that Medicare, Medicaid, Veterans Health Care or other public health insurance are responsible for paying diabetes related health care costs, the taxpaying public is directly affected by the disease in a pecuniary fashion. It is also important to measure the potentially high indirect costs of diabetes to the taxpaying public when severe diabetes related health problems result in lost working days and foregone earnings. The public income and payroll tax share of these foregone earnings affect public budgets, and the magnitude of these costs may be large. Moreover, should diabetes become so severe and disabling that individuals can no longer work and must rely on public income transfers as a means of support, taxpayer costs again increase to cover Social Security Disability Insurance, Supplemental Security Income and related income assistance program costs. Thus, to the extent that diabetes causes loss of earnings and dependence on public income support systems, the taxpayer is doubly affected.

The estimation of direct costs savings to taxpayers from short term diabetes intervention is straightforward, it need only be remembered that public funds used in diabetes intervention projects be included as part of the public direct costs of diabetes. At the baseline measurement points, indirect costs in terms of foregone tax revenues from lost earnings can be estimated by calculating the tax share of the estimated foregone earnings over the period in question. These foregone earnings should be adjusted for "normal" spells of unemployment which workers might be subject to regardless of their diabetic condition. Similarly, indirect costs in the form of public expenditures for diabetes related disability and income support payments can be counted, provided that adjustments are made for non-diabetes re-

lated probabilities of recipients of such benefits, for instance, income support based on low incomes alone. The cost savings from diabetic education programs can then be measured by the lower direct outlays for diabetes treatment from public health care funds and from higher tax revenues due to increased work and earnings resulting from better diabetes control. Sizable gains in this area may be noted over a one to two year period. Indirect taxpayer benefits from reduced outlays for income support might also be realized, but the short-term nature of a diabetes education intervention strategy would likely show little gain in this area, particularly if public income support is related to the long term complications of diabetes which are not likely to be affected by the intervention.

Individuals and Their Families. The final cost-saving perspective is that of the individual diabetic and family members. Reductions in out of pocket medical expenses for diabetes related hospitalization and other medical costs can be estimated and compared to the direct personal dollar costs of an education intervention program (including transportation costs, ancillary costs and patient fees paid out of pocket). One might find that in the case of heavily insured individuals, outlays for the education intervention program would exceed their personal share of diabetes related costs. But such a result would be likely to defer the individual from pursing the program only if the analyst (and the diabetic!) ignored the indirect and personal costs of poor control.

Indirect costs to a diabetic and family can be measured by the net income foregone due to work time lost on account of the illness. This includes lost after-tax earnings net of any compensating public income support received when the diabetic is unable to work, or when another family member must reduce work hours to provide care.[1] In addition, non-market work may be lost due to diabetes related illness, and it is necessary to impute a wage equal to the opportunity cost of time lost from housework in order to measure the value of these services foregone (Cooper & Rice 1976). In the case of diabetes related premature mortality, the present value of lost net earnings summed over the remaining estimated worklife should be counted as an indirect cost of diabetes. Over a short-term treatment intervention the indirect effects of better diabetes control can be estimated for those who realize an increase in net income from added work hours. These indirect benefits will probably be large enough to outweigh the potential net direct personal cost of the education intervention (including lost earnings due to program attendance), providing the program is successful.

The final category of personal diabetes costs deals with quantifying its effects on the diabetic's quality of life. The prevalence of excluding these less tangible costs is particularly critical for chronic diseases which may be non-life threatening, but which impose a substantial quality of life cost on the affected individual.

However, it is possible to estimate these costs and the reduction in such costs from education intervention designed to diminish the effects of diabetes. While poor control can in some cases limit one's ability to work (which will add to indirect costs), often diabetics are able to continue with their jobs, albeit at a below capacity rate. Diabetics who either cannot (or will not) work may also experience diminished functionality and lower quality of life due to diabetes side effects, particularly ketoacidosis. These costs can be measured by survey techniques which indicate (1) the average number of low quality/low effectiveness days per month which diabetics feel they are functioning at less than their normal capacity, and (2) the percentage of normal capacity at which they are able to function during these days. Taking the number of days below full capacity and the percentage of full capacity at which the diabetic functions on those days, and comparing this to the quality of life on days of normal functionality, an estimate of quality of life foregone can be derived. If a diabetic indicates that in an average month functioning is only at 50 percent capacity for roughly five days, the quality of life foregone over a typical month is: five days per month X 50 percent, or three days per month. When divided by thirty days per month, this equals a 10 percent dimunition in overall functioning. In order to quantify these adverse affects, one needs an estimate of the value of this portion of life over and above the portion of the value of life captured by the indirect costs of diabetes which have already been counted.[2]

Perhaps this argument deserves additional discussion. Conceptually an economist would like to measure the value of life by estimating the amount of money which an individual would be willing to pay to extend life for a given period. Since life is a non-marketed good, one is not able to directly observe or derive such values from market behavior. In theory, it seems clear that the overall value of life should exceed foregone earnings if one values non-work time and related consumption activities - the quality of life - above zero. Thus, we would expect the value of life to exceed the indirect costs of illness by the difference between the total value of life and lost earnings. That is exactly the portion of quality of life which we seek to measure as a personal cost to the diabetic, if we are correct, the total value of life should always exceed lost earnings. While direct observation of willingness to pay is not possible, shadow pricing can be employed to estimate these values. Economists who have attempted to estimate the value of saving a life from observed production or consumption behavior, by extrapolating the differences in prices people are willing to pay or are willing to forego to reduce the probability of death, have found this to be universally the case. Blomquist (1981) has found the value of life over and above foregone earnings, as estimated by several studies, to be in a range from $.212 million to $2.401 million in 1979 dollars.[3] The ratio of these values to lost earnings varies, with four studies finding a ratio between 1.5 and 5.9. Taking the average of these studies we find that the percent value of a life saved is roughly 3.2 times as great as lost earnings.

It is fair to assume that the factor of 3.2 is a reasonable ratio of the monetary value of non-working time (value of life over and above earnings) to working hours (foregone earnings) when arguing that at the margin, the value of the last hour's work is equal to the value of the first hour of non-working time. Assuming that the marginal value of the last hour of work (first hour of non-work time) is equal to the average value, one can look at the ratio of non-working hours to working hours over time. If the average work week is forty hours, the ratio of non-working hours to working hours is: 168 - 40, divided by 40, from this perspective, the ratio of 3.2 appears to be a reasonable estimate.

Assuming this relationship holds across a short period of time, and assuming that an X percent reduction in the quality of life is equivalent to an X percent reduction in the value of life over and above earnings loss, the aggregate dollar value of reduced quality of life for diabetics can be estimated. Foregone gross earnings are multiplied by 3.2 and then by the "percentage of life capacity lost" estimate derived from the survey questions posed. For instance, assuming a 10 percent dimunition in overall functioning and an annual earnings loss of $1,136 in 1979, the value of the quality of life lost would be: $1,136 X 3.2 X .10, or $364 for the typical diabetic over a year's time in 1979.[4]

Since we have only estimated the value of diminished life quality for the diabetic alone, there are still costs to be accounted for. If diabetic illness also affects the quality of life for other family members, they may be willing to pay additional amounts to avoid those aspects of the disease which reduce their own quality of life. Such effects could be measured by asking a diabetic's family how much they would be willing to pay for a non-life threatening procedure which would cure diabetes. Additional research will indicate the viability and robustness of estimating personal costs using such experimental measures as these.[5]

In summary, the total direct, indirect and personal costs of diabetes can be broken down according to the group which bears the burden of those costs: third party payers, the tax-paying public and the individual diabetic (and family). Such an accounting framework provides the analyst with cost and net benefit results which can be easily communicated to each of these interested parties who have a financial stake in diabetes related expenditures.

Overall Social Perspective. Most studies of the cost of illness or diabetes do not break down the total costs this way, but concentrate on the differentiation between direct and indirect costs. In these studies direct costs are estimated as in this paper, but indirect costs are measured by gross earnings foregone, plus the value of home production foregone. The reader should note these exact figures can be arrived at by summing across the first two rows of the table in Figure 1. The summation of the first row, direct costs, is straightforward; the second row sums neatly into gross earnings foregone. Net after tax earnings lost and foregone taxes equal gross earnings foregone, and public income support payments received in the third column exactly cancel public income support payments made in the second column. The matrix indicates that while income maintenance payments are usually excluded in measuring the cost of illness they are important from a distributive point of view since they are real tangible benefits to one party (diabetics) while at the same time they are equivalent aggregate costs to another party (taxpayers). By using this differentiation, the net pecuniary cost to the taxpayer and the net income loss to the diabetic are separately and directly measured, rather than being lost in the indirect cost aggregates, as in other studies. The next section of the paper uses these "social cost" estimates to derive a disaggregated estimate of the cost of diabetes.

III. The Aggregate Cost of Diabetes. Using data on diabetics from Entemacher (1982), Marks (1980) and Lipset (1982), along with federal government data on payers for various types of health care expenses and other items, we have derived an estimate of the aggregate economic cost of diabetes for 1980, as shown in Figure 2. The reader is cautioned that these are, in some places, only rough estimates drawn from secondary sources. More accurate figures are being developed from primary data, but are not yet ready for publication.

Direct payments of $5.66 billion were reported by type of expenditure by both Entmacher and Marks, these were distributed across the categories of payers according to national medical care expenditure data which reports these distributions for each type of expense. Data on the indirect costs of diabetes, which totaled $10.03 billion, include earnings losses from morbidity, premature mortality, and diabetes related complications which limit work activity, as well as opportunity cost imputations for time lost in household production. Taxes foregone due to lost market earnings were estimated at 26 percent of earnings, or $2.61 billion. Using data reporting the number of disabled receiving Social Security, Supplemental Security Income, or other related disability benefits, and average annual 1980 benefit levels for such persons, we have calculated that about $.95 billion of income transfers were received for diabetes induced causes in 1980. Personal costs were estimated using figures presented in Section II which are based, in part, on data reported by Lipset and Marks. The $364 1979 individual loss was adjusted to 1980 dollars, then multiplied by the 8.27 million diabetics who were estimated to have experienced these losses to arrive at an aggregate figure of $3.22 billion.

Assuming these figures are accurate in the aggregate, the standard Cooper-Rice framework would have estimated similar direct and indirect costs. The analysis in Figure 2 expands upon the information given in these costs in two ways. First, $3.22 billion (20.5%) is added to the total to account for previously unmeasured personal costs. Thus, the first column showing total social costs of $18.91 billion exceeds the Cooper-Rice estimates by $3.22 billion. Second, the final three columns, which disaggregate total costs according to the responsible party, are added. This reveals that only

about 6.1 percent of the total cost of diabetes is borne by third party payers, 29.4 percent by the government, and 64.5 percent by individuals. Because group and/or individual private insurance premiums are usually determined on a non-experience rating basis, diabetics shift part of the cost of their illness to other private insurance holders in the form of a higher premium. While the incidence of this "tax" is probably different than what governments levy, one could argue that the entire non-personal cost of $6.72 billion is borne by the taxpaying public.

Both the taxpaying public and diabetics and families suffer indirect costs which far exceed their direct costs. As compared to the usual Cooper-Rice estimates, the figures differ in two significant ways. First, while diabetics lose $10.03 billion in foregone earnings, their disposable incomes fall only by $6.47 billion. This differential is due to the $2.61 billion in taxes which they would have paid on earnings and the $.95 billion of compensating transfers. The second difference is that Figure 2 identifies pecuniary taxpayer costs of $3.56 billion due to lost tax revenues and increased transfer payments to diabetes. This exceeds the direct public cost of $2.01 billion in Medicare, Medicaid, Veteran's and other programs for diabetics, by 70 percent! Thus, if one argument for governmental involvement in diabetes eradication is the budgetary costs of the illness, counting only direct costs sorely underestimates these losses by counting only about 36 percent of the total goverment cost.

The final item of concern is the individual diabetic who bears the largest overall share of indirect costs and all of the personal costs, as compared to about 44 percent of the direct costs alone. When presented in this manner, diabetics may be more eager to accept the strict daily regimen necessary for proper control due to the high indirect and personal costs, in addition to the direct medical bills which they must bear.

IV. Summary and Conclusions. This paper serves many masters, however, it is hoped that it has presented at least three useful items for health statisticians, health economists, and health policy analysts: (1) a framework for disaggregating and measuring the cost of diabetes, (2) a method for estimating the net dollar cost of reduced quality of life for diabetics, and (3) an illustration of the relative and absolute dollar importance of each component in the total social cost of diabetes.

The potential usefulness of this cost classification scheme and the methodology for estimating the net dollar cost of reduced quality of life can be easily extended to other chronic or acute health problems. In particular, if the methodology for estimating the net dollar cost of reduced quality of life proves useful, the high personal cost of several non-life threatening, but terribly debilitating and painful chronic diseases can, for the first time, be quantified. We urge others to undertake the research necessary to refine the methodology and improve the estimates presented here.

1. It is probably that some diabetics will receive sick pay, private disability insurance benefits (even severance pay in the case of job loss) while unable to work. In this case, diabetes serves as a tax on the employer and fellow employees whose premiums for disability insurance will increase.
2. When applying this framework in cost-benefit situation of intervention strategies researchers should be careful to net out the personal costs of complying with the daily regimen (e.g., blood testing, insulin taking, dietary limitations). In fact, one might argue that the failure of many preventive health care interventions stems from the fact that the personal costs of compliance are perceived to exceed the personal costs of the illness itself.
3. Another method of ascertaining these amounts is to ask individuals how much they would pay to avoid the risk of death; studies using this approach have resulted in a wide range of estimates.
4. This earnings loss was reported by Marks (1980) as the average per capital indirect cost of diabetes related morbidity in 1979.
5. A Utah Department of Health, Centers for Disease Control diabetes project is currently experimenting with these measures. See Smeeding (1983) for more on this topic.

REFERENCES

1. Blomquist, G., 1981. "The Value of Life: An Empirical Perspective," ECONOMIC INQUIRY, January, pp. 157-164.
2. Cooper, B. and Rice, D., 1976. "The Economic Cost of Illness Revisited," SOCIAL SECURITY BULLETIN, February, pp. 21-36.
3. Entemacher, P., 1981. "Economic Impact of Diabetes Tables," Statistical Bureau estimates, unpublished data. (Metropolitan Life Insurance Co., New York).
4. Geller, J. and Butler, K., 1982. "Study of Educational Deficits as the Cause of Hospital Admissions for Diabetes Mellitus in a Community Hospital," DIABETES.
5. Lipsett, L., 1982. "The Economic Cost of Diabetes," Mimeo for Z. Laren, Ed., PEDIATRIC AND ADOLESCENT ENDOCRINOLOGY (based: S. Karger Publishers).
6. Marks, R.G., 1980. "Diabetes: The Expensive Disease," DRUG TOPICS, September 2, pp. 55-56.
7. Most, R., Sinnock, P. and Alogna, M., 1982. "Evaluation of Diabetes Control Interventions," PROCEEDINGS FROM THE FIFTH ANNUAL DIABETES CONTROL CONFERENCE, pp. 103-106.
8. Neresean, W. and Zaremba, M., 1982. "Impact of Diabetes Outpatient Education Program - Maine," MORBIDITY AND MORTALITY WEEKLY REPORT, June 18, pp. 307-314.
9. Rice, D. and Cooper, B., 1967. "The Economic Value of Human Life," AMERICAN JOURNAL OF PUBLIC HEALTH, Vol. 57, No. 11:1954-1966.
10. Smeeding, T.M., 1983. "Measuring and Valuing the Benefits of Diabetes: The Economic Perspective," Presented to the 6th Annual CDC Conference on Diabetes, May 11, 1983.

FIGURE 1
THE ECONOMIC COST OF DIABETES

DISTRIBUTION OF COSTS
BY INTEREST GROUP

TYPE OF EFFECT	"OVERALL SOCIAL COSTS"[3]	PRIVATE THIRD PARTIES	PECUNIARY TAXPAYER COST	PERSONAL (INDIVIDUAL/ FAMILY) COST
Direct Medical Costs[1]	Total direct medical costs	Private third party payments for medical costs	Public payments for medical costs	Out of pocket payments for medical costs
Indirect Cost	Gross earnings loss[4]	N/A	Foregone tax revenues from earnings loss plus: income maintenance support benefits paid	Net after-tax earnings loss due to morbidity or mortality[2] minus: income maintenance support benefits received
Personal Cost	Cost of diminished quality of life	N/A	N/A	Cost of diminished quality of life for diabetics & their families
Total Costs[5]	Total social cost	Total private third party cost	Total pecuniary taxpayer cost	Total personal cost

FIGURE 2
THE ECONOMIC COST OF DIABETES ($ BILLIONS) IN 1980 AND
THEIR DISTRIBUTION BY INTEREST GROUP

TYPE OF EFFECT	"OVERALL SOCIAL COSTS" 3 (column %)	PRIVATE THIRD PARTIES	PECUNIARY TAXPAYER COST	PERSONAL (INDIVIDUAL/ FAMILY) COST
Direct Medical Costs[1]	$ 5.66[1] (39.9)	$1.15	$2.01	$ 2.50
Indirect Cost	$10.03 (53.0) 0.00[4] 10.03 total		$2.61[6] (foregone taxes) + .95[7] (transfers paid) $3.56 total	$7.42[2] (after tax earnings) - .95[7] (transfers rec'd) $6.47 total
Personal Cost	$ 3.22 (17.0)			$3.22[8]
Total Costs[5] (row %)	$18.91 b (100.0)	$1.15 (6.1)	$5.57 (29.4)	$12.19 (64.5)

[1]Includes "diabetes related costs" only; distributed by source of payment.

[2]Includes imputed value of lost home production, and lost earnings or home production for other family members caring for diabetics.

[3]Row sums for each category of effect.

[4]Note that public income maintenance benefits paid (second column) and benefits received (third column) exactly cancel.

[5]Column sums for each interest group and for overall social costs.

[6]Taxes calculated at 26.0%, average federal, state income taxes plus payroll taxes.

[7]Includes 25,000 blind; 116,260 disabled; 50,000 other assistance recipients, times average benefits for each, summed.

[8]Average cost of $364 per diabetic x 8.3 million diabetics in 1979, adjusted to 1980 (see page 21 for calculations).

Sources of Estimates: Lipset, 1982; Entemacher, 1982; Marks, 1980; Social Security Bulletin, 1982; U.S. Statistical Abstract, 1982.

Health Education

Session E

BEHAVIORAL RISKS OF PREGNANT ADOLESCENTS

Paul A. Hensleigh and Nancy Moss, Stanford University School of Medicine

Toxic substance use and delay in seeking prenatal care are health behaviors which contribute to adverse outcomes of pregnancy. As a first step toward developing an intervention program to improve health behaviors during pregnancy, we have studied the social context of pregnancy in adolescents with particular attention to health knowledge and health behaviors. In this population we anticipated prevalence of certain potentially modifiable and probably adverse health behaviors: smoking of tobacco and marijuana, drinking of alcohol, and delaying onset of prenatal care. We sought to more thoroughly understand the target population and the factors which would affect their pregnancy behavior: their social network, important events in their everyday lives, and their family backgrounds.

Methods

In the first stage of the research intensive interviews were conducted with a non-random sample of 40 pregnant adolescents age 17 and younger, of whom 20 were Anglo and 20 were of Mexican descent. Several factors emerged as important lifestyle characteristics of adolescent pregnancy: school enrollment, actual and perceived stress, social support, information seeking, future orientation, orientation to the infant, preventive health behavior, and parents' substance abuse.

In the study's second phase a survey was administered to the population of 93 adolescents, 17 years and younger, who delivered at five San Jose hospitals in July and August, 1982. Table 1 shows the age, ethnicity, birthweight and time of first visit for this population and also shows the similarity to all adolescents delivering in Santa Clara County in 1981. The low incidence of low birth weight confirms our knowledge from review of the State Health Department records that this county has relatively infrequent adverse pregnancy outcome compared to other California counties.

From the structured interviews of the study population we collected data on the incidences, amount and modifications of substance use during pregnancy and the social contextual factors first identified in the intensive interviews. Table 2 lists and defines the independent variables. For each dependent variable (incidence, amount and change in substance use, and timing of prenatal care), we fitted two multiple regression models. A reduced model consisted only of factors which zero-order correlations had shown to be associated with the dependent variables. A full hierarchial model consisted of these social contextual factors, plus age and ethnicity. Only Anglos and Hispanics were included in these analyses. Age was included because a number of studies of adolescent substance use have shown that of substance use incidence increases with age. Ethnicity was included because an important question addressed by the study was whether health related behavior of pregnant adolescents differed, depending on ethnic factors. Although phrased predictively, the analysis is really descriptive and the results should be regarded as suggestive rather than conclusive.

Results

The incidence and amount of substance use reported prior to pregnancy is shown in Table 3. Knowledge among the adolescent mothers about pregnancy risks associated with substance use is shown in Table 4. Especially with regard to tobacco and alcohol their information base was substantial. About one fourth of the sample knew how drinking and cigarette smoking could harm the infant. However, over half didn't know of an effect or gave an incorrect response. Others simply reported substance use as being "bad for you and bad for the baby." While classes, books and the media are the most prevalent sources of information, a number of teens used personal fables and hearsay to justify their substance related beliefs or actions. Often the fables were examples drawn from their observations of what happens to other peoples' babies. There were no ethnic differences in knowledge levels or in sources of information.

Of adolescents who had used a particular substance, the per cent who decreased their use in relation to pregnancy by at least one amount category are shown in Table 5. Although these data are based on self reports, it appears that most of these girls were motivated to reduce substance use before or during pregnancy even beyond the extent of their understanding the associated health risks.

Incidence of smoking was associated with different predictors than amount of smoking using the full hierarchial regression model. Whether or not a pregnant teen had smoked was affected positively by the extent to which her parents smoked or drank and negatively by the extent to which she had received social support from her partner. Alone these two factors explain 33% of the variance in smoking incidence among pregnant teenagers.

As we would expect the amount of smoking among those who smoked was associated with perceived stress. Additionally, of all the dependent variables studied, amount of smoking was associated with ethnicity. Chicanas smoke less than Anglos when factors such as school enrollment and stress are taken into account.

Marijuana use is smoking for many adolescents in San Jose. Marijuana was widely available and inexpensive. But unlike cigarette smoking, it was associated with social life and partying. Descriptions of marijuana use virtually always referred to a joint being passed around or at least being used in the company of other users. As pregnancy progressed into the second trimester, teens generally began to stay at home rather than party and this almost assured that their marijuana use would be curtailed. There were no apparent ethnic or age differences in incidence and amount of marijuana used. Factors which appeared to make a difference were parents' substance use, perceived stress and active social support. When the variation introduced by parents' substance use was accounted for in a regression model, active support made it more likely that the teenager would use marijuana but also more likely that she would use less and quit sooner. Active social

support should probably be viewed as a surrogate. It probably reflects the close and cumulative attention which some pregnant teens had available to them from family members and partners. Teens who received this kind of attention probably had fewer opportunities and less need to smoke marijuana.

It was much more difficult to identify the variables associated with drinking than cigarette or marijuana smoking. One explanation for the lack of clear results with multiple regressions is that normalizing the sample distribution on drinking measures may have obscurred how pregnant adolescents really use alcohol. From the intensive interviews it appeared that those who acknowledged "drinking" could be divided into two major groups: those who had an occasional drink on a special occasion and those who more regularly drank substantial amounts in the setting of parties. All of the former group and some of the latter stopped drinking when they realized they were pregnant. The few who continued drinking had some interesting characteristics: (1) they were more likely than peers to have been birth control users, (2) they were more future oriented, (3) they were from non-intact families, and (4) they were more likely to drink wine and/or liquor rather than beer. Accounting for age, ethnicity, active support and stress: both use of birth control and future orientation have an impact on drinking. It is particularly important that age was held constant in the analysis because older teens are more likely to have used birth control and also were able to anticipate and plan the future. It is also intriguing that while birth control users were more likely to drink, drinkers who used birth control were likely to stop earlier than non-users.

Timing of the first prenatal visit was found in this adolescent population to be strongly linearly related to age. The younger the adolescent the later she obtained care. This study also shows that when age was accounted for, first visit timing could also be affected by the social support available to the teen from her partner and his family. In contrast, parental support was associated with later, not earlier, care. This latter finding may relate to a number of issues or events which the pregnant teen must deal with before prenatal care is obtained. From the intensive interviews it was apparent that most girls would first seek out a pregnancy test, tell her partner and probably her mother; and only when these "prerequisites" had been accomplished would she seek out prenatal care. Thus, if a younger teen or one whose primary relationship was with her parents were more anxious about communicating with her parents, this could serve as a barrier to her pressing on to the point of seeing a physician. Another factor contributing to these findings could be that teens who were closer to their partners (as opposed to their family) may have had less ambivalence about continuing the pregnancy and, therefore, acted more quickly to arrange prenatal care.

Conclusions

In considering generalizability of this study's findings, we need to take account of two issues: the population from which observations were drawn, and generalizability of the model. Essentially, this should be regarded as a case study of a population with the findings generalizable only to populations exactly like this one.

For policy makers interested in how the findings would apply to pregnant adolescents across the nation, the question is, do these regression models apply elsewhere? Are the inferences drawn from these models applicable to teenagers in Washington, D.C. or New York state or Miami? The model applies only to the extent that the population looks like San Jose's. For example, although we included ethnicity as a variable, no blacks were included in the regressions; with only a handful of blacks, the standard errors of the coefficients would have been very large, and the models would have been difficult to fit.

This study should be regarded as suggestive. It brings to attention factors affecting health behavior in pregnant adolescents which should be included in systematic studies done on larger samples of more diverse populations. Those populations should be more representative of the actual population of pregnant adolescents.

Finally, a striking and recurrent message expressed in various ways by the adolescents in this study is that pregnancy was a time of intense stress and anxiety, usually superimposed on an already stressful existance on the economic margin of the community. One 15 year old put it this way:

"I think the environment that you're in, that's alot of what's important when you're pregnant. You should be in a really good, happy environment, something that you're content with. But if you're all uppidity and all under stress all the time and everybody's always yelling and everybody's always telling you - do this, do that , do this! And you're exhausted and you've been trying to make everything work out right and nothing seems to match up anytime...you just feel like closing the door and saying goodnight."

Our future research interests have been directed by this observation toward a hypothesis testing study of the influence of stress on perinatal events and pregnancy outcome. More complete understanding of the influence of the biophysical as well as the behavioral influences of stress may help to explain and more appropriately approach the pregnancy complications of adolescents and other groups "at-risk" and to plan successful interventions to improve pregnancy outcomes.

A full report of this study and further analysis of the data is contained in a document titled "A Model of Adolescent Perinatal Risk Behavior" prepared for the Maternal and Child Health and Crippled Children's Services Research Grants Program. Copies of this report on Grant #MCR-060466-01-0 may be obtained for a fee from the National Technical Information Service, U.S. Department of Commerce, Springfield, Virginia 22161.

Table 1

Hospital Sample and Santa Clara County Populations: By Age, Ethnicity, Birthweight, and Timing of First Prenatal Visit

		Hospital Sample N=93 (%)	County Population N=947 (%)
Age	13 & under	1	0.1
	14	3	5
	15	16	16
	16	31	31
	17	48	48
Ethnicity			
	Anglo	25	38
	Hispanic	62	49
	black	10	8
	Other	3	5
Birthweight			
	Under 2500	7.6	7.8
	2501-4309	89.1	90.2
	Over 4310	3.3	2.0
Timing of First Prenatal Visit			
	No care	2.2	1.4
	1-3 mos	54.8	51.0
	4-6 mos	39.7	35.0
	7-9 mos	3.3	8.0
	Missing	0.0	4.6

Table 3

Use of Cigarettes, Marijuana and Alcohol Before Pregnancy Acknowledged by Adolescents

	Frequency
Cigarette Smoking	
None	46%
1-3 day	8%
½ pack/day	18%
½-1 pack/day	14%
Pack/day	12%
More than pack/day	2%
Marijuana	
None	44%
Rare	5%
1/week	23%
1-5/week	17%
6-10/week	7%
More than 10/week	4%
Beer and Wine	
None	45%
Rare	29%
1-5/week	24%
6 or more/week	2%
Liquor	
None	65%
Rare	26%
1-5/week	8%
6 or more/week	1%

Table 2

Measures: Independent Variables in Regression Analysis

Construct	Measures
Background Variables	
Ethnicity	
Marital Status	
Age	Age in Years
School enrollment	Quit school prior to pregnancy, during pregnancy, or still enrolled
Instability of living situation	Moved during pregnancy
Use of birth control	Use prior to pregnancy
Parents' substance use factor	Mother's drinking and smoking and father's drinking and smoking
Social Support	
Parents' support	Parent accompanied in labor, visited in hospital, girl lived with parent during pregnancy, named parent as source of support
Partner/family support	Partner and/or family accompanied in labor, visited in hospital, gave financial support, girl still with partner, told partner about pregnancy before mother
Partner only support	(Same as above - partner only)
Active support	Number who accompanied girl in labor
Psychosocial Variables	
Perceived stress	During pregnancy perceived frequent moves, change of friends, discouragement, boredom
Future orientation	Active plan for future in 2 years, specific reference to partner and baby in future
Information seeking	Took prenatal class, hospital tour, obtained written pregnancy information on own.
Baby care orientation	Has definite child care, sick baby and well baby plans

Table 4
Knowledge of Substance Use on Pregnancy

	CIGARETTE	MARIJUANA	ALCOHOL
Correct Response*	25%	14%	24%
Bad for You	18%	14%	25%
Unknown	25%	37%	26%
Incorrect Response	32%	38%	24%

* Cigarette smoking = LBW; Marijuana = effect unknown or none; Alcohol = birth defects or LBW

Table 5

Time When Substance Use Changed
in Relation to Pregnancy

Cigarette Smoking	
Prior to Pregnancy	21%
Trimester 1	56%
Trimester 2	16%
Trimester 3	7%
Marijuana	
Prior to pregnancy	56%
Trimester 1	37%
Trimester 2	7%
Alcohol	
Prior to pregnancy	23%
Trimester 1	77%

TELEPHONE RISK FACTOR PREVALENCE SURVEYS

James S. Marks, Gary C. Hogelin, Jack T. Jones,
Eileen M. Gentry and Frederick L. Trowbridge, Centers for Disease Control

Lifestyle behaviors - those adopted by personal choice - are strongly associated with many of the leading causes of death after infancy. Similarly, clinical and epidemiologic studies have linked several of these lifestyle behaviors to serious illness. Because of this, many of the 1990 Objectives for the Nation regarding health and prevention of disease target reductions in the prevalences of selected risk factors. The participation and contribution of State health agencies are essential if we are to reach these goals since programmatic efforts aimed at risk reduction are often carried out and/or coordinated through these agencies. Yet, particularly at the State level, no system exists to enable States to develop estimates of behavioral risk factor prevalence in their own population or to monitor progress toward the goals of risk reduction. Furthermore, knowledge of the prevalence of these risk factors in their jurisdictions will help State health agencies decide how to best allocate resources for risk reduction.

This need prompted several States to carry out surveys to determine the prevalence of these risk factors. However, few States had the technical expertise to design and carry out such surveys. Furthermore, large differences in methodology and questionnaire design precluded easy comparisons between States or with available national information.

Beginning in 1981, The Centers for Disease Control (CDC) developed a short questionnaire designed for telephone administration concerning major behavioral risk factors. This questionnaire was made available to interested States along with training in survey operations and assistance with data processing and basic analysis. Since that time, 27 States and the District of Columbia have conducted surveys using the CDC questionnaire either in its entirety or with minor modifications.

In this presentation we will discuss findings of behavioral risk factor surveys both in terms of outlining the methods and questions used and in terms of the State-to-State variation in prevalences. Finally, we will discuss the implications of this variation in risk factor prevalence and outline some directions the CDC is taking.

Methods

Because of the perceived needs of the States for an inexpensive, relatively simple method to gather risk-factor data, CDC chose to use telephone interviews of randomly selected households much like a polling organization would. A standard questionnaire was developed using questions from previously conducted national surveys such as the Health Interview Survey and a National Heart, Lung, and Blood Institute survey. Only questions on exercise were developed at CDC since no standard set of questions on this topic existed. The basic philosophy behind the questionnaire was to concentrate on actual behaviors rather than on attitudes or knowledge. It was purposely kept very short, taking less than 10 minutes to complete, in order to permit the individual States to add questions of local interest without overburdening the respondent. The core questionnaire provided data that could be compared between States

This core questionnaire provides a few questions on the following major risk areas: smoking, exercise, alcohol misuse (including drinking and driving), obesity, hypertension, stress, and seat belt use. Basic demographic data are included as well as appropirate transitional wording throughout the questionnaire.

We provided on-site instruction in use of the Waksberg method for selection of the telephone numbers. Basically, in this method a random sample of blocks of 100 telephone numbers is selected from among all possible blocks of numbers within the State. The actual number of blocks selected is based on the desired sample size. The blocks are then screened by calling one number from the block to determine which blocks are residential and which are primarily business phones. Only those blocks in which the screening call reaches a working residential number are retained for final sampling. This prescreening improves the later efficiency of the interviewers by deleting those groups of numbers that are largely business. Next, the actual numbers to be called are obtained from these working blocks by randomly generating the last two digits of the telephone numbers. Usually it is desirable to complete three interviews per block of 100 numbers. From each block as many numbers as needed were generated to yield the desired number of interviews.

The interviewer uses a random selection chart (based on the last digit of the telephone number and the number of adults in the household) to select which adult is to be the respondent. This avoids the selection bias associated with who answers the phone and time of the day when the call was placed. Finally, the number of separate phone lines in a household is determined to permit adjustment for the increased likelihood of selection of households with two or more lines.

The interviews typically were conducted on evenings and weekends. In some States, health department personnel conducted the interviews; in others, students; and in others,

interviewers with survey research firms were used. Usually the interviewers were female. Health agency personnel supervised questionnaire editing and monitored the interviews and survey procedures. Training in these functions was provided by CDC staff.

Surveys in each State followed a given set of procedures for identifying eligible respondents and for assuring that there was an adequate attempt to reach a respondent at each selected telephone number. Only supervisors could replace a number that could not be reached or where the respondent refused an interview. Each interviewer was periodically monitored during interviewing and verification or repeat calls were made on a portion of completed calls to monitor interviewer compliance with the protocol.

Each interview took 8-10 minutes to complete. Considering call backs, no answers, etc., two interviews generally could be completed per hour of interviewer's time.

The overall response rate was 70% to 75% after exclusion of business and nonworking numbers. Primary reasons for nonresponse included respondent refusal and no answer despite repeated calls. This rate of response is somewhat higher than that usually obtained by private polling organizations.

Results

Specific findings from the State surveys are presented in the table. First, in response to the question, "Do you smoke cigarettes now?", a median of 31.9% of persons in the 28 locations indicated that they are current smokers. The range of positive responses was from 23.4% to 37.4% among the States.

In general, two States with small surveys (about 500 people) that are being compared will need a difference in response of about 6% to have statistically significant differences in smoking prevalence. Thus, the differences between States with high and low rates are likely to be statistically significant.

To estimate obesity, persons were asked their height and weight. Other studies have shown that self-reporting of this type is fairly accurate (within 0.5 inches and 1-3 pounds) when compared to objective measurement. For this presentation, we used 120% of the figures in the 1959 Metropolitan Life Insurance tables as our measure of obesity. Overall, the median proportion of persons meeting or exceeding this standard measure of obesity was 23.4% with a range of 16.4% to 28.2%. States with the highest and lowest prevalence are shown. Again, differences of approximately 6% are statistically significant when comparing two States prevalences.

Several questions were asked to determine how much people exercised. First, interviewers asked respondents how often they exercised vigorously. Responses were coded according to the number of times the respondent exercised per week or month and average duration of each time. Next, the interviewer asked about frequency of light exercise such as gardening, bowling and golfing. Finally, the interviewer asked whether work-related activity could be characterized as light (sitting), moderate (walking), or heavy (pushing or carrying heavy objects). Persons who exercised vigorously less than 1 hour per month, and who exercised lightly less than 8 hours per month, and whose work activity was light were considered having a sedentary lifestyle.

Overall, the median percentage of adults with sedentary lifestyles was 12.3% with a range of 5.7% to 17.7%. Differences of approximately 4% are considered statistically significant.

Seat belt use was determined by a direct question as to frequency of use with responses permitted from always to never. The median prevalence of never or seldom wearing seat belts was 60.5% of the adults with a State range of 41.6% to 71.1%. Prevalences of usage differences that are greater than 7% are statistically significant.

Interviewers asked a series of questions on hypertension including when the respondent's blood pressure was last checked, history of hypertension at any time in the past, if the respondent was treated for hypertension, and whether the blood pressure was still high.

The median prevalence of those who stated that their blood pressure was still high was 3.8% with a range of 2.0% to 6.8% among States. For these estimates, differences of approximately 3% are statistically significant.

Finally, a series of questions on alcohol misuse were asked. From the responses, the prevalence of chronic drinking (defined as consuming an average of two or more drinks/day) was determined. A median of 8.3% of the population in these States reported chronic drinking with a range of 3.3% to 14.5%. Differences between States are significant if they are greater than 4%.

An estimate of acute or binge drinking was derived by asking respondents if at any time in the past 4 weeks they had drunk five or more drinks on an occasion. The median prevalence of persons who admitted to drinking five or more drinks on an occasion, one or more times in the last 4 weeks was 22.5% with a range among the States of 7.8% to 30.5%. A difference of between 5% and 6% is statistically significant.

Interviewers also asked respondents how often during the past 4 weeks they had driven after perhaps drinking too much. This obviously subjective measure probably has a built-in bias toward underreporting; yet, surprisingly, a median of 5.2% of adults admitted to this. The range among States was from 1.9% to 10.3% with a difference of approximately 3% being statistically significant.

Discussion

Lifestyles vary greatly according to sex, age, education, and other factors. Many of

the States have examined their individual surveys with regard to these demographic factors. Analyses such as these will help them target high-risk populations for intervention. This is an important first step for developing effective prevention programs.

These surveys demonstrate the feasibility of obtaining State-specific estimates of the prevalence of behavioral risk factors using telephone survey techniques. Furthermore, the relative ease and low cost of telephone surveys has permitted States to use the technique for programmatic purposes. Several States have assisted local health agencies in conducting surveys to provide base-line information before the initiation of local intervention projects.

Telephone surveys have several disadvantages, however, when compared to personal interviews. Telephone interviews usually have a higher refusal rate. While these States surveys had relatively low refusal rates for this survey technique, refusal rates were higher than those for personal interview techniques. Also, not everyone has a telephone and the population that is without a phone is clearly different from that which has one. Those persons without phones are likely to be at higher risk for many illnesses and health problems of interest. However, in the United States overall about 93% of households have a telephone.

Finally, an important question concerning telephone surveys is their reliability. How accurately people will respond over the phone is largely unknown. Evidence, where it exists, has been conflicting. Some authors have shown relatively lower rates of undesirable behavior when telephone surveys are used. Others have suggested that the anonymity of the telephone interview process leads to slightly higher rates of reporting undesirable behavior.

Despite these disadvantages, we believe that the advantages make this an attractive technique for obtaining important health information that is otherwise difficult to determine. These advantages are: 1) the extremely low cost - these surveys can be completed for less than $3,000 in direct costs; 2) the ease of administration because all interviewing is carried out centrally, and 3) the short time needed for completion - - the actual interviewing for a typical survey can be completed in about a week.

Comparison of the State data demonstrates several points that we would like to emphasize. First, and most importantly, there is substantial State-to-State variation in the prevalence of these risk factors. These differences obviously have important implications for later rates of health-related outcomes. Accordingly, individual States could and should come to different conclusions regarding which risk factors should be the highest priority for intervention efforts. National estimates alone are not adequate for setting priorities in State jurisdictions.

The presence of data from other States, especially neighboring ones, can help provide estimates of what levels of risk reduction are feasible. States with low values can become a standard by which others measure themselves - much like the United States compares itself with the Scandanavian countries in the area of infant mortality. With repeated surveys, States can follow their trends in prevalence and their progress toward 1990 goals.

Because of the rapid acceptance of the use of telephone survey methods by State health departments and the continuing need for the States to monitor the trends in prevalence of these risk factors, CDC is establishing a mechanism for coordinating on-going State surveillance of these behaviors similar to the assistance CDC provides in the coordination of surveillance of infectious illnesses.

This surveillance system will use the telephone interview technique. It has the important advantage of being flexible enough for individual States to include additional specific questions about behavioral risks already covered or about other areas of interest.

This system has the potential to stimulate greater efforts aimed at the prevention of the major causes of premature death and disability in the United States. By working with the State health agencies we hope that local concern will become as great about the rate of smoking as it is about measles; that people will become as supportive of exercise programs as they are of immunization programs, and as worried about alcohol misuse as they are about influenza epidemics.

PREVELANCE OF SELECTED BEHAVIORAL RISK FACTORS IN 27 STATES AND DISTRICT OF COLUMBIA

RISK FACTOR	MEDIAN (%)	RANGE (%)
Current Smoker	31.9	23.4-37.4
Obesity	23.4	16.4-28.2
Sedentary Lifestyle	12.3	5.7-17.7
Seldom or never Seat belt use	60.5	41.6-71.1
Known Hypertension Still Elevated	3.8	2.0-6.8
Chronic Drinking (>2 drinks/day)	8.3	3.3-14.5
Acute Drinking (>5 drinks at one time in last month)	22.5	7.8-30.5
Drinking and Driving	5.2	1.9-10.3

MANAGEMENT AND EVALUATION OF A LARGE SCALE HEALTH PROMOTION PROJECT IN THE PUBLIC SECTOR

Kirby L. Jackson, School of Public Health, University of South Carolina
Zora T. Salisbury, Jennie J. Kronenfeld, Keith E. Davis, Steven N. Blair

Health promotion and health behavior modification is becoming an increasingly important aspect of total health care. The benefit of convincing America to adopt healthier lifestyles has become widely accepted, based on the LaLonde Report (1) in Canada and the Healthy People Report of the U.S. Surgeon General (2). These papers emphasize the role of lifestyle factors in contributing to illness and mortality. Interest in and impact on behavior has come from two directions: (1) the public health community, especially in a variety of public health oriented campaigns - antismoking, safety through seatbelt use, responsible use of alcohol, dietary changes, blood pressure intervention; and (2) a harder to document, more individually oriented emphasis on preventive health practices and behavior change exemplified through concerns about nutrition, weight control, cigarette smoking and exercise. Additionally, lay organized groups concerned with alcohol abuse, home childbirth and faith healing have become more prominent. The rapid increase in organized and individual physical fitness activities are one cogent example of a reservoir of self help energy that exists within the population.

As an outgrowth of individual interest and observed success in clinical and community studies and as a response to economic conditions within industries (specifically rising health insurance costs), American businesses and industries are showing increased interest in employee health promotion programs. Some industries view such programs both as fringe benefits for employees that have intrinsic appeal and as benefits that have the potential of increasing worker productivity and decreasing worker health insurance costs. The implementation of health promotion in industry is a salient trend in American business in the 1977-1982 period. Of the seventeen programs described in Parkinson (3), all but two started recently.

In this paper, we describe a health promotion project being conducted with state employees in a two county metropolitan area in Columbia, South Carolina. We describe first the initial design of the program and the overall plan for delivery of interventions. One focus in this discussion will be the interrelationships among staff from the three state agencies involved in planning and developing the project (the Office of Health Education of the State Department of Health and Environmental Control, the University of South Carolina and the Division of State Personnel). The importance of employee volunteers as planners and participants in implementing the project is also emphasized.

The overall evaluation strategies are described with emphasis on the development and use of questionnaires, data management, and other external methods for evaluating the outcome of the program. In addition, two questionnaires used internally to evaluate organizational effectiveness and efficiency are described. These questionnaires were administered to project staff, group leaders and volunteers involved with the presentation of the project.

Carolina Healthstyle - Project Organization

Carolina Healthstyle is an ambitious project that began in July, 1982. It is funded through the Insurance Section of the South Carolina Division of State Personnel and represents a three year commitment to provide support for development of programs designed to stimulate the adoption and maintenance of positive health behaviors of approximately 20,000 state employees in the Columbia metropolitan area. With a modest budget of 100,000 dollars per year, the project is implementing a comprehensive worksite health promotion program that encompasses stress management, nutrition and weight control, exercise and physical fitness, alcohol and drug abuse, special women's health issues, and safety and accident prevention.

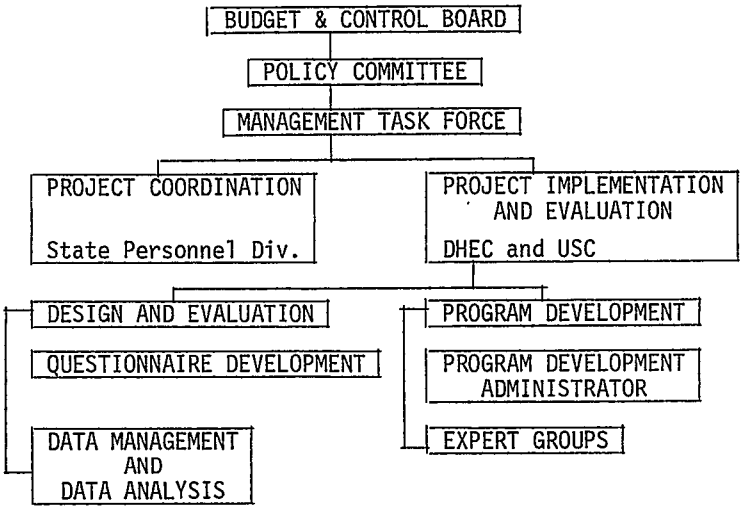
The funding of the project is somewhat unique in that there is no dependency on the usual sources of funds for programs for public employees, i.e., state or federal government. Rather the funds for Carolina Healthstyle were approved by South Carolina Budget and Control Board from the portion of insurance premiums set aside for administration of the health insurance program for state employees. Since the monies are independent of the state agency funding and allocation process, the project is not held hostage to the vagaries of political interest or the uncertainty of the block grant process.

Three separate organizational units set policy for the project: The Division of State Personnel, The South Carolina Department of Health and Environmental Control, and the School of Public Health, University of South Carolina. Together, representatives of these units set the goals and objectives and deal with issues related to the overall project, while certain tasks allocated to specific employees of the various agencies. The Division of State Personnel is responsible for overall policy coordination through its dual role as both a funding source and project participant and because of its direct contact with the Budget and Control Board. The project coordinator is administratively located in the Division of State Personnel. She acts as the contact person for the individual agencies and helps them organize the promotion and awareness phase of the intervention.

The Office of Health Education of the Department of Health and Environmental Control (DHEC) and the University of South Carolina (USC) are responsible for project implementation and evaluation. The program development coordinator and most of the graduate assistants work in the

implementation of intervention programs. University faculty are members of the expert groups described below. The co-principal investigators who are part of the Management Task Force direct the research and evaluation component of the project. Data collection and management is also handled through the University under the direction of the project biostatistician. The organization chart for the project is shown in Figure 1.

Figure 1
Organization Chart: Carolina Healthstyle



In practice, management and specific direction of the project is through the Management Task Force containing representatives from USC, DHEC and State Personnel. This small group meets regularly to set goals and to monitor progress on the project. The project coordinator from State Personnel and the Program Development administrator both are members of the Management Task Force and thus communication is kept open between the different groups. Once policy is set, and after the Management Task Force has received possible effects of a policy or procedure, the co-principal investigators implement the procedure either through the Design and Evaluation group centered mainly on the USC campus or through the Program Development group based at DHEC. These groups develop the actual programs to be implemented.

This management system is complex but with find tuning as the project develops it is proving workable. Problems do occur. Different groups can have substantially different underlying goals for the project so it is extremely important that there be clear understanding and agreement on the priorities, interest and direction. A major strength lies in the inclusion of individuals with different backgrounds and goals into a single project useful for research purposes and for delivery of a health promotion program. The diversity in background and perspectives of the organizers of Carolina Healthstyle adds a strength and comprehensiveness to the program that would not be available otherwise.

Program Design

There are over eighty state agencies with offices in the two counties involved in this study. Organization of these eighty agencies for evaluation and intervention was complicated. Some large agencies are located at multiple sites and, in other cases, a large building might contain several agencies. Interventions were offered on a site-wide basis since individual agencies were often too small to support an individual program and, in addition, if treatment and control agencies were located in the same building, contamination by diffusion could occur confounding the evaluation results.

The first steps in agency contact were to send an announcement of the project to each agency director. A letter of support from the Governor was sent to all agencies describing the project and requesting that each director appoint an agency liaison person to Carolina Healthstyle. This liaison was to function as a contact between the agency and the project office. This individual also was to lead an agency based committee that would maintain and coordinate activities at the agency. These liaisons were invited to a day long conference with speeches by members of the project and by a representative of the Governor's office. Displays and demonstrations organized around each of the intervention areas were also given.

The design for the intervention programs proposed several phases in each site. First, a promotion phase was conducted. This included a promotional seminar at lunch time in which the concept of health promotion was discussed, a film entitled "The Wellness Revolution" was shown, and a general discussion of the concept of Carolina Healthstyle was presented. In the following weeks, short presentations (30-45 minutes) on each health promotion activity were given as a means of encouraging participation in comprehensive intervention programs on lifestyle change. Another general promotion activity was the opportunity for all interested employees to take the Health Risk Appraisal (HRA) Questionnaire (developed by The Centers for Disease Control in Atlanta). The results were returned to the individual and interpreted by a staff member of Carolina Healthstyle.

The next phase of the project included more comprehensive intervention courses developed by the expert groups. Group activities were presented in the intervention agencies by project staff and expert volunteers. Table 1 gives a summary of individual contacts in spring 1983 and participation in a representative state agency. Table 2 shows the more expanded list of long term programs available in the fall 1983.

Table 1
Participation in Carolina Healthstyle Program at one State Agency

Total individuals in the agency (96)

Activities	# Participants	
Wellness Revolution Movie	20	Promotional Activities
Health Risk Appraisal	31	
Nutrition Seminar	12	
Weight Control Seminar	14	
Stress Seminar	15	
Stress Groups (2)	23	area oriented programs
Exercise Group	20	
Walking Group	6	spinoff activities
Running	2	
Aerobics	12	

A cadre of expert volunteers recruited from the University, the Department of Health and Environmental Control and other health oriented groups such as the Commission on Alcohol and Drug Abuse is an important part of the project. The goal was for experts in each of the intervention areas (such as smoking, nutrition, fitness, etc.) to act as Program Planners and Evaluators. Volunteering was made attractive by several means: (1) these expert groups were given major responsibilities and control in diagnosing needs and developing interventions; (2) support was provided in terms of computer searches, articles, graduate research assistant time and limited funds to purchase materials and supplies to support the interventions; and (3) in addition, research plans and student projects and theses were encouraged as activities of the expert groups. This original group of volunteers had approximately 75 people in the various intervention areas.

Table 2
Scope of Activities in Two Sites

Activity	Frequency	Duration
	Site 1	
Aerobics	2 per week	6 weeks
Walking	2 per week	6 weeks
Nutrition/Weight Control	1 per week	6 weeks
Smoking Cessation	1 per week	6 weeks
Co-op Buying	1 meeting	
Health Fair	1 meeting	
	Site 2	
Weight Control	1 per week	6 weeks
Walking	1 per week	6 weeks
"Diplomatic Approach to Smoking"	1 meeting	
Health Risk Appraisal	1 administration	
Health Risk Appraisal Interpretation	1 meeting	
Stress Management	2 per week	6 weeks

Evaluation Design

A quasi-experimental design was considered the most appropriate for evaluation of the project. The 24 program units described earlier were grouped into three categories. These three groups will be selected in turn for receipt of the intervention. This multiple-baseline approach provides a comparison group for evaluating possible changes in treatment agencies relative to control agencies. Studies of this type virtually preclude random assignment of either agencies or individuals. For example, employees in the same office or agency in the same building should not be assigned to treatment and control groups since contamination would result. In addition the logistics of presenting the health promotion program required starting in phases in specific sites.

The importance of evaluation of Carolina Healthstyle was recognized from the inception of the project. There were three main goal areas in which evaluation was considered possible. These were (1) that such health promotion programs may change attitudes, beliefs and knowledge; (2) to demonstrate that through the delivery of worksite health promotion programs, it is possible to change health behaviors in selected areas; and (3) to demonstrate effects on external factors such as job absenteeism or on utilization and costs of health insurance or health care. Three main evaluation methods were planned: (1) use of questionnaires to examine need, participation, and

behavior change of individuals in the project; (2) use of Blue Cross/Blue Shield records to examine costs and types of claims; (3) use of state personnel records to examine absenteeism. In order to analyze behavior and attitude change in the many specific areas within the scope of the project, questionnaire and survey approaches were selected as the main evaluation method. We will also examine both Blue Cross/Blue Shield records and personnel records, however we do not expect to see easily interpretable results from this since effects on actual health costs or abseenteeism are not expected to be dramatic in the short term.

The questionnaire and survey approach was incorporated into the quasi-experimental design through two separate components (1) a core questionnaire administered initially to a stratified random sample of 10% of the state employees in a region and which is administered periodically to those who receive it initially and to a new random sample. (2) A set of specific area oriented questionnaires administered either to an additional random sample in selected agencies or specifically to those individuals participating in a given intervention. The core questionnaire is used to measure the overall effects of the program while the area oriented questionnaire is intended for the measurement of behaviors and practices within one area of health promotion.

The core questionnaire is a twelve page, machine readable questionnaire. It was designed to measure aspects of basic health and well-being with emphasis both on practices and knowledge and attitudes. It was developed through cooperative actions of members of all the expert groups who submitted questions or sets of questions on each area of interest. These were reduced through the actions of a questionnaire committee who worked to eliminate redundant questions but kept questions from each intervention area, to help evaluate specific interventions. The questionnaire is broken down into sections as follows:

(1) General Information: containing questions about demographic characteristics of the respondent and family.
(2) Personal Health: containing questions examining health attitudes and practices.
(3) Personal Health Knowledge: containing questions examining knowledge of various aspects of health practices.
(4) Stress: containing questions that could be used to construct several different indices of stress including a work stress index.
(5) Personal Health Attitudes and Opinions: a series of statements about health and health attitude with a agree-disagree scale for response.
(6) Personal Relationships: containing questions concerning social relationships and social support.
(7) Personal Health Intentions: questions on satisfaction with characteristics relevant to health (such as weight) and whether individuals intend to change the characteristic or not.
(8) Personal Health for Women: containing questions on issues specific to women's health such as breast self examination and pregnancy.

The core questionnaire has a planned administration schedule shown in Figure 2.

Figure 2
Administration of Core Questionnaire

February-March 1983
  Core Questionnaire in intervention sites (10% Random Sample)
February-March 1983
  Core Questionnaire in control sites (10% Random Sample)
March-December 1983
  Intervention in first group of 8 sites
February-March 1984
  Core questionnaire in original 10% sample and 10% new sample from population
March-December 1984
  Intervention in second group of 8 sites
February-March 1985
  Core questionnaire in original 10% sample and 5% new sample from population
March-December 1985
  Intervention in final sites
February-March 1986
  Core questionnaire final administration
  Original 10% sample

The repeated test on the same individuals is the most effective of determining actual change and, in addition, this structure makes it possible to examine relationships among the variables on the questionnaires to determine if there is an underlying pattern to any changes. The addition of a new random sample at several time periods allows examination and adjustment for any effect of the test as a behavior modifier in itself. This effect is not expected to be significant since the results on the core questionnaire are not returned or interpreted to the individuals who take it.

The important comparisons will be the changes observed between individuals in sites where Carolina Healthstyle presented a program and those sites where it did not. In addition the longitudinal aspect of the administration structure allows examination of residual effects after the initial intense intervention program, since agencies in phase I will have almost two years before the third administration of the core questionnaire.

At project headquarters questionnaires were edited for correctness and then sent on for keypunching (in the case of early questionnaires not machine readable) or entry through an optical scanner (for the machine readable questionnaire). At the same time, one of three codes was attached to a computerized listing of the individuals in the sample. These codes indicated the followup status: 1 = filled out questionnaire, 2 = refused to fill out questionnaire, do not followup, 3 = questionnaire not filled out, followed up. Questionnaires with status = 3 were followed up, either with another

group administration of the questionnaire or by sending the blank questionnaire to the liaison person in the appropriate agency with instructions to deliver it to the individual. These were then to be returned by mail to project headquarters. Overall response rate on this questionnaire was 64 percent, less than we desired but nevertheless adequate for examining for changes over the period of the study.

Logistically, the administration of the questionnaire throughout the state agencies was difficult. A 10% random sample of individuals stratified by agency was generated using computerized records from the Division of State Personnel. Name, agency and social security number were given on a tape to the project biostatistician. These were coded with a sequence number which was linked to social security number in the computer files. This number was used on the questionnaire to ensure privacy since there were questions that concerned possible sensitive areas such as sexual or drinking behavior.

In larger agencies it was decided to have group administration of the questionnaire. Arrangement for space and for contacting individuals in the sample from that agency were made by the volunteer liaison person in the agency. Time off during working hours were allowed for individuals taking the questionnaire. The actual administration of the core questionnaire in the larger agencies was handled by two graduate assistants who described the questionnaire to the group, distributed the informed consent form and a questionnaire and remained and collected the questionnaires when all were finished. The questionnaires were field edited as they were collected to eliminate glaring errors or problems due to misunderstanding the questions or the procedures for filling out the questionnaire. Often it took more time than expected to administer the core questionnaire and occasionally individuals would be given an envelope and asked to mail the questionnaire to project headquarters. In agencies with less than 10 people in the random sample, questionnaires were given to the volunteer liaison with a mailing envelope for returning the completed questionnaire to the project, and the liaison distributed the questionnaire to the individuals in the sample, collected, and returned them.

The area specific questionnaires were designed by individuals interested in the specific intervention area and thus were shorter and more specific. Since these were used in much smaller quantities machine readable versions were not developed though some used standard machine readable answer sheets. These questionnaires are used to examine the effects of an intervention program in a specific health promotion area. A before and after design was implemented for all groups, however, latitude was allowed in the actual administration procedures and in the choice of populations to be examined. Efforts were made to standardize the administration of these questionnaires but sampling on a population basis was impractical and introduced too great a burden on the agency liaison person and employees. Possible administration procedures varied from that of the women health group which used a mail administration to a 20% random sample of women in the sites of interest to that of the stress management group which gave a before and after questionnaire only to those participating in the stress management interventions. Except for the Women's Health area, limited population conclusions can be drawn since the questionnaires are given only to those participating in a given program. However, the before and after comparisons allow estimation of the effects of an intervention program. This is especially useful when more than one type of intervention program is used for the specific health area.

The stress questionnaire contained a series of statements with a Likert scale for agreement or disagreement by the respondent. Sections included personal stress, work stress and personal resources. This questionnaire was given to participants in stress management groups before and after the group. Different modes of stress management groups were compared using the questionnaire to evaluate the effectiveness of the program.

The women's health questionnaire investigated practices related to areas of specific interest to women such as breast self examination and education about and attitudes toward it. This questionnaire was administered to a random sample of women in agencies where a women's health education program was to take place. The questionnaire will be readministered at the end of the program.

The exercise questionnaire contained several sections: knowledge, opinions and attitudes, motivation, physical behavior, personal health intentions and a small section on demographic information. This questionnaire was designed to measure both changes in behavior and changes in attitudes and knowledge about exercise and its effects. This was administered before and after the exercise program.

The questionnaires discussed above allow an estimate of changes in attitudes and behavior both in the overall population through the core questionnaire and in a specific intervention group through an area oriented questionnaire. However, in monitoring the project, internal process evaluation is also critical. It was felt that, in addition to the usual progress reports and lists of task accomplished a detailed analysis of relationships among the key project personnel (paid staff plus volunteers) and agency liaisons was critical. Using questionnaires based on Popov's model of organizations and organizational climate (4), the key personnel and agency liaison individuals were surveyed three times over the first six months of service delivery to allow a picture of organizational effectiveness. The questionnaire used for this internal process survey had two forms, one for project personnel and a

second for agency liaisons. These questionnaires contained specific questions relating to the interactions between project staff liaisons and agency personnel. Due to the repeated nature of the questionnaire they were very useful in pointing up areas where change was needed and in evaluating the effects of internal project changes.

Record Linkage and Data Base Management

The number of different foci of the project made data management a complicated procedure. The core questionnaire was managed and administered under the direction of the centralized data management group lead by the biostatistician. The administration of the area specific questionnaire was less centrally controlled and the data management group often received completed questionnaires without having given substantial input in their development or administration. Separate data files are maintained for each type of questionnaire. Requests for tables or analyses are submitted by project staff. These are approved by the biostatistician and graduate assistants who are familiar with the specific data sets then run the required analyses.

For confidentiality purposes, all questionnaire results are kept separate from any direct identifying information on the individual. However, code numbers that can be used to link the questionnaire are kept on the file and a separate file linking social security numbers and code numbers is kept. This allows linkage of questionnaires over time, and if desired by an area group and if they collect social security numbers, the file for the core questionnaire can be searched for individuals participating in a specific intervention.

Problems develop occasionally due to lack of communication between the expert groups who design and administer the area specific questionnaire and the data analysis headquarters. The point of view and the desires of the area oriented individuals may be different from that of the overall project. For instance some area groups have collected social security numbers on the participants and other have not. This is reasonable in the context of a health promotion program but makes linking of records on an individual impossible. Some of these problems are due to lack of resources to collect and enter the data and others to lack of understanding of the possible usefulness of the information. A reasonable goal would be to have a set of files on each questionnaire linkable by individual between all the different questionnaires. However, this is a difficult task given the scope and diversity of the project.

Another reason to keep files with social security number linkages possible is to allow examination of personnel files and Blue Cross/ Blue Shield records as related to participation in Carolina Healthstyle. While access to the files at an individual level is possible and may be done in examination of the results of the core questionnaire, other analyses of the effects of Carolina Healthstyle on absenteeism and health insurace claims will be performed only on an agency wide basis.

Conclusion

Carolina Healthstyle presents a comprehensive health promotion project in the public sector. Strengths of the project are in the communication among different areas of state government, Division of State Personnel, Department of Health and Environmental Control and the University of South Carolina. This communication has allowed the development and implementation of a large scale program. The support from State Personnel has allowed entry into the agencies and time during work hours for administration of the core questionnaire and some time for the Carolina Healthstyle work of the volunteers, liaisons in the agency. The Department of Health and Environmental Control has worked in developing and administering intervention programs and along with the University of South Carolina has encouraged the development of a monitored project with a strong research component.

References

1. Healthy People--the Surgeon General's Report on Health Promotion and Disease Prevention 1979. (DHEW (PHS) Publ. No. 79-55071) U.S. Government Printing Office, Washington, D.C.

2. Lalonde, M. A New Perspective on the Health of Canadians. Ottawa: Trigraphic Pringing, 1975.

3. Parkinson, R.S., & Assoc. Managing Health Promotion in the Workplace: Guidelines for Implementation and Evaluation. CA: Mayfield, 1982.

4. Popov, D. Understanding and Predicting the Effectiveness of Organizations: An Exercise in Process Description. University Associates, 1979. (Copyrighted by Systems Effectiveness Association.)

Statistical Methodologies for Establishing Health Promotion Goals and Priorities

Session F

IDENTIFYING AND MEASURING PROGRESS TOWARD HEALTH PRIORITIES: THE YEARS OF LIFE LOST TECHNIQUE

Susan A. LeBailly, Janet D. Perloff, Phillip R. Kletke, John P. Connelly, & Peter Budetti, American Academy of Pediatrics

I'd like to tell you about some work we've been doing with a measure of years of life lost. We think it can be useful when developing priorities for preventive efforts. A measure which is sensitive to the age at which death occurs may help to establish health priorities and measure progress towards their improvement. A measure of years of life lost expresses deaths in terms of the difference between the age at death and a preselected "expected age at death." It is easy to calculate and requires only readily available mortality statistics. This paper describes how to calculate years of life lost, how its implications differ from the death rate's, and examines the limitations imposed by the way mortality data is categorized, using both national and local data.

The concept of years of life lost is not new, but has received little use in the United States until recent years. As early as 1950, Haenszel proposed measuring the amount of life lost when establishing health priorities.[1] In 1965, Stickle responded to the President's Commission on Heart Disease, Cancer, and Stroke by demonstrating the substantial loss of years of life due to infant mortality, accidents, and other violent deaths.[2] More recently, Kleinman advocated the use of a years of life lost measure in state and local health planning.[3] Last year, CDC began reporting a measure of years of life lost in the Morbidity and Mortality Weekly Report.[4]

Computing the Years of Life Lost

A years of life lost measure simply calculates the number of years between the age of death and a preselected end-point. The index then is the sum of the number of deaths in each age category (Di) weighted by the number of years between the midpoint of the age category and a cutoff age (Pi):

(1) Years of Life Lost $= \sum_i$ Pi Di.

While the concept of years of life lost is simple, the selection of the begin-point and the end-point greatly affect the findings. The first year of life has frequently been omitted in order to avoid giving heavy emphasis to infant deaths.[4,5] Similarly, deaths in the population over 65 or 70 have been excluded by different researchers.[4,5,6]

Our analysis differs from most previous studies by measuring the potentially productive years of life lost. We wanted to include deaths under the age of one because we felt it illogical to ignore infant death in a discussion of premature death. We used the age 70 rather than 65 as the cutoff age since many people in the 65-69 age category are still economically active. The cutoff age could be extended to 75 or even higher, but this may introduce methodological problems since the precise cause of death among the elderly is often difficult to determine.

In our analysis, the deaths of children under the age of 15 are weighted with the loss of productive years, which we have defined as the ages 15 to 70. As the second equation shows, P_i equals the difference between 70 and the midpoint of the five-year age category for individuals over 15, and 55 for children under 15. The value of P_i is zero for deaths at ages 70 and above.

(2) PPYLL $= \sum_i 55\ Di$ for $i \leq 15$

$= \sum_i Pi\ Di$ for $15 < i < 70$

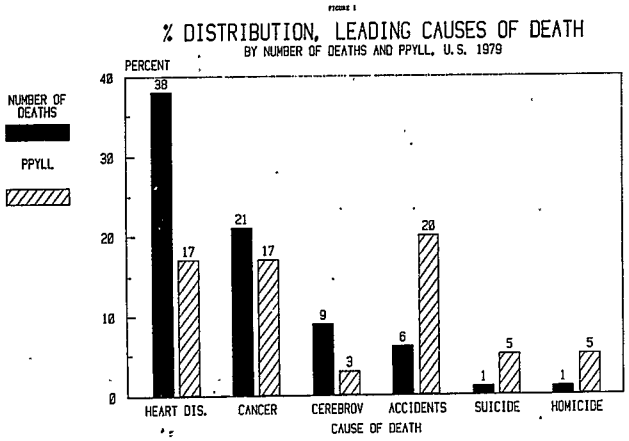
As the third equation shows, we age-standardized the rate of potentially productive years of life lost since differences in the rates for two populations may be partially attributed to differences in their age distributions.

(3) Age-standardized rate of PPYLL $= \sum_i \frac{PiDiKi}{N} \times 1000$

Where Ki is the adjustment factor based on the population. We have examined years of life lost on two data sets: national 1979 mortality data and Chicago 1981 data. The national data reported deaths by 5-year age categories, race, and sex. The Chicago data reported deaths by ICD9 code, age, race, sex, city, and community area.

Ranking Causes of Death

The measure of potentially productive years of life lost gives a different emphasis to various causes of death (See Figure 1). Accidents, suicide, and homicide represent a larger proportion of years of life lost than their proportion of deaths. Accidents are responsible for one-fifth of all years of life lost. On the other hand, heart disease accounts for 38% of all deaths, but represents only 17% of years of life lost. Similarly, cancer and cerebrovascular disease represent a higher proportion of deaths than they do years of life lost.



Ranking the leading causes of death also shows the change in emphasis (See Table 1). When standard death rates are the measure, heart disease, cancer, and cerebrovascular diseases are the three leading causes of death in the United States for people over one year of age. With

potentially productive years of life lost, however, accidents are the largest cause of life lost among people 70 and younger. Cancer ranks second, and heart disease, the leading cause of death using death rates, is the third leading cause of years of productive life lost.

TABLE 1

TEN LEADING CAUSES OF DEATH
RANKED BY NUMBER OF DEATHS AND PPYLL

Cause of death	Rank	
	Deaths	PPYLL
Heart disease	1	3
Cancer	2	2
Cerebrovascular disease	3	8
Accidents	4	1
Chronic obstructive pulmonary disease	5	-*
Pneumonia and influenza	6	10
Diabetes mellitus	7	-*
Chronic liver disease and cirrhosis	8	9
Atherosclerosis	9	-*
Suicide	10	6
Certain conditions in the perinatal period	-*	4
Homicide	-*	5
Congenital abnormalities	-*	7

*Not rated among top ten causes for this measure.

Since accidents result in such a great loss of years, we may want to examine the impact of different types of accidents. Motor vehicle accidents result in 61% productive years lost in accidents, but represent 51% of accidental deaths. Falls, on the other hand, result in 3% of productive years lost in accidents and 12% of accidental deaths. Drowning, poisoning, and firearms accidents result in a slightly greater loss of productive years than number of deaths.

Years of life lost rates, just like age-standardized death rates, can be used to identify subpopulations at risk for different causes of death. The two do not produce the same results when subpopulations are at risk at different ages. Looking at deaths from motor vehicle accidents, as shown in Figure 2 the striped bars here indicate that white females and, to a lesser extent, white males experience a greater loss of life than death rates would indicate. Cancer results in more years of life lost for white females, but fewer years of productive life lost for all other groups (See Figure 3). From this, we would conclude that white females are particularly at risk for premature death due to cancer.

FIGURE 2

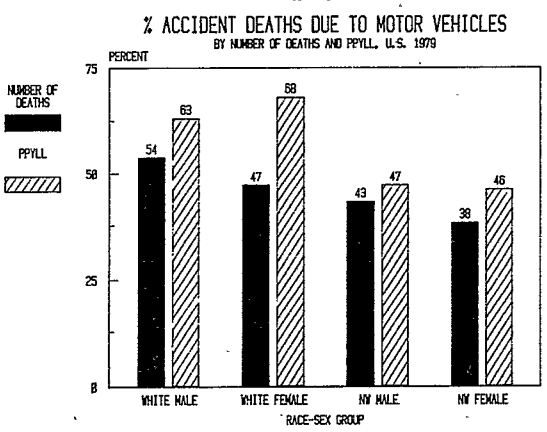
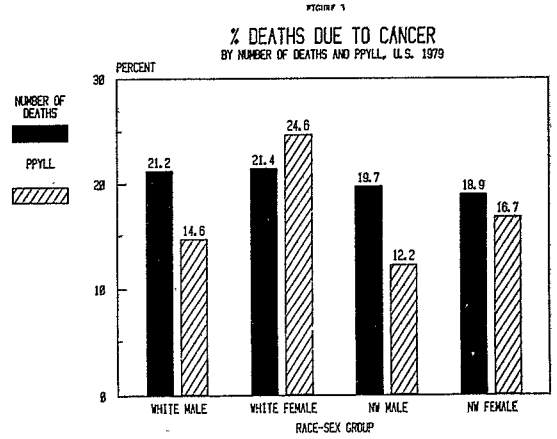
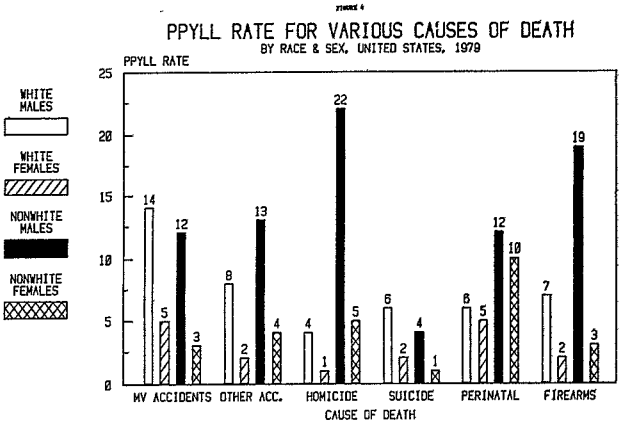


FIGURE 3



Examining race and sex differences in years of life lost, we found that males are more likely to lose productive years of life to auto accidents, other accidents, homicide, suicide, and deaths due to firearms, as well as drowning, fires, poisoning, and falls (See Figure 4). Nonwhites have a higher risk for premature death due to homicide, firearms, and perinatal conditions as well as falls, poisoning and fires. Whites are at greater risk for premature death due to auto accidents and suicide.

FIGURE 4

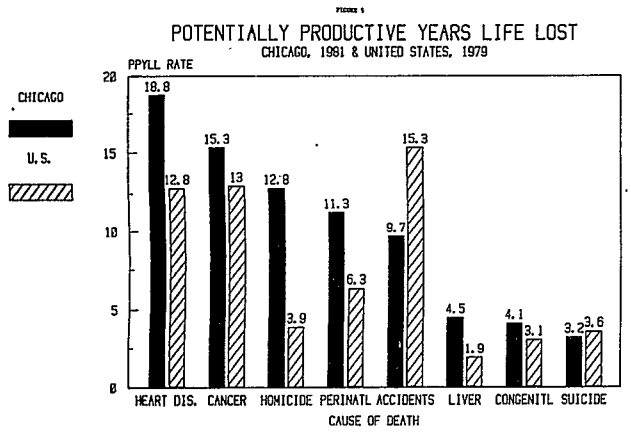


While we cannot empirically state that income level predicts the risks to a population, factors related to low income such as poorer housing, less parental safety education, living in areas with high rates of criminal victimization and lack of access to preventive medical care may be more common among nonwhite families.[7]

Examining potentially productive years of life lost draws attention to the amount of life lost in accidents, homicide, suicide, and perinatal or congenital conditions. We began to realize that conditions which have a strong connection to the individual's lifestyle, or social environment result in a greater loss of productive years than do conditions with a strictly biomedical origin.

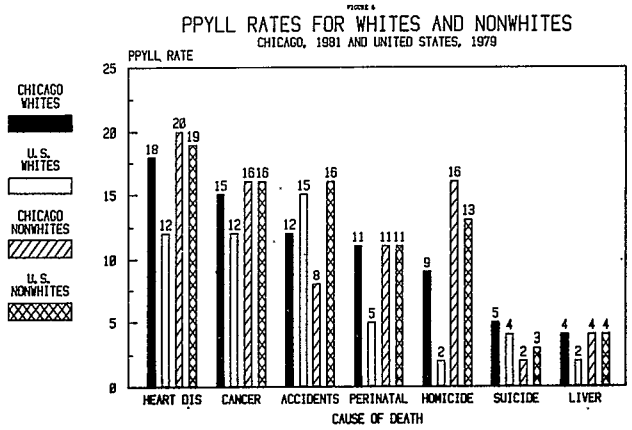
To briefly summarize, a measure of years of life lost can be used to: 1) identify the leading causes of premature death, 2) to examine subsets of causes which contribute to premature death, and 3) to identify subpopulations at particular risk for different types of premature death. The years of life lost analysis also demonstrated the greater impact of causes of death which have a strong relationship to the environment, social structure, and individual lifestyle.

We recently obtained 1981 Illinois mortality data to examine productive years of life lost in Chicago, as an example of what local planners might learn from such an analysis (See Figure 5). While the leading causes of death are the same, their ranking differs from the national. Let me remind you that our data came from two different years and are not strictly comparable.

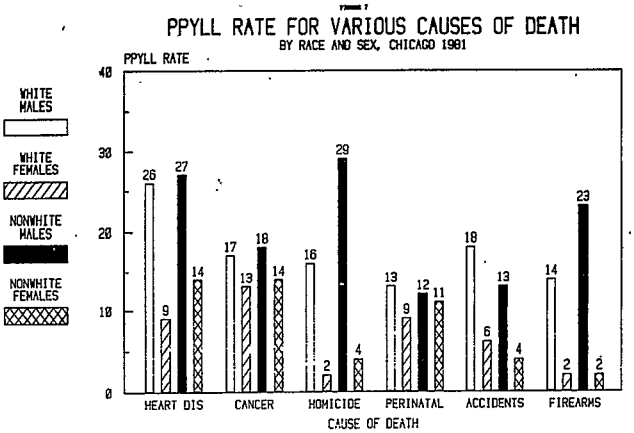


Heart disease is the leading cause of life lost in Chicago, followed by cancer, homicide, perinatal conditions, accidents, chronic liver disease, congenital disorders, and suicide. Remember that accidents are the leading cause of life lost nationally, followed by cancer, heart disease, perinatal conditions, homicide, and suicide.

Chicago's lower ranking of accidents is due to a substantially lower rate of years of life lost for motor vehicle accidents, a trend which does not hold for other accidents. The greater impact of heart disease in Chicago appears to result from the racial composition of the city. When we examine race-specific rates, we find the rate of years of life lost due to heart disease among nonwhites is 20 per 1,000 population in Chicago and 19 per 1,000 in the United States (See Figure 6). Nonwhites in Chicago tend to have years of life lost rates similar to the national, while whites in Chicago have a higher years of life lost rate than do whites in the nation for heart disease, cancer, homicide, perinatal conditions, and chronic liver disease.



As Figure 7 shows, Chicago males lose more years of productive life to heart disease, homicide, accidents, and firearms. Homicide is the leading cause of years of life lost among nonwhite Chicago males, followed by heart disease, cancer, and accidents. Heart disease is the leading cause of years of life lost for white Chicago males, followed by accidents, cancer, and homicide. For both white and nonwhite Chicago females, cancer, heart disease and perinatal conditions are the leading causes of premature death. Homicide is the fourth leading cause of years of life lost for nonwhite females, while congenital anomilies ranks fourth among white Chicago females. Accidents are the fifth leading cause for all females.



As we saw at the national level, the years of life lost measure highlights causes of death which have their roots in either external, social situations or lifestyle. Clearly, the health problems we face develop from social problems. While similar findings could have been deduced from an examination of death rates, the measure of years of life lost brings special focus to deaths occurring at an early age, which often result from accidents, homicide, and suicide.

When we consider the conditions resulting in substantial amounts of premature death, we begin to see that interpersonal violence is common to many of them. Colleagues of ours at Northwestern University recently completed a study of reasons for emergency room visits to a Chicago community hospital.[8] While that study has all the limitations of a single hospital study, it provides an interesting look at health problems in that community. Slightly more than half the ER visits were for injuries. The most common injuries resulted from falls, motor vehicle accidents, interpersonal attacks, being hit by a person, being hit by an object, alcohol/drug abuse, sports injuries, animal bites, machinery injuries, poisoning, and fires. While some of these injuries are the result of environmental hazards, the researchers explain that a large proportion of these injuries result from interpersonal violence. Examining the mortality data in that community, we found that the leading causes of premature death were homicide, perinatal conditions, heart disease, accidents, and cancer, again demonstrating the impact of violence through accidents and homicide, as well as problems of lifestyle and lack of adequate medical care. The health consequences of violence have been recognized by several leading health officials: the Surgeon General recently acknowledged that violence is a great health problem facing the nation, as do the 1990 health objectives.[9,10]

The focus on lifestyle, environment, or interpersonal violence changes our concept of the types of preventive measures needed. It requires that we not direct all our preventive efforts to the individual. Increased funding for medical care or even reorganizing the health care delivery system alone cannot improve lifestyle conditions.[11] For example, researchers helping a community organization inventory health problems in a poor Chicago neighborhood found the local health care system was dealing with social problems such as traffic patterns causing auto accidents, inadequate control of stray dogs resulting in dog bites, and inadequate nutrition resulting in bronchial ailments.[12] To improve health, the community group added stop signs, mounted campaigns to catch stray dogs, and grew food in community gardens. Recognizing the social roots of much illness and premature death requires that we maintain a broad outlook on both the nature of problems and preventive strategies.

Even though some of our leading health planners have recognized the impact of violence and other social problems, most preventive efforts still are focused on the individual. This is due to a variety of reasons, including the fact that social problems are more difficult to define than are individual problems, and it is more difficult to measure progress in alleviating social problems. Program grants requiring documentation of "success" or impact for continued funding and third-party reimbursement for only "medical" care encourage an emphasis on individual, medicalized problems.

The kind of data we collect also plays an important role in shaping our concept of health problems and prevention. Our current concept of prevention revolves around identifying risk factors and then preventing them. The type of data collected about morbidity and mortality, then, plays a key role in identifying risk factors. Let me offer three examples of how collecting data only by medical definition of the condition may limit our understanding. First, the Northwestern researchers could identify the impact of interpersonal violence because they audited the complete medical file -- they didn't rely on official statistics. Second, the absence of income data limits our understanding of underlying causes of death. For example, income similarities may explain why there are fewer disparities in years of life lost between whites and nonwhites in Chicago than nationally. Third, studies of accidental injuries are limited by the classification of the injury as a fracture, contusion, etc. rather than examining the cause of injury, they are not used in any data other than mortality reports.

There are several reasons this type of data has not been routinely collected. Hospital workers may be reluctant to supply information about the cause of injury because it asks them to make judgements about the nature of the problem they don't feel qualified to make. Problems with child abuse or neglect reporting shows this reticence. Secondly, medical professionals are often hesitant to make judgements which have legal connotations. Stating that an accident was due to drug abuse or that an injury actually resulted from child abuse introduces the question of legal culpability. Finally, recognizing the impact of poverty and violence requires coordinating efforts with non-medical professionals and community members.

Nevertheless, if we are to make advances in the study of health problems and identify risk factors involved, we must begin to collect more data about events surrounding the incident, income levels, and contributing factors. We must in some way elevate the importance of this data collection so that record keepers see the value of collecting and coding this information.

A measure of years of life lost will never replace standard death rates. It can supplement death rates in an examination of health problems. It highlights causes of death among younger populations, and focuses on premature death resulting from accidents, homicide, suicide, heart disease, and cancer. These problems draw attention to lifestyle, environmental, and societal problems requiring both medical and nonmedical interventions. The measure can be used nationally and at the local level to identify health problems resulting in premature death. At the same time, we must strive toward collecting better data to help us understand the underlying causes of premature death. Then perhaps we can begin to consider issues of quality of life rather than merely quantity.

FOOTNOTES

1. Haenszel, W. "A standardized rate for mortality defined in units of lost years of life." American Journal of Public Health 40 (January 1950): 17-26.

2. Stickle, G. "What priority, human life?" American Journal of Public Health 55 (November, 1965): 1692-1698.

3. Kleinman, J.C. "Age-adjusting mortality indexes for small areas: applications to health planning." American Journal of Public Health 67 (September, 1977): 834-840.

4. U.S. DHEW, PHS, Center for Disease Control. "Introduction to Table V Premature Deaths, Monthly Mortality, and Monthly Physician Contacts - U.S." Morbidity and Mortality Report, Volume 31, March 12, 1982.

5. Romeder, J.M. and McWhinnie, J.R. "Potential years of life lost between ages 1 and 70: an indicator of premature mortality for health planning." International Journal of Epidemiology 6 (1977): 143-151.

6. Kleinman, J.C. "Mortality." Statistical Notes for Health Planners, NCHS 3 (February, 1977): 1-16.

7. Mierly, M. and Baker, S.P. "Fatal house fires in an urban population." Journal of the American Medical Association, March, 1983, Volume 249, No. 11: 1466.

8. Whitman, S., Coonley-Hoganson, R. and B.T. Desai. "Comparative Head Trauma in two socioeconomic Chicago area communities: A Population Study." Northwestern University Center for Urban Affairs and Policy Research, submitted for publicaiton, 1983.

9. "Koop cites surge in youth violence." New York Times, October 27, 1982, p. 13.

10. Office of the Assistant Secretary for Health: Promoting Health, Preventing Disease: Objectives for the Nation. Washington, D.C.: Gout Printing Office, Fall 1980.

11. Somers, A.R. "The nation's health: issues for the future." Annals of the American Academy of Political and Social Science, January, 1972.

12. McKnight, J.D. "Politicizing health care." Social Policy, November/December, 1978.

THE SENTINEL HEALTH SURVEILLANCE SYSTEM

M. Gudes, S. B. Blount, Detroit Health Department

1. INTRODUCTION

The Sentinel Health Surveillance System was developed in 1979 by the Detroit Health Department as an instrument for setting objective measurement guidelines for identifying unnecessary deaths to Detroit residents. Using a list of preventable diseases established by the Working Group on Preventable and Manageable Diseases (Rutstein, et al., 1976), the system allows us to link vital events with demographic, social, and geographic factors that might be associated with the increased risk of unnecessary disease as a cause of untimely death.

The individual causes of death were used to classify all sentinel deaths into one of three major categories, then into five sub-categories. A sentinel death was classified into one of the three major categories according to whether it was due to a preventable, treatable, or both preventable and treatable disease. If the death was preventable, it was then grouped according to whether it was due to a failure in primary care or due to environmental, life-style, or occupational risk factors (McEvoy, 1980). It is possible that reclassification of a death could occur as a result of technological changes in the medical field. A death that at one time was in the preventable category might have been reclassified when adequate treatment became available. Reclassifications were made in conjunction with the replacement of ICDA 8 with ICD 9 revisions.

In addition to classifying individual causes of sentinel deaths, we attempted to fit simple linear models to the sentinel system data. A comparison of models was made between methods which used parametric distributions and those based on distribution-free methods. These methods were used to predict an overall rate of sentinel deaths as well as individual rates for sex, race, and geographical sub-groups. Models were used to describe the trend in individual diseases that were the major contributors to sentinel death aggregate — specifically somoking-related deaths and infant mortality.

2. METHODOLOGICAL APPROACH

The system can be viewed either as a surveillance system that identifies sentinel deaths as they occur or as an evaluation system that monitors trends over time from annual statistical reports. At this point, the latter system is the functional one, where already existing files are used to identify trends and examine the relationship between these trends and changes in environmental and demographic factors. The use of the system as an evaluation tool rather than a surveillance tool can be explained by the source of the data as well as the intended purpose of the results. A surveillance system would use occurrence of vital events in Detroit as a data source. Inferences regarding the health status of the Detroit community could be based on only that subset of events occuring to Detroit residents. In addition, events occuring outside of Detroit but to Detroit residents would be excluded. This problem of generalization to the Detroit population is not encountered with residence data. However, since the source of this data is the State of Michigan registry of births and deaths, local data is reported on an annual basis with a lag time of one to two years.

Death records were analyzed in order to identify factors that were associated with trends in overall and cause-specific sentinel death rates. Linked birth and infant death records were used to investigate the overall association among a variety of maternal variables related to differentials in infant mortality. A simple linear model based on the X^2-distribution was used to predict the expected value of the overall sentinel death rate, the smoking-related death rate, and the infant mortality rate for Detroit residents. A comparison of this method to techniques based on Exploratory Data Analysis showed no appreciable differences in the resulting estimates. The chosen method allowed distribution-based statistical tests of significance and appropriateness of the model without having to reexpress the data or used the lagged value as a dependent variable. Departures from linearity were explained using residuals of the basic model.

3. ANALYSIS

A. Total Sentinel Deaths

13.2% of all deaths to Detroit residents since 1970 can be classified as sentinel health events. The overall rate has not changed significantly since 1970, when the rate was 13.8% of all deaths compared to13.1% in 1981. The slope of the fitted line estimating the change in the rate estimated a decrease of only five deaths per year for every 10,000 total deaths. Besides this overall rate, an adjusted overall rate was computed which excluded infants in both the sentinel death figures and in the total number of deaths. This resulted in an increase in the sentinel death rate from 10.1% in 1970 to 10.4% in 1981. The positive slope of the line estimated an increase of approximately five deaths per year for every 10,000 total deaths. The conclusion from statistical tests for the appropriateness of a linear model indicated marginal significance (X^2=19.75; p=.03). Rather than search for a more complex model to explain the data, we looked for explanations in the variation among specific disease categories or population sub-groups (Table 1.).

Table 1. Sentinel Death Rate (Adjusted Overall)
% of Total Non-Infant Deaths

	Total	White	Non-White
1970	13.8	9.3	11.9
1971	13.6	9.5	11.0
1972	13.8	9.5	10.9
1973	13.2	9.5	10.4
1974	12.6	9.1	9.6
1975	13.0	9.6	10.2
1976	13.0	10.0	10.3
1977	13.0	9.4	10.6
1978	13.8	10.3	10.4
1979	13.1	9.6	10.2
1980	12.9	9.8	9.2
1981	13.1	10.0	9.0
Slope	+0.050	+0.055	-0.177
X^2(Slope)	5.96*	4.39*	29.84*
X^2(Linearity)	19.75*	8.08	17.71

*p<.05

Comparisons in sentinel deaths between whites and non-whites suggest the source of the greatest increase in the overall rate which occurred between 1974 and 1978 can be explained by the fact that both races experienced an increase in their respective rates. Between 1970 and 1978, the opposite direction in the trends of white vs. non-white sentinel death resulted in nearly equal rates. Between 1979 and 1981, the rates continued to increase for whites and decrease for non-whites to the point where whites had a 1% greater rate in 1981 compared to a 3% lower rate in 1970. The reduction in sentinel deaths among non-whites had its greatest effect within the treatable disease category — specifically, from deaths due to tuberculosis and pneumonia. The effect of the decrease in the number of tuberculosis and pneumonia deaths during this time was a significantly decreasing trend in deaths due to treatable diseases.

B. Smoking-Related Deaths

Deaths due to environmental, life-style, and occupational risk exposure increased significantly between 1970 and 1981. The specific diseases that most influenced the increase were lung cancer and chronic obstructive pulmonary disease. The combination of these diseases with other smoking-related diseases was the largest contribution to the total number of sentinel deaths for each year we examined (Table 2.). Those deaths classified as smoking-related included (1) malignant neoplasms of the trachea, bronchus, and lung, (2) malignant neoplasms of the bladder, (3) pulmonary heart disease, (4) chronic bronchitis, emphysema, and obstructive pulmonary disease, (5) malignant neoplasm of the mouth and/or lip, and (6) malignant neoplasm of the larynx. The inclusion of these causes as sentinel events introduces some of the limitations of using incomplete information as indicators of underlying preventable causes of death. For example, although the greatest proportion of lung cancer deaths are smoking-related, the absence of smoking history prevents an accurate estimate from being made. Also, the effect of a reduction in smoking behavior would not appear as a reduction in smoking-related deaths until many years later.

Smoking-related deaths have increased as a proportion of total deaths for each race and sex group. The rate of increase was higher for non-whites than for whites. Whereas the difference in their initial smoking-related death rate in 1970 was higher for whites, the increase over the subsequent eleven years in the rate for non-whites resulted in nearly equal rates by 1981. The estimated slope was higher among females — where the predicted increase was 2.8 additional smoking-related deaths for every 1,000 non-infant deaths per year. The greatest increase in the female rate has occurred since 1976. The rate prior to that period was 2.9% compared to a subsequent rate of 4.8% of total deaths.

C. Infant Mortality

Infant mortality was included in the sentinel health surveillance system as related to failures in primary care. The reason for inclusion was less related to the belief that each infant death was preventable and more to the idea that the occurrence of an infant death can be described according to the environment surrounding the infant such as prenatal care, maternal characteristics, and postnatal care.

There have been three major shifts in the direction of infant mortality between 1970 and 1982. Although there appears to have been a slight decrease from the rate of 23.4 in 1970 to 21.2 in 1982, this rate is really a reversal of a decreasing trend in infant mortality that ended in 1977 with a rate of 19.6 (Table 3.). Attempts to explain this increase triggered investigations into some of the maternal characteristics mentioned above as well as the influence of race and the difference between neonatal and postneonatal mortality rates.

Table 2. Smoking-Related Deaths
% of Total Non-Infant Deaths

	Total	White	Non-White	Male	Female
1970	5.9	6.2	5.3	8.3	2.5
1971	5.9	6.4	4.9	8.4	2.4
1972	6.3	6.6	5.9	8.9	2.8
1973	6.4	6.7	5.9	8.9	2.9
1974	6.5	6.7	6.1	8.8	3.2
1975	7.0	7.1	6.8	9.5	3.4
1976	7.3	7.7	6.7	10.0	3.5
1977	7.8	7.7	8.0	10.1	4.8
1978	8.1	8.2	7.9	10.9	4.4
1979	8.0	7.9	8.1	10.8	4.4
1980	8.0	8.1	7.8	10.4	4.9
1981	8.3	8.4	8.1	10.5	5.6
Slope	0.2	0.2	0.3	0.2	0.3
X^2(Slope)	196.6**	86.1**	131.7**	85.3**	209.7**
X^2(Linearity)	9.8	4.4	13.8	10.8	14.8

Table 3. Infant Mortality Rates
(Rates per 1,000 live births)

	Total		Race-Specific Rates	
	No.	Rate	White	Non-White
1970	746	23.4	17.7	27.9
1971	747	25.1	16.5	30.8
1972	744	27.5	17.9	33.1
1973	643	25.2	15.8	30.2
1974	588	25.8	17.4	30.3
1975	520	23.5	16.3	27.1
1976	463	22.8	16.6	25.7
1977	402	19.6	13.6	22.3
1978	452	22.4	15.4	25.4
1979	439	21.3	17.4	22.9
1980	430	20.9	11.4	25.5
1981	415	21.9	12.1	26.0
1982	391	21.2	NA	NA
Slope		-0.4	-0.4	-0.7
X^2 (Slope)		28.5**	12.1**	36.9**
X^2 (Linearity)		29.4**	10.8	32.3**

**$p \leq 0.01$

The linear model that we used to describe the twelve-year infant mortality pattern suggested that separate models were needed to include the influence of different mortality experiences between whites and non-whites and the changing racial distribution of the population. Within the twelve-year interval, there was an average difference of 11.6% between the white vs. non-white mortality rates. The decrease in the rates that occurred between 1972 and 1977 was much sharper among non-whites and therefore resulted in smaller differences. However, as the overall rate has increased since 1980, the racial differences widened as well. The influence of the racial differences in the overall mortality rate

can be explained by the fact that there was an almost three-fold decrease in white births, from 14,093 to 5,557 in 1981. In contrast, the decrease in non-white births during this same interval was from 17,787 to 13,396. During the time period between 1979 and 1981, whites experienced a decline in infant mortality where as non-white rates increased. When a linear model was used to describe race-specific rates, the model was appropriate in describing the declining white infant mortality, but resulted in a very poor description of non-white patterns. The predicted values for whites overestimated the true rates in 1980 and 1981 when the rate dropped from 17.4 to 12.1. The fitted line underestimated the non-white rate for each year between 1971 and 1974. It then overestimated the rate for every year between 1975 and 1979. It underestimated the rate again in 1980 and 1981. For 1982, the predicted rate for non-whites is 22.3. Although the race-specific rates are not yet available, since the non-white rates are increasing at a higher rate than the overall rate, the predicted value will underestimate the true rate by approximately 10%.

Between 1976 and 1980, birth and death certificates that were linked through matching information from each record were used to explain trends in infant mortality. The association between infant mortality and different demographic characteristics was investigated by measuring the extent to which the average partial association for all years between 1976 and 1980 was non-zero. In addition, a measure for the relative risk of infant death for high versus normal risk groups was computed for the following risk factors: teenage pregnancy, little or no prenatal care, out-of-wedlock births, mothers with less than a high school education, and low weight births. The X^2 test for homogeneity was used to indicate whether the differences in survival probabilities for these factors was consistent for all years. The weighted average odds ratio was used as an estimate of the avarage relative risk (Fleiss, 1973).

Births to mothers under 14 years of age, as well as births to teenage mothers in general, declined between 1976 and 1980. The proportion decreased from 29.5% to 23.8% for non-whites and from 17.6% to 14.7% for whites. However, there was very little change in the relative risk of infant mortality for teenage mothers. The odds ratio remained constant at 1.5 during this time. The ratio actually increased for whites from 1.4 to 2.5 but decreased for non-whites from 1.4 to 1.2 (Table 4.).

Although total births to mothers with less than a high school education was equal for both whites and non-whites, the numbers were distributed quite differently between those with less than eight years of school and those with 9 - 11 years. Among white mothers, the percentage of total births was almost twice as high for those with less than eight years education. Although births to mothers with less than a high school education increased, the proportion of infant deaths in this group decreased during this time. The relative risk of infant mortality increased between 1976 and 1978, then dropped with the decline in the overall rate between 1978 and 1980. The estimated relative risk increase for white mothers increased from 1.5 to 2.8 during this time while the non-white ratio increased from 1.1 to 1.6 and them dropped again to 1.2. This was the only risk factor studied where the risk of dying was greater for whites than for non-whites.

Total births reporting little or no prenatal care declined during this time. The proportion of total births with less than four prenatal visits decreased from 13.8% to 10.8%. However, the births with no prenatal care decreased for two years and then rose in 1981 to its original rate in 1976. The percent of births in 1976 with no prenatal care was 2.2% compared to 5.5% of infant deaths. By 1980, this ratio had increased to 11.9% of infant deaths compared to 2.2% of the births. The relative risk of infant mortality increased from 3.2 to 7.1 for whites and from 2.3 to 4.6 for non-whites.

Because birth certificates do not report information about marital status, an estimate is computed based on the number of certificates where information about the father is unknown. The percent has increased from 37.9% to 43.0% since 1976. The non-white rate has risen from 49.2% to 53.7% and among whites it has risen from 14.2% to 16.7%. The relative risk of infant death has risen from 1.5 to 2.1 during this time. Although the ratio for whites has remained constant at 1.9, it has risen among non-whites from 1.3 to 1.8 in the last five years.

The percent of low-weight births has varied from 11.9% in 1976 to 11.2% in 1980 and among these low-weight births, the estimated relative risk of infant mortality has increased from 16.4 in 1976 to 23.3 in 1980. The ratio for whites has increased from 24.0 to 37.8 and for non-whites from 14.2 to 19.7.

Table 4. Comparison of Maternal Characteristics (Rates per 1,000 live births)

	1976	1977	1978	1979	1980
Mother's Age					
<20 years	29.6	28.2	27.9	24.9	26.7
≥20 years	19.9	18.9	20.9	19.3	19.0
Odds ratio	1.5	1.5	1.3	1.3	1.4
Prenatal Care					
≤4 visits	59.9	60.0	60.7	57.6	62.9
>4 visits	16.1	14.6	15.1	13.6	13.6
Odds ratio	3.9	4.3	4.2	4.4	4.9
Education					
<HS	25.8	27.5	29.5	22.8	20.2
HS Grad	21.7	17.8	16.3	18.1	17.9
Odds ratio	1.2	1.6	1.8	1.3	1.1
Father Info					
Unknown	28.3	27.1	29.9	26.5	29.3
Known	19.0	17.2	18.1	16.3	14.3
Odds ratio	1.5	1.6	1.7	1.6	2.1
Birthweight					
<2500 gm	123.7	118.5	125.4	108.4	121.3
≥2500 gm	8.5	8.8	7.7	7.1	5.9
Odds ratio	16.4	15.2	18.4	17.0	23.3
IM Rate	22.5	21.2	22.6	20.5	20.7

Odds ratio = Estimated Relative Risk of Infant Mortality

A comparison on neonatal and postneonatal mortality rates was made in order to establish the differential effects of changes in the overall rate upon these two types of infant death. The overall infant mortality rate varied directly with the ratio of neonatal to postneonatal deaths (Ta-

ble 5.). Although there has been very little shift in the age distribution of infant deaths, the distribution by specific causes has shifted. This is primarily due to the decrease in deaths due to pneumonia and other infectious diseases and the increase in Sudden Infant Death Syndrome among infants older than one month. Congenital anomalies have increased as a proportion of neonatal deaths but decreased as a proportion of postneonatal deaths.

Table 5. Neonatal, Post-Neonatal, Infant Mortality Rates

	Neonatal Rate	Post-Neonatal Rate	Neonatal Proportion	Neonatal/Post-Neonatal
1970	16.0	7.4	68.5	2.2
1971	17.7	7.4	70.7	2.4
1972	19.9	7.5	72.7	2.7
1973	17.7	7.3	70.8	2.4
1974	17.8	8.0	69.0	2.2
1975	16.8	6.7	71.7	2.5
1976	15.9	6.9	69.8	2.3
1977	13.6	6.0	69.4	2.3
1978	14.1	8.3	63.1	1.7
1979	14.0	7.3	65.6	1.9
1980	15.0	5.9	71.9	2.5
1981	15.4	6.5	70.4	2.3

Neonatal Rate = Deaths occurring 28 days per 1,000 live births. Post-Neonatal Rate = Deaths between 28 and 364 days per 100 live births. Neonatal Proportion = % of infant deaths per 1,000 live births.

D. Discussion

The comparison of relative risk estimates among high risk groups uncovered differences in the ratios between whites and non-whites that suggests not only a much greater proportion of non-whites in higher risk categories for infant mortality but a much higher proportion of the overall non-white population suffering the higher infant mortality rate. Whereas there is a much greater distinction in the comparative mortality between whites in low versus high risk groups, a non-white infant not subject to any of the risk factors still falls within the high risk profile.

The future goal of the Sentinel Health Surveillance System is to use the results of infant mortality data from 1970 to 1981 and intervene in areas of high maternal risk based on the relationships established between infant death and maternal risk characteristics. A more complex model will be used that will be generated from individual records of matched births and deaths rather than the summary reports we now rely upon. Discrete multivariate logistic techniques will be applied to generate models relating maternal behaviors and characteristics to the expected risk of infant mortality. The analysis will include a geographical factor which will reflect twelve different health areas within the city of Detroit, the boundaries being determined by socioeconomic differences within the areas that are associated with access to health care.

Past data suggests that there is a large variation in the overall infant mortality rate within these twelve health areas. In 1980, the rate ranged from 13.5 in one health area to 48.3 in another. In this same health area, the proportion of low-weight births among blacks was 15.9% compared to a city-wide rate of 11.1%.

Infant mortality has become the major focus of the Sentinel Health Surveillance System not only because it is a major contribution to sentinel deaths but because it is so reflective of the general health status of the community and so affected by the changes in the economic environment and its affect on the availability of medical care. Once infant deaths can be explained by all known elements surrounding their occurrence, then we are provided with a more systematic way of describing the changing health status of the Detroit community through the analysis of mortality statistics.

REFERENCES

Fleiss, J. L.(1973). Statistical Methods for Rates and Proportions. New York: John Wiley and Sons, Inc.

Bishop, Y., Feinburg, S., and Holland, P.(1975). Discrete Multivariate Analysis: Theory and Practice. Cambridge, Massachusetts: M.I.T. Press.

Rutstein, D. D., et al.(March, 1976). "Measuring the Quality of Medical Care: A Clinical Method", New England Journal of Medicine. 294:582-588.

McEvoy, Lawrence, The Use of Sentinel Health Events in a Mortality Surveillance System, Missouri Center for Health Statistics, 1979.

Theodore D. Woolsey, Health Statistics Consultant

1. The goal of the research

First, the goal of the research on an index of preventable mortality, and the rationale behind that particular goal will be presented.

As succinctly as possible, then, the objective is to find ways of analyzing mortality statistics by cause of death so that they can serve as indicators for determining the special health problems of a community. The indicators must be more sensitive than those traditionally used for presenting mortality by cause of death. By "more sensitive" we mean: giving weight to that part of the mortality in each major cause-of-death category that is preventable by currently known means and that should, therefore, be subject to reduction by actions of the public health authorities and the health care system. Thus, we would aim to be able to distinguish between the major preventable health problems of communities and also to reflect changes in the extent of these problems over time. The indicators would be expected to react favorably to any effective actions taken to correct the problems. All of those ideas are intended to be included in the term "sensitive."

2. Why is this goal selected for the research?

As for the reasons for tackling this problem there is a good deal of evidence, I believe, of an increasing need for a health-problem indicator that can be used in jurisdictions smaller than a state. Among the pieces of evidence are: 1) the health planning legislation of 1974 which is, fortunately, still on the books; 2) the current tendency to shift ultimate decision-making responsibility for public health priorities toward state and local authorities; and 3) the ever-present competition for scarce budgets for prevention activities.

Most existing sources of data do not meet the kind of need described here at all, usually because they do not provide indicators in sufficiently fine-grained geographic detail. Furthermore, the data that are available for areas smaller than a state are not, in my view, being analyzed in a way that makes them useful for the purposes cited here.

In short, the health planners and policy makers with responsibility for the health of the people in a particular jurisdiction ought to have but do not have an objective basis for deciding where to put emphasis and resources, and should also have some way of comparing their jurisdiction with others, and of measuring change, or lack of change, following implementation of corrective measures.

Survey methods of providing the base data for such indicators are far too expensive. One would need, I would judge, a survey that would provide comparable estimates at least every five years for each of, say 200 jurisdictions in the United States, some with populations as small as 750,000. In my opinion, the costs of such a system are absolutely prohibitive, and the technical problems very discouraging. Note that I put heavy emphasis on the need for comparability of data over both time and space. Only when such comparisons are possible can the statistics be usefully interpreted.

Since we want a measure of health, and not of the use of health services, statistics of hospitalizations are not appropriate. Statistics of physician encounters come closer to measuring health, but organizing a data collection system of the scope and magnitude required seems currently impractical, except, perhaps, for the elderly, for whom Part B Medicare claims would offer some hope if there was any disposition to use them for such a purpose.

We turn to the statistics of births and deaths which are, at least, gathered and published as statistics for every major jurisdiction in a reasonably comparable manner, year in and year out. To me, these represent the only feasible source for the kind of indicators needed. Infant mortality has long been recognized as a valuable measure, but its applicability to the wide variety of health problems a community might face in the last decades of the twentieth century is limited. However, infant mortality tends to be increasingly an indicator of the socio-economic influences on health and, as such, it is still a useful part of our statistical resources.

One is inevitably pushed, it seems to me, toward the use of mortality statistics by cause of death for the sort of health measures we are trying to find. They meet reasonably well the criterion of comparability over time and space, and the marginal cost of new endeavors in this area is likely to be reasonable because statistics are already being produced at different levels of detail by large cities, some counties, all states, and the federal government.

Nevertheless, statistics of deaths by cause do suffer from several inherent drawbacks:

1) The data are too sparse. For example, our minimum-size jurisdiction of 750,000 population has only about 6,800 deaths a year, and for some causes of death which might have real value as indicators the numbers may be no more than 0.5 percent of the total, or, roughly, 35 deaths a year.

2) There are a number of serious health problems to which a community might well wish to give priority that have only trivial reflection in mortality.

3) Interpretation of cause-of-death statistics is almost completely dependent upon observing the relation of the numbers of deaths to the population of the jurisdiction by age, sex, and race. Yet often such population data are not available between censuses.

4) A lot of the deaths are not preventable by means currently known, and here it must be emphasized that by preventable we mean either primary prevention of the disease or injury, or postponing death by treatment; that is, secondary prevention. For example, two thirds of the deaths from stroke occur at age 75 and above, and 30 percent at ages 85 years and over. While stroke deaths of persons in their 40's and 50's could reasonably be considered preventable, those at these advanced ages are not.

3. Dealing with the drawbacks

Let us take up these inherent characteristics of mortality statistics one by one and suggest the reasons they do not seem to prevent such data from being enormously useful.

Regarding the small frequencies of deaths for particular causes that are likely to occur in a population of 750,000, it is clear that something has to give. Rather than accept greater aggregation in cause-of-death categories or a larger minimum population, it is my view that we should accept less frequent availability of whatever indicator we design. That is, we must combine data for more than one year. For example, if the statistics are pooled for 5 years centering on the decennial census and, hopefully someday, the mid-decade census, then the index will be available twice a decade. Health problems of the type we now consider of priority concern do not tend to show much year-to-year change, so twice a decade should usually be enough at this level of geographic detail.

It must be acknowledged that mortality by cause of death is not particularly useful as an indicator of the extent of some diseases: for example, the arthritides, or visual disorders, or schizophrenia, dental caries or the minor respiratory diseases which cause so much loss of time from work or school. Nevertheless, many other problems of ill health that are important by most any criterion can be measured in terms of the mortality they cause, and it is my contention that more of the information we have gained over the years about the trends and distribution of disease in human populations comes from statistics of mortality than from any other single source.

To interpret cause-of-death statistics one must have the appropriate population denominators. This is true almost, but not quite, regardless of the type of statistical measure of mortality one uses. (I do not intend to go into indices based on indirect adjustment or proportionate mortality in this paper, but will only say that I think they are inherently less satisfactory.) One must have these population denominators for the same geographic jurisdictions and at the same five-year intervals that we have already set as our goal. The only comments one can make about this potential drawback are these: First, one has to have such population data for many kinds of other planning, too, so we are not asking for a resource to be created especially for this purpose. Second, it is true that gradually the increasing demand for post-census estimates of population for sub-state jurisdictions has led to greater effort on the part of the Census Bureau and others to supply these; and, if the Office of Management and Budget had not been so short-sighted as to cut planning money out of the 1982 budget in 1980, we would have an assured supply from the regular mid-decade Census that the law calls for. We'll now have to wait until the issue of a Census in 1995 arises to see whether that law is going to be observed.

As for the fourth characteristic listed, the lack of known means to prevent many of the deaths that are included in our statistics, this is the drawback which I believe can be dealt with by improved analytic methods, and it is methods for such analysis that this research is intended to develop. It is this characteristic, moreover, which has led to the frequently heard objection that indices of mortality are too insensitive to be useful for measuring the health problems with which public health authorities and planning bodies are concerned. With that objection I strongly disagree.

4. Questions to be answered and solutions arrived at to date

a) Quite clearly, the task of forming an index to meet the objective I have described required dealing with data for specific causes of death because the purpose is to pin-point, as far as possible, preventable health problems and measure progress in dealing with them. But what grouping of causes of death should be used? The grouping could not be too detailed because of the problem of small numbers already mentioned. Yet, generally speaking, the greater the specificity, the more useful the analysis would be.

The recode of the International Classification that has been used for experimentation so far has 19 categories which sum to All Causes. This includes 5 malignant neoplasm site groupings; 3 for circulatory diseases; 2 for respiratory diseases; 2 for digestive diseases; 4 for trauma; and one each for diabetes, congenital malformations and diseases of early infancy, and all other diseases.

This has been modified to bring it into line with the Ninth Revision of the International Classification, but this modification has not yet been tried out because there have not yet been enough years of data available.

Of these 19 categories the smallest in terms of numbers of deaths for both sexes was one called "major digestive diseases except cirrhosis of the liver." In 1976, there were about 14,000 deaths in this group in the United States, roughly 0.8 percent of all deaths. For males and females, considered separately, bronchitis, emphysema, and asthma, as well as cirrhosis of the liver, among females, and, of course, cancer of the breast among males were even smaller relative to national totals for the cause grouping. Suicide and homicide among females were also low-frequency categories.

When it is recalled that the goal is to produce indices for areas with a minimum population of 750,000 for five data-years aggregated, it is soon evident that for some disease categories in some areas the numbers will be small and the amount of variability attributable to random variation will be relatively large. The groupings were, therefore, a compromise.

b) The second, and perhaps most critical, question with which the research had to deal was how to sort out the preventable part of the mortality within an age-sex cell for a particular cause-of-death group.

At the outset it had been assumed that this proportion would have to be determined for each age-sex cell because it almost certainly would be different from cell-to-cell.[1] (The standard 11 age groups used by NCHS in its detailed reports were used throughout.)

Two decisions were made early. The first was that all mortality in the 85 years and over age group would be considered non-preventable. The second was that in arriving at what were called achievable target death rates, i.e. the

death rates that could be experienced if we were to apply successfully all that is now known regarding primary and secondary prevention, only mortality in the white population would be used. The reason for this decision was simply that the author believes that, with rare exceptions, any minimum but achievable death rates that can be reached for the white population can be reached for the non-white population as well.

c) The discussion of the tentative conclusion reached on how to determine the preventable part of the mortality in each age-sex cell will be briefly postponed to mention the third question with which the research has dealt, up to this point. That is the question of the form of summary index that should be used to present the data for the population of an area. In some ways this is the least important of the questions, but it is also one that has had much interest for statisticians since the days of William Farr and, later, in the 1920's, Greenwood, Wolfenden, and Yule, in those gentlemanly debates in the Journal of the Royal Statistical Society, which I first read back in the 1940s. The index number had to be one that was independent of the age-sex distribution of the population because the object was to be able to make comparisons over time and space, and differences should not be attributable to demographic shifts alone. Such variations are not within the control of health authorities or the medical care system. But it is also desirable to use an index number that is as statistically sensitive as possible. By this we mean that the ratio of the range of the distribution of the index over a large set of areas to the mean standard error of the index must be at, or near, a maximum. Finally, the index should be one that stresses the underlying purposes of preventive health care which can, perhaps, be summarized as preserving as many healthy person-years of life as possible.

Without going through the experimentation done for a sample of areas (which has been described in the NCHS Vital and Health Statistics Series 2, No. 85 - "Toward An Index of Preventable Mortality") it will simply be reported here that the form favored at this point is the so-called "years-of-productive-life-lost" index, used by Kleinman and other authors, but in this case basing the measure on preventable deaths. Quoting from the earlier report: "The years-of-life-lost form has the conceptual advantage that mortality at the younger ages, considered much more amenable to correction efforts, is weighted a great deal more than is mortality at advanced ages." The statistical sensitivity, analogous to the engineer's information-to-noise ratio, was nearly as high for this form of index as for the traditional standard mortality ratio which does not have that advantage. (I think most people are unaware of the heavy influence upon the usual age-adjusted death rate played by deaths at the oldest ages.)

5. The preventable mortality

Beginning, then, with the premise that a determination is to be made of the preventable proportion of the mortality in each age-sex group up to age 85 years, and for each cause-of-death category, using existing mortality in the white population as a guide, the question becomes how to determine this proportion.

Statisticians are pre-disposed to methods that do not depend upon individual judgments. Consequently, a method was adopted that was a modification of one used by Guralnick and Jackson in a 1967 article in Public Health Reports, "An index of unnecessary deaths," which, in turn, came from an idea put forward by William Farr in the mid-1800's.

This idea originally was that the area with the lowest death rate can be used as a standard against which the experience of other areas can be compared. From that point it is not too much of a jump to the argument that low mortality achievable in one jurisdiction can be achieved in the others if major demographic variables are held constant. While it can reasonably be stated that the lowest mortality experienced in one age-sex-cause group in a set of areas is not always as low as might be achieved by successful application of all we know about prevention, it can also be argued that the lowest mortality for any area may sometimes actually be lower than is achievable solely by methods within our control. Genetic factors, for example, could bring about such a result. Thus, there appear to be counterbalancing factors making the estimate by this method at times too low, and at others too high.

In practice, what has been done so far, has been to tabulate a frequency distribution of white death rates in the nine Census geographic divisions of the United States for each age-sex-cause cell and to use, instead of the bottom end of the range, a statistic based on the mean less twice the standard deviation of the distribution to establish the achievable minimum. (Of course, no probability interpretation should be attached to this statistic since nothing is known or can be assumed about the form of the underlying distribution.) However, the statistic is more stable than the lowest rate for any area. This preliminary achievable minimum was set in ratio to the corresponding U.S. death rate. The age curve of these ratios was then smoothed to give the final proportion of non-preventable to total mortality. The balance, of course, was the preventable part.

While this method is reproducible and has some logic to support it, one wonders how experts in preventive medicine would look at the results. Superficially they appear reasonable. The proportions preventable are larger at the younger ages and, for the most part, the differences among the cause groups are what one might expect, but there are some surprises which are probably due to artifacts in the data.

6. The next steps

Where this unfinished research should go from here depends to a large extent upon who does it. Certainly, the present investigator, though convinced that an index of this sort can eventually be a useful tool, does not expect to be able himself to bring it to that point. So let me conclude by issuing an invitation to anyone interested in collaborating and then taking over where this leaves off.

There are at least four tasks yet to be performed:

1) Establishing a new set of what I have called achievable target death rates (i.e. the

non-preventable part of the mortality) based upon the Ninth Revision of the International Classification and the geographical variation provided by mortality of the white population in the United States during the five years centering on the 1980 Census.

2) Using the figures obtained in that step, calculating years-of-life lost mortality indexes by cause and sex for a number of local jurisdictions and offering the data to the authorities in those areas, free of charge, to market test them for usefulness. Those accepting the offer would be asked to report back in a year about uses to which the data were put and any shortcomings.

3) Sending the basic proportions of deaths preventable in the age-sex-cause of death cells to a panel of experts, along with suitable background data, asking the panel to provide their views as to the reasonableness of the proportions as seen from the standpoint of preventive medicine. In connection with this task it would also be desirable to look at the consistency of these supposedly achievable minimums with the national targets established in the report entitled "Healthy People," published by DHHS.

4) And, finally, altering the methods and achievable target death rates in accordance with the findings of Tasks 2 and 3.

A small start has been made on Task 1 already, and I intend to see this Task through, but beyond that I would expect to be an interested bystander, for the most part.

I might add, in closing, that W.H.O. has expressed interest in having a thorough but nonmathematical exposition of the method, including reference to experience in the application of this index in health planning in the United States.

I shall welcome hearing from health statisticians with an interest in carrying on this research.

Thank you.

---

[1]Incidentally, this is a point at which this research departed from the methods used by Rutstein, et al., in their research using mortality statistics to measure the quality of medical care. In that research, cause groupings in considerable detail were classified as preventable or non-preventable on an all-or-nothing basis in most instances.

Data Needs About Accidents, Violence, and Occupational Injuries

Session G

NATIONAL HEALTH STATISTICS AND INJURY PREVENTION

Janine Jagger
University of Virginia
Susan P. Baker
The Johns Hopkins University

This presentation describes a neglected area of public health - that of injuries, and the role of national data in the prevention of mortality and morbidity from injuries.

Injuries constitute a medical and public health problem of truly vast dimension. Today, injuries are the fourth leading cause of death in the United States (1). But this simple statistic does not do justice to the importance of the problem. Injuries are the number one cause of death from ages 1 through 44. Cancer and heart disease, which gain in prominence in the later years cause far fewer deaths than injuries from 1 through 44 years of age. Another way of looking at the dimension of this human tragedy is to compare the number of potential years of life lost due to injuries in comparison to other major causes of mortality. This figure is obtained by finding the difference between age at death and a potential life expectancy of 70 years. In 1978, more potential years of life were lost due to injuries than for any single disease category (2). There were 4.3 million potential years of life lost in 1978 because of injuries; more than for cancer and heart disease combined.

Each year approximately 1 in 1,400 people dies of injuries, but deaths from injuries represent only the tip of the iceberg. A recent hospital-based study of emergency room visits has found that 1 in 5 people seeks emergency room treatment for injuries in a given year (3). Injuries account for more physician contacts than any single disease and they rank a close third after circulatory and respiratory diseases as a reason for hospital admission (4,5).

A public health problem of this magnitude deserves a level of detail in published national statistics in proportion to its importance; that is, relative to other health problems. The following compares the level of detail presented in the NCHS tables showing the 15 leading causes of death, 34 selected causes of death, 72 selected causes of death, and 281 selected causes of death. Four injury categories were included among the 15 leading causes of death for 1980 (1). Accidents and adverse effects (presented as motor vehicle accidents versus other accidents and adverse effects) ranked 4th, suicide 10th, and homicide 11th. The next most detailed table showing 34 selected causes of death for 1978 adds 6 more categories for cancer, 5 more for heart disease and a handful of individual categories with as few as 321 deaths due to complications of pregnancy and childbirth and 169 deaths for syphyllis and its sequellae (6). No additional categories are presented for injuries despite the large numbers they represent. Likewise, in the even more detailed table showing 72 selected causes of death, still more categories are added for cancer and heart disease, and a new section is added with 13 categories of infectious and parasitic diseases (1). Four of those 13 diseases account for 15 or fewer deaths in 1980. Yet, in the same table there is still no detail added to injury categories.

What further breakdown would prove useful? To be able to distinguish pedestrian deaths from motor vehicle deaths, to distinguish deaths from falls, poisonings, drowning, and burns from the category "all other accidents and adverse effects" and to be able to distinguish suicide and homicide by firearms would contribute greatly to efforts to set prevention priorities and document injury rates and trends. Each of the categories just mentioned includes between 4,000 and 15,000 deaths per year (6). To present injuries in broad, heterogeneous categories obscures detail valuable both to the documentation and control of this public health problem.

A greater amount of detail is provided in the table showing 281 selected causes of death. But this table appears only in the Vital Statistics of the United States volumes. The most current year available is 1978. Greater detail is needed in the more timely Advance Report and the Annual Summary in which data from 1980 and preliminary 1981 data (based on a 10% sample) have already been published. The format and detail in which injury data are now presented originated during a period when infectious diseases were the prominent cause of death. The greater detail we suggest corresponds to the current importance of injuries as a public health problem.

What are injuries caused by? As obvious as the answer may seem, persistent confusion about this point has been a major barrier to progress in the area of injury epidemiology and control. Injuries are caused by physical agents such as heat, mechanical or electrical energy, chemicals, and ionizing radiation (7). Injury results when these agents are transferred to the human body in quantities that exceed human tolerance or when an agent necessary to sustain human life such as oxygen or heat is lacking. This definition of the etiology of injuries covers the full spectrum of injuries whether the injury was intentional or unintentional and regardless of how the physical agent was delivered to the injured person.

Traditionally, the occurrence of injuries has been thought of as a problem of human behavior. The natural extension of this perspective is that the prevention of injuries must involve changes in human behavior (8). This is an unfortunate misconception. Historically, the most successful injury prevention strategies have not involved changes in human behavior but rather the modification in the access to or delivery of a physical agent to a host. For example, child-proof caps on medicine bottles, design standards for gas tanks of motor vehicles to prevent post-crash fires, pedestrian overpasses at traffic intersections, and the regulation of crib slat spacing to prevent infant strangulation. In published national mortality

statistics, the conception that injuries are a problem of human behavior is reinforced by the classification of data into intentional versus unintentional injuries as the major subdivisions. This classification obscures injuries related to certain products or sources of injury for two reasons. First, a different level of detail is provided for unintentional and intentional injuries: i.e., there is far less detail in the case of homicide and suicide. There is a special section in Vital Statistics of the United States presenting seven detailed tables of unintentional injuries (6), but for homicide and suicide no detail is presented beyond five categories for each in the table of 281 selected causes of death (6). This is unfortunate and ironic as the means by which people commit homicide and suicide are frequently the same as those by which people are unintentionally injured and killed.

The second reason that the classification of data into intentional versus unintentional is problematic is that intent cannot always be established. Close to 4,000 injury deaths per year are classified as "undetermined whether intentional or unintentional." A special analysis carried out by Withers and Baker (9) shows that for deaths in this category a significant number are due to firearms, drowning, poisoning and burns; each involving different etiologic agents and mechanisms of injury. Without carrying out special analyses, it is not possible to determine the total number of deaths in a given year due to firearms, drowning, poisoning and burns - each of which requires unique prevention strategies. It is therefore important to injury control efforts that at least as much detail be presented for mechanism of injury as for intent.

Morbidity data concerning injuries can be found in Series 10 and 13 of the Vital and Health Statistics rainbow series. These data point out additional problems in the collection and presentation of injury statistics. The National Health Interview Survey data includes incidence rates for injuries. These data are classified by the event that led to injury such as motor vehicle crash, fire or explosion, or discharge of firearm (10). This classification is similar to that used in mortality statistics which is based on the "E" codes of the International Classification of Diseases (ICD). However, data from the Hospital Discharge Survey present an anatomic description of injuries based on the ICD "N" codes (11). This is because hospitals often do not record "E" codes which describe the event leading to injury but record only "N" codes; or the actual injuries which were the basis for medical treatment. This discrepancy in available data creates a barrier in correlating events leading to injury with the resulting human damage, and represents an enormous loss of information.

Certainly, the National Center for Health Statistics has inherited some of the problems described here. The accuracy of mortality data cannot exceed the accuracy of the death certificates from which the data are abstracted. The different levels of detail provided for intentional versus unintentional injuries is inherent in ICD coding and this discrepancy is, in turn, reflected in mortality statistics. Many hospitals do not record the "E" codes for injuries, hence, this information is not always available.

What improvements, then, might be carried out given the current data structure and limitations? First, additional categories for injuries could be added to the table of 72 selected causes of death. Using the existing major subdivisions "motor vehicle accident deaths" could be further broken down by "motor vehicle occupants", "pedestrians", and "motorcycle crashes". The major subdivision "other accidents and adverse effects" could be broken down into "falls", "poisonings", "burns", and "drowning". Suicide and homicide could be broken down as "firearm deaths" versus other means.

In the table of 281 selected causes of death, additional categories could also be added. The possible mechanisms of injury for suicide and homicide could be presented in roughly the same detail as for accidents and adverse effects. In this way, deaths caused by a single product could be totalled in a given year regardless of intent. Also, the seven table section in Vital Statistics of the United States devoted to unintentional injuries should be extended to include the same amount of detail for homicides and suicides. The present format represents an illogical division of a single public health problem.

These are suggestions which would just begin to bring injuries into a proper perspective among other health problems in national data. The described changes would preserve the existing format of data presentation and continue to be compatable with international conventions for statistical reports.

Additional improvements, although more difficult to accomplish, would greatly enhance the detail, accuracy, and usefulness of injury data. The ICD codes, the most basic element of national statistics, have some fundamental problems in providing useful codes for injury. In particular, the level of detail for homicides and suicides is considerably less than for unintentional injuries. Furthermore, there are no codes to define the etiologic agent of injury or to distinguish between the etiologic agent and the means by which it is delivered to the host. Changes addressing these points have been recommended for the 10th revison of the ICD codes (12). Should changes such as these be adopted, the potential information gain will eventually be passed on to the national data base.

The problems and recommendations outlined here represent only the highlights of an issue that has received little attention. Despite the magnitude of injuries as a public health problem, the potential for gains in the prevention of injuries remains great. However, our ability to set prevention priorities, to implement effective prevention strategies, and to document our progress is, in fact, dependent on the availability of data consistent with modern concepts of injury occurrence and

control. Our national data base is one of our most valuable resources for injury control, yet the current presentation of injury data reflects a conceptual framework originating prior to the 1950's. The time is long overdue to present injuries in a format and at a level of detail appropriate to its significance as a public health problem.

REFERENCES

1. National Center for Health Statistics: Advance Report of Final Mortality Statistics, 1980. Monthly Vital Statistics Report. 32, #4, 1983.

2. National Center for Health Statistics: Years of Potential Life Lost, Deaths, Death Rates, by Cause of Death, and Estimated Number of Physician Contacts, by Principal Diagnosis, United States. MMWR. Vol. 31, December 1982.

3. Barancik JI, Chatterjee BF, Greene YC, Michenzi EM, and Fife D.: Northeastern Ohio Trauma Study I. Magnitude of the problem. Am J Public Health 73:746-751, 1983.

4. National Center for Health Statistics: Vital and Health Statistics Series. Series 13, #33, 1975.

5. Haupt BJ: Detailed Diagnoses and Surgical Procedures for Patients Discharged from Short-Stay Hospitals: United States, 1979. DHHS Publication No. (PHS) 82-1274-1. Washington, D.C., 1982.

6. National Center for Health Statistics: Vital Statistics of the United States, 1978. Vol. II, Part A. Public Health Service, Washington, U. S. Government Printing Office, 1982.

7. Haddon W, Jr.: A Note Concerning Accident Theory and Research with Special Reference to Motor Vehicle Accidents. Am NY Acad Sci 107: 635-646, 1963.

8. Baker SP: Medical Data and Injuries. Am J Public Health 73:733-734, 1983.

9. Withers BF, Baker SP: Epidemiology and Prevention of Injuries in Emergency Medicine Clinics of North America, Special Issue on Multiple Trauma, ed. Kayten CG, Philadelphia, W. B. Saunders Co. (in press).

10. National Center for Health Statistics: Vital and Health Statistics Series. Series 10, #105, 1972.

11. National Center for Health Statistics: Vital and Health Statistics Series. Series 13, #35, 1975.

12. Baker SP: Injury Classification and the International Classification of Disease Codes. Acc Anal and Prev 14:199-201, 1982.

OCCUPATIONAL INJURY DATA: ARE WE COLLECTING WHAT WE NEED FOR INDENTIFICATION, PREVENTION, AND EVALUATION?

Karl Kronebusch, Office of Technology Assessment

Introduction

Occupational injury data collection should serve three purposes: identification of workplace hazards, design of preventive measures, and evaluation of preventive interventions. In the first section of this paper I briefly describe the major sources of information on the magnitude of the occupational safety and health problem. Unfortunately, much of the information from these sources has historically been limited to recording information about the injured employee and the nature of the injury. Although this is useful information for identifying occupational hazards it is insufficient in most cases for designing preventive interventions. Moreover, the evaluation of various efforts to reduce the incidence of occupational injuries is made difficult by a number of factors. In the second section of this paper, I briefly describe some of the factors that influence trends in occupational injury rates. Finally, questions still remain about the accuracy of employer-maintained injury records. In the final section, I compare the number of occupational injuries reported from several different sources and discuss possible explanations for the reported differences among these sources.

Sources of Occupational Injury Information

Since the passage of the Occupational Safety and Health Act of 1970, workplace injury data collection has included mandatory recordkeeping by employers and annual surveys of a sample of employers by BLS. The survey results are used to compute injury and to a limited extent, illness rates by industry, as well as estimates of the total numbers of fatalities, lost workday cases and days, and cases without lost worktime, but which involve medical treatment.[1] BLS has also since the mid-1970s compiled information provided by 25 to 30 state workers'compensation agencies, in a database known as the Supplementary Data System (SDS).[2] Information on occupational injuries is also available from the National Health Interview Survey of the National Center for Health Statistics[3] and a system recently created by the National Institute for Occupational Safety and Health (NIOSH) in cooperation with the Consumer Product Safety Commission (CPSC) that uses reports from from hospital emergency room admissions.[4] In addition, the National Safety Council has prepare estimataes since the 1920s.[5]

For nonfatal injuries, the estimates from different sources differ in part because of differing definitions and differing population universes. The National Safety Council estimates that about 2.1 million disabling injuries occurred at work during 1981. The 1981 BLS Annual Survey provides estimates of 2.4 million lost workday cases and 2.9 million cases of injuries that required medical treatment but did not involve loss of worktime. Thus the BLS total is about 5.3 million injuries.

The National Health Interview Survey estimate for 1981, on the other hand, is that about 11.3 million occupational injuries occur that are either medically treated or that restrict activity. To some extent, the difference between the Health Interview Survey and the BLS estimates is due to injuries incurred by public employees and the self-employed, both of which are excluded from the BLS survey universe. Other possible reasons for this difference are discussed in my third section. NIOSH estimates that about 3.2 million work-related injuries were treated in hospital emergency rooms in 1982.

Current data are probably sufficient to identify major injury hazards. For example, BLS Annual Survey data show that the leading causes of fatal occupational injuries are over-the-road motor vehicles, industrial vehicles, and falls. For nonfatal injuries, the Supplementary Data System can tell us that the leading types of disabling injuries are overexertions, "struck by" injuries, and falls.

But as aids for devising preventive interventions, currently collected data do not provide sufficient detail. OSHA requires that employers maintain certain records of occupational injuries. These include a listing of injuries known as the OSHA Log or OSHA Form 200, as well as certain required suplementary information concerning the circumstances of the injury, the nature of the injury, etc. It appears that most employers keep injury records primarily for workers' compensation purposes and not to aid in the identification of accident circumstances and causal factors.[6]

In addition, employer record-keeping concerning the details of the injury sequence is often incomplete. Indeed, the most commonly used coding system, the American National Standards Institute (ANSI) Z16.2 standard, is structured to record only "one pertinent fact" about the injury in each of seven categories. Injury research ought to emphasize the interactions between workers and their workplaces, including the events, objects, conditions, and employees that lead to injuries. The ANSI standard attempts to collapse these into the categories "hazardous condition" and "unsafe act." In this kind of classification, many interrelationships and detail useful for designing preventive measures are lost. Work is in progress, at NIOSH and other places to develop injury investigation systems that will capture the detail necessary for prevention.

Injury Trends

Next we turn to the topic of evaluating preventive interventions. An appreciable increase in injury frequency rates, revealed by BLS surveys, during the 1960's provided part of the impetus for passage of the OSH Act in

1970. Unfortunately for the task of evaluation, the OSHA act also changed the definition of injuries, while the BLS survey procedures were changed and the universe expanded. Thus a simple before and after comparison is not possible using BLS data.

The post-1970 BLS data for injuries and illnesses reveal declines in the total case rates, especially from 1973 to 1975 and in 1980. This decline continued in 1981, when the total case rate fell from 8.7 to 8.3. The rate for non-fatal cases without lost workdays has declined since 1973. The lost workday rate has, on the other hand, generally increased over the period, with declines only in 1980 and 1981.

There are several factors that confound the injury rate picture and make an evaluation based on simple trend analysis misleading. First is the effect of the business cycle. Since the 1930's, researchers have noted the effect of the level of business activity has injury rates. Other things being equal, increased business activity leads to higher reported injury rates and decreased activity leads to lower reported rates. The general explanation for this phenomenon is that as business picks up, employers hire more young and inexperienced workers, both of whom tend to have higher injury rates than older and more experienced workers. Moreover, as production expands, businesses open new plants and bring on line new machinery leading to a period of adjustment as management and workers learn how to use the machinery safely. Moreover, during a business upturn there will be increases in the pace of production, increases in the amount of overtime worked, less down time, and less time devoted to repair and maintenance, all of which will also lead to increases in injuries. During business downturns, all of these elements are reversed--younger and less experienced workers are laid off while older and more experienced workers are retained, plant operations slow down, and more effort is devoted to repair and maintenance.[7] In addition, during a business downturn with increasing levels of unemployment, workers become more fearful of losing their jobs. Because of that, they may be more hesitant about reporting injuries to their employers.

It is possible that the rise in injury rates during the 1960s was because the economy was being run at full steam with low levels of unemployment. Similarly, the declines in the injury rates in 1980 and 1981 are largely due to the effects of the recession. There are few things in life as certain as death and taxes, but one that comes close is the decline in reported occupational injury rates during recessions. It is therefore safe to predict that the BLS statistics for 1982, which will be released this fall, will show a dramatic decline. I expect that the decline will, at a minimum, be from the 1981 rate of 8.3 per 100 employees to 7.3 per 100. Given the depth of the recession in construction and manufacturing, I would not be surprised if the rates fell even further, e.g. to 7.0 per 100. Such a decline would be the largest year to year percentage change since BLS began collecting these statistics.

Other factors include: shifts in the composition of the workforce among occupations and industries, changes in the practice of medicine, changes in the administration of workers' compensation, increasing recognition of the occupational origins of illnesses and cumulative trauma injuries, increases in the number of two-earner households, and cultural changes that are more tolerant towards persons recuperating from injuries. The effect of these factors might help explain why the workday case rate and the number of lost workdays per 100 workers and per injury increased during the 1970s at the same time that the non-lost workday injury rate was decreasing or remaining virtually constant.

Accuracy of Employer-Maintained Records

One other impediment to evaluation is the question of the accuracy of employer-maintained records. There have been persistent criticisms that these records understate the magnitude of the occupational safety and health problem. The fears that employers do not record all workplace injuries have been magnified by recent OSHA actions, which now effectively exempt employers from inspections if the employer's records show an injury rate less than the national average lost workday rate for manufacturing.

Previous Research

Gordon and colleagues[8] found serious underreporting to BLS in the late 60s, but this was before the creation of the mandatory recordkeeping system in 1970. BLS, in the early 1970s, did conduct a "quality assurance program" that included on-site comparison of the mail survey form with employer records.[9] It appears, however, that no attempt was made to verify the accuracy of the underlying employer records.

A NIOSH-funded study found, in a sample of establishments in two states, that 42 percent of the occupational injuries reported by employees to the researchers were not recorded in the employer records. However, this was based on a small sample, in a limited number of industries, and was conducted between July 1972 and August 1974, relatively early in the life of OSHA.[10] Two other studies[11], examined employer records. But in neither study were any efforts taken to find injuries that had not been recorded by employers. Susan Baker and colleagues[12] have found that for occupational fatalities, workers' compensation records, reports to OSHA, and death certificates are each individually insufficient to identify all fatal occupational injuries.

Comparison of Annual Survey and Supplementary Data System

To look further at this question, I examined information from several different sources. It must be emphasized that no field work was done, but only a comparison of data already available from BLS and NCHS. The first comparison is between the BLS Annual Survey and the BLS Supplementary Data System, both of which are based on information coming ultimately from employers.

Because states vary in the types of cases reported to the SDS, we need to have a distribution of the number of cases by length of

disability. Although, neither the Annual Survey nor the SDS can provide us with such a distribution, a distribution of disability cases is available from the National Council on Compensation Insurance (NCCI) and is presented in Table 1.[13]

Table 2 presents information from both data sources for the states that report to SDS only cases that involve disability days. The last three columns present the ratio of SDS cases to the number of Annual Survey cases for 1978, 1979, and 1980. A ratio equal to one indicates that the two systems are reporting the same number of cases. Ratios less than one indicate fewer SDS cases than Annual Survey cases, while ratios greater than one show a greater number of SDS cases than Annual Survey cases. To judge whether the general decline shown in Table 2 for these ratios is above or below that expected, we must refer back to the distribution on Table 1.

This comparison shows that 8 of the 36 state and year combinations are nearly equal to expectations, 15 are higher than expected and 13 are lower than expected. The pattern seems to be an overall decline in the ratio as the workers' compensation waiting period increases. The state by state ratios are generally consistent from year to year, but are not consistently related to the "expected" number from the NCCI distribution.

A clearer pattern emerges in the comparison for those states that submit medical treatment and disability cases to the SDS. Table 3 presents the data for these states. Because certain states have changed their reporting definitions, comparisons can be made only for some of the years that these states have submitted cases to SDS. In Table 3, asterisks mark the ratios for which a meaningful comparison is possible. Again, these ratios should equal one if the two systems are consistent.

The results are striking. Of the 20 state and year combinations that can be compared directly, 18 are above the expected ratio of 1.0, one is below and one is nearly equal to unity. For these 20 combinations, the SDS has, on average, 33 percent more cases than the BLS Annual Survey reports. This indicates either that the Annual Survey is underreporting the numbers of such cases or that the SDS is overreporting. One possibility is that many injuries arequire medical treatment, for which a workers' compensation claim is filed, but that this medical treatment is labelled by employers as "first aid" and thus does not need to be recorded on the OSHA Log. The states for which this is true, however, may not be representative of the nation. Most of the states in Table 2 are small to medium-sized states which lie outside the industrialized Northeastern and Midwestern regions.

Comparison of Annual Survey and NHIS

Information from the National Health Interview Survey (NHIS) can also be compared to that from the BLS Annual Survey. The OSHA/BLS definition of recordable cases is similar, but not identical, to the NHIS definitons for injuries.[14]

The universes for the two surveys are different. BLS surveys private employers and excludes the self-employed, farms with fewer than 11 employees, and government employees. The NHIS covers the entire civilian, non-institutionalized population of the U.S. Work injuries, however, are limited to those who are currently employed. The NHIS therefore includes nearly everyone who works, covering nearly all of those excluded from the BLS universe.

I adjusted for the different universe sizes by developing a work injury rate for the NHIS. This simply involved dividing the number of people injured while at work by the number currently employed.[15]

Table 4 presents these rates and the rates for the private sector from the BLS Annual Survey. The NHIS injury rates are consistently higher than the BLS rates, and except for 1973, higher than the BLS rates for injuries and illnesses combined. These differences do not appear to be attibutable to the sampling errors of these surveys. The relationship appears to be relatively constant since 1975. Moreover, approximate 95 percent confidence limits, calculated from the relative standard errors for 1981, do not overlap for the two surveys.

The result is particularly striking because with some exceptions, most of the self-employed and public employees who are included in the NHIS universe, but excluded from the BLS universe, are generally perceived to be in low risk employments. Making an adjustment for the large numbers of these workers in low risk jobs would magnify the differences between the two surveys. It is possible that people responding to the NHIS are including as "medically atended," injuries that OSHA/BLS classifies as "first aid only." But it is also possible that the OSHA/BLS definitions of "first aid" are too broad or not clear enough. Furthermore, it is possible that employers are systematically applying the first aid definition to more cases than would their employees. Finally, it is possible that some employers are failing to record cases.

In summary, based on a limited analysis of state by state totals from the BLS Annual Survey and the BLS Supplementary Data System (SDS), it appears that for states reporting only lost workday cases to the SDS, the numbers of cases are not consistently higher or lower in one or the other data source. This adds some confidence that the two data sources are reporting the same kinds of cases. For the states that report all cases involving either lost workdays or medical treatment, consistently more cases were reported through workers'compensation agencies to the SDS.

The comparison between BLS Annual Survey injury rates and rates calculated from the estimates of the National Health Interview Survey seems to show that, in recent years, overall injury rates based on the NHIS are about one-third higher than those from the BLS Annual Survey. This difference could arise from the different methodologies of the two surveys. As such, the differences may not be worrisome. Or it could be that employers are systematically failing to record certain types of injuries or are labelling these injuries as minor injuries involving only first aid treatment, even though

employees consider them serious enough to report them to the NHIS. Finally, it is possible that the discrepancy could be confined to one group of employers.

Conclusion

In conclusion, we should return to the three criteria of identification, prevention, and evaluation. The data currently collected are adequate for describing the overall magnitude of the occupational injury problem and to list the types and sources of injuries. But they are of only limited usefulness for designing preventive interventions. In the process of evaluation, one must be careful to examine all the factors that influence injury rates, especially the apparent effects of the business cycle on reported injury rates. Finally, there are differences among the various sources of data. The comparisons of data from the BLS Annual Survey, the BLS Supplementary Data System, and the National Health Interview Survey, illustrate some of these differences, which include different survey universes, methodologies, and perhaps, different interpretations of what kinds of injuries are significant. These differences may not be enough to alter the kind of preventive strategies or the level of effort that we as a society, choose. But given these problems, we must be cautious in our use of the data.

Notes

1. Bureau of Labor Statistics, Occupational Injuries and Illnesses in the United States by Industry.

2. Norman Root & David McCaffrey, "Providing More Information on Work Injury and Illness," Monthly Labor Review, April 1978, pp. 16-21.

3. National Center for Health Statistics, Current Estimates for the National Health Interview Survey.

4. Pat Coleman, NIOSH, personal communication.

5. National Safety Council, Accident Facts.

6. Safety Sciences, Feasibility of Securing Research-Defining Accident Statistics, NIOSH Publication #78-180, Sept. 1978.

7. Max Kossoris, "Industrial Injuries and the Business Cycle," Monthly Labor Review, March 1938, pp. 579-94; Robert Smith, "Intertemporal Changes in Work Injury Rates," IRRA Proceedings, 1972, pp. 167-74.

8. Jerome Gordon, et al., Industrial Safety Statistics, (New York: Praeger, 1971).

9. John Inzana, BLS, personal communication.

10. David Discher, et al., Pilot Study for Development of an Occupational Disease Surveillance Method, NIOSH Publication #75-162, May 1975.

11. John Mazor, "How Accurate Are Employers' Illness and Injury Reports?" Monthly Labor Review, Sept. 1976, pp. 26-31; Safety Sciences, op cit.

12. Trudy Karlson & Susan Baker, "Fatal Occupational Injuries Associated with Motor Vehicles," Proceedings, Amer. Assoc. for Auto. Medicine, 1978, Vol. 1, pp. 229-41; Susan Baker, et al., Fatal Occupational Inuries, JAMA, August 13, 1982, pp. 692-7.

13. This is a multi-state, multi-industry compilation of data from 1974. The distribution of injuries for individual states or industries may differ from the NCCI distribution. Applying the NCCI distribution to the Annual Survey and SDS data to develop an "expected" number of injuries, thus cannot be conclusive.The universes covered are slightly different. Because the BLS survey covers only private sector employers, information on public employees has been excluded from the SDS data on Tables 2 and 3.

14. OSHA and BLS require that a case be recorded if it involves death, medical treatment beyond first aid, restriction of work or motion, or lost worktime. Work injury cases reported to the NHIS are injuries classified in nature of injury codes 800-999 of the International Classification of Diseases that are medically attended or involve a restriction of activity and that occur while the injured is at work. Some first aid only cases may appear in the NHIS that will be excluded from the OSHA/BLS definition, but injury cases that involve either medical treatment or restriction of work activity should appear in both.

15. This borrows a methodology used by Pat Breslin, in OSHA, Division of Statistical Studies, "A Comparison of Occupational Injury Rates," xerox, Feb. 1978.

Michael Gough, Mark MacCarthy, Cynthia Bascetta, and Raymond Donnelly provided helpful comments on an early draft of this paper.

Table 1
Duration of Disability for Cases with Temporary Total Disability

Disability Period	Cases*	Percent
1 or more days	103,371	100.0
2 or more days	94,398	91.3
3 or more days	86,200	83.4
4 or more days	79,964	77.4
5 or more days	72,887	70.5
6 or more days	66,450	64.3
7 or more days	61,294	59.3
8 or more days	56,440	54.6

Source: National Council on Compensation Insurance (NCCI)

* Cases closed in 1974 and reported to NCCI in their "Special Call for Accident Statistics" issued in June 1974. Forty-one insurance companies responded.

Table 4
Work Injury Rates from the BLS Annual Survey and the National Health Interview Survey

Year	BLS Total Case Rate[1]	BLS Recordable Injury Rate[2]	NHIS Injury Rate[3]
1962-63	N.A.	N.A.	10.8
1963-64	N.A.	N.A.	12.3
1964-65	N.A.	N.A.	12.8
1965-66	N.A.	N.A.	12.8
1966-67	N.A.	N.A.	13.8
1967	N.A.	N.A.	12.2
1968	N.A.	N.A.	12.2
1969	N.A.	N.A.	10.7
1970	N.A.	N.A.	10.0
1971	N.A.	N.A.	12.4
1972	N.A.	N.A.	9.9
1973	11.0	10.6	10.8
1974	10.4	10.0	11.0
1975	9.1	8.8	11.8
1976	9.2	8.9	10.7
1977	9.3	9.0	12.6
1978	9.4	9.2	11.3
1979	9.5	9.2	12.4
1980	8.7	8.5	11.1
1981	8.3[4]	8.1[4]	11.3[5]

1. Incidence rate per 100 full-time equivalent workers for all recordable injuries and illnesses in private sector employment.
2. Incidence rate per 100 full-time equivalent workers for only recordable injuries in private sector employment.
3. Number of medically attended or activity restricting injuries reported to the NHIS divided by the number of persons "currently employed" from the NHIS times 100.
4. Relative standard error is less than 0.5%.
5. Relative standard errors are approximately 6.6%, 5.6%, and 10.8% for the injury rate, restricted activity days, and bed disability days, respectively, in 1981.

Source: BLS, Annual Survey and NCHS National Health Interview Survey.

Table 2
Disability States

State	Type of Case Reported to SDS	Ratio: SDS/Annual Survey 1978	1979	1980
Alaska.....	1 or more lost workdays	N.A.	1.09	1.26
California.....	1 or more lost workdays	.99	1.00	1.00
Massachusetts..	1 or more lost workdays	N.A.	.70	.85
Indiana.....	more than 1 lost workday	.72	.70	.73
Kentucky.....	more than 1 lost workday	1.35	1.41	1.01
Iowa.....	1978-79: 2 or more lost workdays 1980: 3 or more lost workdays	.60	.57	.81
Minnesota.....	3 or more lost workdays	1.00	1.04	1.01
Maryland.....	4 or more lost workdays	.65	.63	.57
Oregon.....	4 or more lost workdays	.94	.85	.97
Michigan.....	7 or more lost workdays	.73	.75	.78
Arizona.....	1979: 7 or more lost workdays 1980: 8 or more lost workdays	N.A.	.54	.64
New Mexico.....	8 or more lost workdays	.35	.31	.41
Tennessee.....	8 or more lost workdays	.60	.64	.66

Source of data: BLS, Annual Survey and Supplementary Data System, unpublished data.

Table 3
Medical Treatment States

State	Type of Case Reported to SDS	Ratio SDS/Annual Survey 1978	1979	1980
Hawaii.....	$\geq$ 1 lost workdays or medical	1.27*	1.39*	1.43*
Idaho.....	$\geq$ 1 lost workdays or medical	1.69*	1.70*	1.73*
Maine.....	$\geq$ 1 lost workdays or medical	1.27*	1.31*	1.35*
Montana.....	$\geq$ 1 lost workdays or medical	N.A.	1.74*	2.13*
Vermont.....	$\geq$ 1 lost workdays or medical	1.50*	1.48*	1.45*
Utah.....	1978-79: medical 1980: $\geq$ 1 lost workdays or medical	.65	.74	.69*
Wyoming.....	1978: OSHA, 1979: medical 1980: $\geq$ 1 lost workdays or medical	1.40	1.71	1.86*
Nebraska.....	1978 OSHA, 1979: lost work, medical or first aid 1980: $\geq$ 1 lost workdays or medical	.81	1.47	1.53*
Missouri.....	1978-79: received during year 1980: permanent disability, $\geq$ 1 lost workdays or medical	.96	1.02	.96*
Virg. Islands..	1978-79: received during year 1980: $\geq$ 1 lost workdays, medical or first aid	1.27	1.21	1.12*
Washington.....	1978-79: received during year 1980: mixture of current & closed cases	1.73	1.84	N.A.
South Dakota...	1978-79: occurred during year	1.27	1.33	N.A.
Alaska.....	1978: lost workdays or medical	1.87*	N.A.	N.A.

*Years for which comparisons are possible.
Source of data: BLS, Annual Survey and Supplementary Data System, unpublished data.

A NEW INITIATIVE FOR THE CENTERS FOR DISEASE CONTROL (CDC): THE EPIDEMIOLOGIC STUDY OF VIOLENCE

Jack C. Smith, CDC

This is my third time to present a new CDC initiative to the Public Health Conference on Records and Statistics. In 1972, I talked to the conference about an important and politically sensitive public health issue that CDC was studying epidemiologically--induced abortion (1). In 1980, I talked to the conference about CDC's involvement in the study of the morbidity and mortality associated with one of the most frequently performed surgical procedures in the Nation--surgical sterilization (2). With assistance from health professionals around the Nation and from State health departments in particular, CDC has been able to conduct epidemiologic surveillance and clinical studies to assess the health consequences of both abortion and surgical sterilization.

Now I am here again to present yet another new CDC initiative--one whose success like the previous two will depend greatly on the availability of statistical information and assistance from State and local health departments. This new initiative is the epidemiologic study of violence.

As with any new initiative, it is reasonable to ask why the initiative was undertaken. As you are all well aware, violence-related mortality such as homicide and suicide and violence-related morbidity such as aggravated assault, attempted suicide, and abuse of children, spouses, and the elderly all present major problems for public health action.

The enormity of the problem can be illustrated by statistics on any one of the violence-related events mentioned above. I will use homicide as an example. When measured in terms of numbers of deaths, more than 20,000 homicide deaths occur annually (3). Homicide is one of the five leading causes of death for each 5-year age group between the ages of 1 and 40 (4). Homicide is the fourth leading cause of premature mortality for persons between the ages of 1 and 64, with only accidents, cancer, and heart disease ranking higher (5). And finally, homicide is an even greater problem for certain subgroups of the population. More than one of every three black males 15-24 years of age who die, die of homicide--more than the number who die from motor vehicle accidents and all other accidents combined (6).

Of course, not only are the health consequences of violence great, the whole issue of violence is complex. Thanks to the success in dealing with infectious diseases and to tremendous advances in medical treatment, many principle causes of morbidity and mortality of yesteryears have declined dramatically. In their place are the more complex degenerative diseases like heart disease and cancer and, of course, the "nondisease" killers and cripplers--accidents and violence. Like most modern health problems, violence-related morbidity and mortality are consequences of a multiplicity of intertwining factors such as environment, personal lifestyle, and human behavior. Thus, a multifaceted approach is required to deal with the problem--an approach with involvement from professionals in many disciplines.

Recently the Surgeon General of the United States in a speech on the subject of violence as a public health issue clearly stated that the epidemiologic expertise of CDC needed to be applied in the area of violence (7). CDC's Director, Dr. William H. Foege, has stated his concern that violence is not just a law enforcement issue nor just a psychiatric issue, but a broad public health issue. He also expressed his concern about the lack of basic epidemiologic information on the subject (8). In response to this need for basic epidemiologic information, CDC's Center for Health Promotion and Education (CHPE) has undertaken a new initiative for the epidemiologic study of violence.

I would like to discuss one particular aspect of our violence epidemiology work that is relevant for attendees of this meeting--namely, the identification and development of data sets for analysis.

As one example, I would like to talk about a new and valuable source of data that is just now being used for health research--data from the FBI's Uniform Crime Reporting (UCR) program. Through this program over 15,000 law enforcement agencies across the nation voluntarily contribute crime statistics. The primary purpose of the program is to generate a reliable set of criminal statistics for use in administering, operating, and managing law enforcement (9). The data emanating from the UCR program has been widely used by criminal justice professionals and other researchers who have an interest in crime.

Because of our study of homicide at CDC, we have had a special interest in a supplement to the UCR data, the Supplementary Homicide Report (SHR). We have acquired and analyzed this data set for the years 1975-1979. This data set has rarely, if ever, been used for epidemiologic purposes, so we are investigating the data set carefully in three ways:

1. File structure

   The file structure for most health data is a fixed-length record with only one observation per record as in the case of vital records. The SHR file structure, on the other hand, is a variable length record with a variable number of observations. Each record in the SHR is, in fact, a record of a crime event--homicide--containing information on the victim or victims of the crime, the perpetrator or perpetrators of the crime, and the circumstances surrounding the crime event.

   In order to process the file in a manner similar to processing homicide as a cause of death on death certificates, we had to restructure the file significantly from a

file of crime events to a file of homicide victims.

2. Definitions
   Criminal justice terminology and health terminology are substantially different. FBI speaks of murder and nonnegligent manslaughter as the crime of the willful killing of one human being by another (9). They, however, do not include in their compilation of crime statistics deaths that are justifiable such as the killing of felons by law enforcement officers in the line of duty or by private citizens. These justifiable killings would be classified as homicide under the International Classification of Diseases (ICD) and would be included in the compilation of health statistics. On the other hand, accidents such as hunting accidents, though not counted as a crime by the FBI are reported on the UCR file as negligent manslaughter.
   The classification of all crimes including murder and manslaughter in UCR are based solely on police investigation as opposed to determination by the court, physician, medical examiner, or coroner.
3. Completeness of reporting and quality of data
   The 15,000 law enforcement agencies that report through the UCR program have wide geographic distribution. Not all agencies report, however, and the rural areas of the country have lower participation than the urban areas (9). Participation can also be sporadic. For one recent year, no homicides were reported for New York City for a 6 month period of time (10). Little information is available to us to determine the completeness of reporting for those agencies that do participate in the UCR program. We are currently investigating the comparability of UCR and National Center for Health Statistics (NCHS) data on homicide. By being aware of definitional differences, we should be able to make reasonable statements about the degree to which the two data sets measure the same event--the incidence of homicide nationally. Furthermore, we have just begun to investigate the quality of data contained on the SHR file.

The value of the UCR's SHR data set lies primarily in the information which it contains on homicide that is not available from the death certificate such as age, race, and sex of both the offender and the victim, relationship of the victim to the offender (husband, boyfriend, stranger, etc.), and circumstances surrounding the crime (killed during a robbery, brawl due to influence of alcohol, etc.). Also, more details of the weapon used are available from the SHR data file than from vital statistics data. Since 1980 ethnicity of the victim (Hispanic, non-Hispanic) has also been included in SHR data.

One other important feature of data available from the UCR program is its timeliness. The 1982 SHR file will be ready for analysis this fall, several years ahead of national vital statistics data.

Though I will not elaborate on it, we are also beginning to work with crime data from local sources such as police departments and coroners' offices. By using local agency data, we are able to obtain information not in the UCR data set or available from vital statistics. An important example is toxicologic information, since substance abuse is often associated with violent behavior.

Now let me move from my discussion of a new data set to tell you of some interesting uses we are making of the traditional source of mortality data used by epidemiologists--the death certificate. (In each of the cases I will mention, we are working with the vital statistics offices of State health departments, not with national mortality data files from NCHS.)

1. Identification of persons at high risk of violence-related mortality.
   With the cooperation of State health departments in five Southwestern States (Arizona, California, Colorado, New Mexico, and Texas), we have used information on ethnicity from death certificates to analyze the suicide and homicide rates for Hispanics. Information on ethnicity was not available for this study from the national mortality data set.
2. Causal factors for violence-related mortality.
   With the cooperation of the Nebraska State Health Department, we are involved in a time series analysis of the relationship between occupation and violence-related mortality over an 8-year period. Information on occupation was not available for this study from the national mortality data set.
3. Surveillance of violence-related mortality.
   CDC is preparing for publication a series of surveillance reports on suicide and homicide. The first reports will provide background data for 1970-1978 for the general population nationally and for two specific subgroups at high risk of suicide and homicide--youths (for suicide) and blacks (for homicide). The Department of Health and Human Services (DHHS) has two stated objectives related to the reduction of suicide and homicide nationally:
   a. By 1990, the rate of suicide among people ages 15-24 should be below 11 per 100,000. (In 1978, the suicide rate for this age group was 12.4 per 100,000.)
   b. By 1990, the death rate from homicide among black males ages 15-24 should be reduced to below 60 per 100,000. (In 1978, the homicide rate for this group was 72.5 per 100,000.)

   CDC is responsible for helping to monitor DHHS progress toward meeting these objectives, but surveillance of suicide and homicide using the national mortality data file will not be timely enough. Thus CDC has begun exploring States' interest and willingness to participate in suicide and homicide surveillance for two purposes:
   a. To produce special State-level surveillance reports that would deal uniquely

with suicide and homicide as a public health problem in individual States. These reports could provide data on a smaller geographic area with more detail and with more timeliness than would be the case with national level reports.

b. To produce for their State on an ongoing basis an annual tabulation of suicide and homicide by age, race, and sex. This tabulation could be included in the State's annual vital statistics report and available to CDC for compilation into a national surveillance data set for monitoring suicide and homicide rates.

In conclusion, let me reiterate that the compilation and analysis of statistical information as a part of CDC's new violence initiative is providing many interesting opportunities and challenges. Hopefully the efforts of the interdisciplinary team of statisticians, social scientists, physicians, health educators, and other health professionals tapped by CDC to participate in this new initiative will result in better knowledge and understanding of violence as a public health problem. The ultimate goal of our epidemiologic studies is to aid in the development of appropriate strategies for reducing violence-related morbidity and mortality.

REFERENCES

1. Smith JC. Specific informational needs of the health service programs--a progress report on the abortion surveillance program of the Centers for Disease Control. In: Proceedings of the Public Health Conference on Records and Statistics, 14th National Meeting, June 3, 1972. Maryland: National Center for Health Statistics, 1972.
2. Smith JC. Surveillance of surgical sterilization: Centers for Disease Control. In: Proceedings of the Public Health Conference on Records and Statistics, 18th National Meeting, December 1980. Maryland: National Center for Health Statistics, 1980; pp. 221-225.
3. National Center for Health Statistics. Monthly Vital Statistics Report DHHS Pub. No. (PHS) 81-1120, Annual Summary for the United States, 1979; 28(13):255. Washington, Government Printing Office, 1980.
4. National Center for Health Statistics. Vital Statistics of the United States, 1978. Mortality. Part B Vol. II, DHHS Pub. No. (PHS) 82-1102, Washington, Government Printing Office, 1982.
5. Centers for Disease Control. Fifteen leading causes of death for metropolitan and nonmetropolitan populations, United States, 1978, Section II, Potential Years of Life Lost-1978; pp. 24, Atlanta, Centers for Disease Control, 1982.
6. Special tabulations produced by the Centers for Disease Control for 1978 from the national mortality data file of the National Center for Health Statistics.
7. Family violence: A chronic public health issue, presented by C. Everett Koop, M.D., Surgeon General, USPHS, November 9, 1982 to the Western Psychiatric Institute, Pittsburgh, Pennsylvania.
8. Atlanta Constitution, August 27, 1981. "What bugs a germ chaser?"
9. Federal Bureau of Investigation: Uniform Crime Reports for the United States, 1979; p. 3, Washington, 1979.
10. Personal communication with FBI, Uniform Crime Reports Section.

Second Plenary

THE BURDEN OF ILLNESS: PAST TRENDS AND FUTURE PROJECTIONS

Dorothy P. Rice, University of California, San Francisco

The burden of illness is a major concern for all societies. Measurement of that burden takes on special significance in the allocation of these resources. The setting of priorities for the allocation of these resources and the evaluation of health service programs should be improved by such information. Planning for the future in terms of the Nation's economic, social and health institutions and services require a good data base, sensible projections of past trends, and alternative assumptions for projecting the future.

The multitude of health problems facing our Nation today varies greatly in the amount, kind, and cost of medical care used and in their indirect effects on society in terms of pain and suffering, and losses in productivity due to illness, disability, and death.

When applied to illness, the burden includes the heavy load borne by society in providing services to prevent, cure, and care for the sick. It also includes the substantial losses of output to the economy due to disease, disability, and death. The burden on the family in caring for and accommodating a sick member of the household can also be severe. Finally, there is the burden of pain, discomfort, and suffering of each sick person and that of anguish and grief of relatives and friends.

The first two indicators of burden -- the provision of direct services and the losses in output -- are more easily measureable. The dimensions of the burden on the family and that of the individual's suffering are not so easily quantifiable (Rice, Feldman, and White, 1976).

The health care industry in the United States encompasses a wide variety of institutions, organizations, and personnel that provide the full gamut of preventive, medical, therapeutic, restorative, and related services. It is a pleuralistic health care system: the organization, delivery, and financing is complex, comprising a mixed medical care economy, with an interdependence of the private and governmental sectors.

It is one of the largest industries in the United States today. About 7.5 million persons -- about 7.5 percent of the employed population -- are health workers (National Center for Health Statistics, 1982). There are more than 450,000 physicians, 1.7 million registered nurses, and 600,000 licensed practical nurses actively providing medical care services, to name a few of the medical profession practitioners. There are about 7,000 hospitals with 1.4 million beds that are mainly nonprofit and 23,000 nursing homes with 1.5 million beds that are for-profit.

The health status of the population has improved considerably in present years. In fact, mortality has declined dramatically since 1950, from 29.2 to 11.2 per 1,000 live births in 1982 -- a 62 percent reduction. Age adjusted death rates have also declined substantially -- a 29 percent reduction in the 30-year period from 1950 to 1980, with a 38 percent reduction in death rates from cardiovascular diseases. These encouraging trends in health status have resulted in improved life expectancy throughout the age range -- a 5 1/2 year gain since 1950 in life expectancy at birth and 2 1/2 years at 65 years of age.

Despite these improvements in health, the burden of illness is still great. There are many persons who suffer from disabling conditions resulting in illness, disability, and limitations of activity. There are 2.2 acute conditions per person per year resulting in 9.4 restricted activity days, 4.1 bed-disability days, 4.8 school loss days, and 3.5 work-loss days. Chronic conditions, especially for the elderly, result in a significant burden on society as well as the individual and his or her family: 44 percent of the non-institutional elderly have arthritis, 27 percent have heart conditions, 38 percent have hypertension (National Center for Health Statistics, 1981).

The volume of use of institutional and medical care reflects the demands for and burdens created in the use of medical care resources: one in 6 persons is discharged from a short-stay hospital in a year but 2 in 5 elderly persons are hospitalized. Among the elderly aged 65 and over, 5 percent are in nursing homes, but 22 percent of those aged 85 and over are in nursing homes.

This paper will examine the burden of illness across several dimensions: (1) direct expenditures for medical care services by source of funds, type of expenditures, by age, sex, and medical condition; (2) indirect costs -- the value of losses to society because of illness, disability, and death; and (3) the future burden in terms of the aging of our society and the impact of these demographic changes on the use of and spending for health services.

Expenditures for Medical Care: Trends

The continued and persistent rise in medical care spending during recent years has generated much interest, attention, and concern as it has consumed a growing share of the Nation's Gross National Product (GNP). Outlays for medical care amounted to $322.4 billion in 1982 and comprised 10.5 percent of the GNP; in 1950, expenditures amounted to 12.7 billion and represented 4.0 of the GNP (Table 1). During this 32-year period, medical care spending increased 25 times while its share of GNP rose 2 1/2 times (Gibson and Waldo, 1983). Health spending in 1982 amounted to $1,365 for every man, woman and child in America rising from $82 per capita in 1950.

What are the factors behind the continuing increase in medical care spending and the growing burden on society? Several broad factors can be identified: population growth, inflation, procuct change, and the increase in utilization of the quantity of care demanded and supplied. This simple list, however, does not capture the enormous changes that have taken place in the organization, delivery, and funding of medical care services. Additional medical and demographic changes also affecting the increased demand for medical care services and the rise in spending include growth in private health insurance and prepayment plans, increased public support of medical care for the aged,

disabled, and poor, increasing population and a rising proportion of elderly, a shift from care of acute to more expensive long-term illnesses, improvements and growth of high-cost technology, and higher wages and salary costs in the health care industry. The growing burden on the economy of medical care spending results from all of the above factors as well as higher medical care prices relative to general funds and a slow down in the general economy with continued growth in the health sector. Thus, in the 5-year period, 1978-82, the GNP rose 41 percent compared with a 70 percent increase in medical spending.

Rising prices for medical care services have been the primary force in the rise of personal health care spending in recent years. Between 1950 and 1965, price inflation accounted for 41 percent of the increase: between 1965 and 1982 it accounted for 61 percent of the increase. The effect of population growth has diminished substantially from the first period to the second. Increases in intensity of use and quality changes accounted for a slightly larger share of the increase in the first period as summarized below:

	1950-65	1965-81
Total Increase (in billions)	$24.9	$219.2
Sources of Increase (percent)	100	100
Price Rise	41	61
Population Growth	20	08
Quality Improvement and Utilization Increase	39	31

Sources of Medical Care Funds

The major portion of health care expenses has in the past been borne by the private sector. From 1950 to 1966, private outlays constituted almost three-fourths of the total. Private health insurance coverage grew rapidly in the fifties. In 1950, only 9 percent of personal health care spending was paid by private health insurance; by 1960, this proportion rose to 21 percent. Direct payments declined from 66 percent to 55 percent of the total. Within the public sector, state and local governments spent more than the federal government Implementation of Medicare and Medicaid, together with increasing coverage of private health insurance, altered these relationships. By 1982, the government's portion rose to 40 percent of the total, and federal spending equaled more than 2 1/2 times that of state and local governments. Private health insurance and philanthropy covered 28 percent further reducing direct private payments to less than one-third of the personal health care bill compared with two-thirds in 1950.

The rise in third party payments tends to reduce the financial burden of serious illness and patient concern about the cost of care received, and removes the restraining influences from the physician to hospitalize patients and use high-cost technologies.

In addition to the implementation of public programs, the federal government also has influenced the health care field through tax subsidies and incentives by treatment of employers' contributions to health insurance plans as a nontaxable form of employee income. This preferential tax treatment represents a tax loss to the U. S. Treasury, estimated at $18 billion in 1980. Recognition of the lack of cost-consciousness in the choice of health care plans by employers and employees has prompted the Reagan Administration to propose that annual employer contributions above $2100 a family and $840 for individuals be taxed.

Type of Expenditures

Hospital care accounted for the largest portion of health expenditures, accounting to $136 billion, 42 percent of total outlays in 1982 (Table 2). Hospital costs have been increasing rapidly due to a number of factors: advances in medical and hospital technology that are often costly; growth in the range of services available such as intensive care units; elaborate and costly equipment with rapid obsolescence rate; rise in labor costs, both in wages and in the number of employees per patient day; and, finally, reimbursement at cost under Medicare and Medicaid lacked incentives for efficiency and economy.

Health program managers and federal policy makers have become increasingly concerned about curbing inflation in hospital and medical care costs. Numerous controls have evolved, many of them in the last decade. Among these are supply limitations (certificate-of-need for capital construction), financial disincentives to the patient (insurance deductibles and coinsurance rates), authorization requirements (authorization for surgery pending second opinion), review programs (utilization and claims processing reviews), legal action (malpractice suits), regulation of rates (state hospital rate setting commissions), and prospective budgeting (payment per admission). The Reagan Administration designed a new Medicare payment system, effective October 1, 1983, establishing separate rates for hospital admissions with different diagnoses (diagnosis-related groups -- DRGs).

Professional services -- physicians, dentists, and other professional personnel -- totaled $87.6 billion in 1982 and accounted for 27 percent of expenditures. Almost half of these costs were paid by individuals out-of-pocket. Private insurance paid 32 percent and public programs 23 percent.

Nursing home care accounted for 8.5 percent of expenditures. Public programs, mainly Medicaid, paid 55 percent of these expenditures and out-of-pocket payments comprised 44 percent; private health insurance covered virtually none of these costs.

Drugs and appliances accounted for 9 percent of expenditures. Most of these costs -- 78 percent -- were paid by individuals out-of-pocket. Other health expenditures include the remainder of personal health expenditures (2 percent), expenditures for prepayment and administration (4 percent), government public health activities (3 percent), and research and construction (4 percent).

Public Health Care Programs

During the last few decades, medical care has come to be recognized as a basic right, along with food, clothing, and shelter. The government has played a significant role in the attainment of this right. The role of government in financing medical care was greatly enhanced by the enactment in 1965 of Medicare (health insurance for the aged and disabled) and Medicaid (health care coverage for low-income persons). Medicare and Medicaid are the two largest public programs supporting health care services and supplies with total expenditures of $50.9 billion and $32.4 billion, respectively in 1982. Since the inception of these programs, the gap in medical care use between low-income and other persons has narrowed.

In 1980, with 27 million elderly persons enrolled in the Medicare program, 39 percent did not receive care paid by Medicare. Six percent received $5,000 or more each in benefits; these persons accounted for 61 percent of Medicare reimbursement. Eight percent received $2,000 to $4,999 each, accounting for 23 percent of reimbursements. Thus, 14 percent of the elderly received 84 percent of the amounts reimbursed under the Medicare program.

Of the 22 million Medicaid recipients in 1979, 16 percent were aged 65 and over; they accounted for 37 percent of the expenditures under the program. Twelve percent of the recipients were disabled and accounted for 31 percent of the expenditures. Sixty-three percent of the recipients obtained benefits through Aid for Dependent Children; accounting for 28 percent of the Medicaid expenditures.

Expenditures by Age and Sex

Although only 11 percent of the population was age 65 and over in 1980, this group accounted for 31 percent of all personal health care expenditures. The aged population are generally sicker and use more medical care services than younger persons. Women use more medical services and incur disproportionately higher expenditures relative to their numbers than males (Hodgson and Kopstein, 1983). Females represented 52 percent of the population but accounted for 58 percent of the expenditures (Table 3). The same pattern exists for the younger as well as the older population. Women under age 65 comprise 51 percent of the population, but 56 percent of the health expenditures are in behalf of women in that age group. Among those aged 65 and over, 59 percent are women and incur 63 percent of the expenditures.

Expenditures by Medical Condition

Health care expenditures by medical condition and type of care for the ten leading disease classes are shown in Table 4. The economic burden varies by disease category. Diseases of the circulatory system were responsible for the highest amount (15.1 percent), followed closely by diseases of the digestive system (including dental care -- 14.5 percent). A second group, with expenditures of about the same magnitude, are mental disorders (9.3 percent), injury and poisoning (8.8 percent), and diseases of the respiratory system (7.9 percent). Neoplasms, diseases of the musculoskeletal system and connective tissue, diseases of the genitourinary system, and diseases of the nervous system and sense organs all required health care services of about the same value (between $12 and $14 billion and 6 percent each of the total).

The relative amounts spent for each type of health care varied among the disease categories and reflects the differences in pathology and consequent needs for each type of service among diseases. Hospital care usually accounted for the largest share of expenditures for a disease class, ranging from 35 percent of expenditures for diseases of the nervous system and sense organs to 67 percent for neoplasms. Professional services were usually the second leading cause of expenditures in a disease category, but ranged from 10 percent of spending for mental disorders to 61 percent for digestive system diseases. It should be noted, however, that out of $19 billion for professional services for diseases of the digestive system, $15 billion were for dental services. Nursing home expenditures were high for circulatory diseases (24 percent), mental disorders (22 percent); and endocrine, nutritional metabolic diseases (15 percent), but contributed small amounts to other diseases. Drugs were most important in terms of expenditures for diseases of the nervous system (18 percent), respiratory system (17 percent), and endocrine, nutritional, and metabolic diseases (13 percent).

Indirect Costs of Illness

In addition to the burden of illness of the use of medical care resources expressed in terms of direct expenditures, the value of losses in output due to disability (morbidity) or premature death (mortality) must be considered (Rice, 1966 and Cooper and Rice, 1976). Morbidity losses are incurred when illness results in absence from employment, prevents the performance of household duties, or results in disability that keeps someone from working. The lost earnings and the dollar value of the unperformed housekeeping services are the morbidity costs. Calculation of morbidity costs involves applying current average earnings by age and sex to work-loss years for those in the labor force, attaching a dollar value to household services and applying it to related bed-days, and applying labor-force participation or work- experience rates and average earnings, by age and sex, to persons in and out of institutions who are too sick to be employed or keep house.

The latest data available on the indirect costs of illness by disease are for 1977 (Rice, 1981). When the annual morbidity losses are combined with the lifetime mortality losses, a total of 38 million person years are estimated to be lost to productivity in 1977, amounting to total indirect costs of $176 billion based on a 6 percent discount rate (Table 5). On the basis of the higher discount rate of 10 percent, the total indirect costs amount to $140 billion. Person years lost for males exceed those for females by 18 percent. Owing to the higher lifetime earnings for men, their indirect costs are estimated to be 86 to 88 percent higher than

for females, depending on the discount rate.

The greatest losses are for those with heart disease, which account for 25 percent of the person years lost and 20 to 21 percent of the indirect costs of all illnesses. Person years lost for males are 26 percent higher than those for females; indirect costs are 86 to 94 percent higher for males, depending o the discount rate. Neoplasms account for 17 percent of the total person years lost and about the same proportion -- 15 percent -- of the total costs of morbidity and mortality. Person years lost for women exceed those for men, but the indirect costs for men are 8 to 13 percent higher. This occurs because women have longer life expectancy and men's earnings exceed those of women at each age. There has been a dramatic downturn in the death rate for heart disease in recent years, attributed to significant decreases in cigarette smoking, the decline in consumption of saturated fat and cholesterol, improvement in life styles, and control of hypertension. The incidence of heart disease is predominantly a male disease, with higher morbidity and mortality among males than females. Stroke, on the other hand, results in a larger number of deaths among women; a total of 910,000 person years are estimated to be lost for men, representing 69 percent of the person years lost for women. Indirect costs for men amount to $3.1 billion, or 87 percent of the costs for women at a 6 percent discount rate, and 90 percent at a 10 percent rate. Significant reductions in mortality from stroke in the United States have also occurred in recent years due in large part to improved treatment of hypertension.

The Burden of Changing Demographics

The final approach to measuring the burden of illness is in terms of the aging of our society and the impact of these demographic changes on the use of and spending for health services. The changing age structure of the American population, with its growing number of elderly, has profound consequences for the nation's economic, social, and health institutions and services. Since 1960 the population aged 65 and over has grown more than twice as fast as the younger population. The elderly increased from 16.7 million 1960 to 25.9 million in 1980 -- a 55 percent increase; for the population under age 65, the increase was only 24 percent. The elderly have also increased as a proportion of the population, from 9.1 percent in 1960 to 11.1 percent twenty years later. The number of the very elderly is growing even more rapidly. In the same time span, those aged 75 to 84 rose 65 percent while the 85 years and over group rose 174 percent.

Declining death rates from heart disease, cerebrovascular disease, influenza, and other causes of death contributed to the growth in the elderly population. The Social Security Administration (SSA) recognized the downturn by building into their population projections the assumption that mortality among the elderly would continue until the year 2005 to decline at a relatively rapid rate and level off thereafter (Faber and Wilkin, 1981). We applied current age-specific rates of activity limitation and utilization patterns to the SSA population projections (Rice and Feldman, 1983) to assess the impact of aging on the health care system.

Table 6 shows the population, limitations in activities of daily living (ADL), medical care utilization, and expenditures in 1980 and projections to 2040. The difference between the rates of growth in the population and in the various indices of burden is a reflection of the aging of the population. In 1980, 3.1 million noninstitutionalized persons were reported needing assistance in ADL; by 2040 the number is projected to be more than double, while the population is projected to increase during that period by two-fifths. The number of physician visits will increase in the future due to the aging of the population, but the increase will be less than for other measures of utilization because age-specific utilization rates do not vary as much for physican visits as, for example, for hospital care. Only 6 percent of the increase in visits from 1.1 billion in 1980 to 1.6 billion in 2040, an increase of 47 percent, results from the aging of the population.

The aging effect is quite different for hospital and nursing home care. Total short-stay hospital days will double, increasing from 274 million in 1980 to 549 million in 2040, with more than half the increase due to the aging of the population. Forty percent of the days of care in 2040 are projected for those aged 75 and over; in 1980 only 20 percent were in that age group.

Again, assuming that current patterns of use prevail in the future, there will be very large increases in the number of nursing home residents. From 1.5 million in 1980, the number is projected to 5.2 million residents in 2040 -- a 3.5-fold increase. The increases are particularly large among residents 85 years of age and older where a 5-fold increase is projected in the number of residents. In 1980 37 percent of the residents were aged 85 and over; by 2040 the proportion will be 56 percent. Adding the projected nursing residents aged 75 to 84, about 87 percent of the total residents will be aged 75 and over. It is evident that the aging of the population has a much greater impact on nursing home residents than on days of hospital care or physician visits.

We made no attempt to forecast future inflation rates; thus, the expenditures are in constant 1980 dollars. As with the use of medical care services, the proportional increase in expenditures is projected to rise at a significantly faster rate at the older ages. Of the total $219 billion spent in 1980 for personal health care, $64.5 billion or 29 percent was spent in behalf of the elderly population aged 65 and over. This amount would rise to $167.5 billion in 2040 -- an increase of 159 percent, due to the aging of the population during that 60-year period. By contrast, for the population under age 65, expenditures are projected to increase by 30 percent. In 1980, 11 percent of the population who are aged 65 and over consumed 29 percent of the expenditures; by

2040 the elderly are projected to comprise 21 percent of the population and almost half of the expenditures would be made in their behalf.

CONCLUSION

Regardless of the measure, the burden of illness is large and multifaceted. The various dimensions of burden presented in this paper, have significant implications for investments in efforts to solve the health, disease, and medical care problems that beset our society. They are legitimate considerations essential for planning to meet the future demand for medical care services and for priorities for biomedical and health services research.

Note: The author presented her paper at the Public Health Conference using slides which could not be reproduced in the Proceedings. The tables show most of the data presented in the slides.

The author very much appreciates the assistance of the following individuals in the National Center for Health Statistics in the preparation of: (a) the data for this paper: Jacob J. Feldman, Thomas A. Hodgson, Andrea N. Kopstein; and (b) graphics used in the presentation: Naomi Forester and Steve Sloan.

REFERENCES

Cooper, B. S. and D. P. Rice. The Economic Cost of Illness Revisited, "Social Security Bulletin, February 1976.

Faber, F. and C. Wilkin. Social Security Area Population Projections, 1981. Actuarial Study No. 85. Social Security Administration. SSA Pub. No. 11-11532, July 1981.

Gibson, M. and D. R. Waldo. "National Health Expenditures, 1982." Health Care Financing Review. Vol. 5 No. 1, Fall, 1983.

Hodgson, T. A. and A. N. Kopstein. "Health Expenditures for Major Diseases in 1980." National Center for Health Statistics. Unpublished paper, 1983.

National Center for Health Statistics. "Current Estimates from the National Health Interview Survey, United States, 1979." Vital and Health Statistics. Series 10. No. 136. DHHS Pub. No. (PHS) 81-1564. April 1981.

National Center for Health Statistics. Health United States, 1982. DHHS Publication. No. (PHS) 83-32. Public Health Service. Washington DC: U.S. Government Printing Office. December 1982.

Rice, D. P. "Estimating the Cost of Illness," Health Economics Series, No. 6, U.S. Department of health, Education, and Welfare, Public Health Service, Pub. No. 947-6, May 1966.

Rice, D. P. Sex Differences in Mortality and Morbidity.: "Some Aspects of the Economic Burden" in The Sex Differential in Mortality: Trends, Determinants and Consequences. Selection of the Papers presented at the ANU/UN/WHO Meeting. Canberra, Australia. 1-7 December 1981.

Rice, D. P. and J. J. Feldman. "Living Longer in the United States: Demographic Changes and the Health Needs of the Elderly." Millbank Fund Memorial Quarterly: Health and Society. Vol. 61 No. 3 Summer, 1983.

Rice, D. P., J. J. Feldman, and K. L. White. "Current Burden of Illness in the United States." An Occasional Paper. Institute of Medicine, National Academy of Sciences, Washington, D.C. 1976.

Table 1: Gross National Product and National Health Expenditures, by Source of Funds, 1950 - 1982

	Gross National Product (billions)	National Health Expenditures						
		Total			Private Funds		Public Funds	
		Amount (billions)	Per Capita	Percent of GNP	Amount (billions)	Percent of Total	Amount (billions)	Percent of GNP
1950	$ 286.5	$ 12.7	$ 82	4.4	9.2	72.8	3.4	27.2
1955	400.0	17.7	105	4.4	13.2	74.3	4.6	25.7
1960	506.5	26.9	146	5.3	20.3	75.3	6.6	24.7
1965	691.0	41.7	211	6.0	31.0	74.1	10.8	25.9
1970	992.7	74.7	358	7.5	46.9	62.8	27.8	37.2
1975	1,549.2	132.7	604	8.6	76.5	57.7	56.2	42.3
1980	2,633.1	249.0	1,075	9.5	143.6	57.7	105.4	42.3
1981	2,937.1	286.6	1,225	9.8	164.4	57.4	122.2	42.6
1982	3,059.3	322.4	1,365	10.5	185.6	57.6	136.8	42.4

SOURCE: Health Care Financing Review, Vol. 5, No. 1, Fall 1983.

Table 2: National Health Expenditures by Type of Expenditure and Source of Funds, 1982
(Billions of dollars)

Type of Expenditure	Total	Private				Public		
		Total	Direct Payment	Insurance	Other*	Total	Federal	State and Local
TOTAL	$322.4	$185.6	$90.4	$84.2	$10.9	$136.8	$93.2	$43.7
Health services and supplies	308.3	179.5	90.4	84.2	4.8	128.7	87.5	41.2
Personal health care	286.9	171.2	90.4	76.6	4.2	115.7	83.7	32.0
Hospital care	135.5	63.5	16.4	44.9	2.2	72.0	54.6	17.4
Physicians' services	61.8	44.8	23.1	21.7	**	17.0	13.4	3.6
Dentists' services	19.5	18.7	13.4	5.2	-	.8	.4	.4
Other professional services	7.1	4.9	3.6	1.3	.1	2.2	1.7	.5
Drugs and medical sundries	22.4	20.4	17.6	2.8	-	1.9	.9	1.0
Eyeglasses and appliances	5.7	4.8	4.4	.4	-	.8	.7	.1
Nursing home care	27.3	12.3	11.9	.2	.2	15.0	7.9	7.1
Other health services	7.6	1.7	-	-	1.7	5.9	4.0	1.9
Program administration and net cost of insurance	12.7	8.3	-	7.7	.6	4.4	2.4	2.0
Government public health activities	8.6	-	-	-	-	8.6	1.4	7.3
Research and Construction of medical facilities	14.1	6.0	-	-	6.0	8.1	5.7	2.4
Research	5.9	.3	-	-	.3	5.6	5.0	.5
Construction	8.2	5.7	-	-	5.7	2.5	.7	1.9

SOURCE: Health Care Financing Review, Vol. 5, No. 1, Fall 1983.

*Spending by philanthropic organizations, industrial in-plant services, and construction privately financed.
**Less than $100 million.

Table 3: Personal health care expenditures according to medical condition, sex, and age: United States, 1980

Medical condition	Both sexes			Males			Females		
	All ages	Under 65 years	65 years and over	All ages	Under 65 years	65 years and over	All ages	Under 65 years	65 years and over
	Amount in millions								
All conditions	$219,400[a]	$146,191	$64,950	$87,618	$63,300	$24,318	$123,525	$82,891	$40,634
Infectious and parasitic diseases	4,498	3,928	569	1,820	1,622	198	2,677	2,306	371
Neoplasms	13,623	8,302	5,322	5,647	3,043	2,603	7,977	5,259	2,718
Endocrine, nutritional, metabolic diseases, and immunity disorders	7,656	4,689	2,968	2,354	1,483	872	5,302	3,206	2,096
Diseases of blood and blood-forming organs	1,205	753	449	427	280	147	775	473	302
Mental disorders	20,301	14,612	5,689	9,330	7,395	1,935	10,971	7,217	3,753
Diseases of the nervous system and sense organs	17,499	13,028	4,471	7,558	5,762	1,796	9,941	7,266	2,675
Diseases of the circulatory system	33,184	13,078	20,105	13,932	6,739	7,194	19,251	6,340	12,911
Diseases of the respiratory system	17,305	13,164	4,141	8,096	6,048	2,048	9,209	7,117	2,093
Diseases of the digestive system	31,755	26,084	5,671	13,428	11,159	2,269	18,327	14,925	3,402
Diseases of the genitourinary system	13,162	10,721	2,441	3,509	2,138	1,372	9,652	8,583	1,069
Diseases of the skin and subcutaneous tissue	6,179	5,036	1,144	2,573	2,147	425	3,607	2,888	718
Diseases of the musculoskeletal system and connective tissue	13,645	9,821	3,824	5,053	4,078	975	8,592	5,744	2,849
Congenital anomalies	1,409	1,321	88	626	590	36	783	731	52
Symptoms, signs, and ill-defined conditions	3,968	3,197	771	1,654	1,310	344	2,314	1,887	427
Injury and poisoning	19,248	15,042	4,206	9,783	8,662	1,121	9,465	6,380	3,085
Other conditions	2,246	2,187	59	331	297	34	1,915	1,890	25
Unallocated expenditures	12,359[a]	1,229	3,036	1,496	548	948	2,769	681	2,088

a. Includes expenditures that could not be distributed by age and sex.

NOTE: Numbers may not add to totals due to rounding.

SOURCE: Computed by the Division of Analysis, National Center for Health Statistics from data compiled by the Health Care Financing Administration, the National Center for Health Statistics, and other organizations.

Table 4: Personal health care expenditures according to medical condition and type of care: United States, 1980

Medical condition	Type of Care					
	All personal health care	Hospital care	Physicians' services	Nursing home care	Drugs	Other professional services
	Amount in millions					
All conditions	$219,400	$100,461	$46,790	$20,593	$19,300	$21,000
Infectious and parasitic diseases	4,498	2,138	1,531	---	805	24
Neoplasms	13,623	9,130	3,163	469	677	184
Endocrine, nutritional, metabolic diseases, and immunity disorders	7,656	3,369	2,055	1,137	1,020	75
Diseases of blood and blood-forming organs	1,205	749	309	---	130	17
Mental disorders	20,301	12,836	2,027	4,363	1,001	74
Diseases of the nervous system and sense organs	17,499[a]	4,409	4,557	387	2,175	870
Diseases of the circulatory system	33,184	16,682	5,813	7,983	2,272	434
Diseases of the respiratory system	17,305	8,438	5,313	435	2,978	141
Diseases of the digestive system	31,755	11,440	3,757	---	968	15,590[b]
Diseases of the genitourinary system	13,162	7,471	4,198	---	1,391	102
Diseases of the skin and subcutaneous tissue	6,179	1,607	3,051	---	1,493	28
Diseases of the musculoskeletal system and connective tissue	13,645	6,220	3,879	833	1,529	1,184
Congenital anomalies	1,409	918	387	---	97	7
Symptoms, signs, and ill-defined conditions[c]	3,968	1,695	1,470	---	789	14
Injury and poisoning	19,248	11,599	4,905	721	1,868	155
Other conditions	2,246	1,760	373	---	106	7
Unallocated expenditures	12,359[d]	---	---	4,265	---	2,094

a. Includes all expenditures for eyeglasses any appliances.
b. Includes all expenditures for dentists' services.
c. Includes complications of pregnancy, childbirth, and the puerperium, and certain conditions originating in the perinatal period.
d. Includes expenditures for other health services, other professional services, and nursing home care that could not be allocated to conditions.

NOTE: Numbers may not add to totals due to rounding.

SOURCE: Computed by the Division of Analysis, National Center for Health Statistics from data compiled by the Health Care Financing Administration, the National Center for Health Statistics, and other organizations.

Table 5: Estimated Person Years Lost and Total Indirect Costs, by Sex and Diagnosis; United States, 1977

Diagnosis	Both Sexes	Male	Female	Sex Ratio
TOTAL PERSON YEARS LOST (thousands)				
Total	38,235	20,665	17,570	1.18
All neoplasms	6,503	3,122	3,381	.92
Malignant neoplasms	6,341	3.068	3,273	.94
Heart disease	9,295	5,199	4,096	1.27
Stroke	2,222	910	1,312	.69
Accidents	4,158	2,885	1,273	2.27
All other	16,056	8,549	7,508	1.14
TOTAL INDIRECT COSTS* (millions)				
Six percent discount				
Total	$175,980	$114,546	$61,434	1.86
All neoplasms	26,272	13,612	12,660	1.08
Malignant neoplasms	25,369	13,270	12,099	1.10
Heart disease	33,560	21,693	11,867	1.83
Stroke	6,642	3,082	3,560	.87
Accidents	26,473	21,038	5,435	3.87
All other	83,034	55,122	27,912	1.97
Ten percent discount				
Total	$140,313	$ 91,632	$48,681	1.88
All neoplasms	20,517	10,880	9,637	1.13
Malignant neoplasms	19,730	10,596	9,134	1.16
Heart disease	27,660	18,105	9,555	1.89
Stroke	5,420	2,573	2,847	.90
Accidents	18,813	14,885	3,928	3.79
All other	67,904	45,190	22,714	1.99

SOURCE: National Center for Health Statistics, Division of Analysis, Health Economics Analysis Branch.

*Includes morbidity and mortality costs.

Table 6: Current and Projected Population, Limitations in ADL, Medical Care Utilization and Expenditures, by Age, 1980 and 2040

Characteristic and Year	All Ages	Under 65	Age 65 and Over		
			Total	65-74	75 and Over
Population (thousands):					
1980	232,669	206,777	25,892	15,627	10,265
2040	328,503	261,247	67,256	29,425	37,831
Persons with Limitation in Activities of Daily Living (thousands):					
1980	3,142	1,362	1,780	648	1,132
2040	7,922	2,002	5,920	1,288	4,632
Physician Visits (millions):					
1980	1,102	936	166	100	66
2040	1,621	1,193	428	187	241
Days of Hospital Care (millions):					
1980	274	169	105	49	56
2040	549	236	312	93	219
Nursing Home Residents (thousands):					
1980	1,511	196	1,315	227	1,088
2040	5,227	248	4,979	425	4,554
Personal Health Expenditures (in constant 1980 billions of dollars):					
1980	$219.4	$154.9	$ 64.5	n.a.	n.a.
2040	369.0	201.5	167.5	n.a.	n.a.

SOURCE: Rice and Feldman, "Living Longer in the United States, Demographic Changes and Health Needs of the Elderly." Milbank Memorial Fund Quarterly, Vol. 6, No. 3, Summer, 1983.

Cost of Cancer Care

Session H

THE COST OF CANCER AND THE MEDICARE HOSPICE BENEFIT

James Lubitz, Health Care Financing Administration

Introduction

The Tax Equity and Fiscal Responsibility Act passed in September 1982 made hospice care a covered Medicare benefit beginning in November 1983. To be eligible for hospice care a physician must certify that a person has less than six months to live. The enrollee may receive Medicare reimbursement for care related to the terminal condition only through the hospice chosen. Among the services not normally covered by Medicare that are covered under the hospice benefit are:

- Homemaker services
- Inpatient hospice care
- Outpatient drugs
- Respite care
- 24 hour nursing care

Data that the Office of Research and Demonstrations of the Health Care Financing Administration (HCFA) developed on the costs of dying for Medicare patients have played an important part in the debate leading to the hospice benefit and will play an important role in the congressionally mandated evaluation of the hospice benefit. In addition, in the course of our efforts to improve data on the costs of dying, HCFA is working with the National Center for Health Statistics (NCHS) and the Association for Vital Records and Health Statistics (AVRHS) in developing a new approach to the acquisition of mortality statistics for large-scale research projects.

The Passage of the Hospice Benefit

Part of the argument in favor of coverage of hospices was not only that hospice care is more appropriate to the needs of dying patients and their families but that hospice care would cost the Medicare program less than care provided in conventional settings. The belief was that Medicare pays for expensive and technically sophisticated efforts to prolong life. It was also felt that conventional care to the terminally ill tends to occur in institutions, while hospice care tends to occur in the patient's home. Hospice ideally could substitute for some conventional care by providing a humane, home-based, and less-costly alternative for terminally ill patients. Thus, data on the costs of conventional care for the dying were key to the assertion that hospice care is indeed less expensive than care in other settings.

The Congressional Budget Office (CBO) used our estimates of the Medicare costs of cancer in their influential report of June 1982 to the House Ways and Means Committee. The cost of cancer patients before death rather than the costs of all patients were used for analysis because the vast majority of hospice patients have cancer. In HCFA's hospice demonstrations 93 percent of the patients had cancer. Data from hospices showed an average stay of about 45 days.

The CBO report took our estimate of Medicare reimbursements in the last 45 days and inflated it to yield $7,080 in 1983 dollars. The CBO report reasoned that if hospice substitute for conventional care in the last 45 days, the portion of the $7,080 due to fixed hospital costs would be reallocated among other payors, including Medicare. Taking this into account, the CBO estimated the gross savings to Medicare would be $6,130. They estimated hospice costs at $5,010. Based on this comparison they found that the net savings to Medicare would be $1,120 per user of hospice care. The hospice benefit was passed in September 1982.

Other data on the experience of Medicare enrollees hospitalized for cancer were also influential in the passage of the hospice benefit. Table 1 shows the distribution of Medicare expenses in the last year of life for cancer patients by time before death. The table shows the increase in the use of Medicare services as death approaches. Of total Medicare expenses in the last 12 months of life, 28 percent occur in the last month and another 17 percent occur in the second-to-last month. Since the average stay in hospices was around 45 days, these data show that there is a large amount of costs that hospice could potentially impact.

Table 1. Percent of Medicare expenses in the last year of life by selected intervals before death for persons hospitalized for cancer, 1976

Days before death	Percent of Medicare expenses in the last year of life
Total (1-360)	100
1-30	28
31-60	17
61-90	11
91-120	8
121-150	7
151-180	6
181-360	23

Data showing that the costs of dying for cancer patients are greater than the costs for patients hospitalized for other reasons also support the idea that there is a large amount of costs that could potentially be impacted by hospices. Table 2 compares hospital use and total Medicare reimbursements for persons hospitalized for cancer, heart disease, and stroke -- the three leading causes of death. Cancer patients averaged 36 hospital days in their last year, as compared to 28 days for both heart disease and stroke patients. Their total 1978 reimbursement in their last year was $8,110, compared to $5,878 and $5,608 for heart

disease and stroke patients.

Table 2. Hospital use and Medicare reimbursement in the last year of life for persons hospitalized for selected diagnoses, 1978

Diagnosis	Hospital days per person	Total Medicare reimbursement per person
Cancer	36	$8,110
Heart disease	28	$5,878
Cerebrovascular disease	28	$5,608 .

Comparison of Hospice and Conventional Care

The law that provided for coverage of hospice also called for a report to Congress by the end of 1986 on the effectiveness of the hospice benefit. The evaluation, being conducted by HCFA, will be of particular importance since the hospice benefit will sunset in October, 1986. Data we have produced in the course of our research on costs in the last years of life will be valuable for comparison with the experience of Medicare patients under the new hospice benefit. Let me cite an example of the kind of data on patterns of non-hospice care that will be interesting to compare to the patterns found under the new benefit.

In earlier studies we found that as age increased, Medicare reimbursements before death decreased for all decedents. Table 3 shows that this same pattern holds for persons hospitalized for cancer.

Table 3. Medicare reimbursements in the last 6 months of life for persons hospitalized for cancer by age, 1980

Age	Total Reimbursement Per Person
65 and over	$8,607
65-74	$9,303
75-84	$8,201
85 and over	$7,136

Total reimbursements for cancer patients in 1980 decreased from $9,303 for persons 65-74 to $7,136 for persons age 85 and over. We have speculated that this pattern may be due to three factors:

1. Nursing home services, not reimbursed by Medicare, may increasingly substitute for Medicare-reimbursed hospital care with advancing age.
2. The time between onset of illness and death may decrease with increased age.
3. Physicans may be less likely to apply large amounts of costly treatments to very old persons whom they may feel have poor chances of survival.

These factors may not apply to hospice care because a whole range of services from home health to nursing home to hospital are reimbursable and because of the supportive and caring rather than curative orientation of hospices. Thus, the pattern of use of hospice services by age will shed light on the nature of the hospice program.

A New Approach to Acquiring Mortality Statistics

The last topic I would like to mention is what we think is a major step forward in the efficient acquisition and use of mortality statistics. For our studies on Medicare use in the last year of life and for comparison data for the hospice evaluation, we plan to analyze Medicare use by cause of death, rather than by hospital diagnosis. This would solve a number of problems associated with use of hospital diagnosis. For example:

1. About 25 percent of Medicare enrollees do not have a hospital stay in their last year and thus we have no diagnosis for them.

2. One half of all persons hospitalized for cancer in their last year also have a hospital stay with a non-cancer principal diagnosis. Cause of death data would help us attach one diagnosis to these deaths.

We decided to obtain cause of death for about 70,000 deaths identified in our 1979 Medicare files. This was a 5-percent sample of all Medicare deaths that year. The usual way to obtain death certificates is shown in Table 4. A researcher could use the National Death Index (NDI) of NCHS to obtain state of death, death certificate number, and fact of death, if not already known. Then the researcher would contact each State to obtain hard copy certificates. The data from the certificates would be coded and entered into machine readable form.

Table 4. Obtaining death certificate information, usual procedure vs. procedure in HCFA-NCHS-AVRHS project.

Usual Procedure

1. NDI run.
2. States contacted individually.
3. Data coded and entered.

HCFA-NCHS-AVRHS Project

1. NDI run.
2. AVRHS obtains permission from States for direct computer link.
3. Computer link between HCFA and NCHS files.

In order to save time and money, HCFA and NCHS, together with the Association for Vital Records and Health Statistics (AVRHS), devised a new approach never attempted before. We first did a NDI run of our 1979 sample of Medicare deaths. The match rate was 94 percent. We learned the exact number of records by State for which we wanted death certificate data and the state death certificate numbers.

In May of this year we signed a contract with the AVRHS to obtain permission from the States for a direct link between our Medicare file and the NCHS Mortality Statistics File to obtain cause of death and other information. Permission is needed from the States since each State has the right to approve uses of the mortality data they furnish to NCHS. We have obtained permission for the link from 40 States and are optimistic about the rest. The direct computer link will not only save time and money, but will take advantage of the expertise of NCHS in coding cause of death and will make the analysis we do by cause of death comparable to other analyses with the NCHS file. We hope that this project will set a precedent that will enable other researchers to obtain mortality statistics in an efficient way for large scale studies. We are very eager to begin analysis of the linked file. We will be able to examine Medicare use for up to five years before death. We also plan to study the relation between hospital diagnosis and the cause of death.

Summary

Data on Medicare use by the dying entered into the debate leading to the passage of the hospice benefit. The data will be used in the congressionally mandated evaluation of the new benefit. Finally, in the course of efforts to obtain better comparison data to evaluate hospice, we have developed what we believe is a more efficient approach to obtaining mortality statistics for large-scale studies.

References

1. Congressional Budget Office, "Cost Estimate of H.R. 5180," June 28, 1982.
2. Lubitz, James and Prihoda, Ronald "The Use and Costs of Medicare Services in the Last Years of Life," forthcoming in Health, US, 1983. USDHHS, Public Health Service.

PATIENT AND PROVIDER REPORTS OF CANCER CARE AND COSTS: A PILOT STUDY

Andrew C. Montgomery and Elizabeth Eastman
Survey Research Laboratory, University of Illinois
Patricia N. Royston, National Center for Health Statistics

INTRODUCTION

In 1979, the National Cancer Institute (NCI) requested assistance from the National Center for Health Statistics (NCHS) in designing a national study to determine the direct and indirect costs of cancer care. The NCHS proposed that a network sampling approach, implemented by means of a supplement to the National Health Interview Survey (NHIS), be used to identify a national probability sample of cancer patients. These patients would be asked to participate in a one-year panel study to determine the direct and indirect costs of their cancer care. If necessary, a provider survey would be conducted to supplement and verify the information on medical care and costs collected from the patients.

NCHS was then asked to design a pilot study to test the proposed methodology. The pilot study was conducted in 1979-1981 in Illinois under contract with the Survey Research Laboratory (SRL), University of Illinois. The first two stages of the four-stage study were designed to test the feasibility of developing a national probability sample of cancer patients from a general health interview survey in which patients were reported in their own households and in households of close relatives. The third stage tested methods of collecting medical care and cost information from cancer patients. Finally, similar cost and utilization data were collected from the cancer patients' medical care providers for purposes of comparison.

The results of the first two stages have been reported elsewhere [1,2,3]. The purpose of this paper is to present a comparison of the patient-reported medical care and cost data with the associated provider-reported data, and to examine the conditions under which patient data and provider data are consistent.

The use of patient surveys for collecting medical care data has been criticized on the grounds that such data are unreliable [4]. Patients are expected to underreport medical care and costs, with the level of underreporting directly related to the length of time since the care was received. Also, patients may be unable to report some cost data in cases where third party payors are billed directly by the medical provider and patients do not see the bill.

Unfortunately, provider verification surveys can seldom correct the deficiencies of patient surveys. Since only patients know all of the providers who have treated them, the sample of providers must be obtained from patients. Thus the problem of patient recall is not solved with a provider survey. If patients cannot be relied upon to remember all care from the providers they report, what guarantee is there that they will report all providers from whom they received care? Also, provider records are not necessarily accurate or complete. Providers often record only those services for which they bill; if several visits are included in a single bill, the number of visits is likely to be underreported. Finally, provider surveys, like all surveys, are expensive.

SURVEY DESIGN

The findings presented in this paper are based on data from the last two stages of the pilot study, a panel survey of cancer patients and a survey of their medical care providers. Following is a brief summary of the sample design and data collection procedures for each of these surveys. Details of the design have been reported elsewhere [5].

The Patient Survey

The sample for the patient cost survey included 201 cancer patients who were reported as having cancer in a general health interview survey conducted earlier in the pilot study. Three types of households were interviewed in the general health survey; (1) households of cancer patients selected from two Illinois cancer registries, (2) households of relatives of the registry cancer patients, and (3) households selected from the general population which served as "decoy" households, so that neither the patients nor the interviewers were aware of the special nature of the sample. Interviews were completed with 264 households of registry cancer patients, 162 households of their relatives, and 363 "decoy" households.

General health interviews were conducted in these households using a current version of the National Health Interview Survey core questionnaire and a cancer screening supplement developed for the pilot study. Two hundred thirty-six of the registry patients and 55 additional cancer patients were reported during these interviews. A subsample of 201 patients was selected for the subsequent cost survey.

Given the source of this sample, the cost data collected in this survey cannot be extrapolated to the general population of cancer patients. However, the purpose of this paper is to examine the willingness and ability of cancer patients to report cost and utilization data. We have no reason to believe that cancer patients selected from these tumor registries would differ from any other cancer patients in their willingness or ability to report these data.

Data collection for the patient survey included two rounds of interviewing. In the first round, face-to-face interviews were conducted with the patients or with proxy respondents if the patient had died or was otherwise unable to be interviewed. Proxy respondents were selected by asking to speak with the person who was most familiar with the patient's medical history. Thirty-seven interviews were attempted with proxies and 75 percent (28) of these proxy respondents completed the interview. Ninety percent of the 164 interviews attempted with patients were completed.

One of the major challenges during the design phase of the patient survey was development of the instrument. The approach ultimately adopted was modeled after that used in the National Medical Care Utilization and Expenditure Survey. Respondents were first asked about the amount of care received from four general sources, and then they were asked for specific information about each visit or event. The four general sources of medical care were (1) hospitals or nursing homes, (2) emergency rooms, (3) hospital outpatient departments and clinics, and (4) private physicians and other individual medical care providers. For each medical care event, respondents were asked to report (1) date, (2) name and address of the provider, (3) type of treatment or medical care received, (4) condition or problem treated, (5) total cost, and (6)amount paid by the patient or the patient's family.

Two types of memory aids were used in the patient survey. First, a specially designed calendar and cover letter were mailed to prospective respondents three to four weeks after the general health interviews. The cover letter contained instructions for noting medical care visits on the calendar and stressed the importance of keeping records of medical care bills and statements. During the subsequent cost interviews, interviewers recorded information about the respondents' use of bills, receipts or records and the overall utility of the calendar.

The second memory aid was a Health Care Summary that was mailed to each respondent after the first cost interview. The summary contained key information about each of the medical care events that the respondent had reported during that interview. A cover letter asked respondents to check over and correct any misinformation on the summary and to fill in missing information, particularly cost data that were not available at the time of the interview.

The second round of interviews for the patient survey was conducted by telephone about three months after the first round. This round of interviews had two purposes, (1) to provide estimates of panel attrition rates, and (2) to permit the respondent to correct or complete information collected in the first round interviews. Thus, the major portion of the round two interview involved reviewing the Health Care Summary and filling in information that was missing at the time of the round one interview. The data presented in the following tables are based on updated information obtained during the second round of the patient survey. All data refer to care received prior to the round one interview.

The Provider Survey

The sample for the provider survey included 219 medical care providers who were reported by 154 of the cancer patients interviewed in the patient survey. A total of 175 respondents were interviewed in the patient survey, however, ten reported no medical care during the specified recall period, ten refused to sign the consent forms permitting us to contact their medical care providers, and one refused to give the providers' names.

Special data collection procedures were used for the two hospitals that maintained the cancer registries from which the initial sample of cancer patients was selected. All other providers named in the patient survey were contacted by mail. Several methods were used to ensure an adequate response rate in the mail survey. The forms were as clear and as brief as possible, asking only for the five essential items of information about each medical care event: date; condition treated; provider seen; total cost; and amount billed to the patient. Specialized forms were developed for each of the four types of providers: hospitals and nursing homes; emergency rooms; outpatient departments and clinics; and individual health care providers. Providers were given the options of providing the information by phone or having an abstractor visit them to collect the survey data. They were phoned periodically to encourage participation and to answer questions. These methods proved effective; 94 percent of the providers returned the completed forms within the ten-week survey period.

FINDINGS

The patient survey estimates presented here exclude the care reported by the ten patients who refused to sign consent forms and the one who refused to provide the names of the providers, since provider data are not available for these cases. Also excluded are the patient reports of care from the providers who failed to respond to the survey. The provider survey estimates discussed below are based on all provider data, including those from the two registry hospitals.

Table 1 shows the three types of estimates that were selected for comparison as indicators of the completeness and the comparability of the patient and provider survey data. These are number of events reported, percent of events with total cost not reported, and mean cost per event.

Two types of errors are expected in patient and provider surveys, those resulting from the patient's failure to report all of the medical providers who provided care in the reference period, and those resulting from the patient's and the provider's failure to accurately report the episodes of care during the reference period. Very little information is available from the pilot study on the first type of error, which is the number of providers that were completely omitted from the patient reports. An indication of the magnitude of the second type, which is the completeness of reporting events by patients and providers, can be obtained from studying the degree of agreement between the two surveys on the number of events that occurred in the reference period. Overall, the two surveys yielded comparable estimates of the number of events; 2010 for the patient survey and 2117 for the provider survey. Patients and providers also agreed on the numbers of hospital and nursing home stays and emergency room visits. This is not surprising, since these are generally very memorable events for patients. Also, they are relatively rare events, compared with the number of clinic and office visits reported. There is less agreement on clinic visits and contacts with individual health care providers. Patients report fewer outpatient department and clinic visits, and more contacts with health care

professionals than do providers. Further study is needed to see whether these differences are due to different record keeping practices; for example, some of the events reported as office care by patients may have been entered into clinic files if the physician routinely saw the patient in both locations.

The percent of events with total cost not reported provides an assessment of completeness of reporting that is not dependent on comparison of the two surveys. For most types of care, patients failed to report total costs for about 20 percent of the events. Only one exception was noted; patients failed to report costs for nearly half of the emergency room visits. Providers provided virtually complete cost data for all types of care except for inpatient care reported by individual providers. This was expected, since providers must maintain complete records for business and tax purposes. Also note that since part of the hospital data was provided by the two cooperating registry hospitals, the completeness of cost reporting in the pilot study may be somewhat better than could be expected in a general medical provider survey.

Finally, a comparison of the mean cost per event derived from both surveys indicates roughly the degree of agreement between patients and providers on the cost of reported events. First, a note on the method used to calculate the means; the mean costs shown for all care are based on all medical care events for which total costs were reported. The calculations by source of care, however, are affected by some data collection problems. Both patients and providers sometimes reported that care from more than one source was included in a single non-itemized bill, or flat fee. Flat fees cannot be included in calculations of mean costs by source of care, since the costs cannot be disaggregated by source. Thus, all calculations of mean costs in Table 1, except the "TOTAL" line, exclude flat fees.

Mean costs per event for all events are in close agreement, $258 and $271 for patient and provider surveys, respectively. However, large discrepancies are seen for mean costs of inpatient care by health care professionals, and smaller differences are seen for hospital or nursing home visits, outpatient department and clinic visits, and office care provided by health care professionals. One hypothesis is suggested by the figures for hospital and nursing home stays and inpatient care by health care professionals. If inpatient care by a health care professional is sometimes billed through the hospital, it is conceivable that these costs would be reported by hospitals and not by the individual providers. This would partially explain why providers report lower mean costs for inpatient care by a health care professional and higher mean costs for hospital stays. This could also explain the incomplete reporting of cost data (column d) by health care professionals for inpatient care.

Table 2 compares the same three data items; number of events reported, percent with total cost not reported, and mean cost per event, for selected characteristics of the patient and the patient interview. The first two characteristics, vital status of the patient and type of respondent, are related, since information on all deceased patients was provided by proxy respondents. Table 2 shows that vital status and type of respondent appear to have little effect on the comparability of patient and provider reports of number of events or mean costs per episode. Completeness of cost data is affected, however; proxies more often failed to provide information on the total costs of the reported events.

The third characteristic presented in Table 2 is length of recall. By design, about one-third of the sample for the patient survey (70) was diagnosed during the year preceding the start of the pilot study, and the remaining two-thirds were diagnosed one to three years before the study. Patients diagnosed as having cancer one to three years before the start of the pilot study were assigned a relatively short recall period of three to five months, bounded by the earlier general health interview. Patients diagnosed less than one year before the start of the pilot study were asked to report about a much longer period, in some cases up to 15 months, extending back to one month before the diagnosis. In a national survey, this long recall period would be needed for newly diagnosed patients so that the survey could provide stable estimates of costs around the time of diagnosis.

Table 2 shows the effect of length of recall period on reports of care and costs. It appears that long recall patients provide more complete cost data and are in closer agreement with providers on the average cost per episode than short recall patients. It is possible that the emotional and financial impact of high medical bills and frequent medical care is felt more strongly by recently diagnosed patients, and consequently these patients are more interested in carefully researching and reporting their medical care and costs than are patients diagnosed some time earlier. This is supported by data presented in another report [6] showing that about 60 percent of the long recall patients used memory aids frequently during the interview, compared with about 30 percent of the short recall patients. Furthermore, the Table 2 data on use of memory aids indicates that respondents who used memory aids frequently were in close agreement with providers on number and average cost of episodes, and provided nearly complete data on total costs for the episodes.

Table 3 looks further at the effects of type of respondent and use of memory aids on reporting of care and costs. It appears that both patients and proxies who frequently used memory aids during the interview reported costs for nearly all care, and were in fairly close agreement with providers on the number of events that occurred and on the average cost per event. Proxies who occasionally or never used memory aids provided the least complete cost data, and showed the least agreement with providers on the number of events that had occurred in the reference period. Thus, it appears that memory aids significantly improve the quality and completeness of data on medical care and costs, especially when the information is obtained from a proxy respondent.

CONCLUSIONS

The results of the pilot study show that patients and providers are in fairly close agreement on the overall number of health care events and the mean cost per event for all events with reported costs. However, patients are not able to report costs as completely as their providers. Patients consistently failed to report costs for approximately 20 percent of the events they mentioned, while providers failed to report costs for approximately 5 percent of the events. Patients reported lower costs only for hospital or nursing home stays. These costs are such a large proportion of all costs that the overall effect was that patients reported a lower mean cost per event. It is not apparent whether the observed differences by source of care are the result of erroneous reporting or the result of inherent difficulties in defining types of health care events in a standard way to patients and to providers.

The data indicate that two techniques may help to minimize the discrepancies between patient and provider reports. Whenever possible, patients rather than proxies should be interviewed, since the patient respondent provided more complete cost data than the proxy respondent. The difference between the patient-reported and provider-reported mean cost per episode was also lower when the patient was the respondent. Supplying memory aids for the patients and encouraging their use by respondents will also result in more complete data. The frequent use of memory aids was associated with the smallest differences between patient-reported and provider-reported data.

REFERENCES

[1] Sirken, M., P. Royston, D. Bercini, R. Czaja, E. Eastman, and R. Warnecke, "Completeness of Enumeration of Cancer Cases in Health Surveys," 1981 Social Statistics Section, Proceedings of the American Statistical Association, pp. 223-227.

[2] Czaja, R., R. Warnecke, E. Eastman, P. Royston, M. Sirken, D. Tuteur, "Locating Patients with Rare Diseases Using Network Sampling: Frequency and Quality of Reporting," Presented at the Fourth Conference on Health Survey Research Methods, Washington, D.C., 1982.

[3] Bercini, D., P. Royston, M. Sirken, R. Warnecke, "Errors in Reporting Date of Cancer Diagnosis: The Results of an Experimental Household Health Survey," Presented at the Annual Meeting of the American Public Health Association, Montreal, Canada, 1982.

[4] Andersen, Ronald, Judith Kaspar, Martin R. Frankel, and Associates, Total Survey Error, San Francisco, Jossey-Bass Publishers, 1979.

[5] Sirken, M., P. Royston, R. Czaja, E. Eastman, and R. Warnecke, "Pilot of the National Cost of Cancer Care Survey," 1980 Section on Survey Research Methods, Proceedings of the American Statistical Association, pp. 579-584.

[6] Survey Research Laboratory, University of Illinois, "Household Network Surveys of Cancer Care Costs: A Research Pilot Study," Final Contract Report prepared for the National Cancer Institute and the National Center for Health Statistics, 1982.

Table 1. Number of Medical Care Events Reported, Completeness of Cost Reporting, and Mean Cost per Event for Patient and Provider Surveys, By Source of Medical Care

Source of care	Number of events reported		Percent with total cost not reported		Mean cost[1] per event		
	Patient Survey (a)	Provider Survey (b)	Patient Survey (c)	Provider Survey (d)	Patient Survey (e)	Provider Survey (f)	Percent difference [(e-f)/f]·100 (g)
TOTAL	2,010	2,117	19%	5%	$258	$271	-5%
Hospital or Nursing Home	157	158	20%	0%	$2,348	$2,801	-16%
Emergency Room	54	53	48%	4%	$82	$78	5%
Hospital Outpatient Department or Clinic	1,051	1,417	18%	5%	$63	$49	29%
Health Care Professionals							
Inpatient Care	230	184	11%	18%	$357	$196	82%
Office Care	518	305	22%	0%	$31	$21	48%

[1]Total mean costs are based on all events with reported costs. Mean costs for specific sources of care exclude events billed as a flat fee, which is a nonitemized bill including more than one source of care.

Table 2. Number of Medical Care Events Reported, Completeness of Cost Reporting, and Mean Cost per Event for Patient and Provider Surveys, By Selected Characteristics of the Patient and the Patient Interview

Selected characteristics of the patient and the patient interview	Number of events reported		Percent with total cost not reported		Mean cost[1] per event		
	Patient Survey (a)	Provider Survey (b)	Patient Survey (c)	Provider Survey (d)	Patient Survey (e)	Provider Survey (f)	Percent difference [(e-f)/f]·100 (g)
TOTAL	2,010	2,117	19%	5%	$258	$271	-5%
Vital status of patient at first panel interview							
Alive	1787	1873	19%	5%	$211	$212	0%
Deceased	223	244	26%	2%	$667	$704	-5%
Type of respondent							
Patient	1326	1436	17%	6%	$195	$190	-3%
Proxy	684	681	24%	2%	$395	$434	-9%
Length of recall period							
Short	630	615	26%	7%	$185	$210	-12%
Long	1380	1502	17%	4%	$288	$295	-2%
Patient use of memory aids during interview							
Frequently	1012	1045	7%	7%	$264	$287	-8%
Occasionally	542	476	33%	3%	$190	$213	-11%
Never	456	596	32%	4%	$319	$290	10%

[1]Mean costs are based on all events with reported costs.

Table 3. Number of Medical Care Events Reported, Completeness of Cost Reporting, and Mean Cost per Event for Patient and Provider Surveys, By Patient and Proxy Use of Memory Aids During the Interview

Type of respondent and use of memory aids during the interview	Number of events reported		Percent with total cost not reported		Mean cost[1] per event		
	Patient Survey (a)	Provider Survey (b)	Patient Survey (c)	Provider Survey (d)	Patient Survey (e)	Provider Survey (f)	Percent difference [(e-f)/f]·100 (g)
TOTAL	2,010	2,117	19%	5%	$258	$271	-5%
Use of memory aids by PATIENT respondent							
Frequently	649	709	9%	9%	$189	$217	-13%
Occasionally	321	302	26%	3%	$139	$150	-7%
Never	356	425	25%	5%	$257	$243	6%
Use of memory aids by PROXY respondent							
Frequently	363	336	3%	2%	$391	$426	-8%
Occasionally	221	174	44%	3%	$290	$323	-10%
Never	100	171	54%	1%	$696	$560	24%

[1]Mean costs are based on all events with reported costs.

THE NATIONAL HOSPICE STUDY

David S. Greer and Vincent Mor, Brown University

The National Hospice Study originated in the public sector and has been associated with the formulation of public policy throughout its history. The Study was initiated in 1978 at the request of Congress which sought information on the cost and effectiveness of hospice care, to be used in developing legislation for federal support of hospice. In mid 1982, Congress passed legislation including hospice in the Medicare program, despite the fact that the Study which Congress itself had requested was not yet completed. At present, amendments to the hospice legislation, partially based on preliminary data from the Study, are being considered in Congress. The Health Care Finance Administration, which is one of the granting agencies and also the channel for transmitting Study data to Congress, has requested that no data be publically released while the Congressional debate proceeds. It is with regret and apologies, therefore, that I must tell you I can give you no cost data in my presentation today. I can tell you, however, how we planned to respond to the request of Congress and how our experiences in the field modified that plan; I can describe our current methodology; and I can give you some information on what we have learned about terminal cancer care in the United States in the course of this large Study. The data presented are based upon preliminary analyses conducted in the spring of 1983. A more detailed description of the study methodology is presented elsewhere.[1]

In our initial design, we formulated the information need of Congress and the granting agencies into four major research questions:

Table 1
Research Questions

1. What is the differential impact of hospice, demonstration or non-demonstration, on the quality of life of terminally ill patients and their families, as compared to conventional care?

2. What are the differential costs of caring for comparable terminally ill patients in demonstration hospices, non-demonstration hospices, and conventional care settings?

3. What is the likely impact of Medicare reimbursement on the organizational structure, staffing pattern, and costs of hospices?

4. What is the likely national utilization and cost of hospice care?

There was debate on the need for a fifth research question, i.e., what is the difference in the intervention in the hospice and conventional care systems. Initially it seemed that this information might be superfluous since we had been asked to study cost and outcome rather than process; but our experience in the field convinced us to include this question since both hospice and non-hospice care were changing so rapidly (and, in many instances, converging), that outcome information might seem vacuous without some insight on process.

Initially, we were presented with 26 Demonstration (D) hospices which had been chosen by HCFA from an applicant pool of 233 hospices. The criteria used by HCFA were:

Table 2
Criteria for Choice of D Hospices

1. Comprehensiveness of the hospice intervention.

2. Soundness and thoroughness of the service plan in the proposal.

3. Operational at time of review.

4. Sufficient distribution of major hospice types.

5. Representation in each of the DHHS federal regions.

The hospices were distributed nationally and classified into three types:

1. Hospital-based.
2. Home care agency based.
3. Free standing.

The Demonstration hospices were provided with generous reimbursement to enable them to give optimal hospice care to terminally ill cancer patients.

In our quest for controls, we chose 14 similar hospices and 12 "conventional care" sites. The latter were systems on oncological care of good quality with accessible data systems. Our control sites were clustered in three regions of the country since we felt that national projections based on regional denominators would be easier to develop and would yield more accurate results. There were sufficient numbers of Demonstration hospices in each region to make comparisons.

Our plan was to select comparable cohorts of patients and families and to follow these cohorts through the three systems of care: demonstration hospice, non-demonstration hos-

pice, and conventional care. In reviewing the plan and with experience gained in the field, several questions arose:

1. As we observed the behavior of hospices, it became apparent that the tripartite classification of hospices might obscure important differences. It appeared that, behaviorally and to some extent organizationally, hospices could be better classified into those which owned or were in some way responsible for inpatient beds and those which did not have beds. There were important site-of-care and process differences in these two types of hospices, regardless of whether they were hospital based, freestanding or home care agency hospices. We therefore chose to analyze our data on the basis of two hospice types: those with beds which we called "hospital based" (HB) whether or not they were organizationally sponsored by a hospital, and those without beds which we called "home care" (HC) hospices.

The behavior of these two hospice types is quite different as can be seen from Table 3 and Figure 1. Home care hospice (HC) patients die at home 62.1% of the time, while only 19.7% of patients in hospices with beds die at home. (The national average for cancer patients is 13.2%). HC cancer patients spend 12.2% of their hospice stay in an inpatient setting, whereas HB hospice cancer patients are inpatients 43.2% of their days. Some of this difference may be due to case mix: HB patients tend to be less independent functionally (Figure 2) and also are less likely to have available informal support at home (Figures 3 and 4).2

Table 3
National Hospice Study
Days Of Service Type Received By Demonstration Hospice Cancer Patients By Hospice Type

	Home Care Based	Hospital Based	Total
Percentage with an inpatient stay	38.9%	78.0%	54.4%
Average number of inpatient days during hospice stay	5-6 days	17.7 days	10.0 days
Percentage of hospice stay in an inpatient setting	12.2%	43.2%	22.7%

Figure 1
Site of Death For Demonstration Hospice Cancer Patients and National Average

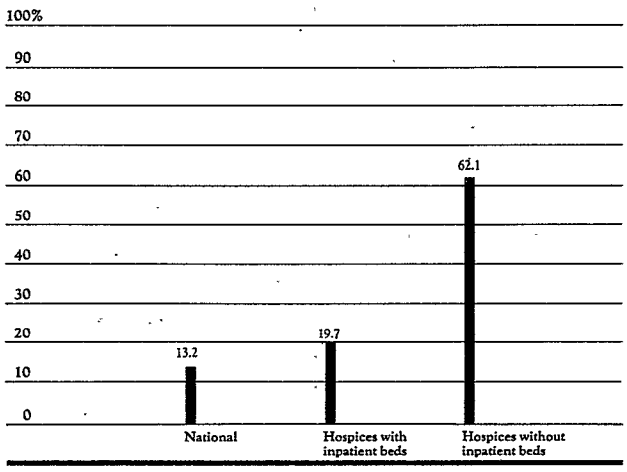


Figure 2
Proportion Of Patients Able To Independently Perform Selected Activities Of Daily Living At Admission By Hospice Inpatient Status

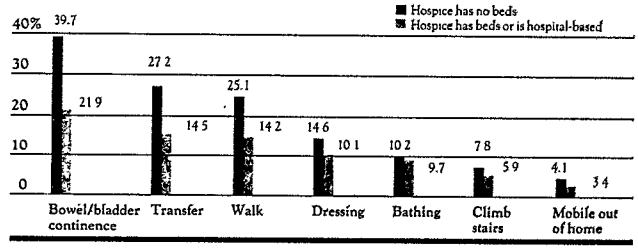


Figure 3
Patient Living Arrangement At Admission By Hospice Inpatient Status

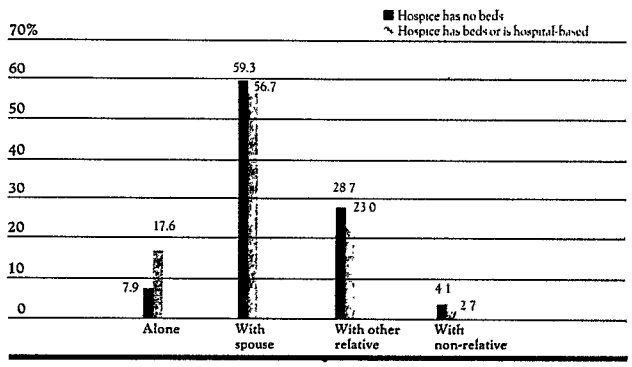
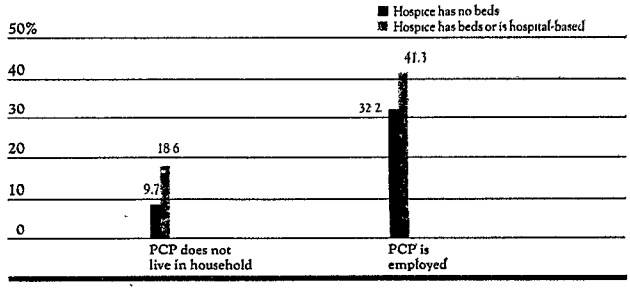


Figure 4
Primary Care Person (PCP) Availability By Hospice Inpatient Status



2. We had anticipated that there would be differences in the samples of patients and families which chose to enter different types of hospices or to remain in the conventional care system. This was indeed our experience. The conventional care sample was younger and more functionally impaired ("sicker") than the hospice sample, for example. To adjust for these differences, we adopted two approaches:

a. Selective inclusion of participants from the sample pool, where possible.
b. Multiple regression techniques which I will describe in more detail later.

3. Since our sample consisted of 40 hospices chosen from an idiosyncratic pool of several hundred, the question arose whether the sample N was 40 or the several thousand patients/family units we were following through the system. That is, was there such a thing as "hospice care" or were we measuring the outcomes of various individual hospices and a complex variety of care systems? Again, we chose to use multiple regression techniques to separate these phenomena.

Strategy to Adjust for Sample and Facility Differences: The technique used to adjust for possible differences in the facilities and the samples exposed to the three interventions is based upon multiple regression models. Two forms of models are employed, depending on the kind of outcome variable being studied. Continuous outcome variables assuming values on an interval scale are described by multiple linear regression models. Categorical outcomes are described by multivariate multiple logistic regression models. In either case the models relate the observed outcomes to characteristics of the patient, the facility, and the regional health care system.

The four main steps in the analytical strategy are:

1. The specification of a linear or logistic regression model relating response (outcome)variables to patient, facility and area characteristics.
2. Fitting the specified models to the observed data and validating the models.
3. The performance of hypothesis tests, specified in advance of inspection of the data, and the estimation of confidence intervals to assess the differences in effects between alternative interventions.
4. Systematic use of fitted models to learn which differences in characteristics of patients or facilities are associated with differences in observed effects.

Modeling a response variable Y that assesses cost of care or quality of life, the equations are of the general form:

$$Y = B + B X_1 + B X_2 + \ldots + B X_n + C F_1 + \ldots C F_n + A_{x1} + \ldots + A_{xn} + eps$$

The subscripted X's denote patient-level independent variables, the subscripted F's denote facility-level independent variables, and the subscripted A's denote area-level independent variables.

Three such equations are used for each dependent variable: one for HC hospice patients, one for HB hospice patients, and one for conventional care patients. Separate equations are required because we anticipated interactions between the intervention and other patient-level independent variables. Prototype patients are passed through each equation to determine the adjusted value of the dependent variables.

In these models, "eps" is the random "error" term and the assumption is that "eps" is normally distributed. The independent variables are selected so that these assumptions will be legitimate, and their validity is tested as the models are fitted to the data.[3]

A different mathematical model is used for each outcome measure. Based upon the expert judgment of professionals intimately familiar with the care of terminal cancer patients and a systematic review of the literature, independent variables that are conceptually related to the outcome measures were identified. The array of independent variables from which those included in the models were chosen is large, ranging from standard demographic data to attitudinal information possibly related to choice of hospice. Among the domains of independent variables were included:

Patient Demographics: Age, sex, race, religion, education, marital status.

Family Situation: Family size, patient and PCP living arrangements, PCP employment status, PCP relationship to patient, family income, other support obligations of PCP, patient's insurance coverage.

Prior Service Utilization: Type of prior cancer therapy, prior hospital or institutional service use, prior compliance with physician's orders.

Patient/Family Attitudes: Knowledge of diagnosis and prognosis, importance of religion, patient's belief in afterlife, characterization of the aggressiveness of patient's treatment.

Extrinsic Factors: These include characteristics of the adjacent catchment area (e.g., prevalence of hospital beds, home health services, skilled nursing homes), characteristics of the Study site, hospice or conventional care.

As a complement to the linear regression approach, logistic regression techniques were used when the outcome variable of interest is categorical in nature (e.g., receipt of a given type of service intervention or not).[4]

Cost Calculations: In the NHS, costs are calculated on the basis of utilization. The basic equation is: Cost = Utilization x Cost Coefficient. Utilization data was available from a variety of sources: e.g., Principal Care Person report, hospice cost reports, Medicare bill files. We believe the utilization figures are complete and accurate.

There was much controversy, however, concerning the appropriate cost coefficients to be used. One could utilize actual hospice costs as reported to HCFA. Alternatively, national or regional cost averages for each service could be used. A confounding variable in the determination of cost coefficients is the cost-to-charge ratio since hospices and hospitals frequently bill charges which are later adjusted by third party payers to approximate actual costs.

Measures of the cost of hospice care should accurately reflect differences among patients in the intensity of resource use (e.g., inpatient room and board, inpatient ancillaries and home services). In addition, to provide results that can be generalized, costs should reflect national pricing patterns rather than the pricing

practices of the providers serving NHS demonstration patients. Cost estimates also should account for adjustments to charges used by Medicare to compute reimbursable cost.

After considerable debate and consultation, the following methodology for the determination of cost coefficients was utilized:

$$\text{Hospital Inpatient Cost} = \text{Inpatient Days} \times \text{National Avg. Medicare Per Diem Cost} + \text{Adjusted Ancillary Cost Per Day}$$

$$\text{Adjusted Ancillary Cost} = \text{Ancillary Charges to Patient Per Day} \times \frac{\text{Avg.* Ancillary Charge (U.S.)}}{\text{Avg. Ancillary Charge (Patient's Provider)}} \times \frac{\text{Avg. Cost (U.S.)}}{\text{Avg. Charges (U.S.)}}$$

$$\text{Home Cost Care} = \text{Reimbursements}$$

*This term was included for hospital costs but not for hospice costs.

The formulas are based on utilization measures that reflect each patient's service intensity. Inpatient costs were computed similarly for both hospice and non-hospice inpatient care. The cost of each inpatient stay equals the product of its number of inpatient days multiplied by the appropriate inpatient cost weight. Each inpatient stay cost weight is the sum of two terms that account separately for room and board and for ancillary use during that stay.

The first term in the inpatient cost weight is a constant measure that accounts for routine inpatient (room and board) per diem. For each hospital stay, this term is equal to the 1982 national average Medicare hospital-routine per diem cost ($154). For hospice stays, the constant inpatient cost weight was inflated to account for the average difference between the NHS hospices' and their affiliated hospitals' routine per diem to yield a routine hospice per diem of $180.

The second term varied to account for each patient's unique ancillary use. These ancillary charges were adjusted in two steps. Charges were converted to Medicare reimbursable costs using the national average community hospital Medicare ancillary charge-to-cost ratio. Then, for hospitals used by NHS patients in conventional care and HC hospices, the influence of each provider's ancillary pricing levels relative to the nation was removed. Hospice ancillary costs were not similarly adjusted because there was no national data source that reflects the experience of all hospices. These calculations yield a measure of a patient's ancillary costs per day which varies only according to differences in intensity of ancillary utilization among patients. Other charges (for example) for skilled nursing facility care) were entered without adjustment to compute total costs. These other inpatient services accounted for one percent or less of average inpatient costs.

Hospice home care costs were derived from cost reports submitted to HCFA and the NHS Study group by the participating demonstration and non-demonstration hospices. Non-hospice provided home care costs were set equal to charges since analyses found charges equal to Medicare reimbursement. Since a national home health data base was not available, provider specific adjustments to a national average were not made.

The TEFRA Legislation and the NHS

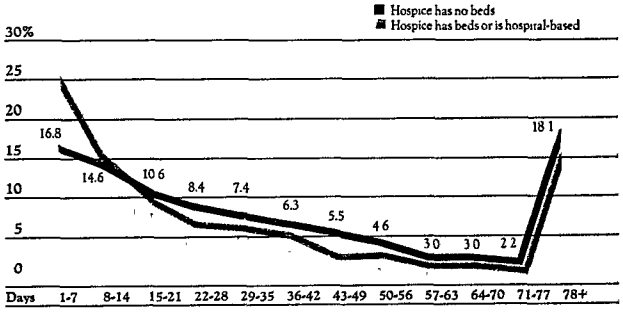
The TEFRA legislation sought to maintain the home-based tenor of hospice care by requiring that, in the aggregate, only 20 percent of all patient days in any individual hospice be provided in an inpatient setting. If home care services for hospice patients are less costly than inpatient care, this provision could be cost saving. As shown above, based upon preliminary data, hospital-based hospices in the NHS failed to meet that criterion (43.2% inpatient days). A change in the behavior of hospital-based hospices would appear to be necessary if they are to satisfy this criterion for participation in the Medicare program.

The TEFRA legislation also mandates a cap on hospice costs. Average per patient costs cannot exceed 40 percent of the regionally adjusted costs to Medicare of the last six months of life of Medicare beneficiaries dying of cancer. This cap was predicated on the assumption that most health care expenditures during the last six months of life occur during the end of the period and that hospice patients have a length of stay of 45 days. Combining these two assumptions, it was estimated that 40 percent of the six-month terminal care expenditure would be unexpended at the time of transfer of the patient to hospice care.

As can be seen, based upon preliminary data, approximately 25 percent of patients served by NHS hospices with beds are dead within seven days of admission. Over 50 percent die within 28 days of admission (Figure 5). The obvious question which emerges from these data is, what percentage of the last six months expenditures have already been incurred by those patients who survive seven days or less in hospice? Or those who survive 21 days? The assumptions underlying the TEFRA legislation seem to be, at the very least, oversimplifications.

Implications for the Health Care System of the Hospice Mode of Care: It is necessary to

Figure 5
Patient's Length Of Hospice Stay Distribution By Hospice Inpatient Status



begin by defining "hospice" since it does not represent a uniform mode of care. As noted above, the National Hospice Study started with a tripartite classification of hospices: hospital-based, home care agency based, and free-standing. It became apparent that, behaviorally, there are really two kinds of hospices; those with and those without their own inpatient facilities. Beds, even in hospices, follow a modified Parkinson's Law; when present, they tend to be filled. Patients in bedded hospices (HB), therefore, spend more time in the inpatient environment and less time at home than patients in hospices without inpatient facilities (HC).

There are confounding features, however. Patients in HB hospices tend to be more functionally impaired and to have weaker social support systems than patients in HC hospices. Adjusting for these differences with the regression techniques I have described does not affect substantially the differences in inpatient utilization, however.

The reduction in inpatient utilization does not necessarily result in reduced cost. The intensity of home care in HC hospices results in very high costs in many instances; and, since the HCFA demonstration paid costs rather than the lesser of costs or charges, costs were higher than charges in many demonstration hospices. Depending on length of hospice stay and local factors, HC hospices may be less expensive than HB hospices, despite the greater use of inpatient facilities in the latter.

Hospice does alter the pattern of care delivered to terminally ill cancer patients, whether HB or HC. In hospice, there is significantly reduced utilization of intensive therapies (e.g., chemotherapy, radiation therapy), and diagnostic testing (e.g., X rays, blood tests), as compared to conventional care. This remains true even when patients are matched by clinical as well as demographic, functional and diagnostic criteria.

Whether this difference in therapeutic approach results in significant alterations of length of life cannot be determined in a quasi-experimental study of this kind. Since we are dealing with the last few weeks of life, differences would probably have to be measured in hours or days and even a rigorous experimental design might fail to detect such small increments.

The central question remains: Will hospice provide better care at less cost than the conventional system? The answer is, "That depends". It depends on the incentive built into third party systems of financing hospice care. The National Hospice Study clearly demonstrates that it is possible to give satisfactory care to a significant subset of terminal cancer patients while reducing the utilization of expensive hospital beds and costly therapeutic interventions. Patient and family satisfaction with care can be maintained; quality of life in the terminal period is no worse and sometimes better.

The TEFRA approach to incentives appears inadequate, however. It includes patient populations and facilities on which there is no reliable data, e.g., non-cancer patients and for-profit facilities. Where TEFRA reflects available data, it also appears frequently to miss the mark: the forty percent cap is based on the erroneous assumption of a forty-five day average stay and a last-year cost curve which is at best an oversimplification. The twenty percent cap on inpatient days assumes erroneously that home care is invariably less expensive; and, in any event, will probably be seldom applied since it is a condition of participation rather than a reimbursement cap.

The implied assumption in TEFRA that terminal patients, whether cancer or non-cancer, have similar cost and care patterns is also belied by NHS data. Non-cancer patients in the NHS have longer lengths of stay in hospice and are more likely to be discharged alive, reflecting the uncertainty of prognosis in this group. Intensive study of the non-cancer group will be necessary before predictions can be made of the cost and/or effectiveness of hospice as an alternative for their care.

References

1. Greer DS, et al. National Hospice Study Analysis Plan. Journal of Chronic Diseases (in press).

2. Mor V, Birnbaum H. Implications of National Hospice Study Data. Health Affairs, 2 (2) : 80-90, Summer, 1983.

3. Allan DM, Cady FB. Analyzing Experimental Data by Regression. Belmont, CA : Lifetime Learning Publications, 1982.

4. Kleinbaum DG, et al. Epidemiologic Research. Belmont, CA : Lifetime Learning Publications, 1982.

Systems Design and Program Administration

Session I

THE SHIFT FROM INPATIENT TO OUTPATIENT CARE FOR FOUR SELECTED SURGICAL PROCEDURES UNDER THE PHILADELPHIA BLUE CROSS: EVALUATION OF A PROGRAM

Monroe Lerner and David S. Salkever, The Johns Hopkins University

As emphasis in this country shifts to the private sector for containment of rapidly rising health care costs, studies of the Blue Cross and Blue Shield Plans' cost-containment programs assume increasing importance for public policy. Many of these programs are intended to reduce inpatient utilization, and therefore Plan costs, by shifting care for selected types of surgical procedures and other cases to an ambulatory setting. The present study is an evaluation of one program of this type involving four selected surgical procedures, that undertaken by the Blue Cross of Greater Philadelphia, Inc., in the latter part of the 1970s.

In Philadelphia some of these ambulatory settings, meeting criteria specified by the Plan, were designated as Short Procedure Units (SPUs). The four surgical procedures studied here, intended by the Plan to be performed in SPUs, were: 1) D and C (dilation and curettage)/D and E (dilation and evacuation), 2) myringotomy, 3) surgical removal of impacted wisdom teeth, and 4) tubal ligation (laparoscopy). For the present study, the data were derived from Blue Cross claims' files and tabulated by quarter-years in a study period covering all of 1979 through the second quarter of 1982 (14 quarters); however, SPU data were available only since the third quarter of 1980.

Analysis of the data from this study (see Table I) showed a very substantial shift from inpatient care to SPU. In the third quarter of 1980, not long after the program was first implemented, about one-fourth of the four procedures were performed in SPUs, but by second-quarter 1982 the comparable proportion was about three-fourths. The timing of this shift, and its speed, varied among the four procedures, as did also their potential for further shift.

The drop in inpatient admissions was not completely matched by the rise in SPU cases. Thus the average quarterly drop in admissions was 321, while the comparable increase in SPU cases was 197, about three-fifths as much.

One implication of this discrepancy between decreasing admissions and increasing SPU cases is that some cases were "lost" to Blue Cross, and this could have happened in one of three possible ways: 1) they were performed in physicians' offices, rather than as inpatient or SPU cases, and for these cases a Blue Shield claim may have been filed; however, if so, Blue Cross would not have known about the claim since the Philadelphia Blue Cross Plan is not affiliated with a Blue Shield Plan, 2) they were performed elsewhere, perhaps in a free-standing surgical clinic not covered by Blue Cross, so that no claim was filed here either, or 3) they were not performed at all; i.e., foregone, or perhaps merely delayed, to be performed at some time in the future. The disposition of these cases was not determined in the present study.

With the shift to SPU has come a very substantial dollar savings to the Plan and its members. For the four procedures combined, over $15 million was saved during the study period, about 80 percent of it ($12 million) since mid-1980. (These data are not shown here due to space limitations.) The savings were most substantial for D and C and D and E, followed in that order by myringotomy, surgical removal of impacted wisdom teeth, and tubal ligation (laparoscopy).

One hypothesis of this study was that changes in provider behavior for Blue Cross patients would be followed by a similar change for non-Blue Cross patients. This would occur because it would be difficult for providers to distinguish among patients on the basis of their insurance coverage; it would simply be easier to treat all patients alike, and we hypothesized that the providers would reason that if the change were to do no harm to Blue Cross patients, it should do no harm to the others as well. Therefore one aim of the study was to test the presumed spill-over to non-Blue Cross patients.

Because data on non-Blue Cross patients were thus required and could only be obtained from the internal records of hospitals, a sample of hospitals (N=11) was selected for this phase of the study. The data from these hospitals showed clearly that the hypothesized spill-over had in fact occurred, i.e., that the shift to SPU had occurred for both Blue Cross and non-Blue Cross patients. However, as anticipated, the shift occurred first for Blue Cross patients, generally one year before the shift for non-Blue Cross. In addition, the Blue Cross shift occurred at a faster rate. Table II shows the decrease in inpatient admissions for Blue Cross and non-Blue Cross. Blue Cross admissions dropped from the beginning of the study period, while non-Blue Cross admissions actually rose for one year before starting to decline. The bottom of the table shows rates of change. Blue Cross decreased 9.2 percent every six months while non-Blue Cross decreased more slowly, 7.0 percent.

Table III shows the corresponding rise in SPU utilization. In 1982, 65.1 percent of Blue Cross patients had these procedures performed in the SPU, compared to 58.1 per-

cent for non-Blue Cross. Thus, Blue Cross patients were shifted to the SPU faster than non-Blue Cross. However, the impact of Blue Cross efforts clearly benefited non-Blue Cross patients as well.

Another objective of the study was to determine the background of the shift from inpatient to SPU. The Blue Cross policy of encouraging the shift was first implemented by providing technical assistance to establish SPUs and by educational activities directed at hospital administrators and medical staffs. At the beginning of 1979, 29 of the 64 hospitals in the area already had established SPUs. The number increased gradually and accelerated during 1981, and by mid-1982 nearly all hospitals had them. However, until mid-1980, relatively few inpatient cases were questioned or denied as SPU-appropriate.

Beginning in mid-1980, an intensive effort was undertaken to identify SPU-appropriate procedures performed as inpatient, and actively to encourage their shift to SPU. First, the aggregate utilization experience of each hospital was reviewed to determine its frequency of admissions for SPU-appropriate procedures. Hospitals with a high frequency were informed that a change would be welcome, and that the Plan would review their claims case-by-case.

Concurrently, medical-review screening procedures (individual claims review) conducted by Plan personnel began to identify all inpatient cases deemed medically appropriate for SPUs. The hospitals involved were informed of the determinations, and educational visits were made to them. It was suggested during these visits that unless a shift occurred in future cases of a similar nature, claims would be denied. Where such shifts did not occur, preliminary written notices were sent and, later, formal denials were made. The peak numbers of denials occurred during the second quarter of 1981.

While inpatient admissions for SPU-appropriate procedures were still occurring in mid-1982, the SPU had in fact even before that time become the preferred site of care for such procedures among most hospitals in the area. It was accepted as such among most of their physicians, and the SPU had become the actual site of care for most such procedures. Only a minority of such cases were still inpatient, and that number was decreasing rapidly. Although both educational visits and denials were still occurring in mid-1982, they were less frequent than formerly. By mid-1983 almost no such cases were inpatient except those with a medical complication.

Current practice is that Plan personnel conduct semi-annual audits of each hospital's performance. Hospitals are then informed of the results and provided with a comparison of their performance with regional "norms". Hospitals identified as varying significantly from these norms are "focused" for more frequent case-specific reviews and subsequent contact by Blue Cross staff.

On an a priori basis, it seemed reasonable that hospitals would vary in their responses to this program according to occupancy levels, although other factors might also be important. However, the present study focused on occupancy level, since such data were readily available.

Hospitals were classified into three categories on the basis of their occupancy levels during the last six months of 1981 in accordance with data from the Delaware Valley Hospital Council. Hospitals classified as high-occupancy were those with overall occupancy ratios higher than 86 percent, medium-occupancy 86 to 82.5 percent, and low-occupancy lower than 82.5. Hospitals in each category had approximately the same number of Blue Cross admissions during the six-month period but, because high and medium-occupancy hospitals had larger bed-complements, only 18 hospitals were high occupancy and 17 medium, while 29 were low. High-occupancy were more likely to be teaching hospitals or suburban, while low-occupancy were community and inner-city.

Table IV shows the decrease in inpatient admissions by occupancy level. The average quarterly percent decrease for all occupancy levels was only 3.3 percent in the first six quarters, but as high as 14.4 percent in the last eight quarters. For all 14 quarters combined the decrease was 12.4 percent. The relatively small drop in the first six quarters was not statistically significant; however, the relatively large drop in the last eight quarters and the drop for all fourteen quarters were both highly significant.

Against this backdrop, the high-occupancy hospitals clearly began the decrease in admissions first. Thus their decrease in the first six quarters was more rapid than the comparable decrease for the others. The decrease for low-occupancy hospitals was more substantial than the comparable decrease for the medium, but this difference was not large enough to be at conventionally accepted levels of statistical significance.

The decreases for the last eight quarters of the study period were significant at each level of occupancy, and this was true for the entire study period as well. While these decreases were greater at the high and low-occupancy levels than at the medium, it was hard to distinguish between high and low here. While the low-occupancy hospitals may have started their decreases after the high, they quite clearly caught up during the later period.

Table V shows that the number of SPU cases increased during the last eight quarters of the study period. The changes were clearly relatively the most rapid among the low-occupancy hospitals, followed by high and medium. Statistical significance for the increase was attained by low and medium, and almost by the high.

Table VI shows that the low-occupancy hospitals started at a relatively low figure for SPU cases as a percent of total cases for these four selected procedures, but made up the difference by the second and third quarters of 1981, after which they were indistinguishable from the others.

DISCUSSION

The Philadelphia Blue Cross program to shift care for selected surgical procedures among Blue Cross members from inpatient care to SPUs was clearly successful; further, it spilled over to non-Blue Cross patients as well. Although high-occupancy hospitals were the first to shift, the others soon followed suit. Thus, a major behavioral change was induced among providers; the mechanism here was that the Blue Cross Plan was able to exert pressure on hospitals, over whom it has considerable leverage in the Philadelphia area because of its large market share, while the hospitals, in turn, were able to convince their staff physicians, after considerable initial resistance, to conform to the new policies. Perhaps an important aspect of their willingness to conform was that physicians were financially unaffected by the change, neither gaining nor losing; presumably they were eventually agreeable because they saw the change as beneficial to the welfare of the hospitals, a matter of some considerable concern to themselves.

The pressure exerted by Blue Cross on hospitals was both educational and financial, the latter by threat of, or actual, payment denial. Low-occupancy hospitals in precarious financial condition might anticipate a significant financial loss as a result of the shift to SPUs, and therefore might be expected in response to fill their beds with other patients to the degree possible; unfortunately the data were not available to test this hypothesis, although it should be tested as a matter of significant public-policy concern. Here a "filling effect" could be neither confirmed nor denied. Both total patient-days and occupancy ratios were rising in the aggregate for all hospitals in the Philadelphia area during the Blue Cross program, so that possibly even low-occupancy hospitals might have considered action unnecessary since their losses for the four selected surgical procedures studied here were small as a percent of their total patient-days, while their occupancy ratios may have been rising in any case. High-occupancy hospitals may well have welcomed the shift to SPUs, seeing it as a mechanism to increase their revenues without increasing their bed complements, and therefore as an enhancer of their efficiency levels.

Can the experience of the Philadelphia Blue Cross be generalized to other areas? Presumably "yes" where Blue Cross Plans have relatively large market shares and where they also have good relationships with hospitals and providers. In addition, if the Philadelphia experience can be taken as a guide, a spill-over to non-Blue Cross third-party payers can be anticipated in these areas. Large market shares for Blue Cross exist primarily in the Northeast and East North Central states, but it is also in these states that the need for cost-containment programs is most acute, since inpatient stays are longer there and per-capita expenditures on hospital care higher. Public policy should encourage Blue Cross Plans to follow Philadelphia, for example by establishing institutional arrangements like SPUs elsewhere, but also by expanding the list of surgical procedures covered for outpatient care and covering as well many other services not often now covered on an outpatient basis. Just a short list of these services might include diagnostic laboratory and X-ray services, home care, nursing home care, hospice care, various forms of preventive care, multiphasic screening, and various categories of psychiatric, drug abuse, and alcoholism services. These programs should be undertaken in concert with Plan involvement in community health planning activities, including providing technical assistance to health planning agencies, establishing penalties for disapproved hospital investment, and establishing incentives to closure or consolidation of hospitals where these are appropriate.

TABLE I

Number of Inpatient Admissions and SPU Cases
Four Selected Procedures, Total
1979 Through Second Quarter 1982
Blue Cross of Greater Philadelphia

Quarter	Inpatient Admissions	SPU Cases	Total	Inpatient as Percent of Total
1979, 1st	3,423			
2nd	3,410			
3rd	3,558			
4th	2,715			
1980, 1st	2,968			
2nd	3,105			
3rd	3,425	1,205	4,630	74.0
4th	2,850	1,688	4,538	62.8
1981, 1st	3,010	2,437	5,488	54.8
2nd	2,875	1,956	4,831	59.5
3rd	2,466	2,098	4,564	54.0
4th	2,063	2,319	4,382	47.1
1982, 1st	1,480	2,310	3,790	39.1
2nd	1,014	3,151	4,165	24.3

TABLE II

Admissions to Eleven Sample Hospitals
For Four Selected Procedures and Total Admissions
Blue Cross and Non-Blue Cross Patients
By Half-Year Periods, 1979 Through 1981

Year and Half-Year Period	Four Selected Procedures		Total Admissions	
	Blue Cross	Non-Blue Cross	Blue Cross	Non-Blue Cross
1979 1	1,864	1,951	18,909	40,855
2	1,739	1,975	19,220	40,081
1980 1	1,411	2,020	19,136	42,890
2	1,506	1,872	19,382	41,907
1981 1	1,461	1,542	19,562	42,883
2	1,036	1,380	18,568	42,308
Average Half-Year Numerical Change (Linear)	-139.4	-122.9	-12.4	+419.7
Average Half-Year Percent Change (Logarithmic)	-9.24	-7.03	-0.07	+1.02

TABLE III

SPU Cases in Eleven Sample Hospitals
All Four Selected Procedures
Blue Cross and Non-Blue Cross Patients
Second Half 1980 Through Second Half 1981

SPU Cases, Total Cases, and Half-Year Period		Blue Cross	Non-Blue Cross
SPU Cases			
1980	2	1,129	1,146
1981	1	1,894	1,636
	2	1,929	1,916
Total Cases†			
1980	2	2,625	3,018
1981	1	3,354	3,178
	2	2,965	3,296
SPU as Percent of Total			
1980	2	43.0	37.9
1981	1	56.5	51.5
	2	65.1	58.1

†SPU plus inpatient.

TABLE IV

Inpatient Admissions for Four Selected Procedures
Hospitals by Levels of Occupancy Ratios
Blue Cross of Greater Philadelphia
1979 Through Second Quarter 1982

INPATIENT ADMISSIONS FOR ALL FOUR PROCEDURES COMBINED

Year and Quarter		All Occupancy Levels	High Occupancy	Medium Occupancy	Low Occupancy
1979	1	3,422	1,208	1,040	1,174
	2	3,413	1,183	985	1,245
	3	3,560	1,209	1,085	1,266
	4	2,714	942	860	912
1980	1	2,968	971	992	1,005
	2	3,105	997	1,032	1,076
	3	3,424	1,084	1,176	1,164
	4	2,851	1,008	901	942
1981	1	3,009	946	1,000	1,063
	2	2,875	940	994	941
	3	2,467	832	880	755
	4	2,064	635	750	679
1982	1	1,478	495	512	471
	2	1,014	330	333	351

TABLE V

SPU Cases for Four Selected Procedures
Hospitals by Levels of Occupancy Ratios
Blue Cross of Greater Philadelphia
Third Quarter 1980 Through Second Quarter 1982

SPU CASES FOR ALL FOUR PROCEDURES COMBINED

Year and Quarter		All Occupancy Levels	High Occupancy	Medium Occupancy	Low Occupancy
1980	3	1,205	443	541	221
	4	1,688	629	642	417
1981	1	2,437	899	899	639
	2	1,956	599	821	536
	3	2,098	608	795	695
	4	2,319	722	813	784
1982	1	2,310	726	848	736
	2	3,151	1,096	1,083	972

TABLE VI

SPU as Percent of Total Cases
(Inpatient Admissions Plus SPU Cases)
Hospitals by Levels of Occupancy Ratios
Blue Cross of Greater Philadelphia
Third Quarter 1980 Through Second Quarter 1982

PERCENTS FOR ALL FOUR PROCEDURES COMBINED

Year and Quarter		All Occupancy Levels	High Occupancy	Medium Occupancy	Low Occupancy
1980	3	26.0	29.0	31.5	16.0
	4	37.2	38.4	41.6	30.7
1981	1	44.0	48.7	47.3	37.5
	2	40.5	38.9	45.2	36.3
	3	46.0	42.2	47.5	47.9
	4	52.9	53.2	52.0	53.6
1982	1	60.9	59.5	62.4	61.0
	2	75.7	76.9	76.5	73.5

THE DEVELOPMENT AND IMPLEMENTATION OF DRG SYSTEMS IN AN ACUTE CARE INSTITUTION

David Spivack, Mount Sinai Medical Center of Greater Miami

Since its inception the most radical change in the Medicare Program occurred April 20, 1983 when President Regan signed Public Law 98-21, the Social Security Amendments of 1983. Title VI of the Social Security Act of 1983 provides legislation to prospectively pay hospitals for all Medicare patients by Diagnostic Related Groups (DRGs). DRGs are a method of grouping patients into 467 diagnostic related categories or groups based on resources consumed during a hospital stay (1).

With the Department of Health and Human Services regulations for the Tax Equity and Fiscal Responsibility Act of1982 (TEFRA) and the regulations to be released September 1, 1983 implementing the Prospective Payment System (PPS) hospitals are faced with a situation requiring a massive increase in the collection and generation of statistical data by DRGs. These regulations place hospitals in a situation where they must be able to use their existing information systems to provide data of a different nature than currently available. The monitoring and control of hospital costs will be dependent upon the institution's ability to generate DRG charge and cost information on a daily basis.

There are several policy questions that, unfortunately have yet to be settled and makes this task more difficult. This paper will review a number of areas in which decisions have yet to be formulated by HCFA and the impact on hospitals developing case mix management systems using Mount Sinai Medical Center, a 699 bed non-profit teaching hospital where over 150,000 patient days in 1982, or 73% of the institutions patient days, were Medicare.

This paper will also explore to date the incorporation of DRGs and associated data systems and reports into the medical center's existing management information system. The implementation of DRG information systems has and will continue to be an immense challenge for hospitals as they cope with the new regulations and the ability to better contain health care costs as mandated by TEFRA and the PPS. This data will be of the utmost importance to hospitals in their attempts to document the intensity and utilization of services provided to patients and consequently justify adequate reimbursement to an institution. Such information will not only be crucial to documenting the costs of care to patients with acute illnesses but also patients with chronic illnesses who are frequently readmitted and generally have a longer length of stay. The PPS DRGs will have a far reaching policy and organizational impact on acute care institutions.

Teaching hospitals and large community hospitals that already have management systems will be immediately at task to establish DRG information systems. While several policy decisions that remain unanswered at this time may create immediate difficulties for hospitals with a national average Medicare patient day load of 35%, these issues have a potentially severe impact on Mount Sinai's operations.

Since Medicare provisions in TEFRA will be in effect for the next several years in addition to the PPS, the TEFRA regulations serve as a starting point for institutions establishing DRG systems. There are several problems in TEFRA however that have not yet been resolved to provide comfort in data systems design in anticipation of the PPS. First TEFRA presumes to establish utilization by Medicare patients at an institution from a 20% sample 1980 MedPar data. While HCFA is allowing a one time adjustment, hospitals are under pressure to maintain the same case mix as recorded three years in the past in 1980 without any significant changes into the three years in the future (1983, 1984 and 1985, the years TEFRA will be effective).

Second, has the mix of services, as represented by the 1980 sample and the case mix index derived from the sample, changed since 1980? HCFA has not made any allowances for such changes. In fact, HCFA considers any increases in Medicare admissions as being promoted only by incentives to take advantage of increased Medicare reimbursement offered by the regulations and will be adjusted downward accordingly:

> Under the reimbursement system established by P.L. 97-248, a hospital may have an incentive to increase its number of Medicare patients. For example, a hospital that has costs less than the target amount will receive an increased payment per discharge above its actual costs. We are concerned that some hospitals may promote the increase admissions of Medicare patients to take advantage of this aspect of the reimbursement system. Such action would be contrary to the intent of the legislation, which was to reward efficient operation, not to stimulate increase[d] hospital admissions (2).

This does not allow for the increase in elderly population in a given service area or for that population's aging and subsequent requirement for more inpatient hospital services.

Third, the TEFRA regulations do not adequately account for multiplicity of diagnosis in patients, those patients who have more than one diagnosis during a hospital stay, and only accounts for, to a limited degree, complications that arise during the patient's stay. In addition, the DRGs do not address the age of the patients except on a greater or less than 70 basis. This could also be a severe problem. Assigning one DRG to a patient on discharge, which dictates the reimbursement an institution receives, may not consider or be sensitive to the fact that different patients of varing ages with the same principal diagnosis may be considerably sicker (with additional comorbidities and/or complications) and will require more intensive utilization of resources. Low income elderly particularly fall into this category more often than others. The National Association of Public Hospitals is currently exploring this in more detail.

Fourth, HCFA believes that the case mix index utilized in TEFRA accounts for the severity of illness of cases and the requirement of these cases for more intensive services.

> Since the new limits on total inpatient operating cost will be applied on a per discharge basis, and will be directly adjusted to reflect individual hospital differences in case mix, we do not believe the exception for hospitals with higher per diem cost resulting from more intensive routine care should be available to hospitals covered by those limits (3).

Tertiary care centers and teaching hospitals will be the most affected by the TEFRA and PPS regulations as expensive services, e.g. cardiac catheterization, are required in such institutions to fulfill the needs of the community the institutions serve. While these concerns are specific to TEFRA, these are also germane to PPS. Whether the federal DRG PPS will compensate hospitals adequately as the New Jersey DRG system remains to be seen. The New Jersey system includes provisions for the increased costs associated with teaching programs, charity care, bad debt, and outliers which comprise 30% of the entire system as compared to 5 to 6% in the federal system.

In the context of this vacuum and that Mount Sinai still has to proceed to develop a new patient case mix management information system, the following will present the considerations the institution has and will continue to review to implement the PPS.

Proper information and records are requisite to set the DRG PPS in place within an existing management information system. Sufficient data recording, retrieval and recapture mechanisms must be set in place as soon as possible not only for admitting, medical records and billing services but also with respect to the hospital's management and finance services, planning, marketing and regulatory functions. The value of such mechanisms in the case of management and finance is:

oTo determine the financial consequences of care delivered to each individual;
oTo determine resources used to deliver a unit of care;
oTo determine hospital expenses per admission (or unit of care);
oTo more accurately budget for admitted units (by diagnosis);
oTo identify points where productivity could increase;
oTo have a data base relating rates and charges to costs and revenues by DRG;
oTo more effectively involve physicians in controlling hospital costs associated with each physician; and
oTo control costs and volume of services used (i.e. maintain an appropriate mix of patients in all modalities offered by the hospital to insure stability).

In the case of Regulation:

oTo provide rate regulators a more accurate reflection of the costs of providing care to the Mount Sinai case mix;
oTo provide Peer Review Organizations a more accurate reflection of utilization and quality of care provided by Mount Sinai;
oTo better plan, document and report charges for prospective reimbursement systems; and
oTo determine and provide reliable documentation of reimbursement shortfalls or overruns in the Mount Sinai case mix.

In the case of Planning:

oTo accurately determine service utilization and patient origin information, cost, access, and quality factors for inpatient and outpatient care.

In the case of Marketing:

oTo accurately determine service utilization and patient origin information for the identification and targeting of special program needs;
oTo encourage patient referrals;
oTo determine the location and magnitude of market growth.

The impending federal PPS required Mount Sinai Medical Center to review all computerized information systems from data entry to reporting and the consequent impact and need for revised or additional data elements and systems analysis capabilities to adapt to the DRGs. Some of the organizational issues involved in the implementation of DRG systems and data collection requirements included restructuring admission/discharge, billing and medical record systems. On-line computer services require additional expansion and development to insure data capture and, therefore, retrieval to analyze patients, services and physicians by DRGs. To perform this function, expecially critical with the anticipated emphasis on planning and marketing by hospitals under DRGs, institutions have been approached by a virtual plethra of consulting firms offering software packages providing grouper, DRG reporting, utilization review and complete resource/patient care management systems.

Planning considerations for the PPS required a review of several areas: resources utilization by DRGs; cost data by DRG; medical staff profiles; and appropriate on-line systems. Data/information requirements imposed by the PPS system require an institution to better document internally the utilization of resources by DRG; to review revenue and cost information to determine certain characteristics of services - profitability, resource consumption, market attractiveness and patient care units are terms that come to mind - with the intention of establishing and assigning a financial value to the institution to maintaining services in that specific DRG. Should a service be continued? This is a difficult question to answer at this time as there may be overriding factors to be weighed in the forthcoming regulations such as statistical allowances/exceptions to teaching hospitals regional referral centers or specialized care units such as rehabilitation or psychiatric units. To arrive at the point of determining the appropriateness of services there is a need to identify the actual costs of patient care resources - information most hospitals are not cognizant of. While charge or revenue information is readily apparent, hospitals often do not know the actual costs of services provided. Cost accounting systems that are in place have generally responded to initiatives of the

government, insurors and other special information needs rather than monitoring actual cost data for specific procedures and/or treatments administered during the course of a patient's hospital stay. Direct and indirect cost allocations are also adjusted differently by each institution creating a lack of a standard practice and, therefore, difficulties in peer group comparisons.

As a first step to handle the cost data problems, Mount Sinai became involved in a South Florida regional data base established by the South Florida Hospital Association Health Data Network to focus on direct costs. This data base, comprised of the majority of the hospitals in Dade and Broward Counties, will provide a hospital on-line capability to access, compare and evaluate detailed cost and revenue data by DRG within the region.

Another area of importance in institutional planning is the profile of the Medical Staff - the Medical Staff's utilization of resources and length of stay by DRG. In-house reporting of physicians by DRG will be crucial to determing variances and, therefore, profitability of that physician to the hospital. Reports have been generated at Mount Sinai that depict these physician characterization.

Finally, Mount Sinai has insured that the availability of appropriate on-line data entry capabilities in Medical Records, admission and discharge billing information is available to document patient care delivered. Medical Records should have sufficient space in on-line abstracts to record a principal diagnosis and 8 secondary diagnoses (ICD-9-CM codes). The Medical Record is in turn given to the Business or Billing Office who prepare a billing record (UB-16 or UB-82 or 1453) submitted to the Medicare Intermediary. The billing record only contains room for reporting a principal diagnosis and one secondary diagnosis. This elimination of up to 7 ICD-9-CM codes (Mount Sinai's medical record abstract has sufficient space for 8 secondary diagnosis codes) could alter the final DRG assignment. Organizational issues abound within the institution as to which departmental entity, medical records, business office or other areas, should be responsible for selecting the ICD-9-CM secondary diagnosis that will have impact on the final payment rate to the institution.

Clearly Medical Records is in the best position to review the data for completeness and accuracy insuring the proper secondary diagnosis code or codes in the future, is selected that accurately reflects services provided to that patient during the course of his/her stay and to maintain the lines of communication to the key parties affected: medical staff, administration, finance, nursing and ancillary staff (4). The end result is healthier reimbursement reflecting the case mix severity of the institution. This is important to hospitals in such cases where an acute mycardial infarction may be reimbursed at 6,700 dollars versus congestive heart failure which may be reimbursed at 2,400 dollars - a difference of 4,300 dollars - and reflects the need for proper coding and reporting.

All these issues reviewed above impact on the hospital records and information systems. The data generated will be of the utmost importance to accurately interpret and document a hospital's case mix.

In conclusion, these represent major concerns of acute care institutions and how one institution, Mount Sinai Medical Center, has proceeded in the implementation of a radically different payment system - a system that will have significant impact to hospitals devoted to a large degree serving patients over 65 years of age. The future impact of the PPS will be major in terms of what data is collected, reported and analyzed to document resource allocation by DRGs. Trends in health policy will likely result in prospective payment for all payors, providing hospitals with strong incentives to better manage and record data documenting case mix severity.

Footnotes

1. New Jersey Hospital Association,"Reimbursement under S-446:Elements and Effects; Update 1982 (Princeton, 1982) p2.
2. 42 CFR Part 405, Federal Register. (September 30, 1982) p43285
3. Ibid. Severity of illness and patient care classification systems have been addressed more thoroughly by, e.g., Susan Horn, "Measuring Severity of Illness: Comparisons Across Institutions". American Journal of Public Health. Volume 73, No. 1 (January 1983). pp 25-31.
4. Hans Boerma, The Organizational Impact of DRGs, Health Research and Educational Trust of New Jersey (January 1983) p21.

USE OF MEDICARE HISTORY IN DESIGNING MEDICARE CAPITATION

Jennifer J. Anderson, University Health Policy Consortium

SUMMARY

A small group of Medicare recipients, typically people with serious and chronic health problems account for most Medicare expenses. The control of expenses for this subgroup of individuals, through their enrollment in capitated programs, for example, could have a dramatic effect on overall Medicare expenses. Current formulae for HMO reimbursement, using adjusted average per capita costs (AAPCC) provide disincentives for bringing such high-risk individuals into a capitated system, however. In fact it has been demonstrated that current financing arrangements have resulted in the enrollment of healthier than average individuals into HMOs. We have designed a variety of risk-based capitation formulae that would not have this drawback and that would protect both government and provider against the problems of adverse selection.

Our studies, using 1977 current Medicare Survey data, and 1974-1977 Medicare History File data for randomly selected cohorts of Medicare recipients in California, Texas and Massachusetts demonstrate the value of incorporating health items such as disability status and levels of prior use of medical services such as hospitalization into the construction of a revised AAPCC formula. We show that in predicting subsequent year reimbursement from Medicare History File data, the prior use models have considerably greater explanatory power compared with models using only demographic data that are analogous to the current AAPCC.

This type of revised AAPCC would focus attention on the recipients with greater need for medical services, and would encourage the private sector to take on responsibility for the medical care of these people. If the use of a risk-based AAPCC were coupled with a relaxation of the constraints on open enrolment this would even allow provider organizations to limit their enrolment to frail high-risk elderly where major opportunities for cost containment can be found.

INTRODUCTION

In any one year about half of all HCFA's expenditure for Medicare is for a small proportion, around five percent, of all recipients. These people typically are disabled or have other serious and chronic health problems, so that recurring expenses are somewhat predictable. The control of expenses for this subgroup of individuals, through their enrollment in capitated programs, for example, could have a dramatic effect on overall Medicare expenses.

The current formula for HMO reimbursement, termed the AAPCC (Adjusted Average Per Capita Cost) does not take health status into account so that there is no incentive for HMOs to enroll the chronically ill or disabled. Recent work of Eggers[1] has shown that HMOs using this formula have for the most part enrolled individuals with prior use of medical services that is lower than average. Formulae that explicitly account for disability or chronicity of disease could be used to encourage HMOs to specialize in the care of those elderly who have greater than average medical needs.

The construction of such formulae is difficult, however, since the historical data bases available, whether for development or for implementation, do not include explicit health status information. They do include proxies for health status and chronicity of need, but the use of proxies in a formula could lead to adverse selection and manipulation, perhaps worse than may arise from the use of a formula based on demographics alone.

For example, the current AAPCC[2] takes into account only age, sex, welfare status and institutional status and has relative cost factors ranging from 0.60 up to 2.60, which is a less than fivefold range. A formula that takes into account the number of times a person was hospitalized in the previous year as well as their age and sex could easily have eightfold variation in relative cost factors from say, 0.40 for a non-hospitalized woman under seventy, up to 3.20 or greater for an older man who was hospitalized more than once in the previous year. Although the number of times a person has been hospitalized in a given year, or the number of days they have spent in the hospital, is related to their current health status and to their future need for services it is still a proxy and it is possible that an HMO could obtain a higher capitation rate for its enrollees based on a high hospitalization rate in the previous year for conditions that were self-limiting and curable rather than the chronic, recurring conditions that genuinely require subsequent high use of services.

Health status has another advantage over proxies based on use of services in that since it is intrinsic to the enrollee it is less affected by the type of medical system used by the enrollee than are those proxies. Beebe et al.,[3] in their paper on incorporating prior use into an AAPCC, comment that for people whose medical care system prior to enrollment in the HMO was not Medicare fee for service based, but either private insurance or an HMO, there would be problems in implementation of a prior use formula. Such problems would be reduced, or even eliminated if health status were the basis of the capitation formula used.

We have used ordinary least squares regression with a subsequent year's reimbursement as the dependent variable, to construct relative cost factors for a variety of categorizations of prior use, including some

that directly reflect chronicity of disease, in an effort to avoid the possible adverse selection problems associated with more distant proxies for future need for services. The analysis uses 1974-1976 Medicare History File data for a randomly selected cohort of Medicare enrollees from Massachusetts (one percent of all 1974 recipients), Texas (one percent of 1974 recipients) and California (one-half percent of 1974 recipients). Only those individuals resident in the same state in both 1974 and 1975 have been included. The data set is further restricted to include only those who were sixty-five or over in 1974 and who had both Part A and Part B coverage in 1974. This results in a total of 12,645 cases with data for 1974 and 1975, and 11,970 cases with data for 1974 and 1976.

PRIOR USE MODELS

Since statewide data from three rather different states has been included, the states themselves and urban as distinct from rural residence have been included as factors in the models developed. Table 1 shows the coefficients of some one year predictive models. The 'demographic only' model is a stand-in for the current AAPCC. A relative cost factor for a particular category of recipients may be computed by adding relevant coefficients since the dependent variable used was the subsequent year reimbursement divided by the overall average subsequent year reimbursement. (The coefficients in Tables 1-4 should not be interpreted as underwriting factors for an AAPCC, however, because state and urban/rural factors are also included additively in the models.)

For example, in rural Texas the one year 'demographic only' model indicates a relative cost factor of $0.65 = 0.53 + 0.12$ for a male under seventy, and $0.71 = 0.53 + 0.09(x2)$ for a female in the third age group, namely seventy-six to seventy-nine years of age, while the 'number of hospitalizations' model indicates a relative cost factor for 1975 of $2.43 = 0.12 + 0.10 + 2.21$ for a male under seventy with more than one hospitalization in 1974, and $0.84 = 0.12 + 0.05(x2) + 0.62$ for a woman seventy-six to seventy-nine years old who had Part B in 1974, and was not hospitalized. Note that the percentage of variance explained by the prior use models, though low, is considerably higher than that explained by demographic variables alone. The prior use factors, whether days in the hospital or number of hospitalizations, have greater impact on subsequent use of services than do age and sex.

In the two year models of Table 2, in which 1974 factors are used to predict 1976 reimbursement, the prior use factors still have substantial effects, though not as large as in the one year model, and the demographic factors of age and sex have more effect than before. It is apparent, however, that the prior use model is a considerable improvement over the 'demographic only' model for both one and two year ahead predictions. Note also that the state and urban/rural differences are of the same order of magnitude as age and sex factors, and are stable across the different types of models.

CHRONICITY OF DISEASE MODELS

Specific disease information that could be used to measure chronicity is available on the Medicare History File for the hospitalized patients only. We have categorized each hospital discharge diagnosis according to whether hospitalization for the condition could be expected to result in substantial further medical costs or not, that is, as chronic or non-chronic, in a rather loose sense. A complete list of these "chronic" diagnoses is contained as an appendix to Anderson et al.[4]

Chronicity of disease may also be indicated by repeated hospitalizations for the same condition. In categorizing patients according to whether or not they had chronic conditions we have incorporated repeated hospitalization for certain conditions as an additional factor. These conditions include cancers and cardiac and musculoskeletal system diagnoses. Thus the chronicity measure has three categories with each hospitalized patient belonging in a single category. The categories are:

1. no chronic hospitalizations;
2. one or more chronic hospitalizations, but no more than one cardiac, cancer or musculoskeletal diagnosis; and
3. chronic hospitalizations, including repeated cardiac, cancer or musculoskeletal diagnoses.

These hospitalization categories have been used to replace the quantitative measures of prior hospital use, namely hospital days and number of hospitalizations, in one and two year predictive models. Tables 3 and 4 summarize the results. The first column of coefficients in each table, headed Model A, is for this chronicity model, while in the column headed Model B are coefficients for a slight variant on Model A, in which each hospitalization category has been split into two according to whether an individuals's final hospital admission in 1974 occurred in the first half of 1974, or later in the year. This additional factor of the timing of hospitalizations increases the explanatory power of the models a little, and the values of the coefficients are for the most part consistent with a priori expectations.

These chronicity models have R^2 values that are comparable to those obtained for the other prior-use models. The location, age and sex coefficients have values that are similar to those obtained in those other models. The coefficients associated with the seriousness of the hospitalizations are such as to result in higher cost factors for individuals with repeated hospitalizations in the prior year. For example the coefficient 2.94 in the one year Model A of Table 3 is considerably greater than

the 2.21 associated with two or more hospitalizations in the one year model of Table 1. As with prior use in two year models in Table 2, the effects of chronicity though still strong are not as marked in the two year models of Table 4.

The Model B coefficients in the one year model (Table 3) are lower if the second half of the year is free of admissions, for the first two hospitalization categories. In the two year model the differences between the two subcategories of these hospitalization categories are not as marked but there is a substantial difference between first and second half of year coefficients in the most serious hospitalization category. This anomaly points to an issue of limitations of the data and of the analysis methods chosen. Although these regression analyses have been performed using more than 10,000 cases, no more than one percent of Medicare recipients fall in the most serious hospitalization category. The estimate of coefficients for its two subcategories are based on relatively small numbers of cases, and are consequently, as their relatively large standard errors indicate, rather unstable, compared with estimates associated with factors with greater representation in the data set. Such estimates could perhaps be improved by including relatively more data for high users of services in the data sets analyzed or by use of empirical Bayes or other shrinkage estimation techniques to solidify the estimates for these very small and very variable groups of individuals.

DISCUSSION

It is apparent that the employment of prior history and disease chronicity information as proxies for health status has provided considerable improvement over the use of demographic variables alone in explaining subsequent reimbursement to Medicare recipients. The effects persist to a second follow-up year, which adds credence to the relevance of such factors and also may be of practical importance, since when capitation rates for a particular year are set data from the year just completed may not be uniformly available. The disease chronicity model does not have appreciably greater explanatory power than a model that includes numbers of hospitalizations but since chronicity is the more directly linked to real need for services in the future and less dependent on the prevailing system of medical care its use in capitation formulae should lead to greater stability in the face of adverse selection.

The models presented here are an example only of ways in which chronicity could be included in the construction of an AAPCC formula. However, we would expect other data based approaches to yield similar results in terms of explanatory power and relative sizes of cost factors for those with chronic need for medical services.

Efforts should be made to incorporate even more direct measures of health status in such formulae so that incentives to HMOs work to the advantage of both the elderly and the Medicare program. Table 5, using current Medicare survey data, shows how a prior use measure, the number of hospitalizations, would be inadequate as the basis of a capitation formula, if enrollee selection were then biased by disability status. At each hospitalization level, and overall, there is a progressive increase in subsequent year costs with increasing impairment of the individual. Unless an HMO were to provide facilities designed to attract enrollment by disabled elderly, such as transportation to clinics or home visits, they would be less likely than others to enroll in the HMO, and the consequent adverse selection would be to HCFA's disadvantage, if health status were not explicitly part of the capitation formula used.

REFERENCES

1. Eggers, Paul, and Prihoda, Ronald. "Pre-enrollment Reimbursement Patterns of Medicare Beneficiaries Enrolled in 'At-risk' HMOs." Health Care Financing Review, Volume 4, Number 1, September 1982.

2. Trieger S, Galbaum TW, and Riley G. "HMOs: Issues and Alternatives for Medicare and Medicaid." Health Care Financing Administration, 1981.

3. Beebe, James, Lubtiz, James and Eggers, Paul. "Using Prior Utilization Information to Determine Payments for Medicare Enrollees in HMOs." Working Paper, Health Care Financing Administration Office of Research, June 6, 1983.

4. Anderson, Jennifer, Resnick, Abby and Gertman, Paul. "Prediction of Subsequent Year Reimbursement Using the Medicare History File. II. A Comparison of Several Models." University Health Policy Consortium Discussion Paper 45-b, January 1983.

TABLE 1

COST FACTOR COEFFICIENTS FOR ONE-YEAR PRIOR USE MODELS

Variable	MODELS: Demographic Only	Hospital Days	Number of Hospitalizations
Intercept	0.53(.06)*	0.27(.06)	0.12(.06)
Massachusetts	0.13(.06)	0.12(.06)	0.18(.06)
California	0.23(.05)	0.22(.05)	0.22(.05)
Urban Residence	0.25(.05)	0.23(.04)	0.24(.04)
Male	0.12(.06)	0.11(.06)	0.10(.06)
Age†	0.09(.02)	0.05(.02)	0.05(.02)
Male x Age†	0.04(.03)	0.05(.03)	0.04(.03)
Part B Use		0.47(.05)	0.62(.05)
Days in Hospital		0.06(.003)	
One Hospitalization			1.01(.07)
More Than One Hospitalization			2.21(.10)
R^2	0.008	0.053	0.060

TABLE 2

COST FACTOR COEFFICIENTS FOR TWO-YEAR PRIOR USE MODELS

Variable	MODEL: Demographic Only	Hospital Days	Number of Hospitalizations
Intercept	0.54(.06)*	0.34(.06)	0.22(.06)
Massachusetts	0.18(.06)	0.17(.06)	0.22(.06)
California	0.24(.05)	0.23(.05)	0.22(.05)
Urban Residence	0.17(.05)	0.16(.05)	0.17(.05)
Male	0.17(.07)	0.16(.06)	0.16(.06)
Age†	0.12(.02)	0.09(.02)	0.09(.02)
Male x Age†	-0.01(.04)	-0.005(.03)	-0.01(.03)
Part B Use		0.37(.05)	0.48(.05)
Days in Hospital		0.05(.003)	
One Hospitalization			0.86(.07)
More Than One Hospitalization			1.68(.10)
R^2	0.007	0.034	0.038

*Coefficient standard error in parentheses.
†Age is measured in units of 5 years with age 67 set to zero. Male x Age is an additional age effect for males.

TABLE 3

COST FACTOR COEFFICIENTS FOR ONE-YEAR CHRONICITY MODELS

Variable	Model A	Model B
Intercept	0.12(.06)*	0.13(.06)
Massachusetts	0.17(.06)	0.17(.06)
California	0.21(.05)	0.21(.05)
Urban Residence	0.24(.04)	0.24(.04)
Male	0.11(.06)	0.11(.06)
Age†	0.05(.02)	0.05(.02)
Male x Age†	0.04(.03)	0.04(.03)
Part B Use	0.62(.05)	0.62(0.5)
No Chronic Hospitalizations	0.99(.07)	0.72(.10)
		1.21(.09)
Chronic Hospitalizations Without Repeats	1.70(.09)	1.41(.13)
		1.92(.11)
Chronic Hospitalizations Including Repeats	2.94(.20)	3.04(.37)
		2.90(.23)
R^2	0.059	0.061

*Coefficient standard error in parentheses.
†Age is measured in units of 5 years with age 67 set to zero. Male x Age is an additional age effect for males.

TABLE 4

COST FACTOR COEFFICIENTS

FOR TWO-YEAR CHRONICITY MODELS

Variable	Model A	Model B
Intercept	0.23(.06)*	0.22(.06)
Massachusetts	0.21(.06)	0.21(.06)
California	0.22(.05)	0.22(.05)
Urban Residence	0.17(.05)	0.17(.05)
Male	0.16(.06)	0.16(.06)
Age†	0.09(.02)	0.09(.02)
Male x Age†	-0.01(.03)	-0.01(.03)
Part B Use	0.48(.05)	0.48(.05)
No Chronic	0.75(.07)	0.64(.10)
Hospitalizations		0.84(.10)
Chronic Hospitalizations	1.49(.09)	1.29(.13)
Without Repeats		1.66(.12)
Chronic Hospitalizations	2.36(.23)	1.58(.40)
Including Repeats		2.76(.29)
R^2	0.041	0.042

*Coefficient standard error in parentheses.
†Age is measured in units of 5 years with age 67 set to zero.
Male x Age is an additional age effect for males.

TABLE 5

VARIATIONS IN 1977 MEDICARE COSTS

BY 1976 HOSPITAL EPISODES AND DISABILITY LEVEL*

	NUMBER OF HOSPITAL EPISODES IN 1976			
Disability Status	0	1	2 or more	All Cases
Unimpaired	0.64 (n=2881)	1.73 (n=353)	2.40 (n=122)	.82 (n=3357)
Moderately Impaired	1.06 (n=518)	1.94 (n=132)	3.41 (n=70)	1.45 (n=720)
Severely Impaired	1.65 (n=177)	2.12 (n=61)	3.94 (n=33)	2.03 (n=272)
All Cases	0.75 (n=3576)	1.83 (n=547)	2.94 (n=227)	1.00 (n=4349)

*Cost figures are all standardized, i.e., costs are divided by the average cost for all Medicare enrollees.

The Design of the National Medical Care Utilization and Expenditure Survey (NMCUES)

Session J

SURVEY DESIGN AND ADMINISTRATION OF THE NATIONAL MEDICAL CARE UTILIZATION AND EXPENDITURE SURVEY

Robert A. Wright, National Center for Health Statistics

I'd like to look with you at one of the newest and most complex surveys the National Center for Health Statistics (NCHS) has undertaken, the National Medical Care Utilization and Expenditure Survey (NMCUES).

The cost of health care has been an issue for most of this century, and probably before. In the late twenties and early thirties, it became enough of an issue that the first nationwide survey of illness and medical care utilization and expenditures was conducted. Since then, several other national surveys have been completed.

In the early seventies, the Center began a series of studies of new methods for obtaining information on the cost of medical care. The studies culminated in two surveys: The National Medical Care Expenditure Survey (NMCES), sponsored by the National Center for Health Services Research (NCHSR) and the National Center for Health Statistics in 1977, and the National Medical Care Utilization and Expenditure Survey, sponsored by the Health Care Financing Administration (HCFA) and the National Center for Health Statistics in 1980.

These surveys were undertaken because of several related phenomena: first, total health care expenditures rose from $42 billion in 1965 to $247 billion by 1980 and were projected to rise to $821 billion by 1990; second, national health expenditures as a percent of the Gross National Product or GNP (the total goods and services produced in the U.S.) rose from 6% of GNP in 1965, to 9.4% by 1980, and were projected to be 10.8% by 1990; third, per capita health expenditures rose from $211 in 1965 to $1,067 in 1980 and were projected to rise to $3,309 by 1990; fourth, both private and public health insurance plans and programs were becoming increasingly important in the financing of health care; and, fifth, a large variety of different methods were proposed at that time to provide Federal help in paying for medical care. People considering these proposals did not have a reliable set of data from which to estimate the costs to the government and to various segments of the population, if any particular proposal was enacted.

In 1977, as NCHS and NCHSR were conducting the NMCES, NCHS began discussing a permanent version of that survey. It was to be a periodic panel survey to collect data on medical care utilization and expenditures. At about the same time, the newly formed HCFA was looking for a survey mechanism to enhance the Current Medicare Survey which it had taken over from the Social Security Administration. The new survey was to cover not only Medicare recipients, but Medicaid recipients also.

Due to the efforts of Dorothy Rice, then Director, NCHS, and Cliff Gaus, then Director, Office of Policy, Planning, and Research, HCFA, discussions began between the two agencies in mid-1977. On September 20, 1979, these discussions culminated in a contract between the two agencies and three private research firms. The prime contractor was Research Triangle Institute (RTI), Research Triangle Park, NC. The two subcontractors were National Opinion Research Center (NORC) of Chicago and New York City and SysteMetrics, Inc. of Santa Barbara, CA and Washington, D.C. (SMI). The Co-Project Officers on the contract were Robert Fuchsberg, Director, Division of Health Interview Statistics, NCHS, and Allen Dobson, Ph.D., Director, Division of Beneficiary Studies, HCFA. Principle persons on the contractor side are Dan Horvitz, Ph.D., Vice President of RTI and Project Director during the major data collection period; Esther Fleishman, Associate Project Director for NORC; Pearl Zinner, of NORC, Executive Advisor to the Project; Jim Lubalin, Ph.D., Vice President of SysteMetrics, Inc. and Associate Project Director; and Barbara Moser of RTI, the current Project Director. Other people, including Bob Casady, Ph.D., of NCHS; Larry Corder, Ph.D., of HCFA; Gordon Bonham, Ph.D., then from NCHS, now with the University of Louisville; Brenda Cox, Ph.D., from RTI; and Embry Howell from SysteMetrics, Inc., had key roles in making the survey a success.

After some initial cutbacks in sample size, the current survey design was set. A detailed discussion of the survey design, instruments, and field procedures is presented in "Procedures and Questionnaires . . ." (Bonham, 1983). The following discussion summarizes the major features.

The NMCUES consists of three major survey components: a national household survey, which I will refer to as the HHS; a household survey of Medicaid recipients in four States, which I will refer to as the SMHS; and a survey of Medicare and Medicaid Administrative Records, which I will refer to as the ARS.

In the HHS and SMHS, an interviewer contacted each household in the survey sample five times at approximately 3-month intervals during 1980 and early 1981 to collect information about calendar year 1980. These contacts are referred to as interview "Rounds."

The first two contacts were personal visits by the interviewer to the household; most of the third and fourth contacts were made by telephone; and the final contact was a personal visit. A nominal incentive of $5 was paid to the person who answered the questions for the household during the first two rounds, and a $10 incentive was paid to the respondent at the last round.

Table 1 presents the average interview lengths for the HHS and SMHS components. In the National Survey, the interviews ranged in length from an average of one hour and 24 minutes in Round 1 to 48 minutes in Round 4. In the State surveys, the average times were somewhat longer, ranging from an hour and 48 minutes in Round 1 to one hour in Round 4.

Since the reference period for the survey was calendar year 1980, and since the first interviews began in February, the final one-third of the cases to be assigned in Round 4 actually were scheduled to be conducted after 1980 ended. Because Round 5 had special supplements, was to be conducted in person, and was to end with an incentive payment of $10, the final one-third of the Round 4 interviews were rescheduled as the first one-third of the Round 5 interviews. This procedure avoided having one third of the Round 5 interviews occur in April 1981 with no part of the reference period in 1980.

The response rates in NMCUES were high; not as high in the first round as the 95%-96% or more experienced by the National Health Interview Survey, but high nonetheless. The response rates for Round 1, presented in Table 2, are based on the number of occupied, eligible dwelling units and the number of reporting units found at the sampled addresses. The response rates for Rounds 2-5 are preliminary rates based on the number of persons enumerated in Round 1, who were also interviewed in later rounds. The overall response rate for the National survey can be estimated as 87.9% which is the product of the Round 1 reporting unit response rate, 91.1%, and the Rounds 2-5 combined person response rates, 96.5%.

The NMCUES National and State surveys used the same basic set of instruments to obtain and record data from the household respondents. These are listed in Figure 1.

The control card was used in all rounds. It provided the interviewer with basic assignment information, and it provided a place to record certain common information; information used throughout the interview.

The core questionnaire was also used in all rounds and contained questions on a wide variety of health topics. These questions were asked each round and used the beginning of the year, January 1, 1980, or the date of the previous interview as the reference date. The Supplements contained sets of questions for which an answer was sought only once; either because the data would not change or was assumed to be unchanging during 1980.

Two memory aids, a summary and a calendar, were used in the interview. The summary was a computer generated report of the responses to certain key questions in all previous interviews. It was sent to both the interviewer and the respondent prior to the interview in Rounds 2-5. The calendar provided a place for the respondent to record visits, charges, etc. and had a pocket where bills, receipts, and other items could be collected to aid in respondent recall during subsequent interviews.

The NMCUES was designed to produce estimates for evaluation of the impact of legislation and programs on the health status, costs, utilization, and illness-related behavior of persons and families in the medical care system. To this end, the survey instruments contained batteries of questions about a wide variety of topics, as summarized in Figure 2.

After the data were collected, they had to be converted to useful, computer readable form for making national and, for the SMHS, State estimates. The information collected in the five interviews was first keyed, edited, regrouped, and reduced to 18 data files. Since then, through several iterations, they have been more thoroughly edited, cleaned, weighted to represent the total civilian noninstitutionalized population of the United States, and further reduced to seven analytic data files, six of which are currently available as Public Use Data tapes.

Missing data have been imputed for some variables and some of the raw data have been recoded and grouped to allow for greater ease of use.

The ARS collected data from three main sources: (1) the four SMHS State's Medicaid claims and eligibility files; (2) the National Medicare claims and eligibility files; and (3) the Mediaid eligibility files in the remaining States in the HHS sample.

The ARS data, regardless of source, have been gathered together, reduced to a more or less common set of data tapes, and matched (mostly by hand matching) to the HHS and SMHS data collected in the households. RTI is using these data to prepare "Best Estimates" of use and expenditures for the SMHS sample.

The NMCUES is three component surveys: the HHS, SMHS, and ARS; sponsored by two Federal health agencies with somewhat different needs for data and purposes and motifications for supporting the survey: NCHS and HCFA; contracted to and conducted by three private research firms with somewhat different perspectives, experiences, and skills: RTI, NORC, and SMI. Yet, out of this diversity, we have created a data base. A data base which we hope will be useful in the effort to understand the use and financing of health care services in this country.

REFERENCE

Bonham, G.S., National Center for Health Statistics: Procedures and Questionnaires of the National Medical Care Utilization and Expenditure Survey. National Medical Care Utilization and Expenditure Survey. Series A, Methodological Report No. 1. DHHS Pub. No. 83-20001. Public Health Service. Washington. U.S. Government Printing Office, Mar. 1983.

Table 1: Average Length of Interview in the National Medical Care Utilization and Expenditure Survey by Survey Component and Interview Round: 1980

Interview Round	Length of Interview	
	HHS	SMHS: New York, Michigan, Texas, California
1	1.4 hrs.	1.8 hrs.
2	1.2 hrs.	1.5 hrs.
3	1.0 hrs.	1.2 hrs.
4	0.8 hrs.	1.0 hrs.
5	1.3 hrs.	1.5 hrs.

Table 2: Response Rates in the National Medical Care Utilization and Expenditure Survey by Survey Component and Interview Round: 1980

Interview Round	Survey Component				
	National	New York	Michigan	Texas	California
	Percent Responding				
1*	91.1	79.8	82.6	96.9	87.3
2**	99.5	98.8	99.2	99.3	98.6
3**	97.9	96.5	96.3	96.8	94.6
4**	97.1	94.8	94.8	94.4	92.8
5**	96.5	94.2	93.9	93.8	91.6

*Base: Eligible Reporting Units
**Base: Key Persons in Initially Responding RU's

Figure 1: Interview Rounds in WHich Major Data Collection Instruments Were Used in the National Medical Care Utilization and Expenditure Survey: 1980

Instruments	Rounds Used
Control card	All
Core questionnaire	All
Summary	2-5
Supplement 1	All*
Round 3 supplement	3
Round 5 supplement	5
Calendar	All

*For new persons only after Round 1

Figure 2: Selected Topics About Which Data Were Collected in the National Medical Care Utilization and Expenditure Survey: 1980

- Disability days
- Visits to dentists, physicians, outpatient departments, emergency rooms and other medical providers
- Hospital stays
- Expenditures associated with visits and hospital stays
- Expenses for medicines and certain medical appliances
- Health insurance coverage
- Limitation of activity
- Functional limitation
- Access to care
- Barriers to care
- Demographic characteristics (race, income, employment)

THE SURVEY DESIGN AND ESTIMATION METHODOLOGY FOR THE NATIONAL MEDICAL CARE UTILIZATION AND EXPENDITURES SURVEY

Robert J. Casady, National Center for Health Statistics

SURVEY BACKGROUND

The National Medical Care Utilization and Expenditure Survey (NMCUES) was designed to collect data about the U.S. civilian noninstitutionalized population during 1980. During the course of the survey, information was obtained on health, access to and use of medical services, associated charges and sources of payment, and health insurance coverage. The survey was co-sponsored by the National Center for Health Statistics and the Health Care Financing Administration. Data collection was provided under contract by the Research Triangle Institute (RTI), and its subcontractors, National Opinion Research Center (NORC), and SysteMetrics, Inc.

The basic survey plan for NMCUES drew heavily on two surveys; the National Health Interview Survey (NHIS) which is conducted by the National Center for Health Statistics and the National Medical Care Expenditure Survey (NMCES) which was co-sponsored by the National Center for Health Services Research and the National Center for Health Statistics.

The NHIS is a continuing multipurpose health survey first conducted in 1957. The primary purpose of NHIS is to collect information on illness, disability and the use of medical care. Although some information on medical expenditures and insurance payments have been collected in NHIS, the cross-sectional nature of the NHIS survey design is not well suited for providing annual data on expenditures and payments.

NMCES was a panel survey in which a sample of households were interviewed six times over an 18 month period in 1977 and 1978. NMCES was specifically designed to provide comprehensive data on how health services were used and paid for in the United States in 1977.

NMCUES is similar to NMCES in survey design and questionnaire wording, so that analysis of change during the three years between 1977 and 1980 is possible. Both NMCUES and NMCES are similar to NHIS in terms of question wording in areas common to the three surveys. Together they provide extensive information on illness, disability, use of medical care, costs of medical care, sources of payment for medical care, and health insurance coverage at two points in time.

SAMPLE DESIGN OF NMCUES

General Plan - The NMCUES sample of housing units and group quarters, hereafter jointly referred to as dwelling units, is a concatenation of two independently selected national samples, one provided by RTI and the other by NORC. The sample designs used by RTI and NORC are quite similar with respect to principal design features; both can be characterized as stratified, four-stage area probability designs. The principal differences between the two designs are the type of stratification variables and the specific definitions of sampling units at each stage. The salient design features of the two sample surveys are summarized in following sections.

Target population - All persons living in a sample dwelling unit at the time of the first interview contact became part of the national sample. Unmarried students 17-22 years of age who lived away from home were included in the sample when their parent or guardian was included in the sample. In addition, persons who died or were institutionalized between January 1 and the date of first interview were included in the sample if they were related to persons living in the sampled dwelling units. All of these persons were considered "key" persons, and data were collected for them for the full 12 months of 1980 or for the proportion of time they were part of the U.S. civilian noninstitutionalized population. In addition, babies born to key persons were considered key persons, and data were collected for them from the time of birth. Relatives from outside the original population (i.e., institutionalized, in the Armed Forces, or outside the United States between January 1 and the first interview) who moved in with key persons after the first interview were also considered key persons, and data were collected for them from the time they joined the key person. Relatives who moved in with key persons after the first interview but were part of the civilian noninstitutionalized population on January 1, 1980, were classified as "nonkey" persons. Data were collected for nonkey persons for the time that they lived with a key person but because they had a chance of selection in the initial sample, their data are not used for general person-level analysis. However, data for nonkey persons are used in family analysis because they do contribute to the family's utilization of and expenditures for health care during the time they are part of the family.

Persons included in the sample were grouped into "reporting units" for data collection purposes. Reporting units were defined as all persons related to each other by blood, marriage, adoption, or foster care status and living in the same dwelling unit. The combined NMCUES sample consisted of 7,244 eligible reporting units of which 6,599 agreed to participate in the survey. In total, data were obtained on 17,123 key persons. The RTI sample yielded 8,326 key persons and the NORC sample 8,797.

RTI SAMPLE DESIGN

Primary Sampling Units (PSU's) - A PSU was defined as a county, a group of contiguous counties or parts of counties with a combined minimum 1970 population size of 20,000. A total of 1,686 nonoverlapping RTI PSU's exhaust the land area of the 50 states and Washington, D.C. The PSU's were classified as one of two types; the 16 largest Standard Metropolitan Statistical Areas (SMSA's) were designated as self-representing PSU's and the remaining 1,670 PSU's in the primary sampling frame were designated as non-self-representing PSU's.

Stratification of PSU's - PSU's were grouped into strata whose members tend to be relatively alike within strata and relatively unlike between strata. PSU's derived from the 16 largest SMSA's were of sufficient 1970 population size to be treated as primary strata. The 1,659 non-self-representing PSU's from the continental United States were stratified into forty-two approximately equal sized, primary strata. Each of these primary strata had a 1970 population size of about three and one-third million. One supplementary primary stratum of 11 PSU's, with a 1970 poplulation size of about one million, was added to the RTI primary frame to include Alaska and Hawaii.

First Stage Selection of PSU's - The total RTI primary sample consisted of 59 PSU's of which 16 were self-representing PSU's. The non-self-representing PSU's were obtained by selecting one PSU from each of the 43 non-self-representing primary strata. These PSU's were selected with probability proportional to 1970 population size.

Secondary Stratification - In each of the 59 sample PSU's the entire PSU was divided into non-overlapping smaller area units called secondary sampling units (SSU's). Each SSU consisted of one or more 1970 Census defined enumeration districts (ED's) or block groups (BG's). Within each PSU the SSU's were ordered and then partitioned to form approximately equal sized secondary strata. Two secondary strata were formed in the non-self-representing PSU drawn from Alaska and Hawaii and four secondary strata were formed in each of the remaining 42 non-self-representing PSU's. Thus the non-self-representing PSU's were partitioned into a total of 170 secondary strata. In a similar manner the 16 self-representing PSU's were partitioned into 144 secondary strata.

Second stage selection of SSU's - One SSU was selected from each of the 144 secondary strata covering the self-representing PSU's and two SSU's were selected from each of the remaining secondary strata. All second stage sampling was with replacement and with probability proportional to the SSU's total noninstitutionalized population. The total number of sample SSU's was 2 x 170 + 144 = 484.

Third stage selection of areas and segments - First each SSU was divided into smaller nonoverlapping geographic areas and one area within the SSU was selected with probability proportional to 1970 total number of housing units. Next, one or more nonoverlapping segments of at least 60 housing units (HU's) were formed in the selected area. One segment was selected from each SSU with probability proportional to the segment HU count. In response to the sponsoring agencies request that the expected household sample size be reduced, a systematic sample of one-sixth of the segments was deleted from the sample. Thus, the total third stage sample was reduced to 404 segments.

Fourth stage selection of housing units - All of the dwelling units within the segment were listed and a systematic sample of dwelling units was selected. The procedures used to determine the sampling rate for segments guaranteed that all dwelling units had an approximately equal overall probability of selection. All of the reporting units (RU's) within the selected dwelling units were included in the sample.

NORC SAMPLE DESIGN

Primary Sampling Units (PSU's) - The land area of the 50 States and Washington, D.C. was divided into nonoverlapping PSU's. A PSU consisted of SMSA's, parts of SMSA's, counties, parts of counties, or independent cities. Grouping of counties into a single PSU occurred when individual counties had a 1970 population of less than 10,000.

Zoning of PSU's - The PSU's were classified into two groups according to metropolitan status (SMSA, Non-SMSA). These two groups were individually ordered and then partitioned into zones with a 1970 census population size of 1,000,000 persons.

First stage zone selection of PSU's - A single PSU was selected within each zone with a probability proportional to its 1970 population. It should be noted that this procedure allows a PSU to be selected more than one time. For instance, a SMSA PSU with a population of 3,000,000 will be selected at least twice and possibly as many as four times. The full general purpose sample contained 204 PSU's. These 204 PSU's were systematically allocated to four subsamples of 51 PSU's. The final set of 76 sample PSU's was chosen by randomly selecting two complete subsamples of 51 PSU's; one subsample was included in its entirety and 25 of the PSU's in the other subsample were selected systematically for inclusion in NMCUES.

Second stage zone selection of SSU's - Each of the PSU's selected in the first stage was partitioned into a nonoverlapping set of SSU's defined by BG's, ED's, or a combination of the two types of Census units. SSU's were selected from the ordered list of these SSU's. The cumulative number of households in the second stage frame for each PSU was divided into eighteen zones of equal width. One SSU had the opportunity to be selected more than once as was the case in the PSU selection. If a PSU had been hit more than once in the first stage, the second stage selection process was repeated as many times as there were the first stage hits. 405 SSU's were identified by selecting five SSU's from each of the 51 PSU's in the subsample included in its entirety and six SSU's from each of the 25 PSU's in the group for which only one half of the PSU's were included.

Third stage selection of segments - The selected SSU's were subdivided into area segments with a minimum size of 100 housing units. One segment was then selected with probability proportional to the estimated number of housing units.

Fourth stage selection of housing units - Sample selection at this level was essentially the same as for the RTI design.

COLLECTION OF DATA

Field operations for NMCUES were performed by RTI and NORC under specifications established by the co-sponsoring agencies. The sample dwelling units were interviewed at approximately 3 month intervals beginning in February, 1980 and ending March, 1981. The core questionnaire was

administered during each of the five interview rounds to collect data on health, health care, health care charges, sources of payment and health insurance coverage. A summary of responses was used to update information reported in previous rounds. Supplements to the core questionnaire were used during the first, third and fifth interview rounds to collect data that did not change during the year or that were needed only once. Approximately 80 percent of the third and fourth round interviews were conducted by telephone, all remaining interviews were conducted in person. The respondent for the interview was required to be a household member, 17 years of age or older. A non-household proxy respondent was permitted only if all eligible household members were unable to respond because of health, language or mental condition.

IMPUTATION

Nonresponse in panel surveys such as NMCUES occurs when sample individuals refuse to participate in the survey (total nonresponse), when initially participating individuals drop out of the survey (attrition nonresponse) or when data for specific items on the questionnaire are inadvertantly not collected (item nonresponse). In general response rates for NMCUES were excellent: approximately 90 percent of the sample RU's agreed to participate in the survey and approximately 94 percent of the individuals in the participating RU's supplied complete annual information. Even though the overall response rates are quite high, survey based estimates of means and proportions may be biased if nonrespondents tend to have different health care experiences than respondents or if there is a substantial response rate differential across subgroups of the target population. Furthermore, annual totals will tend to be underestimated unless allowance is made for the loss of data due to nonresponse.

Two methods commonly used to compensate for survey nonresponse are data imputation and the adjustment of sampling weights. For NMCUES, data imputation was used to compensate for attrition and item nonresponse, and weight adjustment was used to compensate for total nonresponse. The calculation of the weight adjustment factors will be discussed in the section on sampling weights.

Attrition imputation - A specialized form of the sequential hot deck imputation method was used for attrition imputation. First, each sample person with incomplete annual data (hereafter referred to as a "recipient") was linked to a sample person with similar demographic and socioeconomic characteristics who had complete annual data (hereafter referred to as a "donor"). Secondly, the time periods for which the recipient had missing data were divided into two categories: imputed eligible days and imputed ineligible days. The imputed eligible days were those days for which the donor was eligible (i.e., inscope) and the imputed ineligible days were those days for which the donor was ineligible (i.e., out of scope). The donor's medical care experiences such as medical provider visits, dental visits, hospital stays, etc., during the imputed eligible days were imputed into the recipient's record for those days. Finally the results of the attrition imputation were used to make the final determination of a person's respondent status. If more than two-thirds of the person's total eligible days (both reported and imputed) were imputed eligible days, then the person was considered to be a total nonrespondent and the data for the person was removed for the analytic data file.

Item Imputation - The data collection methodology and field quality control procedures for NMCUES were designed so that the data would be as accurate and complete as possible subject to real world cost consideration. However, individuals cannot report data that are unknown to them, or they may choose not to report the data even if it is known. This latter situation is especially true for data relating to expenditures, income and other sensitive topics. Due to the size and complexity of the NMCUES data base it was not feasible, from a cost standpoint, to replace all missing data for all data items. The twelve-month data files, for example, contain approximately 1,400 data items per person. With these facts in mind, the NMCUES approach was to designate a subset of the total items on the data base for missing data imputations. Thus, for five percent of the NMCUES data items, the responses were edited and missing data imputed by a combination of logic and hot-deck procedures to produce revised variables for use in analysis. Items for which imputations were made cover the following data areas:

1. visit charges
2. source of payment codes and amounts
3. annual disability days
4. health insurance premium amount
5. length of hospital stay
6. total weeks worked in 1980
7. average hours worked per week
8. educational level
9. hispanic ethnicity
10. income
11. age and birthdate
12. race
13. sex
14. health insurance coverage
15. visit dates

These items were selected as the most important variables for statistical analyses.

WEIGHTING AND ESTIMATION

For the analysis of NMCUES data, sample weights are required to reflect the complex sample design and to adjust for the potential biasing effects of systematic nonsampling errors related to total nonresponse and sampling frame undercoverage. Data imputation procedures as discussed in the preceding section, were used to compensate for attrition and item nonresponse.

Basic sample design weights - Development of weights reflecting the sample design of NMCUES was the first step in the computation of person level analytical weights. The basic sample design weight for a dwelling unit is the product of four weight components which correspond to the

four stages of sample selection. Each of the four weight components is the inverse of the probability of selection at that stage, when sampling was with-out replacement, or the inverse of the expected number of selections, when sampling was with replacement and multiple selections of the sample unit was possible.

Two sample adjustment factor - As previously discussed, the NMCUES sample is comprised of two independently selected samples. Each sample, together with its basic sampling weights, yields independent unbiased estimates of population parameters. As the two NMCUES samples were of approximately equal size, it was decided to use a simple average of the two independent estimators for the combined sample estimator. This is equivalent to defining an adjusted basic weight by dividing each basic sample weight by two. Hereafter we will consider only the combined sample and the adjusted basic weights.

Total nonresponse/undercoverage adjustment - This weight adjustment factor is computed at the RU level. As every RU within a dwelling unit is included in the sample, the adjusted basic weight assigned to an RU is simply the adjusted basic weight for the dwelling unit in which the RU is located. As noted earlier an RU was classified as responding if the RU initially agreed to participate in NMCUES and as nonresponding otherwise.

Initially 96 RU weight adjustment cells were formed by cross-classifying the following RU variables: Race of RU head (2 levels), Type of RU head (3 levels), Age of RU head (4 levels) and Size of RU (4 levels). These cells were then collapsed to 63 cells so that each cell contained at least 20 responding RUs.

The formula for computing the total nonresponse/undercoverage adjustment factor, for RU's in cell C was

$$A_1(C) = CPS(C) / \sum_{k \in C} \phi(k)\, W_1(k)$$

where

$CPS(C)$ = March 1980 Current Population Survey estimate of the number of RU's in cell C;

$$\phi(k) = \begin{cases} 1 & \text{if } k^{th} \text{ RU was classified as responding} \\ 0 & \text{otherwise,} \end{cases}$$

and $W_1(k)$ = the adjusted basic weight for the k^{th} RU.

The nonresponse/undercoverage adjusted weight for the k^{th} RU, denoted by $W_2(k)$, was then computed as the product of the adjusted basic weight for k^{th} RU and the nonresponse/undercoverage adjustment factor for the cell containing the RU.

Post stratification adjustment - This weight adjustment factor is computed at the person level. As each person within an RU is included in the sample, the nonresponse/undercoverage adjusted weight for a sample person is the nonresponse/undercoverage adjusted weight for the RU in which the person resides. Each person was classified as responding or nonresponding as discussed in the section on attrition imputation.

Initially, 60 post-strata were formed by cross-classifying the following three variables: Age (15 levels), race (2 levels) and sex (2 levels). One post strata (black, males over 75 years) had less than 20 respondents so it was combined with an adjacent post-stratum (black, males, 65-74 years) resulting in 59 post-strata.

Estimates, based on the 1980 Census, of the U.S. civilian noninstitutional population by age, race and sex for February 1, May 1, August 1 and November 1, 1980 were obtained from the Bureau of the Census. The mean of the mid-quarter population estimates for each of the post-strata was computed and used as the 1980 average target population in calculating the post-strata adjustment factors.

Similarly, survey based estimates of the average post-strata population were developed using the nonresponse/undercoverage adjusted weights. First, a survey based estimate of the target population of post-stratum p at mid-quarter q was computed as follows:

$$S(p,q) = \sum_{j \in p} \delta(q,j) W_2(j)$$

where $$\delta(q,j) = \begin{cases} 1 & \text{if survey respondent j was in scope at mid-quarter q} \\ 0 & \text{otherwise} \end{cases}$$

and $W_2(j)$ = nonresponse/undercoverage adjusted weight of respondent j.

The survey based estimate of the 1980 average population for post-stratum p was computed as the mean of the four mid-quarter estimates, or

$$S(p) = \left[\sum_{q=1}^{4} S(p,q)\right]/4.$$

The post-stratification adjustment factor for the p^{th} post-stratum was then computed as

$$A_2(p) = C(p) \,/\, S(p)$$

where $C(p)$ = Mean 1980 population for post-stratum p based on Bureau of Census data.

The post-stratified weight for the j^{th} respondent, denoted by $W_3(j)$, was then computed as the product of the nonresponse/undercoverage

adjusted weight for the j^{th} respondent and post-stratification adjustment factor for the post-strata containing the respondent.

Adjustment for proportion of days eligible - For many analyses estimates of the average 1980 population are required. Since some respondents were eligible for only a portion of the year the aggregation of the W_3 weights over all respondents will estimate the total 1980 population ever inscope and will over estimate the average 1980 population. Therefore an adjustment factor was calculated for each respondent to reflect the proportion of time during 1980 the respondent was eligible to report NMCUES data. This adjustment factor for respondent j is

$$A_3(j) = E(j) / 366$$

where $E(j)$ = number of days during 1980 respondent j was inscope.

ESTIMATORS

Weighted linear estimators - This type of estimator is used for estimating population and population subdomain aggregates. Suppose, for example, an estimate of the parameter "total doctor visit charges for persons 65 years and older" is desired.

The estimator of this parameter denoted by $\hat{\theta}$, is given by

$$\hat{\theta} = \sum_{j \in A} W_3(j)\, x_j$$

where A is the collection of all NMCUES respondents 65 years and over and X_j is the total doctor visit charges reported by the j^{th} respondent during their eligible period.

Ratio estimators - This type of estimator is used for estimating population and population subdomain parameters such as means, proportions and rates. As will be illustrated in the following examples care must be taken in determining the appropriate weights to be used in the denominator of the ratio estimator.

Example 1 - The NMCUES estimator for the proportion of doctor visits attributable to persons 65 years and older is given by

$$\hat{\theta} = \sum_{j \in A} W_3(j)\, y_j \Big/ \sum_{All j} W_3(j)\, y_j$$

where y_j is the number of doctor visits reported by the j^{th} respondent.

Example 2 - The NMCUES estimator for mean annual doctor visit charges for persons 65 years and older is given by

$$\hat{\theta} = \sum_{j \in A} W_3(j)\, x_j \Big/ \sum_{j \in A} W_3(j) A_3(j)$$

where x_j is the total doctor visit charges reported by the j^{th} respondent during their eligible period and $A_3(j)$ in the time adjustment factor for the j^{th} respondent. The time adjustment factor is used in this situation to adjust for the fact that the j^{th} respondent contributed doctor visit charges to the numerator only during their period of eligibility.

COMPENSATING FOR MISSING DATA IN THE NMCUES

Brenda G. Cox, Research Triangle Institute

If nonrespondents as a group answered survey questions similarly to respondents, missing survey data would not present a problem for analysis. With similar response patterns for the two groups, respondent data could be analyzed directly and used to make inferences to the total population. However, this is seldom the case. Hence data analyses must make adjustments to account for the differential characteristics of nonrespondents and respondents. With 1980 health care data collected in five interview rounds, the National Medical Care Utilization and Expenditure Survey (NMCUES) encountered three forms of nonresponse: completely missing annual data, partially missing annual data, and missing item data. To compensate for nonresponse, NMCUES used weight adjustments, logical editing, statistical imputation, and administrative records. This paper will describe the nonresponse that occurred for the NMCUES and the procedures that were used to compensate for the nonresponse.

1. Description of the NMCUES

The rapidly rising cost of medical services in the United States in recent years, together with a continuous effort to improve the quality, effectiveness, and availability of health care, has lead to a continuing need for comprehensive data for individuals and families on health status, patterns of health care utilization, charges for services received, and payers and amount paid. Sponsored by the Health Care Financing Administration (HCFA) and the National Center for Health Statistics (NCHS), the National Medical Care Utilization and Expenditure Survey (NMCUES) was the second of a series of national medical care surveys planned to provide data on a regular basis. These surveys will permit in-depth statistical descriptions of the utilization of health care services and the associated costs for various population segments, including the nation as a whole. They will also provide valuable data for the evaluation of current public programs such as Medicare and Medicaid, for the assessment of inequity in access to the health care delivery system and other unmet needs, and for the comparison of alternative solutions to health policy issues. The findings from these studies will ultimately have an impact on public policy concerning health care for the entire nation.

Conducted by the Research Triangle Institute (RTI) in conjunction with the National Opinion Research Center (NORC) and SysteMetrics, Incorporated, the NMCUES National Household Survey (HHS) was based upon a stratified cluster sample of 7,600 dwelling units selected so as to represent the civilian, noninstitutionalized residents of the United States in 1980. Repeat interviews were conducted with the initial panel of 6,600 responding households at approximately twelve-week intervals beginning in early 1980 and ending in mid-1981. In five rounds of data collection, information was collected on demographic characteristics and 1980 health care utilization and expenditures.

In addition to this household survey, NMCUES also included four State Medicaid Household Surveys (SMHS) of Medicaid beneficiaries. Using administrative record data provided by California, Michigan, New York, and Texas, a clustered list sample of Medicaid cases was selected from each state. The procedures developed for selecting Medicaid cases yielded aid-category balanced samples of 1,000 cooperating Medicaid households per state. Using the same instrument and data collection procedures as the national household survey, 1980 health care data were collected for the four state samples of Medicaid recipients and their households.

2. Procedures for Total Nonresponse

Total nonresponse refers to the situation in which the entire set of survey data is missing for an analysis unit. In NMCUES the units of analysis are persons and families. In both cases, total nonresponse was chiefly the result of the loss of entire households in Round 1 through refusal, nonavailability, or physical/mental barriers to response. From eligible households in Round 1 of the NMCUES, a 91 percent response rate was obtained in the national household survey and rates ranging from 80 to 97 percent in the State Medicaid Household Surveys. Table 1 summarizes other details of the Round 1 data collection experience for the five survey components. These rates are for reporting units. For data collection convenience, a reporting unit (RU) was defined to be a family unit living within the same dwelling unit. A family with a college student living away from home had interviewing occurring at two locations, for instance, so that data collection was easiest when a reporting unit was defined for each of the two locations.

Total nonresponse is best handled through a weighting procedure. For analysis of NMCUES data, sample weights were used to reflect the complex sample design and to adjust for the potential biasing effects of nonresponse and undercoverage. Undercoverage errors result when the list of units comprising the sampling frame do not provide access to all the eligible target population members. The NMCUES analysis weights were computed as inflation factors to represent the number of units in the survey population that were accounted for by the sample unit to which the weight was assigned. The initial weight for each sample individual was defined as the inverse of the individual's overall selection probability. This initial weight was adjusted in later steps to account for undercoverage and nonresponse.

Both nonresponse and undercoverage occur at differing rates for demographic domains defined by individual characteristics such as age, race, and sex and household characteristics such as size and type of family. To compensate for the differential impact of nonresponse and undercoverage on domain estimates of means and totals, adjustment factors were calculated to inflate the weights of respondents to known population totals for these demographic domains.

Census data and Current Population Survey data were used in making these adjustments.

3. Procedures for Partial Nonresponse

The second type of missing data that the NMCUES encountered was partial nonresponse in which individuals initially responded but failed to complete all of the data collection rounds. Over the five rounds of data collection, attrition of initially responding individuals occurred at a rate of approximately five percent for the national household sample (HHS) and at a somewhat larger rate for the four State Medicaid Household Survey (SMHS) samples. Even this low a level of missing data could not be ignored since many NMCUES analyses will use data aggregated over the full year. For example, estimating the distribution of 1980 medical care expenditures requires that complete annual data be available. In performing an analysis requiring annual data, two options are available. One could analyze only the complete data after reweighting the complete data records to compensate for the removal of partial data records. The second option is to perform the analysis using the full data set after imputing for the partially missing data.

The simplest solution to implement restricts analyses to individuals with complete data with reweighting of the complete data records to compensate for the removal of partial data records. This can be wasteful when many individuals have data that are almost complete. Further, since death and institutionalization result in high levels of survey attrition, valuable information may be lost which cannot be adequately compensated for by sample weighting.

Another option is to perform analyses using the full data after imputing data for time periods for which data are missing. This approach has the disadvantage that nearly the entire year's data may be missing for some individuals and hence almost all their data will be imputed. There would seem to be little information gained by including individuals with almost but not all their data missing. Further, imputation can have a greater variance inflation potential for survey estimates than weight adjustment procedures.

NMCUES decided that a combination of reweighting and imputation was the best approach. The two types of analysis units used in analyzing NMCUES data - individuals and families - have required two sets of sample weights. Each weighting task defined the concept of respondent so that units with data for less than one-third of the year are considered to be total nonrespondents and the data that they did provide are not used. Individuals with data for one-third or more of the year were classified as respondents and any missing annual data were replaced in the attrition imputation task.

In order to replace data missing due to survey attrition, it was necessary to determine the dates for which each individual responded, the dates for which they were ineligible, and the dates for which they failed to respond. This information was used to determine whether or not individuals had provided data for all of the 1980 time period in which they were eligible for data collection.

NMCUES sample data were analyzed to obtain insight into the variables that were important with respect to modeling partial nonresponse. Estimates of the average number of annual medical visits were made to identify variables that had levels that were correlated with annual utilization. This analysis is summarized in Table 2. The variables that appeared to be most related to utilization were age, sex, number of medical visits in the first quarter, health plan coverage and for the SMHS samples, State Medicaid Aid category. Annual complete data rates were also calculated for each of the variables. These annual completion rates are summarized in Table 3. Overall, the rates were quite high; 94 percent of the individuals from responding Round 1 reporting units provided complete data in the HHS and from 86 to 92 percent in the SMHS.

The variables shown in these tables were thought relevant to the number of medical events that would be missed for a person with incomplete data, and hence potential classing variables in the imputation. The imputation for missing annual data was set up so that it occurred within groups or classes of people who were thought to be similar with respect to health care utilization and expenditures. In forming imputation classes, the overall goal is to form classes for which responses are homogeneous within each class, heterogeneous between classes, and for which the rate of missing data varies. Further, the characteristics used to define the classes have to be known for both respondents and nonrespondents.

Attrition imputations were made using a weighted hot deck imputation procedure (Cox, 1980). This imputation strategy may be thought of as utilizing two data files, a data file of respondents (donors) and a data file of nonrespondents (recipients). The imputation procedure uses data for responding individuals to substitute for missing data for nonresponding individuals. The imputation occurs within imputation classes so that the distribution of means and proportions is preserved within each class over repeated imputations. The number of times that the data for a donor is accessed to impute to recipients is defined as a function of the sampling weight for the donor and of the recipients to which the information can potentially be imputed. For time periods for which the recipient had data missing, the visits (if any) reported by the donor for the same time period were imputed to the recipient. If the donor was ineligible to provide data for any part of the time period for which the recipient had data missing, the recipient was imputed to be ineligible during that part of the time period.

Unlike item-level imputation where particular item responses were created, the attrition imputation procedure created complete records for the recipient person based upon the records of the donor person. Health provider visits, health service charges, and other health care experiences were imputed as "blocks" of data from the donor to the recipient. The data files involved in the imputation were those associated with medical provider visits, dental visits, hospital stays, prescribed medicines, and other

medical expenses.

4. Treatment of Item Nonresponse.

Health care estimates obtained using household data are known to be subject to bias due to the inability of many respondents to accurately recall past health care events. Household-reported expenditure and utilization data are especially prone to response errors since individuals may not be able to recall the charges associated with visits or may fail to recall all visits that were made. To minimize such sources of response error, NMCUES used a calendar/diary and computer-generated summaries mailed to each household prior to each data collection round. In addition, NMCUES also used a relatively short recall period of twelve weeks so that basically one can say that the survey data would appear to be as accurate as the households could provide. However, the respondents could not report data that were unknown to them. For example, respondents who were enrolled in public programs such as Medicare and Medicaid frequently did not have access to provider bills or other documents that contained cost data associated with their visits. In other instances, the respondents did not know (or chose not to state) the charges associated with a visit.

Because of the size and complexity of the NMCUES data base, it was not feasible, from a cost standpoint, to replace all missing data for all data items. Neither was it reasonable to use only those records with absolutely complete data; over the 1,400 NMCUES data items only a few individuals could be expected to provide complete data. With these facts in mind, the NMCUES approach was to designate a subsample of the total items on the data base as important enough to merit missing data imputations. For five percent of the NMCUES data items, the responses were edited and missing data replaced in order to produce imputation-revised variables for use in analysis. Items for which imputations were made cover the following data areas:

- age and birthdate,
- race and Hispanic origin,
- sex,
- educational level,
- employment status,
- disability days,
- nights hospitalized,
- health insurance,
- income, and
- health care charges.

These items were selected as the most important variables for statistical analyses.

Because the number of items subject to imputation was large, the items were divided into sets and imputations performed within those sets. Every imputation and editing operation was conducted independently within the HHS and SMHS samples. For each sample, the data records were usually partitioned into classes with the imputation occurring independently for each class. Within each class the data records were sorted for additional control over the imputation process.

For each group of data items, the most cost effective imputation strategy was selected. The large number of items for which imputations were required and the costs associated with data processing for the large files involved necessitated that alternative approaches be evaluated in terms of the quality of the imputations and the associated costs. By reducing the number of passes through the data, some approaches could drastically reduce data processing costs while producing results that were essentially comparable in quality to the best approach.

Item imputations were made using three procedures: logical editing, hot deck imputation, and weighted hot deck imputation.

Logical editing were used whenever similar information were available in the data record. For example, when sex was missing, gender-specific relationship to head responses were used for imputation.

Hot deck imputation were usually used when the level of missing data was small. This procedure partitions respondents into imputation classes by characteristics related to the item being imputed. Within each class, the records are sorted by variables related to the item being imputed. An initial value is determined for each class based upon previous or current data. As the new data are processed, the imputation class to which each individual belongs is determined. If the record being processed is complete, then that record's response is supplied for the cell of the hot deck. When a record is encountered with a missing item, the response in the cell of the hot deck is imputed for the missing response.

A weighted hot deck procedure was also used for item imputation in the NMCUES. This procedure is the same as described for the attrition imputation except single items were imputed rather than an entire record. The weighted hot deck imputation procedure was used whenever the level of missing data was large. This procedure is designed so that within imputation classes the means and proportions estimated from the imputation-revised data will be equal in expectation to the means and proportions estimated using respondent data only. Variances, covariances, correlations, regression coefficients, and other higher order population parameters estimated from the imputation-revised data will also equal the corresponding estimator obtained from the respondent data alone.

5. Use of Administrative Records

Longitudinal surveys of medical utilization and expenditures are subject to several types of measurement errors. The reporting of medical events is subject to both overreporting and underreporting. Events are sometimes telescoped into the survey reference period by respondents, resulting in overreporting of medical events. Also, respondents often do not accurately remember the dates of their visits and the number of visits. This phenomenon can lead to either underreporting or overreporting. Moreover, when Medicare and/or Medicaid pays the charge for an individual's medical expense, the individual often does not know the amount of the charge for the medical service. As a result, there can be a high level of missing medical charge data, particularly for the State Medicaid Household Surveys (SMHS).

To help alleviate these problems, NMCUES included a "best estimation" task that created

augmented data files containing claims data supplemented by survey data for those events not covered by the claims. The Administrative Records Survey (ARS) was performed to obtain Medicaid and Medicare paid claims records and eligibility data from the Health Care Financing Administration and the State Medicaid agencies. These data were obtained by locating records that matched Medicare and Medicaid identification numbers provided by survey respondents. Round 1 Medicare identification numbers were used to extract the Medicare claims data. Due to the unstable nature of the population of Medicaid beneficiaries, Medicaid identification numbers from all five rounds of the survey were used to obtain Medicaid claims. Medicaid identification numbers from the sampling frame were also used for the originally selected Medicaid beneficiaries. Additional Medicaid matching was also done using name, address, and demographic data.

Respondent reported survey data were then matched to the claims data. The matched survey and claims data were used to produce augmented data files containing three components: (1) a survey component, (2) a claim component, and (3) a best estimate component. The primary variables in the best estimate component are the total charge, sources of payment, and amounts paid. The best estimates are formed as a judicious combination of survey and claims data.

For each SMHS State, four event-level augmented files were created based upon the Medicaid claims: (1) a medical visit file, (2) a dental visit file, (3) a hospital stay file, and (4) a prescribed medicine and other medical expenses file. For the SMHS States, the augmented hospital stay file will have a fourth component, a Medicare claim component. Best estimates for the prescribed medicine and other medical expenses file will consist of person-level best estimates, since only person-level matching was feasible for prescribed medicines.

Two person-level best estimate files are based upon the Medicare claims for each SMHS State and for the national household survey (HHS): a medical provider visit and doctor in hospital visit file, and a prescribed medicine and other medical expenses file. Dental visits are not covered by Medicare, so there are no Medicare best estimates for dental visits. The above Medicare best estimate files were produced at the person level since Medicare claims cannot generally be matched to survey data at the event level. A single Medicare claim generally covers a series of medical events over a specified time span, rather than a single medical event. A combined file of medical visits and doctor in hospital visits was created since it is not possible to separate these medical events in the Medicare claims data. The facility claims for hospital stays could be matched at the event level, however. Hence, the Medicare claims were used to produce event-level Medicare best estimates of hospital facility charges for the national household survey (HHS).

The major benefit of the Medicare best estimation is to obtain a more accurate estimate of the burden of medical expenses for Medicare beneficiaries. The person-level best estimates will improve the estimate of total charge, amount paid by Medicare, and amount paid by other sources for Medicare beneficiaries. The event-level Medicare best estimates for hospital stays will produce more accurate charge and expenditure data for individual hospital stays. Data for hospital stays that were not reported by the household, due either to oversight or to attrition, will be obtained from the claims data.

6. Concluding Remarks

Missing survey data result from a variety of sources, each source having implications for the way the missing data should be treated. The NMCUES attempted to deal with all sources of missing data in order to build a complete data set which could be used by analysts both inside and outside the government. In analyzing NMCUES data, the researcher is advised to determine in what ways, if any, the methods used to replace missing data will affect the analysis. This paper has attempted to provide some of the required information to make that determination.

7. Reference

Cox, Brenda G. (1980). The Weighted Sequential Hot Deck Imputation Procedure. Proceedings of the American Statistical Association, Survey Research Methods Section, 721-726.

8. Acknowledgements

The work described in this paper was done for the National Center for Health Statistics (NCHS) and the Health Care Financing Administration (HCFA) under Contract No. HRA-233-79-2032. The views expressed in this paper are those of the author and no official endorsement by NCHS or HCFA is intended or should be inferred.

Table 1. Reporting Unit Interview Final Results in Percentages for Round 1 of the NMCUES

Final Result	HHS	SMHS California	SMHS Michigan	SMHS New York	SMHS Texas
Total Sample	100.0	100.0	100.0	100.0	100.0
Ineligible Reporting Units	13.3	8.8	14.4	5.8	15.0
Eligible Reporting Units	86.7	91.2	85.6	94.2	85.0
Ineligible Reporting Units	100.0	100.0	100.0	100.0	100.0
Vacant Dwelling Unit	49.3	0.0	0.0	0.0	0.0
Demolished Dwelling Unit	12.2	0.0	0.0	0.0	0.0
Merged Dwelling Unit	2.3	0.0	0.0	0.0	0.0
Not a Dwelling Unit	13.5	0.0	0.0	0.0	0.0
Vacation/Second Home	16.1	0.0	0.0	0.0	0.0
Entirely Armed Forces	1.5	0.0	0.0	0.0	0.0
Entirely Ineligible Students	1.4	0.0	0.0	0.0	0.0
Entirely Institutionalized	0.8	82.8	88.4	51.6	89.8
Entirely Deceased	0.7	14.1	9.1	20.9	9.2
Other Out of Sample	2.2	3.1	2.5	27.5	1.0
Eligible Reporting Units	100.0	100.0	100.0	100.0	100.0
Questionnaire Complete	91.1	87.3	82.6	79.8	96.9
Questionnaire Incomplete	8.9	12.7	17.4	20.2	3.1
No One Found at Home	1.0	2.3	2.4	3.9	0.6
Moved and Not Located	0.1	5.4	9.3	0.8	1.0
Language Barrier	0.0+	0.0+	0.0	0.0+	0.0
Physically/Mentally Incompetent	0.2	0.3	0.1	0.8	0.0+
Refusal	7.2	3.9	3.5	3.8	0.9
Breakoff	0.1	0.4	0.2	0.4	0.0+
Other In Sample	0.3	0.3	1.7	0.9	0.4
Percentage Base					
Total Sample	8,359	1,449	1,683	1,560	1,440
Ineligible Reporting Units	1,115	128	242	91	216
Eligible Reporting Units	7,244	1,321	1,441	1,469	1,224

Table 2. Average Number of Medical Visits for the Domains Defined by the Levels of the Potential Classing Variables

Variable	Sample Type HHS	SMHS CA	SMHS MI	SMHS NY	SMHS TX
Total	5.11	7.29	6.82	9.16	4.70
Age					
0-16	3.73	3.54	3.88	4.56	2.27
17-29	4.43	4.83	5.38	8.03	3.60
30-44	4.92	7.48	8.43	14.96	5.60
45-54	5.54	10.03	10.10	12.36	6.69
55-64	6.48	11.44	8.93	13.23	8.17
65+	7.96	11.83	9.75	13.78	7.77
Race					
Black	3.90	7.73	5.25	7.56	3.84
Non-Black	5.26	7.20	7.80	9.91	5.13
Sex					
Male	4.32	5.34	5.75	8.52	3.43
Female	5.83	8.74	7.64	9.60	5.56
Education of Head					
0	6.57	8.23	4.92	8.91	5.54
1-8	5.22	7.79	7.32	10.44	4.76
9-12	4.83	6.85	6.21	8.01	4.26
13-18	5.42	7.63	9.37	10.78	5.23
Number of Medical Visits in 1st Quarter					
0	1.98	2.09	2.09	2.30	1.62
1	4.18	4.23	4.57	4.77	3.76
2	6.53	7.08	7.08	7.04	6.46
3-4	9.53	11.39	10.64	10.75	10.20
5-6	14.56	16.10	15.82	17.27	14.22
7-8	19.02	22.00	21.17	23.14	19.06
9+	38.45	44.62	40.43	70.92	40.32
Self-Reported Health Status					
Excellent	3.64	3.70	3.48	3.51	2.14
Good	5.27	5.42	5.51	6.37	2.86
Fair	9.05	9.66	10.47	13.10	6.84
Poor	13.54	19.55	13.83	23.40	12.31
Health Plan Coverage					
Medicare	8.75	8.32	7.88	10.60	5.84
Other Public	5.45	5.86	7.84	8.82	6.37
Private Plans	4.59	4.32	3.66	3.56	2.67
Uninsured	2.54	2.62	2.80	3.73	1.86
State Medicaid Aid Category					
SSI Aged	N.A.	9.61	8.90	12.85	5.66
SSI Blind/ Disabled	N.A.	10.02	8.86	10.80	6.66
AFDC	N.A.	5.14	5.08	5.46	3.15
State Only	N.A.	6.33	6.28	10.35	N.A.

Table 3. Annual Complete Data Rates for Key Individuals

Variable	Sample Type HHS	SMHS CA	SMHS MI	SMHS NY	SMHS TX
Total	94.0	85.6	90.3	89.2	92.3
Age					
0-16	94.7	86.1	90.4	92.2	91.6
17-29	93.1	81.4	89.6	87.0	88.4
30-44	94.7	84.7	89.8	92.0	91.3
45-54	94.4	85.5	91.5	85.1	93.1
55-64	94.5	90.1	88.5	86.4	95.8
64+	92.7	87.4	91.5	88.0	96.2
Race					
Black	92.2	81.3	85.5	87.7	90.5
Non-Black	94.3	86.5	93.5	89.9	93.2
Sex					
Male	93.9	84.5	89.9	88.3	91.6
Female	94.1	86.4	90.6	89.7	92.8
Education of Head					
0	82.7	96.5	95.6	89.2	97.1
1-8	94.4	83.3	89.7	89.8	93.9
9-12	93.9	86.4	90.1	89.3	90.1
13+	94.7	85.0	92.5	87.0	83.6
Number of Medical Visits in 1st Qtr.					
0	93.2	84.6	89.2	88.0	92.0
1	95.1	84.0	93.2	89.0	93.5
2	94.4	87.9	90.1	89.0	92.2
3-4	95.6	88.4	91.9	93.2	93.2
5-6	95.9	85.4	91.9	92.4	87.6
7-8	94.4	92.3	91.3	85.7	92.4
9+	93.6	84.9	83.6	87.6	93.8
Self-Reported Health Status					
Excellent	95.1	86.5	91.1	90.7	90.6
Good	93.9	86.5	92.4	91.0	93.4
Fair	92.9	87.5	88.4	87.2	92.9
Poor	91.0	84.0	86.6	87.7	93.8
Health Plan					
Medicare	92.9	85.8	91.4	89.5	93.4
Other Public	92.9	81.0	83.9	80.0	93.6
Private	95.3	87.0	91.0	91.1	90.6
Uninsured	87.2	81.5	81.5	82.2	89.6
State Medicaid Aid Category					
SSI Aged	N.A.	85.2	91.2	87.7	91.6
SSI Blind/ Disabled	N.A.	87.1	88.5	89.3	95.9
AFDC	N.A.	85.2	91.5	91.8	91.2
State Only	N.A.	85.2	87.5	86.3	N.A.

Automation in Vital Registration and Statistics

Session K

VITAL STATISTICS SYSTEMS ANALYSIS FOR COMPUTER TECHNOLOGY OF THE '80s

Bee Biggs, Vital Statistics Idaho
Andy Desilet, Systems Management Idaho
George Myers, Prodata, Inc. Idaho

INTRODUCTION

On January 1, 1984, the Idaho Registrar of Vital Statistics will launch a new, fully automated Vital Statistics System that revolutionizes Idaho's Vital Records System, responding to data needs in a matter of hours rather than months, as is the case with the existing system.

This paper describes the type and value of the analytical approach used by the Idaho Vital Statistics Unit as a system development methodology to determine if full computerization was needed, useful, or cost effective in order to gain increased efficiencies and improved quality and accuracy of Idaho's Vital records. The system analysis was begun in 1982, and completed in early 1983.

What A System Analysis Is

The analysis must examine systematically, the purpose, the inputs, and the desired outputs of the system that will use the data base as a repository of information. Once the analysis describes what is, what is needed now, and what will/may be needed in the future, the system analysts can build a conceptual model that is logical and describes the underlying structure of a rational, hierarchical or network data model. Lawrence Peters in Software Design: Methods and Techniques describes system software designs in four phases:

> "1. System Analysis -- The objective of the analysis phase is to demonstrate that the customer's problem is understood and to document it in a manner that will aid the design phase . . . It is during this phase that the customer's problem is externalized, organized, and played back to him to ensure that the problem is understood.
>
> "2. System Design -- During this phase, the statement of the problem is addressed through the use of software design methods and techniques to obtain a logical or abstract model of the system solution. Implementation issues are not considered as the goal is a clear perception of a solution concept.
>
> "3. System Implementation -- This phase begins with packaging of a logical design, followed by implementation of the package design and the target programming language and operating system environment, testing the results and installation.
>
> "4. System Operation -- This phase includes maintenance of the system performance of original tasks, and enhancement to meet changing requirements; the phase leads to the eventual phase-out and replacement of the system."[1]

II. SYSTEMS ANALYSIS

The Beginning

The State registrar selected a mechanism for planning, scheduling, and controlling the analysis and development process into several steps or phases, providing for decision points that fit into an overall development process that describes what is to be accomplished, what the time requirements are for all participants, and what other steps are necessary. This approach, according to Peters, provides project management with milestones and positive feedback during the project.[2]

Using PRIDE System Development Methodology (PRIDE is a proprietary product marketed by M. Bryce and Associates of Cincinnati, Ohio), the Idaho Vital Statistics Unit undertook a nine-phase system evaluation during 1982 and early 1983.

Objectives of the Vital Statistics Unit

The objectives of PRIDE Phase I Analysis were:

1. To determine if computer assistance at the vital records intake point would speed up the availability and improved accuracy of vital records;
2. To determine if computer assistance printing of followup letters, certified copies and statistical reports would be more efficient and decrease the time required to prepare the vital data elements required in the issuance of certified copies; and,
3. To determine if additional health statistics could be made available more rapidly to system users upon request while maintaining confidentiality.

Analysis of Existing Vital Statistics System

The systems analyst team was made available from the Bureau of Systems Management to the State Registrar. This began a critical partnership that was to last throughout the analysis, design and implementation of the new system. Interviews were conducted with all staff members of the Vital Statistics Unit. The legal basis, regulations, and internal procedures were reviewed. Documentation for the current batch environment for data capture and reporting was reviewed. Conceptual models were constructed to communicate the actual document and data pathway within the current system. Frequent meetings between vital records

managers, the State Registrar, and the systems analysts were necessary in order to clarify goals and objectives as they related to the analysis. Written reports of the analysis were compiled and agreed upon by both the analysts and Vital Statistics management.

Why Electronic Data Systems Fail

There are at least seven reasons why systems fail. They are: incorrect requirements; missing, incomplete or inadequate facts about the program requirements; unclear or ambiguous information; inconsistent and incompatible data items; new or changing specifications; requirements that are outside the scope of the project, and finally, typographical errors. PRIDE methodology consciously seeks to avoid or correct for these possibilities throughout each stage of analysis, design, testing and implementation.

By providing for major effort to be directed "up front" toward system study and design, less effort is required for programming, maintenance and modification. Deadlines are predictably met.

III. SYSTEMS DESIGN

Design Overview

The use of the computer is an integral part of the new Vital Statistics System, which is comprised of administrative, manual, and computer procedures necessary for the processing of Vital Statistics documents. The existing batched system with documents centrally processed is complex, expensive, and many of the components have very little capability. With each step of analysis, design took shape. The analysis indicated that there should be one integrated system to service the Vital Statistics Unit, rather than separate systems for births, marriages, divorces, induced terminations of pregnancy, foreign-born adoptions, and deaths. A single system will enable more cross-training of personnel, simplify production scheduling, and development of reports that include cross tabulation of data from several sources such as infant birth/death reports. Using the common data base approach, the single system will enable comparisons of data items, regardless what certificate type the data originated from, and will embrace and process documents and record keeping for all vital records.

The system will be designed to enable the Vital Statistics Unit to retain physical possession of all certificates and vital data at all times.

Processing Time - Public Need

The system will reduce certificate processing time from three months for births and one week for deaths, to not more than one working day from the time a properly completed certificate is received by the Vital Statistics Unit, until a computer-generated version of the document is available. This expedited turnaround is particularly important with death certificates, since generally the decedent's family needs a certified copy of the death certificate as quickly as possible to claim insurance benefits. Most data will be available for reporting and certification one day following receipt of the document by the Vital Statistics Unit.

The present system produces over 80 automated statistical reports to which the new system will add several management reports.

Common Data Base

The system will utilize a common data base approach to eliminate the need for duplicate file updating procedures as currently used. As an example, presently a multi-year magnetic tape of death certificates is updated as well as the death master file.

Since all the certificates have common characteristics, a common data base approach should result in a less complex system than that of the present system, and a system that is more economical to develop and maintain. A common data base should enable relatively easy correlation of data. As an example, the system should provide the capability to access, correlate, and report data pertaining to births without having to initiate a major project.

Six logical data base modules have been defined to support the system. The six modules and their contents are:

1. Event--This is the statistical/code master file for the system.
2. Person--This file contains the names and addresses of every individual associated with each vital event document.
3. Facility--This file contains the names and addresses of all facilities (e.g., hospitals) associated with each vital event document.
4. Official--This file contains the names and addresses of all officials (e.g., physicians) associated with each vital event document.
5. Location--This file contains the code and narrative for geographic locations, races, occupations, industries, and codes for causes of death.
6. Management--This file contains the daily account of all transactions with the public.

Data compression techniques will be used to minimize the overall physical data base system and decrease storage costs.

It will not be practical to maintain all of Vital Statistics data within the data base,

since approximately 51,000 new events will be added per year. Periodically it will be necessary to "unload" older records from the data base. This will allow flexibility to maintain high use data on the data base for a longer period than other data that is needed less often. The system will also have the capability to "reload" all the data when needed if special reporting needs should arise.

The Record Flow

The proposed system will not increase the amount of manual work required of the Vital Statistics Unit personnel; but, in fact, the proposed system should substantially reduce the current amount of manual work associated with the collection of data, data entry and reporting.

The data base will be updated "on-line" by use of CRT terminals in the Vital Statistics Unit. All editing will be performed "on-line." Data which the computer rejects because of errors that cannot be corrected by the operator can be added to the data base. However, if the error concerns the primary key (event record type, event date, and state file number) to a data base module, the record cannot be added to the data base.

Hardware Requirements

The minimum hardware requirements will be three CRT type terminals and a printer. These must be compatible with IBM 3270 technology. The printer will be used to print certified copies of Vital Statistics documents and should be letter quality. This equipment will be located in the Vital Statistics Unit and connected to the controller located in the Bureau of Systems Management.

It is highly desirable to be able to print certified copies of documents throughout the State to improve the services to the public. The function will be added as the hardware is purchased for location in major communities.

Audit Trail on Data

The system must have a comprehensive audit trail physically attached to all data to identify its source and subsequent revisions to the data. The audit trail will tie the automated data to its paper or microfilm counterparts. The audit trail should enable reconstruction or play-back of the sequence of events that may have occurred concerning changes and amendments to a given certificate. The audit trail will contain identification of the CRT operator who entered the data along with the date entry was made.

The system will contain, on-line, the amendments to documents whose data are stored off-line. This will allow an efficient and automated update to the multi-year files. It will also allow less costly storage of less frequently used data.

Microfilm Requirements

Microfilm will be used for archival storage. The system will also print an index to be stored with the archives. The Idaho Vital Statistics Unit contracts with the National Center for Health Statistics to provide that federal agency with an exact duplicate copy of the microfilm for births, deaths, marriages and divorces. Microfilm will be duplicated regularly and accompanied by a computerized transmittal report prepared to NCHS specifications.

Data Base Software

In consultation with the Data Base Administrator, we decided that the current design would be most workable on the product marketed by Software AG "ADABAS" and "Natural." This decision was made based on the complexity of the linkage between the different files required in order to tie the files and associated information together and also for the purpose of file maintenance.

System Crash Recovery

The system will be designed in a manner to eliminate the need to re-key data into the system should a system failure occur that destroys the integrity of the "data base." Copies of all automated files are to be maintained in a secure location other than the computer center. These back-up copies will be updated each night to reflect the previous working day's "data base."

System Security

The system is designed in a manner to prohibit on-line retrieval of data by personnel other than authorized Vital Statistics personnel. System security is composed of many different features that are designed to complement each other and collectively provide an environment that provides system security.

IV. SYSTEM IMPLEMENTATION

Implementation Concepts

Lawrence J. Peters states that, designing software requires both patience and bravery: patience is needed to keep from rushing towards a solution that may, due to haste, be incomplete, and bravery is required because many discoveries will be made as we proceed from problem to solution.

Peter's message for system designers is that "the means of defining a problem should lend itself to the defining of the solution."[3] The following common ideas pervade Peter's recommendations or model of design:

"1. Design is a 'good' thing to do prior to implementation.

"2. Design involves abstraction, including

the use of graphics, mock-ups, prototypes, and physical methodologies to strip away the detail and to get at the essential character of the System.

"3. Some rationale is necessary to focus design activity, make it more effective and assure that its successors will understand what is done.

"4. Design is inexact in that it does not lend itself to the use of formulae or precise estimates.

"5. Design is a creative act, uniquely suited to people rather than automata in that people can bring their entire experience to bear on a new problem.

"6. Design is a discovery process in that as we define our understanding of the problem, and enrich our design to address this new knowledge, we often discover subtle nuances.

"7. Finally, design and analysis (or specifications) are inextricably linked and only artificially inseparable."[4]

V. CONCLUSION

Through a critical partnership, we developed a modern design that responds positively to the objectives of the State Registrar. Using a common data base approach, a single system is proposed that will process documents and record keeping for: live births, stillbirths, induced terminations of pregnancy, adoptions, marriages, divorces, and deaths.

Start-up costs will be approximately $90,000 and thereafter annual cost savings to maintain the system should be about $15,000 per year for at least the first three years.

Cooperative, objective-based, professional analysis and design made this project exciting, satisfying and reproducible.

BIBLIOGRAPHY

1. *Peters, Lawrence J., Software Design: Methods and Techniques, Yourdon Press, New York, 1981, p. 8, 9.*

2. *Ibid, P. 16.*

3. *Ibid, P. 34, 35.*

4. *Ibid, P. 34, 35.*

Other references, not quoted:

1. *Martin, James, Systems Analysis for Data Transmission, Prentice Hall, Inc., New Jersey 1972.*

2. *PRIDE Detail Manual, Proprietary Product by M. Bryce and Associates, Cincinnati, Ohio.*

3. *U.S. Department of Health, Education and Welfare; Public Health Service, National Center for Health Statistics, Model State Vital Statistics Act and Model State Vital Statistics Regulations, 1977, Revision, Hyattsville, Maryland, May 1978.*

4. *Atre, S., Data Base: Structured Techniques for Design, Performance and Management, A. Wiley-Interscience Publication, John Wiley & Sons, New York, 1980, P. 124.*

5. *Ibid, P. 129.*

6. *Model State Vital Statistics Act and Model State Vital Statistics Regulations, National Center for Health Statistics, U.S. Department of Health, Education, and Welfare, Public Health Service, Hyattsville, Maryland, May 1978.*

DEVELOPMENT OF AN IN-HOUSE VITAL STATISTICS MICROCOMPUTER SYSTEM

Edward J. Martin, Rhode Island Department of Health

Introduction

During the last several years, microcomputers have played an increasingly more important role in data processing systems throughout the country. To quote an article from the October 1982 issue of Datamation, titled "Attack of the Micros": "More than ten years have passed since computer power was first compressed into a chip. In that time span, microcomputers have changed the face of the computer industry like nothing else before them..."(1) This same article also cites a market research study which concluded that the market for microcomputer-based management workstations will grow at 40% a year through 1986, with sales of more than $4.5 billion between 1982 and 1986.(2) The main impetus for this microcomputer "explosion" has been the desire to boost production and save money.(3)

It is against this background that I would like to describe a project undertaken by the Division of Vital Statistics in the Rhode Island Department of Health to install an in-house microcomputer information processing system. This new system replaced a main-frame system maintained and operated by a private health data consortium. Although the new system does have some limitations, we have been very satisfied, both with the performance of the new system and with the relative ease of its installation and operation.

By describing the work that has been done in developing and operating the system, I hope to encourage relatively small agencies similar to ours to consider the feasibility of developing comparable systems. In this description I will, by necessity, discuss some technical aspects, but my principal focus will be on considerations for the manager/decision-maker.

At this time, I would like to acknowledge the contributions of the following people in the development and implementation stages of the system: Eva Landy, previously Chief of Data Operations in the Health Department and now Chief of the Division of Fiscal Assistance; Roberta Chevoya, Principal Research Technician in the Division of Vital Statistics; and Roger Goulet, programmer at Rhode Island Health Services Research, Inc. (SEARCH).

Background

The new system actually represents the third phase in the ongoing development of the Division's electronic data processing capacity. Beginning in the mid-1960's, vital statistics data processing was performed by the State Information Processing Division (IPD) at its headquarters. Data was entered at IPD onto IBM cards by IPD keypunch operators. This system was primitive by today's standards. For example, the statistical tabulations that were produced required many hours of further human tabulation and typing before they could be reproduced for publication.

After several years, dissatisfaction with this arrangement peaked. Since data entry was not being performed in-house, the vital records had to be physically transferred to IPD each month, thereby creating both a security problem and an inconvenience for people requesting certified copies of those records. In addition, it was impossible for IPD keypunch operators to do a high-quality job because they did not have access to Division staff for solving problem entries. Finally, IPD did not appear to be developing satisfactorily either from the point of view of hardware or — even more seriously — on the basis of its programming and systems analysis staff. As a result of these problems, alternative systems began to be investigated.

In 1974, the Division's information processing system entered its second phase when a contract was signed with a private health data consortium for data processing. In addition, an in-house data entry operation was begun at the same time. In those early days with the consortium, there was an immediate overall improvement in the operation of the system. High-quality edit programs were written for each file, resulting in extremely accurate data files, and statistical tables were created which were camera-ready for publication in the Division's Annual Report.

Within a few years, however, problems began to arise around four issues: (1) cost-effectiveness, (2) timeliness, (3) confidentiality, and (4) control of the system. Although the consortium was a non-profit organization and had close ties to the Department of Health, its initial cost for providing services was high and its fees were increasing. In addition, the logistics involved in working with an outside agency, which used a main-frame computer at a remote location, often resulted in delays in receiving reports and other outputs. Also, despite the fact that under Rhode Island law information from individual vital records is strictly confidential, the Department did not have physical possession of its own computer files. Finally, the Division had no control over the actual operation of the system and was completely dependent upon the consortium in this respect. This became an immediate concern when doubts arose as to the financial viability of the consortium following federal funding cutbacks. Because of all these concerns, it was decided to try to bring information processing into a third phase: complete and independent operation and control by the Division itself on an in-house basis.

Implementation

Informal discussions among the Division's staff had been going on internally and with the consortium's staff regarding the possibility of bringing the system in-house for about twelve months before final approval of the project was made by the Director of Health in January 1982. Several alternatives had been investigated during this period and are worth mentioning for management purposes. Two approaches which were

offered by the consortium were rejected because the Division would still be too dependent upon it. In addition, a self-contained mini-computer vital statistics system which was on the market at the time was considered too expensive at an approximate cost of $100,000.

Since the Department had already developed expertise with Radio Shack TRS-80 equipment and software in several other divisions, it was decided to develop a vital statistics in-house capacity through a Radio Shack microcomputer system. The final configuration decided upon was a TRS-80 Model 16 microcomputer with 512K memory, a Daisy Wheel printer, and two hard disk units with about 16 megabytes capacity. In addition, a TRS-80 Model II microcomputer was leased to provide for additional data entry capacity on floppy diskettes.

It is important to mention here that the type of system developed could have been supported by the hardware and software of any number of microcomputer manufacturers. At the time this equipment was purchased, Radio Shack was very competitive on price, but that was not the primary factor in the decision. Rather,the deciding factor was the expertise already present in the Department. Another important factor was the availability of Radio Shack equipment in the building for back-up purposes, if the need should have arisen.

There were two levels of goals set for the new system. The primary goal was to reproduce exactly the outputs of the old system (i.e. a monthly tape for the National Center for Health Statistics and indexes and other reports for State use). The secondary goal was to convert most of the Division's manual coding and editing functions from a manual to a computerized basis.

The development of the edit and code assignment programs was contracted out to the consortium, along with several other special programs. In order to ensure that the programs be developed correctly, it was decided that Division staff would draw up exact and detailed specifications for the edit and code assignment routines. In addition, the record formats and test data were produced by the Division itself, utilizing Radio Shack software. Finally, Department staff met weekly with consortium programmers to closely monitor progress and resolve problems.

Data entry presented one of the most difficult aspects of the project. Under the old system, data entry was performed by two keypunch operators dedicated entirely to that task. These operators had been using IBM 3742 floppy diskette data entry machines for eight years, and they expressed some resistance to the new equipment. In fact, there had already been much consideration given to eliminating a separate data entry function and assigning the data entry task for each file to the clerk responsible for the other aspects of that file. It was felt that this would be more effective by providing for complete control over each file by the individual clerk from the point of receipt of the original document to the maintenance of the computer master file and production of reports. Radio Shack software made it very easy to set up screens for data entry which greatly facilitated this task. Because of the combination of these factors, it was decided to release the two data entry operators and to implement this vertically integrated method of data entry.

A training program was begun for the clerical staff responsible for the different records. Plans were made to train each person on the new equipment, and instruction manuals were produced for each file. Full-scale data entry was begun with the file of January 1982 marriage records, and by June all of the files had been brought over to the new system.

Results

Although there were certainly some difficulties and limitations encountered during implementation, it is felt that the new system is an unqualified success. First of all, the four concerns which served as the catalyst for the new system -- cost-effectiveness, timeliness, confidentiality, and control — have been resolved. For a one-time investment of about $30,000 in hardware procurement and systems development, the Division stands to save approximately $50,000 per year (the annual cost for two data entry operators, IBM 3742 machine rental, and the data processing contract with the consortium). Although there will be some additional expenses related to the new system (e.g. supplies and maintenance), there is no doubt that it is much more cost-effective than the old system. In addition, because of the in-house location of the hardware, turn-around times are much shorter, and report production is much more timely. Finally, by successfully implementing the new system on an in-house basis, the closely related concerns of confidentiality and control of the Division's computer files are no longer a problem.

In addition to responding to these four major concerns, the new system has successfully reached its two sets of goals: the outputs of the old system have been reproduced exactly, and much of the demographic coding and error and inconsistency checking are now done by the computer. Thus the new system not only accomplishes as much as the old system at a lower cost, it actually does more.

There have been additional benefits as well. Because of the vertical integration of the data entry function, staff members now have increased responsibility and accountability for all aspects of their respective files; they receive direct feedback; and they have increased flexibility. Because of the ease of use of the software, new files and new reports can be easily generated. Finally, a word processing software package was also purchased, and this entire area has become available for potential exploitation.

There have been problems with access to the computer for tasks other than data entry. This has happened for two reasons. First, some programs and reports currently require a great deal of processing time. Second, the computer is currently capable of performing only one task at a time. The solution to this particular problem demonstrates yet another benefit of microcomputer technology: advances in hardware and software are introduced on an almost continuous basis. In fact, a new operating system has been

introduced for the Model 16 microcomputer which will not only increase its processing speed substantially but will also allow it to perform multiple tasks at the same time.

A Software Demonstration

In order to fully appreciate the ease of use of current software packages, it would be useful to review an example. The particular data management package that we use (PROFILE) is licensed by Radio Shack, but it is comparable to software packages used on most other microcomputers.

In order to utilize PROFILE, the user begins with a menu (Figure 1) which offers various options to choose from, such as "Define Data Formats", "Print Reports", "Print Labels", etc. Option "1" ("Define Data Formats") is utilized for constructing data files (Figure 2). The computer assigns a field number and file location for each field; the user merely gives a name to the field and indicates the number of characters it requires (field length). In this example, State File Number is Field No. 1 and is 6 characters long; and Hospital Code is Field No. 8 and is 2 characters long. After defining all of the fields, the user then selects the expansion option (Figure 3) to create a file with the desired number of records by simply typing in how many records are needed. Of course, the file can be expanded more than once, if necessary.

The file at this point would consist of blank records containing no data. Before data entry can begin, a screen for data entry must be constructed (Figure 4). This is one of the most exciting features of this type of software package: data entry screens can be constructed very easily which closely resemble the source document, thereby facilitating data entry. For example, the screen in Figure 4 is quite similar in appearance to Rhode Island's birth certificate. In order to position each field on the screen, the user need only indicate its corresponding field number in the desired location. Thus, State File Number is Field No. 1, Child's First Name is Field No. 2, and so on. Data entry then becomes a process of merely typing the information onto the screen (Figure 5); the computer assigns the data which has been entered by the clerk to the correct location in the actual computer file. Alternate screens can also be constructed for looking at the file in different ways. For example, the screen in Figure 6 can be used to update the file and contains not only all the information on the data entry screen but also the demographic codes generated by the computer. For example, the State of Birth Abbreviation is Field No. 22 and is typed by the clerk; Field No. 23 is the State of Birth code assigned by the computer.

After a number of records have been entered into the file, a user will generally want to produce some type of report. Once again, this is a fairly simple process. First of all, a report format must be constructed in a manner similar to the data entry screens (Figure 7). The title of the report and column headings are typed in the appropriate locations. To position the data items required for the report, the user need only type in the corresponding field number in the desired location. The report format is then saved and can be used anytime that particular report is needed by choosing the "Print Report" option (Figure 8). The user can choose one field to sort on and one or two fields to select on. In this example, the report would consist of a birth index sorted by the surname of the child (Field No. 4) for a number of records determined by a range of state file numbers (Field No. 1). The same report format could be used again to produce a different index sorted and/or selected by some other fields.

As simple as this all seems, PROFILE has been improved recently to make the running of reports and other tasks even easier. This has been accomplished through "User Menus" (Figure 9). These are menus which are created for a specific purpose, and all the user need do to run a specific job is select the appropriate option. For example, Option "5" ("Run COS Death List") would be selected to produce a list of deaths to out-of-state residents sorted by state of residence.

It should now be clear that this software is fairly simple to use. However, the assumption should not be made that — since the software is easy to use — no time, effort, or skill is required to set up a system with it. On the contrary, each of these factors is necessary if a system is to operate effectively and efficiently. The point is that computer programming or other technical expertise — although helpful -- is not required to use software to construct data files and to produce outputs from these files.

Conclusion

At the beginning of this paper, I indicated that my major emphasis would be on considerations for the manager/decision-maker. With that in mind, I would like to close with some suggestions, based on our experience, for anyone considering the implementation of a system similar to ours. First, be sure to get the support of the clerical staff who will be given the job of running the system on a day-to-day basis; without their support, successful implementation will be very difficult. Second, if possible, utilize hardware and/or software that is familiar to someone on your staff; if that is not possible, select a vendor willing to assist you as necessary. Third, if any of the systems development work is contracted to an outside firm, maintain as close supervision over their work as is possible; the more direct control maintained over their work, the greater the chances for success. Fourth — and perhaps most important -- try to get upper management concerned and involved in the project; their active participation can make success more easily attainable. I hope that these suggestions will be useful to anyone interested in developing a microcomputer system similar to ours.

FOOTNOTES

(1) Edith Myers, "Attack of the Micro", Datamation, vol. 26, no. 11, (Oct. 1982), p.84.

(2) Ibid., p.80.

(3) "Business that Defies Recession", Business Week, Oct. 25, 1982, p.30.

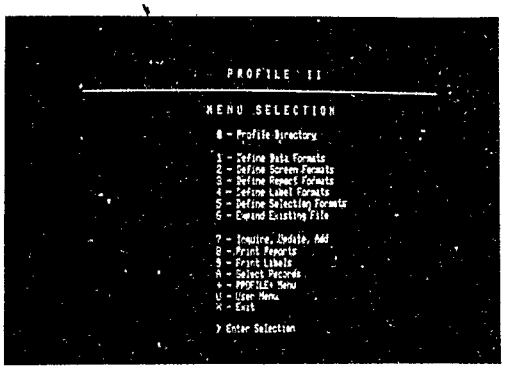


Figure 1

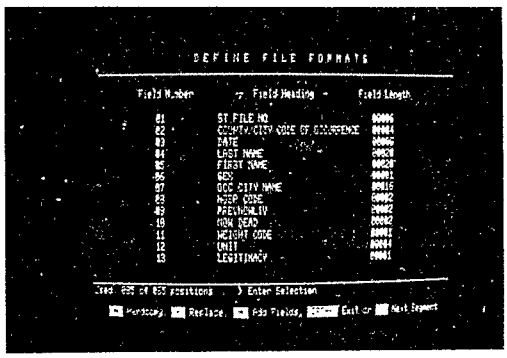


Figure 2

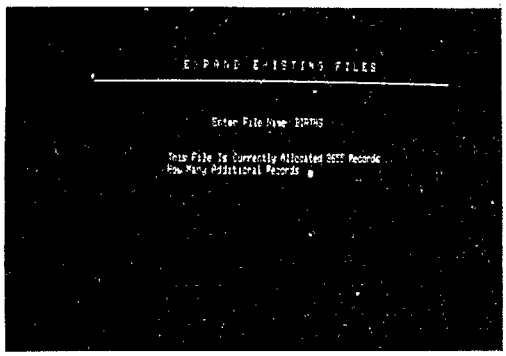


Figure 3

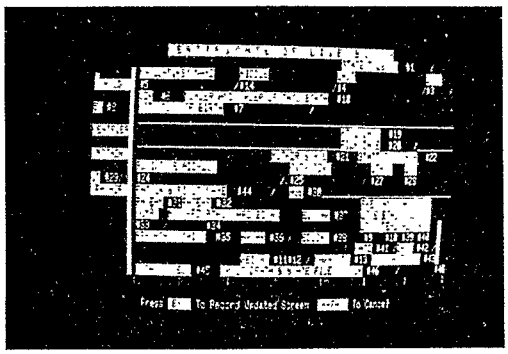


Figure 4

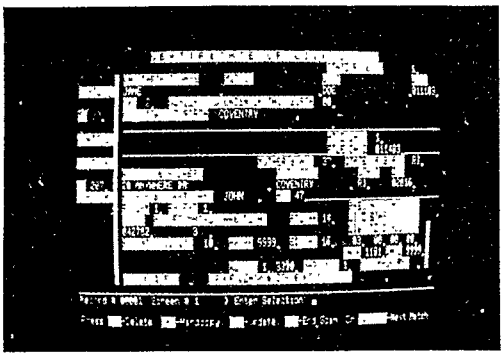


Figure 5

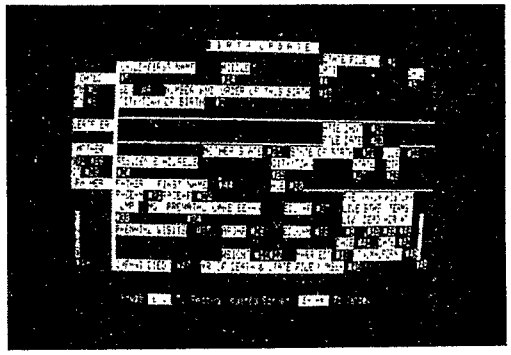


Figure 6

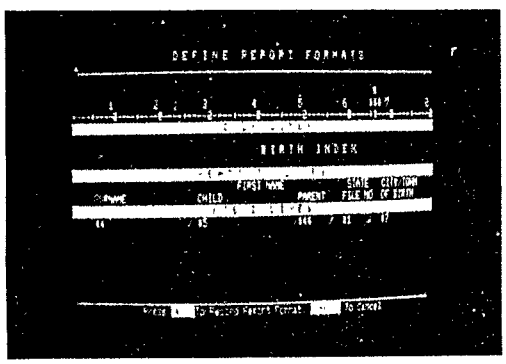


Figure 7

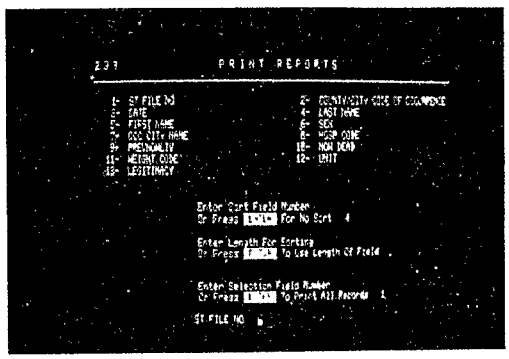


Figure 8

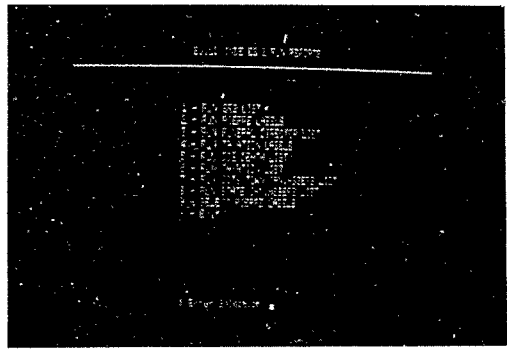


Figure 9

AN AUTOMATED VITAL STATISTICS SYSTEM

Ronald L. Williams, University of California, Santa Barbara, CA
John A. Marinko, University of California, Santa Barbara, CA
Merle L. Shields, Department of Health Services, Sacramento, CA

INTRODUCTION

Birth and death certificates provide the cornerstone of national data sources for monitoring trends in public health. Yet the completion of a nationwide birth and death registry is a relatively recent development. Indeed, the year 1940 marked the first time in the history of the U.S. that census population figures and national vital statistics data from all states were concurrently available. During the last four decades the system has evolved into a reasonably efficient means for information acquisition and dissemination. Now involving the cooperative efforts of federal, state, and local health agencies, vital statistics provide the raw material for a large and diverse number of research activities. These activities have further stimulated the collection of additional information, thus greatly expanding the possibilities for scientific research.

It was from this perspective that we first became interested in the mechanics of registering vital events, particularly in the area of data collection, editing, and management. Working with vital records in California we found that, in spite of the richness and size of California's matched birth/death cohort files, there remained several deficiencies. We wish to briefly describe some of these limitations and then report on a method that we have developed to provide a solution.

By California law, births and deaths are reported within 10 days to the local registrar by means of prescribed paper forms. For purposes of expediency we limit our discussion here to the birth certificate. Nearly 99% of all California births occur in hospitals, thus nearly all birth certificates are typed by hospital personnel using manual typewriters. Required local file copies are produced and maintained and the original certificates are transmitted to the State Registrar of Vital Statistics, where they are indexed and filed as permanent records. Selected items are coded and manually key-entered by the State Registrar, but several larger counties also do independent key-entry. A variety of edit and validation checks are performed for quality assurance of the key-entry process.

In our opinion there are five major shortcomings to the present system:

1. TIMELINESS --- The most acute problem is the long delay now experienced by local agencies before the information processed by the State Registrar is available in the form of tabulations or as machine-readable media.

2. REDUNDANCY --- These delays have resulted in much redundant key-entry by local agencies, who prefer to incur the expense of independent data processing rather than wait for state-processed information.

3. ACCURACY AND COMPLETENESS --- Although computer edits performed by the State Registrar can detect systematic key-entry errors, it is now very difficult to detect and correct source document errors made by hospital personnel.

4. FEEDBACK --- The principal providers of medical information (i.e. hospital staff and physicians) do not usually receive summary data of the information they provide. They consequently often perceive the completion of vital records as a bit of a nuisance, and they are frequently skeptical about their usefulness. This is, of course, a cycle that is self-sustaining: if the providers of the data are skeptical about its quality and usefulness, then poor quality data is likely to be the result.

5. OUTDATED --- In view of the remarkable recent advances in data processing, the current means of collecting and managing vital records is relatively archaic. When nearly every office now or will soon have access to "word processing" for managing the simplest forms of communication, it is remarkable that some of society's most important documents continue to be generated by manual means.

To provide a solution to these shortcomings, we have developed and intensively tested an Automated Vital Statistics System. The resultant product, which we have dubbed AVSS is a complete automated system for the collection, management, and reporting of vital records information. AVSS is both revolutionary and evolutionary:

* It is revolutionary since it establishes a public-private interface by computerizing birth certificates (and other vital records) at their original sources. It thus by-passes many of the time-consuming and error-generating intermediate steps in the current system; in other words, it is a "direct-entry" system.

* It is evolutionary since it adapts to the present system of paper certificates, primarily using computers as sophisticated information management adjuncts. By taking a stepwise approach, the system can be phased in gradually with techniques optimized and refined as they are implemented.

DESIGN REQUIREMENTS

From the beginning it was agreed that AVSS should:

* Be interactive
* Be oriented toward public health
* Have multiple confidentiality safeguards
* Be easy to use by inexperienced personnel
* Have powerful on-line edit capabilities
* Be cost effective

* Be easily transportable
* Be extremely reliable and "foolproof"
* Be flexible with potential for growth

These design criteria virtually dictated that the American National Standard computer language, MUMPS, be the logical choice for AVSS programming development. MUMPS originated at the Massachusetts General Hospital Hospital, hence the name: Massachusetts General Hospital Utility Multi-Programming System. MUMPS was designed for on-line, interactive applications where immediate data retrieval and fast response time are essential, it therefore promotes accurate data entry and its interpretative nature results in rapid and efficient program development. It facilitates the simultaneous usage of computer equipment and it can support multiple data bases within the same computer. Because MUMPS is an American National Standard language, its applications software can be implemented on a wide variety of computers of all sizes with virtually no program modifications. Most MUMPS implementations come equipped with many security measures and, because it has evolved in a medical care setting, it is inherently a very reliable operating system. MUMPS was specifically designed for data base applications involving textual data of variable length, and it thus very efficient at managing the type of non-numeric information found in vital records. Finally, since MUMPS is a modular language, its applications software can evolve to meet changing needs.

IMPLEMENTATION OF AVSS

Having chosen MUMPS as the vehicle for AVSS development, the principal AVSS design concepts may be summarized as:

* To provide an efficient and reliable means for direct entry of vital records.
* To use distributed data processing techniques using a variety of computers.
* To insure transportability through the use of an American National Standard Language.
* To use the ASCII coding scheme; i.e., the American Standard Code for Information Interchange.

These design concepts were established to meet the principal AVSS objectives:

* To improve the accuracy, completeness, and timeliness of vital records.
* To minimize redundant key-entry.
* To promote cooperation between local, state, and federal agencies (and to enhance public-private communication as well).

AVSS was patterned after COSTAR, a versatile MUMPS-based ambulatory care medical record and information management system. Many of the basic COSTAR conventions were used, but the AVSS software was developed from "scratch" to meet the specific requirements of vital records. As a result, AVSS is more efficient, compact, and uses more advanced programming techniques.

AVSS was first implemented on a Digital Equipment Corporation PDP 11/45 computer located at the University of California at Santa Barbara. Local hospitals and the Santa Barbara County registrar gain access to the computer by means of telephone.

Hospital personnel view AVSS as an easy-to-use but powerful word processor, allowing them to use a computer video terminal as an "electronic scratchpad" to produce a perfect copy of the birth certificate on state-approved forms with an attached computer printer. But transparent to the user, AVSS is interactively performing edits and other validity checks so as to detect errors the moment they are entered. Moreover, the simple process of typing characters on a computer terminal rather than on a typewriter automatically places all information contained on the birth certificate into machine readable form.

After the certificate is printed, the required signatures are obtained as usual and the certificate is forwarded to the local registrar in the same manner as are conventionally typed certificates. When the certificate arrives at the county health department, the local registrar uses AVSS to retrieve the computerized record, to validate the accuracy and completeness of each data item, to complete any items on the certificate that were filled in manually, to assign a local file number, and to register both the paper and electronic certificate. Additionally, automated address matching under AVSS takes place at this time, thus yielding the census tract code for mother's residence. Since all information is then machine-readable, AVSS can produce ad-hoc reports or it can generate an IBM computer tape which can be used by mainframe equipment for more conventional reports.

OPERATIONAL FEATURES

AVSS was initially developed specifically for the California birth certificate, but it was soon generalized to adapt to virtually any paper form that satisfies two simple criteria:

1. The form must be case-specific.
2. Each data element on the form must have a unique alphanumeric identifier.

Therefore, AVSS can collect and manage a wide variety of source documents.

How does the typical user interact with AVSS? After clearing the required security procedures, the AVSS user is presented with a choice of options; this is the so-called "menu-driven" approach. When a user response is required, one may invoke a listing of the appropriate choices by simply entering a "?" This general rule applies everywhere in AVSS: if the user is uncertain about the proper response, typing a "?" will produce a set of instructions on how to proceed. This so-called "self-tutorial" approach has greatly minimized the amount of training time necessary for novice users.

A typical beginning session with AVSS is as follows (underscored items reflect user response):

```
                    WELCOME TO AVSS

          AUTOMATED VITAL STATISTICS SYSTEM AT UCSB

GOOD EVENING RON, YOU HAVE MAIL

SYSTEM OPTION > ?

ENTER ENOUGH CHARACTERS TO SELECT THE FOLLOWING:

     BIRTH CERTIFICATE
     DEATH CERTIFICATE
     FETAL DEATH CERTIFICATE
     MAILBOX
     REPORT GENERATOR
     SYSTEMS MAINTENANCE

SYSTEM OPTION > BIRTH CERTIFICATE

BIRTH CERTIFICATE OPTION > ?

ENTER ENOUGH CHARACTERS TO SELECT THE FOLLOWING:

     ALPHA LIST
     COMPLETE INCOMPLETE FORM
     DISPLAY
     EDIT
     PRINT
     REGISTER
     TRANSFER FORM FROM HOSPITAL

BIRTH CERTIFICATE OPTION > REGISTER
```

The REGISTER option is the most complex and important AVSS tool. It sequentially questions or "prompts" the user for each data item, with the user response for the current item determining the next item to be prompted. In addition to extensive tutorial information available for each item, automatic range checking and editing is performed as each data element is entered:

```
BIRTH CERTIFICATE OPTION > REGISTER

1A. CHILD'S FIRST NAME > JOHN

1B. CHILD'S MIDDLE NAME > PAUL

1C. CHILD'S LAST NAME > JONES

2. CHILD'S SEX > MALE

3. THIS BIRTH'S PLURALITY <SINGLE> ?

ENTER ENOUGH CHARACTERS TO SELECT THE FOLLOWING:

     SINGLE
     TWIN
     TRIPLET
     QUADRUPLET
     QUINTUPLET

3. THIS BIRTH'S PLURALITY <SINGLE>
```

To promote efficiency and reduce the time required to produce a certificate, a number of data items have values pre-entered; for example, since most babies are singletons, there is a so-called "default" value of SINGLE for item 3A. One can, of course, easily change this value to TWIN or TRIPLET, but one simple keystroke enters in the default value as SINGLE for 98% of births. When SINGLE is entered, AVSS does not bother to ask for birth order (Item 3B) since it must necessarily be one; for multiple births, however, the birth order item is prompted.

Just a few keystrokes also enters the birth date, but then it is spelled out in complete English on the AVSS-produced birth certificate. Another AVSS timesaver is the batch entry mode that allows a number of sequential data items to be entered in a single string separated by semi-colon delimiters:

```
4A. DATE OF BIRTH > 8/23
8/23/83 <--- CONVERTED TO AUGUST 23, 1983
```

Similarly, a single keystroke enters in the hospital name, address, city, and county; a saving of 50 or more keystrokes for the typical hospital:

```
5A. PLACE OF BIRTH <COTTAGE HOSPITAL>
```

For data items having standardized choices, e.g. the parents' state of birth, the user can select the appropriate data value by simply entering a enough alphabetic characters to uniquely identify the proper choice. A standard abbreviation is then automatically entered. Not only does this save time and reduce the number of keystrokes, but it also greatly increases the accuracy and uniformity of data. Indeed, this is a key AVSS design feature: that is, every data item can potentially have a standard list of choices, which can be easily and accurately entered by the user:

```
7. FATHER'S STATE OF BIRTH > ?

ENTER STATE NAME OR ABBREVIATION:

ALABAMA .... AL  ALASKA ..... AK  ARIZONA ... AZ
ARKANSAS ... AR  CALIFORNIA . CA  COLORADO .. CO
CONNETICUT . CT  DELAWARE ... DE  FLORIDA ... FL
GEORGIA .... GA  HAWAII ..... HI  IDAHO ..... ID
                      (etc.)

7. FATHER'S STATE OF BIRTH > GEORGIA  GA
```

Here, for example, AVSS performs a pattern match using the letters "GE" and completes the spelling of "GEORGIA", then provides the two letter abbreviation of "GA".

As described earlier, each data item is examined by AVSS at the instant of entry to determine if it is a reasonable value. As a simple example, if the mother's or father's age is outside a particular numeric range, AVSS asks the user to confirm the value:

```
8. AGE OF FATHER > 76

76 IS AN IMPROBABLE AGE OF FATHER.
ARE YOU SURE? N

8. AGE OF FATHER > 26
```

As another example of a standard list that produces much time saving for hospital staff, AVSS

stores the names, addresses, and medical license numbers for the attending physicians at a particular hospital. The user can then enter all this information, again involving about 50 characters, by simply typing the first few letters of the physician's last name:

```
13D. CERTIFIER'S NAME AND ADDRESS > ?

BRADLEY, J GLENN, MD, 215 PESETAS LANE, SB
CORLETT, ROBERT C, MD, 5333 HOLLISTER AVE, SB
FISCHER, RICHARD L, MD, 2440 FLETCHER AVE, SB
HERRALD, GORDON A, MD, 215 NOGALES AVE, SB
JOSEPH, DANIEL M, MD, 425 W JUNIPERO, SB
LINBLAD, DONALD E, MD, 5333 HOLLISTER AVE, SB
REID, ROBERT A, MD, 301 W PUEBLO ST, SB
SECORD, DAN B, MD, 2330 BATH ST, SB
TURNER, DUNCAN J, MD, 301 W PUEBLO ST, SB

13D. CERTIFIER'S NAME AND ADDRESS > SECORD,
DAN B, MD, 2330 BATH ST, SANTA BARBARA
```

Perhaps the most thorny item we had to deal with in developing AVSS was race/ethnicity. Here we were faced with two opposing perspectives: the state's need for uniform reporting, and the individual's desire for uniqueness. We believe we have devised a solution that is acceptable to both parties. For race/ethnicity AVSS requires the user to first make a selection from the standard list provided by the State Registrar, which has choices patterned after those used by the Census Bureau:

```
21. MOTHER'S RACE/ETHNICITY > ?
ENTER ENOUGH CHARACTERS FOR ONE OF THE FOLLOWING:
     WHITE
     BLACK
     AMER. INDIAN
     ASIAN
     CHINESE
     JAPANESE
     KOREAN
     VIETNAMESE
     CAMBODIAN
     THAI
     OTHER
     INDIAN
     FILIPINO
     HAWAIIAN
     ESKIMO
     ALEUT
     REFUSED
     UNKNOWN

21. MOTHER'S RACE/ETHNICITY > WHITE
SPECIFY TYPE IF YOU WISH > CAUCASIAN

21. MOTHER'S RACE/ETHNICITY <CAUCASIAN>
```

After the initial choice is made, for example WHITE, the user is then allowed to specify a more precise term, for example CAUCASIAN. AVSS then writes two items into its electronic file: first, a numerical code corresponding to WHITE and then the word CAUCASIAN, which will appear on the printed certificate. Both parties are thus well served: automatic race-coding is performed for the state, and the individual may feel satisfied that the birth certificate reflects his or her preferences for race identification. A similar approach is used for the collection and coding of the SPANISH/HISPANIC question.

From the perspective of perinatal epidemiology, one of the most important data items on the birth certificate is birth weight; thus we paid particular attention to it when designing AVSS. Like all other numeric variables, AVSS performs a range check of birth weight values when they are entered, and if a value is outside the specified range (2500 to 4500 grams), the user is queried. Additionally, since many hospitals still weigh babies in pounds and ounces, and the California birth certificate requires grams, AVSS allows either to be entered. The user indicates pounds and ounces by any type of delimiter, for example a #-sign or simply a space between the pound and ounce data values, and the English to metric conversion is automatically performed by AVSS:

```
26. BIRTH WEIGHT > ?
REQUIRED ENTRY OF BIRTH WEIGHT. IF WEIGHT NOT
KNOWN ENTER "UNK".IF WEIGHT IN GRAMS ENTER NUM-
BER. IF WEIGHT IN POUNDS AND OUNCES, ENTER THE
NUMBER OF POUNDS, A SEPARATOR, AND THE NUMBER OF
OUNCES FOR EXAMPLE "7 0.5" INDICATES 7 POUNDS
AND ONE HALF OUNCE. OR "7-.5" OR "7/.5", ETC...
NOTE HOWEVER THAT "7 .5" IS INCORRECT BECAUSE
THE INPUT CLEANING ROUTINE WILL ELIMINATE BLANKS
OCCURRING BEFORE PERIODS CONVERTING IT TO 7.5
GRAMS.

26. BIRTH WEIGHT > 3#7
WEIGHT OF 3 POUNDS AND 7 OUNCES CONVERTED TO
GRAMS: 1559 INDICATES A LOW BIRTH WEIGHT BABY.
ARE YOU SURE? N

26. BIRTH WEIGHT > 7#3
WEIGHT OF 7 POUNDS AND 3 OUNCES = 3260 GRAMS

26. BIRTH WEIGHT <3260>
```

Medical conditions on the California birth certificate are entered by means of pre-defined numerical codes that are to be collected on detached worksheets. There is of course opportunity for error here. We attempt to minimize errors by incorporating the worksheet conditions into the list of standard AVSS choices for each medical data item. Thus a "?" entered in response to a medical item prompt, results in the display of the state-specified list. When a code corresponding to a particular condition is entered, for example a "02" for Item 28 (Cesarean Section), AVSS will reply with the corresponding English equivalent, for example ELECTIVE PRIMARY; the user may therefore detect an erroneous entry at that point:

```
28. CESAREAN SECTION <NONE> ?
ENTER THE APPROPRIATE CODE OR KEYWORD:

     01 ELECTIVE PRIMARY
     02 ELECTIVE REPEAT
     03 NONELECTIVE PRIMARY
     04 NONELECTIVE REPEAT

28. CESAREAN SECTION <NONE> 02 ELECTIVE REPEAT
```

The newly-revised worksheet for congenital malformations is more lengthy and complex, having a total of 56 possible entries. Again, AVSS will display these choices if a "?" is entered in response to the prompt:

```
32. CONGENITAL MALFORMATION <NONE> ?
ENTER THE APPROPRIATE KEYWORDS OR CODES

01  ANENCEPHALY (740.0), INCLUDES SPINA BIFIDA
02  OPEN SPINA BIFIDA WITH HYDROCEPHALUS (741.0)
03  ENCEPHALOCELE (742.0)
04  MICROCEPHALUS (742.1)
05  HYDROCEPHALUS (742.3)
06  OTHER ANOMALIES OF CENTRAL NERVOUS SYSTEM
07  EYE ANOPHTHALMOS (743.0)
08  EYE CONGENITAL CATARACT (743.3)
09  EYE COLOBOMA (743.4)
10  EYE-OTHER ANOMALIES OF EYE
11  EAR-BRANCHIAL CLEFTS (744.4)
12  EAR-OTHER ANOMALIES OF EAR, FACE AND NECK
13  CARDIAC-TRUNCUS ATERIOSUS (745.0)
14  CARDIAC-TRANSPOSITION OF GREAT VESSELS (754.
15  CARDIAC-TETRALOGY OF FALLOT (745.2.)
                    (etc.)
```

Additionally, however, AVSS allows a key-word search of the list. For example, entering the word CLEFT, will result in four possible malformations. A more specific entry, for example PATENT DUCTUS, will result in an exact match. Again this feature of AVSS holds for all data items, and there is no practical limit to the size of the standard list:

```
32. CONGENITAL MALFORMATION <NONE> CLEFT

     11   EAR- BRANCHIAL CLEFTS (744.4)  ?
     27   CLEFT PALATE (749.90)  ?
     28   CLEFT LIP (749.1)  ?
     29   CLEFT PALATE WITH CLEFT LIP (749.2)  ?

ENTER CODE OR MORE SPECIFIC SEARCH ENTRY

32. CONGENITAL MALFORMATION <NONE> 28
    CLEFT LIP (749.)1).
32. CONGENITAL MALFORMATION > 744.4
    11 EAR- BRANCHIAL CLEFTS (744.4).
32. CONGENITAL MALFORMATION > PATENT DUCTUS
    21 PATENT DUCTUS ARTERIOSUS (747.0)
32. CONGENITAL MALFORMATION >

32. CONGENITAL MALFORMATION <28,11,21>
```

Indeed, we have experimented by adding all possible congenital anomalies and their English descriptions from the ICD-9-CM manual list, and found that it takes only a few seconds for AVSS to perform a keyword search of that rather lengthy list.

Recall that we are still in the REGISTER CERTIFICATE option of AVSS, and we have just completed the last item on the certificate (Item 32). At this point, AVSS clears the video screen and displays a facsimile certificate showing the just-completed data items.

The automatic validation procedures then begin. As opposed to range-checking, which occurs at the instant of entry and evaluates each data item independently of all other items, the validation process compares data values between a number of items and evaluates their consistency, for example:

```
GESTATIONAL AGE COMPUTED FROM DATE OF BIRTH (4A)
AND DATE OF LAST MENSES (26A) IS 42 WEEKS

BIRTH WEIGHT AND GESTATIONAL AGE ARE NOT COMPAT-
IBLE, PLEASE VERIFY BIRTH DATE (4A), DATE OF
LAST MENSUS (26A), DATE OF LAST LIVE BIRTH (25E)
OR TERMINATION (25F), AND BIRTH WEIGHT (27)
```

The user then has an opportunity to perform edits on questionable data items before printing and filing the certificate. Displays and edits can be performed as many times as necessary to obtain a satisfactory product:

```
F(ILE), E(DIT), V(ALIDATE), D(ISPLAY), P(RINT)?P

PLEASE TURN ON PRINTER WITH PROPER FORMS LOADED.
```

An attached computer printer then rapidly types the data values onto pinfeed forms provided by the State Registrar. The entire process of entering data values and printing the certificate requires from 3 to 5 minutes for experienced users. Even novice users can complete certificates in less than 8 minutes.

If an error is detected after the certificate is printed, which frequently happens when the informant or certifier is presented the certificate for signature, the user can return to the AVSS terminal, invoke the EDIT CERTIFICATE option, perform the necessary changes, and produce a revised certificate in a matter of 2 to 3 minutes. This is, of course, a great time saver compared to manual means, since birth certificate errors cannot be erased and correcting even minor errors demands that the entire certificate be retyped, a frustrating and laborious task with a manual typewriter. As discussed earlier, after the certificate is printed, electronically filed, and mailed to the county health department, the local registrar invokes an AVSS option to transfer the electronic certificate in the hospital's file into the registrar's data base.

At this time automatic census tracting takes place. Because MUMPS excels in the processing of character strings, this procedure has proved to be quite successful. There are instances, however, where human intervention is required. For example, if the mother fails to report the entire street address, AVSS presents the user with some possible matches, and it is usually possible to identify the correct address and thus retrieve a census tract code:

```
CITY: LOMPOC; STREET: 413 N 1ST; ZIP: 93436;
STREET NAME CLEANED UP INTO 'FIRST NORTH'
THERE ARE 4 POSSIBLE MATCHES FOR '413 N 1ST',
```

#	STREET NAME	LOW	HIGH	ZIP
1	FIRST PLACE NORTH	1200	1499	93436
2	FIRST PLACE SOUTH	200	299	93436
3	FIRST STREET NORTH	300	599	93436
	FIRST STREET NORTH	600	799	93436

4 FIRST STREET SOUTH 100 499 95436

STREET #3
THE CENSUS TRACT IS RECORDED AS 27.03

To briefly review just a few additional AVSS options:

DISPLAY: retrieves and displays a form
EDIT: same as DISPLAY except it also allows the form to be edited.
ALPHA LIST: alphabetically lists certificates within a user-specified date and alphabetic range. This option is useful for indexing, and can also be used by the local registrar to monitor the ongoing registration activities in hospitals.
COMPLETE INCOMPLETE FORM: allows the user to complete an AVSS form that was filed in a partially completed state.
GENERATE VITAL STATISTICS TAPE: creates a fixed block IBM standard label tape containing all items on each birth certificate within a given date range.
AMEND CERTIFICATE: allows the local registrar to update registered AVSS certificates to reflect information contained on legal affidavits.
MISSING FILE NUMBER REPORT: searches for gaps in local file numbers.
NON-CONFIDENTIAL DISPLAY: displays only the upper portion of the birth certificate.
BIRTH/DEATH LINKAGE: automatically matches infant death certificates with birth certificates using up to 19 common data elements.

Additionally, AVSS contains a rudimentary, but easy-to-use and versatile report generator. The user simply identifies the item number of the data element that is to be tabulated, and specifies date ranges. AVSS then lists the unique values found for that data field. The REPORT GENERATOR applies both to numeric and non-numeric data, and has proven to be a very useful tool for insuring the integrity of the AVSS data base.

Finally, since AVSS can be adapted to virtually any paper form, it can be used for a variety of data collection and management activities outside of vital statistics. For example, we have designed an AVSS data base tailored for Maternal and Child Health applications, with forms for neonatal intensive care, high risk newborn followup, and sudden infant death syndrome.

FUTURE IMPLEMENTATIONS

AVSS is currently being implemented in four California counties, including Los Angeles, which accounts for one-third of all vital events in the State. Because of the American National Standard feature of MUMPS, AVSS has the potential to adapt to the centrally-controlled hierarchical distributed data processing system as suggested by the Model State Vital Statistics Act. Indeed, it goes a step further than envisioned in the Act by directly incorporating hospitals into the system. Under such a statewide system, the state registrar would be responsible for setting AVSS standards and for promoting electronic data telecommunications between the counties and the state and between the counties themselves. Again, the common MUMPS language and file structure will simplify these activities.

Perhaps the most exciting AVSS development is the recent availability of low cost but powerful micro-computers using large scale integrated circuits. Remarkable advances have occurred over the last two years during which AVSS was developed and tested: our original PDP 11/45 minicomputer was purchased new in 1974 and upgraded with MUMPS in 1981, at a total investment of approximately $150,000. The MUMPS operating system alone cost in excess of $15,000. In contrast, we recently took delivery of an LSI 11/23 microcomputer having roughly one-quarter the data processing capacity of the PDP 11/45, but exactly equivalent in terms of the ability to support MUMPS and AVSS software. The cost of this five user microcomputer, including the MUMPS operating system, is considerably less than the cost of MUMPS alone for the 11/45, in fact, it is in the order of $10,000.

Rather than filling an entire room and requiring special air conditioning and electrical power, these smaller machines fit on a table top, require only a conventional 110 volt electrical outlet, and will operate perfectly in a ordinary office environment. Such a system would be capable of managing vital records for most small to medium sized counties in the U.S. For larger counties, larger machines will be required of course, but again, the cost of the central processor is not prohibitive --- in the order of $20,000-$30,000 for all except the largest of counties. Of course, the so-called "peripherals" --- the communications gear, video terminals, and printers --- would cost an added amount roughly equivalent to the cost of the central processor.

We are thus optimistic that the MUMPS based automated vital statistics system described here will provide a cost effective means by which state and local agencies can begin the much needed activity of modernizing the collection and management of vital records throughout the United States.

REFERENCES

1. Linder FE, Grove RD: Vital Statistics Rates in the United States 1900-1940. Washington DC: National Office of Vital Statistics, 1959.

2. Barnett GO: The modular hospital information system. Computers in Biomedical Research 4:243, 1974.

3. Barnett GO, Justice NS, Somand ME, et al: COSTAR --- A comuter-based medical information system for ambulatory care. Proceedings of the IEEE 67:1226, 1979.

4. Williams RL, Marinko JM: Toward a national automated vital statistics system. MUG Quarterly 12(3):16, 1983.

National Focus on Public Programs and Public Records

Session L

ADDING UP THE NATION'S HEALTH BILL:
The Whole Is Not the Sum of the Parts

Daniel R. Waldo, Health Care Financing Administration

When my colleague, Robert Gibson, first proposed this presentation, he chose as its title, "Adding Up the Nation's Health Bill -- The Whole Is Not the Sum of the Parts." His choice reflected the pervasive misconception that, just as where there is smoke there is fire, so where there are aggregate data there is detail. I will address that misconception. I will also address another misconception, one that arises from the clarification of the first. I like to call this second misconception "the Twinkie Metaphor."

First, let me describe, with broad strokes, the construction of our estimates of national health expenditures. I will then talk briefly about those things for which the results are useful, and those things for which the results are not useful.

The basic approach to estimation of our nation's spending for health is to compile estimates both of the money coming from the various sources of funds and of the money going to the various types of providers of health goods and services. The two sets of estimates, each of which must be reconciled with the other, serve as mutual checks and balances. In this respect, the process is quite similar to that used to derive estimates of the Gross National Product. That is not really surprising, either, for national health expenditures were designed to be "the GNP of health"; furthermore, several of our professional staff, myself included, are alumni of the Commerce Department bureau responsible for GNP estimates, chosen perhaps for our towering intellects, but more likely for our training in national economic accounting.

Let me direct your attention to the two "sides" of the accounts:

First, we have to quantify the funds spent for health, sorted by the supplier of those funds. We start with several data sources. My own agency, HCFA, provides data on the Medicare and Medicaid programs, drawn from administrative financial records. We also obtain data directly from the various Federal agencies involved in the delivery of care: the Veterans Administration, the Defense Department, elements of the Public Health Service, and so on. Until last year, these agency data were available through the Executive OMB as attachments to the Federal budget submission. Beginning with our estimates for calendar year 1982, however, EOMB no longer required that exhibit, and we became dependent upon whatever records the individual agency kept. Our estimates of state and local government spending for health (outside the Medicaid program) are based upon Census Bureau surveys of a sample of such governments. We obtain information from Blue Cross and Blue Shield and from the Health Insurance Association of America on the financial experience of their members, and we survey directly a sample of independent health insurance plans -- HMO's, self-funded employers, an so on.

Once in possession of the data, we must fit those data into the definitional framework of our health accounts. The process is still more an art than a science. In some cases, little need be done, except to convert from a fiscal year to a calendar year. In most cases, however, we have to modify the agency report of how the money was spent, in order to line those dollars up with our notion of what constitutes hospital care, physician care, and so on. In altogether too many cases, we know only how much money was spent, and nothing of what it was spent on. Here, we rely upon historical and anecdotal evidence, and upon our analysts' judgment and careful examination of chicken entrails, to determine what that money purchased.

While all this is going on, other analysts are gathering information on the aggregate amounts spent for each type of service. Where possible, this is done independently of the search for sources of funds. Again, our data come from elsewhere: We use provider records such as the AHA annual survey of hospitals. We use IRS tabulations of health professionals' income tax returns, sketchy as those tabulations are. We use secondary economic evidence, such as consumer price inflation, workhours, physician visits, and GNP estimates of consumer spending for medical durables and nondurables. Again, analyst judgement plays a large role in the transmutation of these data from mere numbers into spending estimates that conform to our definitions.

At this point, the two sets of estimates are unveiled, with cries of "ta dah" and other, more raucous, sentiments, and we begin a reconciliation process. The first check is of the implied estimate of consumer direct spending. Because we do not use a direct survey in preparation of our estimates, we do not have an explicit measure of how much consumers pay directly for health care. Instead, we back into that number, starting with total spending for a good or service and subtracting all estimated third-party payments. Another term often used for direct consumer expenditure is "out of pocket" spending, usually abbreviated o.o.p. If you pronouce that abbreviation aloud, you'll see how unwittingly appropriate it is in our case, for aberrations in the size or trend of the direct consumer share of spending are often the first sign of trouble with an estimate. Is that share consistent with the evidence? Is it consistent over time? If not, either the total or the third-party programs needs a fix.

Next, we examine the service items and the sources of funds items. Does the service share of the national total seem right? When compared to estimates change in prices and population, does its year-to-year growth make sense? How about the sources of funds? Does the change in an agency total seem consistent with what we know about that agency?

Juggling all these considerations, we adjust totals and distributions, following Russell Myers' observation that "history is like a mixed drink. If it doesn't suit you the way it is, just keep adding things 'til you get it like you want it." When deciding between two conflicting sets of results, we must consider the quality of the data involved in each set, the amount of kneading we used to bring them into line, and the ramifications of the proposed adjustments on the rest of the estimates. Our

goal, when forced to make changes, is to change the least solid number in the most defensible way.

All this reconciliation and change is effected during staff meetings marked by congeniality, fraternity, goodwill, and endless reserves of tact, after which everybody takes a few days' vacation.

So -- now you have estimates of national health expenditures. What do you do with them? (Actually, that is not a very good question, because you can do anything you want with them, and I've seen some pretty bizarre applications. Perhaps I should ask, "What can you legitimately do with them?") First, you can compare them to the nation's total productive capacity, to identify the aggregate resources committed to health. Hence, "national health expenditures in calendar year 1982 totaled $322 billion, an amount equal to 10½ percent of the Gross National Product." Second, you can identify the sources of funds on a consistent basis, whence "the federal government provided 29 percent of spending for health in 1982, and state and local governments provided 13½ percent. Private health insurance accounted for another 26 percent, and the remainder was paid directly by consumers." Third, you can identify aggregate trends in the use of services and source of funds. For example, "hospital spending increased 14.9 percent overall in 1982, while Medicare hospital outlays on an NHE basis rose 16 percent." And you can quantify the contributors to increases in health spending: "Price inflation accounted for 78 percent of the increase in personal health care spending in 1982. Population growth accounted for another 8 percent, and 14 percent was due to other factors, such as increased use of services, increased intensity of care, and so on."

Now the bad news. What are the shortcomings of our estimates? For one thing, it is hard to get at estimates of spending by state. Some data are available, but not many. Those that are available need a lot of nurture before they bear fruit. One of the big problems is migration of patients across state lines, a problem especially pronounced here in Washington DC, where the medical marketplace spans three states. None the less, I am happy to report that state estimates through 1982, consistent with national health expenditures, may well be available by next June.

Another drawback of our estimates is that they cannot be decomposed to family or individual consumption patterns. Recall that in all of national health expenditures, we sample directly only a small group of insurers, and no consumers. Not many of our data contain information on the age of the beneficiary or recipient, so that, as with state estimates, we must work up proxy measures of consumption. Individual consumption patterns cannot be isolated using macroeconomic data, any more than I could infer your family's food consumption using Giant Foodstore's annual report. At this time, we cannot trace money back to the factors of production, so we cannot create an input/output table for health in the U.S. Nor, sadly, have we the material with which to relate the economic accounts of health spending to measures of health status.

However, lest you plunge into the abyss of despair at this report, I will tell you that we are working on many of these, with varying degrees of success. We have begun a two-year benchmark of national health expenditures, to modernize the definitions and concepts, to incorporate the NMCES and NMCUES survey results, and to attempt an I/O table. Your advice and comments are welcome; please send them to me or to Mr. Gibson. Estimates of state spending are being brought up to date, and we hope to do the same with estimates of spending by age. If you will allow me to mix a metaphor, "Yes, Virginia, there is a light at the end of the tunnel."

So much for Mr. Gibson's theme. Now to address the Twinkie Metaphor.

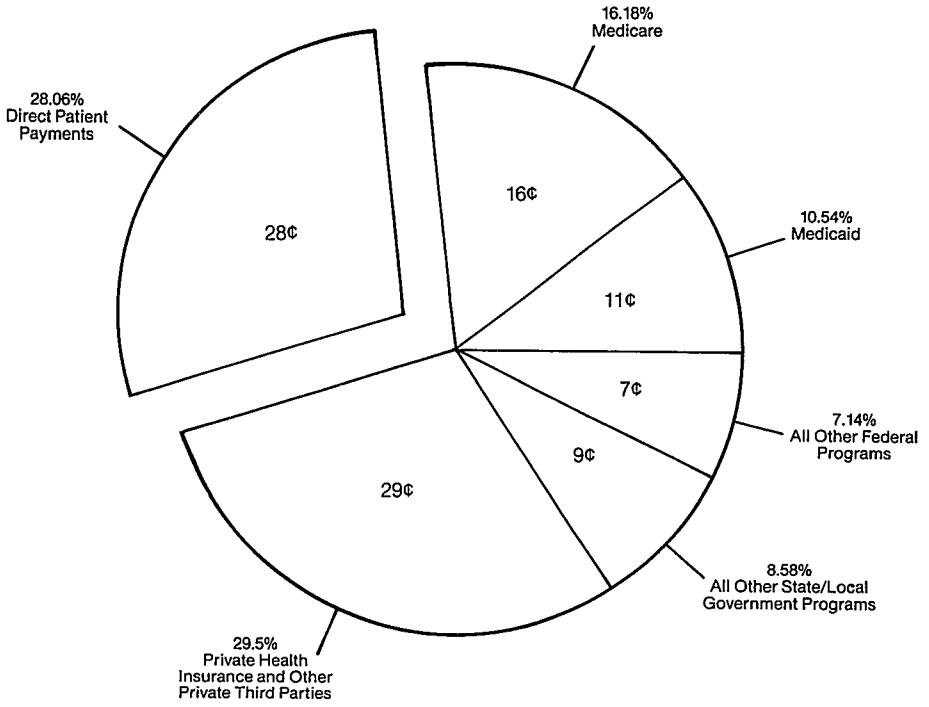
My address so far has been spent showing how we construct the national health accounts, and discussing their strengths and limitations. I have done so with some trepidation, because I have had experiences like this before: people have asked me questions about my work, and I have answered their questions, and their reaction reminds me of an experience from my own life:

When I was growing up, I loved Twinkies. They were the right shape to fit in your hand, they had that nice golden color, they went down really smoothly. Then, one day, I read the side of the package, and found out what they were made of. I have not eaten a Twinkie in over ten years.

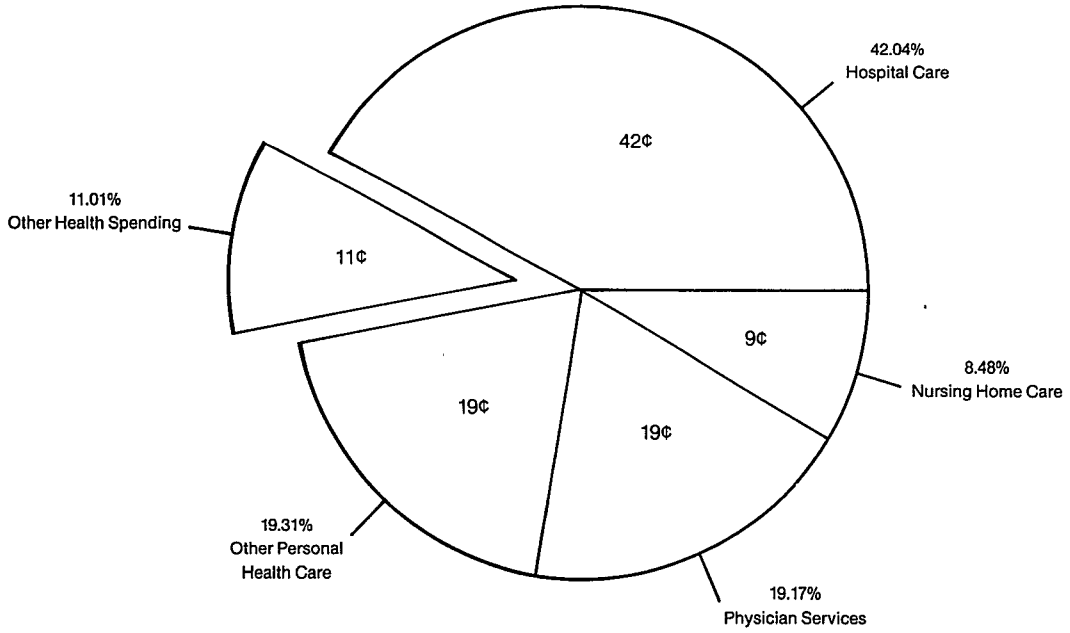
Please do not think of national health expenditures as the junk food of health data. We may not be all-natural; we may not be delicate; petit fours we ain't; but -- I assure you -- we are <u>not</u> Twinkies.

The Nation's Health Dollar in 1982

Where it Came From...



...And Where it Went



Source: Bureau of Data Management and Strategy, Health Care Financing Administration

ANALYZING THE COST EFFECTIVENESS OF KIDNEY TRANSPLANTATION

Paul W. Eggers, Health Care Financing Administration

With the enactment of the Social Security Amendments of 1972 (P.L. 92-603, Section 2991) Congress extended Medicare coverage to most of the persons suffering from end-stage renal disease (ESRD). Since the implementation of the original ESRD law, the program has experienced rapid growth both in the population served and in program costs. In 1974, Medicare expenditures for the 16,000 persons covered under the program were $250 million. By 1979, costs had risen to $1 billion and enrollment to 51,000. The 1982 expenditures are expected to be $1.8 billion and by 1986, costs are projected to reach $2.8 billion (HCFA, 1983). There has been one major change to the program, the End-Stage Renal Disease Program Amendment of 1978 (P.L. 95-292). This amendment was designed to promote efficiency and economy in the delivery of services by encouraging home dialysis and transplantation for the maximum number of suitable patients. Changes implemented through this amendment included extension of eligibility from 1 to 3 years post transplantation, increased coverage of kidney acquisition costs, 100 percent reimbursement for home dialysis equipment and expanded coverage of home dialysis supplies.

Since 1978, there has been one additional legislative change which was a part of the Omnibus Reconciliation Budget Act (OBRA) of 1981 (P.L. 97-35). OBRA required the Health Care Financing Administration (HCFA) to develop an incentive reimbursement for renal dialysis facilities based on the composite costs of home and facility dialysis. The rate would apply to all patients regardless of place of dialysis thereby providing an incentive to have patients dialyze in the least expensive manner, presumably at home.

The debates on the ESRD program and efforts to remodel the program along more efficient lines have taken place largely in the absence of good quantitative data on actual incurred costs. For instance, estimates of program expenditures under alternative proposals usually assume 156 dialysis sessions per patient per year times an average Medicare reimbursement of $110 per session (80 percent of the typical fee screen of $138). By failing to account for the fact that many patients do not dialyze three times a week, this procedure tends to overestimate dialysis costs. Costs of hospitalization for ESRD patients are often based on expert opinion or on small samples of patients. Similarly, cost estimates of transplantation often fail to include subsequent hospitalization for rejection episodes or back-up dialysis costs.

Perhaps the best cost-effectiveness comparison between dialysis and transplantation was performed by Stange and Sumner (1978). They compared the cost effectiveness of facility dialysis, home dialysis and cadaver transplants over a 10-year time period. Their estimates were based on survival data taken from the National Dialysis Registry and cost figures based on estimates of what various treatment therapies should cost. Thus, while the survival figures were based on actual experience, the cost figures were largely guesses.

The present study compares the cost effectiveness of transplanation with dialysis. It differs from the Stange and Sumner analysis in that actual Medicare reimbursements are used rather than estimates. This study estimates the point at which lower post-transplantation costs balance the high initial costs and transplantation becomes cost-effective on a per capita basis.

Methods and Data

The data for this study were taken from three sources; the ESRD Medical Information System (ESRD-MIS), the Medicare Statistical System (MSS) and the Kidney Transplant Histocompatibility Study (KTHS). The ESRD-MIS is an information system maintained by HCFA for the purpose of tracking ESRD patients and providing patient profiles for mandated reports to Congress. Date of onset of renal failure, age at onset, type of therapy and date of death were taken from this data base. The MSS is a by-product of the basic Medicare claims processing functions. All Medicare reimbursements for 100 percent of the ESRD patients were taken from this data base. Reimbursements include short stay hospital discharges, physician payments, outpatient billings (most of dialysis reimbursements are included here), home health care and skilled nursing facility stays. The KTHS was a multi-year clinical study of kidney transplantation performed under the auspices of the National Institutes of Health. Basic patient and kidney graft survivals for live related donor (LRD) and cadaver (CAD) transplants were taken from this study.

The development of the model involves two preliminary steps. First, probability estimates were developed for various patient outcomes. For dialysis patients there are two basic outcomes; survival or death and a probability associated with each. For transplant patients, there are three possible therapy outcomes; survival with a functioning graft, graft failure with return to dialysis and death. Each patient who survives through a year is subject to a new set of probabilities based on his/her therapy status at the beginning of the next year. The probability estimates are derived from survival analyses using the standard modified life table method (Cutler and Ederer, 1958). For transplant patients, these estimates came directly from the KTHS. For dialysis patients, a survival analysis was done using HCFA-MIS data. Because very few patients are transplanted over the age of 55, the dialysis survival analysis was limited to persons under this age. Survival estimates reflect the actual survival of transplant and dialysis patients during the period 1973 through 1979. Medicare reimbursements were calculated for each group of patients by therapy outcome.

For instance, Medicare reimbursements were calculated separately for dialysis patients who survive for 1 year and for those who die during a year. Similarly, Medicare reimbursements for transplant patients differ depending on the outcome of the graft. Reimbursements were calculated on a person year basis to account for persons with less than 1 year of Medicare coverage. Reimbursements are for the calendar year 1979.

Results

Table 1 presents the 5-year patient survival estimates for dialysis patients and transplant patients by type of donor. Transplant patients with live related donor (LRD) transplants experience the best survival rates; 91 percent can expect to survive 1 year after transplantation and 81 percent will survive for 5 years. Recipients of cadaver (CAD) grafts have the second best survival with 59 percent alive after 5 years. Dialysis patients have a first year survival of 86 percent which is slightly higher than first year survival for cadaver graft recipients. However, by the end of 5 years, only about one-half can be expected to survive.

Graft survival by donor source is shown in Table 2. For both types of graft, the greatest failure rate is in the first year after transplantation. One-fourth of LRD grafts and almost one-half of CAD grafts do not survive the first year. After that, the failure rate declines noticeably so that after 5 years 60 percent of LRD grafts and 31 percent of CAD grafts can be expected to be functioning.

TABLE 1
ESRD Patient Survival by Type of Therapy

Year	Dialysis2/	Transplant1/ Live Related Donor	Transplant1/ Cadaver Donor
	Percent Surviving		
1	86	91	83
2	73	89	77
3	64	86	72
4	58	82	66
5	52	81	59

1/ Transplant patient survival taken from KTHS

2/ Dialysis patient survival taken from ESRD-MIS

TABLE 2
Kidney Graft Survival by Type of Donor

Year	Live Related Donor	Cadaver Donor
	Percent Surviving	
1	75	51
2	71	46
3	67	41
4	63	35
5	60	31

Source - KTHS

Average per capita Medicare reimbursements by patient outcome are presented in Table 3 1/. A successful transplant (graft still functioning at the end of the year) incurs, on average, $32,067 in Medicare reimbursements. This includes the transplant stay, all back-up dialyses, physician fees, kidney acquisition costs and other hospitalizations including rejection episodes without loss of graft. Not included are coinsurance and deductibles and non-covered services such as outpatient drug costs. If the graft fails in the first year, then the cost to Medicare is $44,639, or $12,572 greater than a successful transplant. Most of this can be attributed to the hospitalization costs associated with a graft failure. The third outcome in the first year of transplant is death and, on a per capita basis, the costs are $62,886. This is a person year equivalent figure. In reality, dying patients have an average of one-half a year of life in the year they die. The model corrects for this by taking one-half of $62,886 for dying patients.

TABLE 3
Average Medicare Per Capita Reimbursement by Patient Outcome, 1979

Patient Outcome	Medicare Reimbursements Per Person Year
Transplant-1st Year-Graft Functioning	$32,067
Transplant-1st Year-Graft Failed	$44,639
Transplant-1st Year-Death	$62,886
Transplant-2nd & 3rd Year-Graft Functioning	$ 4,074
Transplant-2nd Year and Over-Graft Failed	$30,189
Transplant-4th Year and Over-Graft Functioning	$ 0
Maintenance Dialysis	$18,127
Death on Dialysis	$28,253

If the graft continues to function, the second and third year Medicare reimbursements are $4,074 per person. At the end of 3 years post transplant, an ESRD patient loses entitlement. Therefore, Medicare reimbursements fall to zero for successful transplants. If, however, the transplant fails in any year subsequent to the transplant year, the return to dialysis costs are $30,189. This figure is a combination of maintenance dialysis and graft rejection costs.

1/ Reimbursements were also calculated separately for LRD and CAD transplants. By outcome, reimbursement did not greatly vary so they were combined for simplicity of presentation.

Maintenance dialysis costs for a full year without death are $18,127. This figure is estimated for never transplanted patients and patients returning to dialysis after transplant rejection. For a death on dialysis, the costs are $28,253. As with a transplant death, patients have an average of one-half a year of survival in the year of death.

Model of Five Year ESRD Costs

Comparison of dialysis and transplant outcomes were done by estimating the 5-year costs and person years of survival by combining the observed survival characteristics with the Medicare reimbursements by outcome.

Table 4 presents the cumulative Medicare reimbursements and person years for a hypothetical 1,000 person cohort for each of the three treatments. First year costs for 1,000 dialysis patients amount to $17.6 million. This rises steadily so that, by the end of 5 years, the total cumulative cost of this cohort is $67.1 million. After 5 years, the dialysis cohort will have accounted for 3,570 person years. For transplant patients, the distribution of costs by year are much different. Over one-half of the 5-year cumulative costs are expended in the transplant year (55 percent for CAD and 61 percent for LRD). Total 5-year costs for the CAD and LRD cohorts are $65.6 million and $55.9 million, respectively. CAD cumulative costs are thus two percent lower than dialysis costs and LRD costs are 17 percent lower. The major difference in cumulative costs between the CAD and LRD cohorts are the much higher rejection and return to dialysis rates among CAD transplanted patients.

TABLE 4
Cumulative Medicare Reimbursements and Person Years by Modality (Cohorts N = 1,000)

	DIALYSIS		TRANSPLANT			
			Cadaver		Live-Related	
Year	Person Years	Costs	Person Years	Costs	Person Years	Costs
1	930	$17.6M	915	$36.0M	955	$34.0M
2	1,725	$32.6M	1,715	$44.5M	1,855	$40.7M
3	2,410	$45.5M	2,460	$52.7M	2,730	$47.5M
4	3,020	$56.9M	3,150	$59.5M	3,570	$51.7M
5	3,570	$67.1M	3,775	$65.6M	4,385	$55.9M

The actual cost effectiveness of transplantation is greater than suggested by total cumulative Medicare reimbursements. Because of the better survival rates for transplant patients, the total number of years of life is greater for the same size cohort. For CAD transplants, the cumulative number of person years is six percent greater (N = 3775) than for dialysis patients, and for LRD transplants cumulative person years is 23 percent greater (N = 4385). The effect of this increased survival is illustrated in Table 5.

The average per capita costs for dialysis patients do not change much from year to year, remaining at slightly under $19,000. This is due to the fact that the mortality rate is relatively constant from year to year. Each year the mix of costs for living patients and dying patients remains about the same. However, for transplant patients, the high first year costs drop precipitously in following years, particularly for successful transplants. Thus, the overall average costs continue to drop. For the CAD transplant cohort, the cumulative per capita average costs equal the dialysis cumulative per capita costs at about 4 years. For LRD transplants, this point is reached before the end of the third year. By the end of 5 years, CAD and LRD transplant per capita costs are 7 percent below and 33 percent below dialysis per capita costs, respectively.

TABLE 5
Cumulative Medicare Per Capita Costs by Modality

		Transplant		
			Live	
Year	Dialysis	Cadaver	Related	Combined
1	$18,900	$39,300	$35,600	$38,200
2	$18,900	$25,900	$22,000	$24,700
3	$18,900	$21,400	$17,400	$20,100
4	$18,800	$18,900	$14,500	$17,400
5	$18,800	$17,400	$12,700	$15,800

The mix of transplants is heavily weighted toward CAD donors. About 70 percent of the transplants performed in the U.S. are with CAD donor organs. Because of this, the actual combined cost effectiveness of transplantation will be more skewed toward the CAD transplants. This is shown in Table 5. Because 70 percent of transplants are from CAD donors, the overall cumulative transplant per capita costs are 16 percent lower than the cumulative per capita costs for dialysis patients for the 5-year period.

Conclusions

The results of this analysis show that transplantation in general is a more cost-effective treatment for ESRD patients than dialysis, although the net savings do not appear until 4 years after transplantation. However, given the direction of the trend in costs, each additional year beyond 5 years will continue to increase the cost-effectiveness of transplantation. It should be noted though, that since 1979, increases in costs have been greater for transplanation than for dialysis. Dialysis costs have been held relatively constant due to the Medicare fee screen whereas transplant costs have not been regulated. Therefore, it is likely that the current cost effectiveness of transplantation is less than suggested by this paper.

It should not be assumed from this analysis that transplantation is always the best option for any ESRD patient. One major constraining factor is age. Very few patients are transplanted over the age of 55, yet 47 percent of all ESRD patients are 55 years or over (Eggers, 1983). Thus, almost one-half of ESRD patients are not good candidates for transplantation on the basis of age alone. Another

consideration is quality of life. Whereas most patients would probably agree that a successful transplant results in a much improved quality of life, there are risks involved, as with any operation. The decision to transplant should always be made on the basis of the individual patient's background and prognosis.

This decision making will be greatly influenced by advances in ESRD treatment technology. Two recent advances which are likely to influence the relative desirability of types of treatment are continuous ambulatory peritoneal dialysis (CAPD) and the immunosuppresant cyclosporin. CAPD has the potential of reducing the cost of dialysis while improving the quality of life by freeing patients from a dialysis machine. The major concern with CAPD is the increased chance of infection resulting in peritonitis. Cyclosporin has the potential of increasing transplant graft retention rates, particularly for CAD transplants. As data become available on the these and other changes in ESRD therapy, the relative cost effectiveness, or desirability, of any treatment option is likely to change.

References

Health Care Financing Administration, End-Stage Renal Disease Report to Congress, 1983, in press.

P.V. Stange and A.T. Sumner, "Predicting Treatment Costs and Life Expectancy for End-Stage Renal Disease," New England Journal of Medicine, February, 1978.

U.S. Department of Health and Human Services, The Kidney Transplant Histocompatibility Study, NIH Publication No. 80-2164 (May, 1980).

S.J. Cutler, and F. Ederer, "Maximum Utilization of the Life Table Method in Analyzing Survival," Journal of Chronic Diseases, (8) p.p. 699-712, 1958.

P.W. Eggers, R. Connerton and M. McMullan, "The Medicare Experience with End-Stage Renal Disease: Trends in Incidence, Prevalence, and Survival," in press. Health Care Financing Review, 1983.

THE SHAPING OF PUBLIC POLICY FROM THE PERSPECTIVE OF A DATA BUILDER: THE HOSPITAL EXPERIENCE

Allen Dobson and Ronald Bialek, Health Care Financing Administration

Introduction

The degree to which data analysis and health services research studies influence the development of health care policy has often been debated. While the extent of such influence cannot be easily observed or measured, it is clear that politicians, political appointees, and policy analysts must get their knowledge from some source; after all, these individuals are not born with an innate set of public policy options. Health care policy development is formulated by policymakers after exposure to a wide variety of sources such as: newsletters, magazines, verbal information, briefings, and occasionally from the original sources of data and research.

The decisionmaker is assisted by data analysis and research studies on three levels: identifying the problem, developing and testing a range of solutions, and evaluating the policy's eventual effectiveness. The process is dialectic in nature, with data being used progressively for problem identification, solution development, and policy assessment. Evaluation of the policy ultimately sets the stage for the identification of new problems (or better resolution of old problems).

This paper contends that data analysis and research studies in recent time periods have been instrumental in shaping public and private health care policy. As will be demonstrated, the knowledge necessary for making policy decisions is obtained by policymakers through exposure to and examination of data generated through research studies, experimentation, and analyses.

There are two assumptions underlying our analysis throughout this paper. First, we assume that health care data reflect reality, in that the wealth of information generated in the area of health services research is more right than wrong in the way it portrays underlying health status and related activities. Secondly, we assume that data are used for policy development, but not necessarily for policy choice. Once the problem is defined and policy alternatives are developed, the actual policy is chosen based on political and philosophical considerations as well as economic constraints.

Development of the Argument

For the purposes of this paper, the term "data" is broadly defined to include: survey and other data; data analyses and research studies; and results from experimentation and demonstrations. The term "policy development" refers to activities associated with the design and implementation of public and private health care programs.

As data builders are prone to do, we have built a data base to test our hypothesis that "data matter." First, we reviewed a variety of books, periodicals, and reports used throughout the past 60 years for identifying public health care problems, developing policy solutions, and evaluating program impacts. An investigation related to all public and private health care policy development would have been too vast, so we decided to concentrate on hospital care policy development. We then proceeded to examine the data base looking for time periods that possessed two basic components: 1) a similar theme dominating the health care policy debate; and 2) a similarity in the degree to which data were used for problem identification, policy development, and program evaluation. This examination resulted in the designation of five phases of health care policy development occurring during the past six decades.

Due to space limitations, we were unable to include tables showing the numerous instances of the historical relationship between data and policy development which have occurred throughout the past 60 years. (Copies are available from the authors upon request.) What we have attempted to do, however, is to indicate the substance and implications of our data base tables in the following text.

Our analyses show that policy determination in successive phases relies increasingly on research studies and data analysis. While no single study or data source, in and of itself, can be expected to set the stage for policy development, data in aggregate and their subsequent interpretations, clearly have served to identify and focus the issues that have culminated in health care policy over the past 60 years. We make no claim that our presentation is all inclusive, but rather that it is representative of the events taking place during each phase and is adequate to support our claim that "data matter."

The first phase of the five we have identified began around 1920 when the ability to obtain hospital care became an important health care policy issue. The period culminated in 1946 with passage of the Hill-Burton Act. During the second phase, 1947 to 1965, access to hospital care improved for some sub-groups, yet still lagged behind for others. This period resulted in the enactment of legislation for indirectly improving health care access by increasing supply through manpower legislation and directly improving financial access through Medicare and Medicaid program legislation. The third phase began when statistical data showed rapid and continuing increases in expenditures and at the same time indicated residual access problems. Throughout the period of 1966 to 1972, regulatory programs were designed for controlling costs while federally-funded health care programs were expanded to improve access for other sub-groups. Public interest increased for some form of national health insurance (NHI). During the fourth phase, 1973 to 1979, the results of medical care spending analyses dominated public policy discussion along with a dampening of the debate over the merits of NHI.

This period witnessed implementation of regulatory programs to control hospital expenditures.

The inability of the regulatory efforts to control costs prompted calls for other approaches to health care policy development in the fifth cycle. This period began in 1980 and still is evolving. States have been given greater flexibility to operate Medicaid, the Medicare Hospital Prospective Payment System has been enacted, and competitive solutions are gaining favor over regulatory options. Analysis of demographic data and cost projections indicate that without cost containment and/or reduced benefits, the Medicare trust funds will be insolvent by the end of the decade. This has led to much of the urgency around the rapid succession of newly implemented programs during the fifth phase.

The remainder of this paper discusses these policy development phases and speculates on how data will be used in the future as new phases emerge.

The Phases of Hospital Care Policy

Phase I: 1920 to 1946 - Hospital Care Becomes a Public Policy Issue

By the twenties, the development of antibiotics improved surgical techniques, and technological advances lead to increased demand for hospital care. At the same time, the supply of hospitals was insufficient to meet increased demand, resulting in hospital cost increases. During this phase, data were important for identifying the problems of limited supply (access) and rising costs, but played only a small role in developing policies or evaluating their impacts.

As early as 1924, data were presented by L. Mayers and L.V. Harrison indicating that physician demand was increasing while per capita supply was decreasing.1/ In 1934, the privately funded Committee on the Cost of Medical Care issued preliminary reports indicating that due to rising cost and supply shortages, low-income groups were having difficulty obtaining access to health care.2/

The policies developed from the presentation and analysis of these data were aimed at providing direct aid to patients for improving access to health care. In the early 1930s, National Health Insurance (NHI) was being debated as a mechanism for improving access to health care - the debate was, in part, fueled by the economic and social misery of the Great Depression. The proponents of NHI claimed that similar programs were working in Europe. However, the medical profession claimed that such overwhelming Government involvement in health care would hinder the quality of services. As a means for aiding their own financial burdens, hospitals began underwriting private health insurance so that they would be able to receive more payments for their services. This eventually evolved into the Blue Cross insurance program. Finally, in 1935, the Federal Social Security Act (P.L. 271) was passed as a modest response to pressures for NHI.

The new law made funds available to States on a matching basis for maternal and infant care, rehabilitation of crippled children, general public health work, and aid to dependent children under 16. This legislation also created the Social Security Board which was mandated to perform research and data gathering tasks involving health care.

In 1936, the U.S. Public Health Service reported that the poor were sicker more often and received less adequate care than other economic groups. The report also indicated that the majority of the population had no financial cushion to pay for hospital care.3/ In a series of reports in the early 1940s by J.W. Moutin, E.H. Pennel, and V. Nicolay, data indicated that the supply of hospitals should be expanded.4/ In 1944, the American Hospital Association Commission on Hospital Care began surveying the nation's hospital needs. Preliminary results showed that facility supply was insufficient.5/ Increasing the supply of hospitals was seen as a mechanism for improving access to the entire population.

Private health insurance began improving access to the middle- and upper-income families. But, the final thrust of policy during this phase was to increase the number of hospitals. In the early 1940s, the number of veteran hospitals was increased and treatment was expanded to veterans with non-military related ailments. Finally, in 1946, the Hospital Survey and Construction Act (Hill-Burton, P.L. 725) was passed. This was a major piece of legislation providing funds for hospital construction and planning. In addition, the Act mandated further study of hospital care access problems. As pointed out by Paul Starr in his recent book, The Social Transformation of American Medicine, "Advocates of Hill-Burton originally argued that the program would help provide access to hospital care for families and communities that otherwise could not afford the cost."6/ The first phase culminated with legislation supporting the traditional economic view that increasing supply would not only increase access but decrease costs as well.

Phase II: 1947 to 1965 - The Widening Access Gaps of the Poor and Elderly

Data during this phase continued to be important for identifying problems and started to become important for evaluating the impacts of policies. To improve Government research and aid the administration of programs, the Department of Health, Education and Welfare (DHEW) and the National Center for Health Statistics were established during this period. Data were showing that access was improving for some groups, but was not improving for others.

Hospital costs continued to rise, but private health insurance helped to reduce some of the negative impact of cost increases by improving access for the middle- and upper-income, as pointed out in surveys and studies during the late 1940s and early 1950s by Odin Anderson.

Access was also improved somewhat for the poor by the Social Security Act Amendments of 1950 (P.L. 809). The Amendments expanded

coverage to 10 million more persons by increasing the income cutoff for the disabled poor and raising the existing payment levels. Access was further improved for the middle- and upper-income groups through the 1954 Internal Revenue Tax Code ruling allowing employer provided health insurance to be tax exempt. "Experience rating" developed toward the later part of 1950s also encouraged the purchase of private health insurance by lowering premiums and making policies more comprehensive for low-risk groups.

As the supply of hospitals grew and demand continued to increase, the supply of physicians was becoming a problem. In 1954, the Presidential Commission on Health Needs report stated that there was a physician shortage and that more training was necessary.8/ The Surgeon General in 1959 reported that there existed a severe shortage of medical personnel.9/ Cost of hospital care also was rising rapidly. DHEW data showed that costs had doubled between 1950 and 1960.10/

In response to growing gaps in health care access of different economic and age groups, the Federal Government initiated new research and direct aid programs. In 1956, the National Health Survey Act (P.L. 654) was passed which required data to be collected and research to be conducted on the nation's health needs. A program also was enacted by the military providing a form of health insurance for armed forces dependents. A year later, the Public Health Service initiated the Health Interview Surveys. However, there was growing sentiment that more direct aid was necessary to a seemingly neglected sector of the population, the elderly. In 1960, Social Security Act Amendments (Kerr-Mills, P.L. 86-778) were passed. The Amendments were designed to fill the gap of access to the aged. Kerr-Mills provided matching funds to States for aiding the medically indigent elderly. The Amendments also authorized more research in the areas of health expenditures and the impacts of the new legislation. In addition, the Migrant and Refugee Assistance Act (P.L. 87-510) was enacted in 1962 to improve access for another disadvantaged group, American migrants.

Still, the shortage of physicians remained and was becoming worse. In 1963, the Health Professions Educational Assistance Act (P.L. 88-129) was passed to address this problem. The Act provided loans to undergraduate students and funded the construction of undergraduate institutions. In 1965, the Act was amended to provide capitation payments to medical schools for increasing enrollment.

Meanwhile, the Kerr-Mills Act was being implemented and evaluated. In 1963, the Presidential Commission Report on the Kerr-Mills Act was less than flattering to the program. The report noted that the Act was not being implemented by most States and that the program was not serving the intended population.11/ The 1963 Survey of the Aged by the Social Security Administration reaffirmed the Commission's findings by showing that only 50 percent of the elderly population had any form of public or private health insurance coverage.12/ In addition, hospital inpatient costs per day had risen at an annual average of 10.4 percent between 1955 and 1964, making access even more difficult for the poor and elderly.

On the finding that the Kerr-Mills Act did not work, a new direct aid program was enacted in its place by the 1965 Social Security Act Amendments (P.L. 89-97). The Amendments provided aid to the elderly under Medicare and aid to the poor under Medicaid. Medicare and Medicaid reimbursement methods were based upon the private industry practice of the day--essentially retrospective cost reimbursement models. In addition, the Amendments mandated evaluation of the program's impacts on access to health care, health services, and health expenditures.

Throughout the second phase, access problems for the poor and elderly were compounded by increasing health care costs. The Federal Government responded to earlier data results through expanded research in the mid-1950s. The Federal role grew even greater through enactment of more supply initiatives and direct aid programs.

Phase III: 1966 to 1972 - Emerging Conflicts in Hospital Care Policy: Access vs. Costs

Throughout the third phase, there was a growing sophistication of data gathering techniques and computer assisted analyses. Demonstrations also became mechanisms for experimenting with alternative policy options, departing from the private sector reimbursement models upon which the original Medicare and Medicaid programs were based. These developments increased the decisionmaker's reliance on data for identifying problems, developing solutions, and evaluating impacts.

Almost as soon as Medicare and Medicaid were implemented, research showed a sharp increase in access to health care for the poor and elderly. However, access remained a problem for over one million elderly, children of poor families, and the nearly poor.13/ To improve access and to continue identifying health care problems, the Federal Government initiated additional programs. In addition, the Medicare Statistical System was designed for evaluating the Medicare program and the National Center for Health Services Research was established for other evaluative and problem-solving purposes. (Because the Medicaid program was essentially run by the States, no Federally-centralized data base was developed for the program. This situation still stands today.) In 1966, comprehensive health centers were established for improving the supply of health care facilities. Also in that year, the Comprehensive Health Planning and Services Act (P.L. 89-749) was enacted giving States authority and funding to form voluntary health care planning agencies. In 1967, Social Security Act Amendments (P.L. 90-248) were passed establishing the Early and Periodic Screening, Diagnosis and Treatment program. The Amendments also authorized DHEW to study the impacts of extending Medicare coverage to the

disabled and persons with end-stage renal disease. In addition, studies and experiments were mandated for developing reimbursement systems that would help to control the growing health care expenditures. Data were indicating that the private industry model for medical care reimbursement was leading to a rapid increase in hospital care expenditures. Alternatives were sought.

Hospital costs started growing out of control upon the enactment of Medicare and Medicaid. Between 1965 and 1970, Federal hospital expenditures went from $2.35 billion to $9.6 billion.14/ As a result, studies were initiated for finding ways to control costs. In 1967, the Presidential Commission on Health Manpower report stated that costs could be held down through improved hospital efficiency. A form of peer review was recommended to achieve this goal. In the early 1970s, L.B. Lave and J.R. Lave conducted a number of hospital rate control studies. They began looking at hospital case mix rather than length-of-stay. Hospitals were beginning to be viewed as multiple product firms.15/ Researchers were attempting to better define the hospital product so that payment mechanisms could be developed that would be equitable across hospital settings yet effective in controlling the growth of hospital expenditures.

At the same time that costs were rising and controls were being sought, additional access problems were being identified. In 1967, the Bureau of the Budget Report of the Committee on Chronic Kidney Disease recommended the establishment of a national treatment benefit program.16/ DHEW reports were showing that the disabled also experienced significant access problems.17/

The multitude of research studies on access and cost problems resulted in a wide range of debate and legislation throughout the later part of this phase. As costs were rapidly increasing, new calls were being mounted for NHI. In the early 1970s, States such as New Jersey, New York, and Maryland were implementing forms of hospital rate controls. Four new programs were enacted for improving supply of health care facilities and manpower: the Health Training Improvement Act of 1970 (P.L. 91-519); the Medical Facilities Construction and Modernization Amendments of 1970 (P.L. 91-296); the Manpower Omnibus Bill of 1971 (P.L. 91-667); and the National Health Services Corp. legislation of 1972 (P.L. 92-585). The Federal Government finally took a dramatic step for controlling costs when the Economic Stabilization Program (ESP) was placed into affect in 1971. But this program was only a temporary measure for controlling costs, with the freeze scheduled to be lifted in 1974. Finally, in 1972, a new set of Social Security Amendments (P.L. 92-603) were passed.

The 1972 Amendments addressed both expanding access and reducing costs. The past research studies and demonstrations addressing access and cost problems were relied upon for developing many of the programs contained in the Amendments. Medicare eligibility was extended to the disabled and persons with end-stage renal disease. Also, the elderly not covered by Medicare were permitted to pay a premium for coverage. On the cost control side, Professional Standards Review Organizations (PSROs) were mandated since utilization review appeared somewhat successful in a few States. Benefits were extended to Health Maintenance Organizations (HMOs) in anticipation that they would reduce per capita inpatient hospital admissions as demonstrated by experimental group payment plans. Authority was given for the Medicaid program to impose copayments. The DHEW was authorized to withhold payments to hospitals and physicians for unauthorized and unreasonable expenses. In addition, the Amendments mandated the continuation and expansion of research and experimentation for improving health care access and controlling costs.

This phase culminated with comprehensive amendments to the Social Security Act that attempted to strike a balance between two competitive forces: the desire to expand services and coverage under Medicare and Medicaid; and the growing concern over the runaway costs of these programs. This dilemma was not an easy one to solve. Decisionmakers were now relying more heavily on data for not only identifying problems and evaluating programs, but for developing the solutions.

Phase IV: 1973 to 1979 - The Shifting Emphasis of Hospital Care Policy: Regulating Cost

During this phase, Federally-sponsored research and demonstrations expanded in a dramatic fashion. Toward the end of the last phase, the Congress was realizing that cost was the major problem in the health care system. In response, new attempts were made to experiment with cost control programs. Data also were being relied on more heavily for evaluating the impacts of cost control programs and developing new policies. In addition, The Health Care Financing Administration (HCFA) was established during this period for administering Federal health care financing programs and assuring their quality, as well as for conducting research to assist in the development of solutions for alleviating public health care problems.

To support new research and data analysis, a variety of large-scale public and private studies and surveys were conducted. Data were becoming available from sources such as the Medicaid Management Information System, the National Ambulatory Medical Care Survey, the National Medical Care Expenditures Survey (NMCES), the National Medical Care Utilization and Expenditures Survey (NMCUES), and the Hospital Rate Setting Study.

Cost data indicated that the ESP was temporarily effective for keeping hospital increases down, but as soon as the freeze was lifted in 1974, hospital inpatient costs rose rapidly.18/ Congress realized that the ESP was only a temporary measure and had continued its search for effective cost control. In 1973, the Health Maintenance Organization Act (P.L. 93-222) was passed, providing additional bene-

fits to HMOs, such as loans and grants, to spur formation of these potentially money-saving institutions. In 1974, the National Health Planning and Resources Development Act (P.L. 93-641) was enacted. This legislation authorized the establishment of 200 Health System Agencies and planning agencies, required Certificates of Need (CON) for hospital capital expenditure reimbursement, and permitted DHEW to set limits to hospital charges.

The States also were becoming more involved in hospital cost controls. In the late 1970s, eight States had mandatory rate setting programs. Federal grants were encouraging establishment of other State programs and also were funding further cost control research. Policymakers were recognizing that hospital costs were dependent upon case mix and attempts were being made to more clearly define and measure case mix. One result of this work was that, by the mid 1970s, Yale University researchers, supported by DHEW, had developed 383 diagnosis related groups (DRGs) that could be used for categorizing hospitalized patients based upon case mix.19/ This DRG system would later be tested in State demonstrations.

In 1975, State Rate Setting Studies being conducted by the Social Security Administration, Office of Research and Statistics, were showing that rate setting reduces the magnitude of hospital cost increases.20/ In 1978, the Congressional Budget Office released a report showing that State rate setting reduced hospital expenditures increases by 3 to 4 percent.21/ Still, on a national scale, hospital expenditures were increasing by approximately 18 percent per year.22/

Immediate controls seemed necessary for dealing with the rapid hospital expenditures increases. In 1977, the Hospital Cost Containment (HCC) bill was introduced in Congress. The bill was to limit hospital charge increases significantly below present rates. At first, it appeared that the bill was likely to be passed. However, two developments occurred that led to the HCC bill defeat in 1979. The first development was the mounting evidence that Federal regulations for controlling costs often were not effective. Government and private studies were showing that the PSRO23/ and CON programs were having little, if any, positive impacts. The second development was the introduction of voluntary hospital rate controls in 1978. These voluntary controls in 1978 reduced hospital expenditure increases at a time when the second OPEC boycott was causing rapid inflation throughout the economy.24/ The seeming success of voluntary controls and the failure of regulations was enough to convince Congress that the voluntary hospital program was more desirable than the HCC bill.

Prior to entering this phase, the problem of increasing costs already had been identified. During this phase, programs were implemented for controlling costs, but the Federal regulations seemed ineffective. Some State rate setting was successful, but not enough was known to implement such a program on a national scale. Alternatives were sought, yet none were effectively enacted. As the phase came to an end, an even greater reliance by decisionmakers was being placed on the use of research studies, experiments, demonstrations, and data analysis for solving the cost problem.

Phase V: 1980 to Present - The Dominance of Fiscal Constraints and the Emergence of Competitive Solutions

Further evaluation of voluntary controls showed that they were not working. Costs were rising even faster than before.25/ Demographic data were showing that future population trends coupled with unrestrained hospital expenditures would eventually lead to insolvency of the Hospital Insurance Trust Fund. Emphasis was placed on using data for developing policies to solve the problem.

HCFA-sponsored rate setting studies continued to show that State controls are effective for reducing cost increases.26/ However, the programs still were not ready for nationwide application. As the voluntary hospital control programs appeared to be failing, Congress reacted to the urgent need for holding down expenditures with a barrage of legislative programs. In 1980, the Omnibus Reconciliation Act (P.L. 96-499) lifted the Medicare restrictions on home visits and authorized the use of "swing beds." In 1981, the Omnibus Budget Reconciliation Act (P.L. 97-35) cut the Federal share of Medicaid funding to States, gave States more flexibility over their programs, and permitted higher copayment rates. The 1981 legislation represented a new approach to health care expenditures control through explicit budget control.

Along with the fiscal constraints, policymakers were opting for changes in the reimbursement system for controlling costs. As a result, the DRG system continued to evolve during this phase and serves as an excellent current example of our hypothesis that "data matter."

By 1981, under HCFA sponsorship, Yale University had refined its initial set of 383 DRGs into 467 groups.27/ This refinement addressed earlier concerns identified through demonstrations that the data base for developing the original DRGs was too narrow and that too few medical care experts were consulted over the development of clinical categories. As they were used and improved, DRGs were beginning to be recognized as a practical basis for program development--a recognition that evolved into HCFA's Prospective Payment System (PPS). In addition, data from NMCUES and the National Health Accounts, combined with the American Hospital Association rate survey data, were reaffirming the need to control costs. Congress responded in 1982 with the passage of the Tax Equity and Fiscal Responsibility Act (TEFRA, P.L. 97-248). Under TEFRA, hospital reimbursement case mix cost limits were to be established, constraints were to be placed on budgets, incentives were provided to operate below budgets, peer review organizations (PROs) replaced PSROs, and the development of a national Medicare PPS was mandated. TEFRA, for the first time, placed limits on per case rather than per diem costs and introduced the Medicare

case mix index.

The number of States engaging in some form of hospital cost control had risen to 17 by 1982. DRGs were being successfully used by some States for categorizing patients in clinically meaningful groups and as the basis for payment. Following their successful demonstration, DRGs were proposed for use as the basis for a national PPS. In April 1983, Congress passed another set of Social Security Act Amendments (P.L. 98-21). The Amendments provide for the phased-in implementation of prospective DRG rates for Medicare hospital reimbursement and mandated a large number of studies for adjusting DRG rates, evaluating prospective payment program impacts, developing refinements to the DRG system, and expanding the Medicare prospective payment program to other settings and to physicians in the inpatient setting. The PPS represents the most recent example of data identifying a problem (rising costs), data being used as a tool for developing a solution to the problem (DRGs), and data being used for evaluating and refining an enacted program (mandated PPS studies).

If successful, the DRG program might be able to extend the 1982 Social Security Board of Trustees Annual Report projected date for the insolvency of the Hospital Insurance Trust Fund.28/ The rate of hospital inflation also might come closer to the general economy's rate of inflation instead of exceeding that rate by 9 percent as in 1982.29/ Congress is very likely to expand this program or implement others for holding down health care expenditures.

Future Data Needs and Health Care Issues

If the recent past is any indication, data will be relied on most heavily for evaluating existing policies and developing new ones. To the extent that current prospective payment and competitive approaches are successful, they will be expanded to other providers. If not, regulatory programs may once again be employed, perhaps even to the extent of explicit control of supply in the various health care sectors.

When the cost problem appears to be solved, a new post cost-concern phase may emerge emphasizing quality and access to care issues. Underlying future activity will be the increased rate of technological innovation. Which technologies should be used and who should benefit from them clearly will be the allocative questions of the future.

As an example of future data collection and research requirements, Congress has mandated in the most recent Social Security Act Amendments that a variety of studies be conducted for evaluating existing programs as well as for developing new policies. These studies include examining the: 1) impacts of the PPS on services, providers, patients, and technologies; 2) possibility of including capital expenditures and inpatient physician fees in DRG rates; 3) differences between rates in rural and urban areas; 4) need for severity of illness and intensity of services adjustments to DRG rates; 5) impacts of DRGs on admissions (volume); 6) impacts of DRGs on State reimbursement systems and Medicare and Medicaid; and 7) possibility of including more hospitals and other facilities in the DRG program. In addition, Congress has authorized and encouraged the expansion of State demonstrations.

Studies for establishing and adjusting DRG rates will continue to be conducted. These studies will use data sources such as the Medicare Provider Analysis and Review File, Medicare cost reports, and Medicare discharge files.

Data analysis and research studies also will be necessary for examining emerging concerns over health status, technology, and efficacy. Data will be needed to answer such questions as: What constitutes necessary care? What is high quality care? Which organ transplants (and other procedures) are effective and who should receive them? How much does health care improve health status?

Finally, another issue that needs resolution is explaining the variation in hospital use across geographic areas. DRGs, in their capacity as hospital product measures, might be particularly well suited for investigating this issue.

Final Observations

The cycles of public health care policy appear to be occurring in shorter time frames. This is due in part to the increasing pace in which data are collected and analyzed and to the rapidly changing demographic characteristics of the U.S. population. The data have heightened concerns for developing effective policies and providing immediate as well as long-term solutions to the health expenditures problem.

The increasing frequency and complexity of Congressional mandates for new studies reflect a heightened awareness of the degree to which research, demonstrations, and data analysis can and should impact on the public policy process. In each cycle, we have observed the Federal Government's increasing involvement in health services research. Large numbers of grants have been awarded, Federal data gathering and analysis institutions have been established, and studies have been increasingly mandated by Congress.

The mechanisms are in place for providing ever increasing and more reliable research studies and data analysis for assisting future problem identification, policy development, and program evaluation. However, as the amount of data grow, the form in which decisionmakers receive the data will, by necessity, continue to evolve. More analyses will be presented in summary form. Even fewer original sources will be read by policymakers. Instead, policymakers will increasingly rely on briefings and issue papers which summarize and integrate study findings. This clearly places a heavy responsibility on those who condense, summarize, and combine study results.

As has been demonstrated throughout this paper, "data matter." Data have been instrumental in the examination of the entire spectrum of public health care issues. It is.

likely that research studies and data analysis will continue to play a key role for decision-makers in identifying problems, developing appropriate and effective policies, and evaluating program impact.

Acknowledgements

The authors wish to express their appreciation to Michael Fitzmaurice, Marian Gornick, Bryan Luce, Herbert Silverman, and Sherry Terrell for their reviews, comments, and suggestions.

Footnotes and References

1. L. Mayers and L.V. Harrison, The Distribution of Physicians in the United States. New York, New York: General Education Board, 1924.
2. Committee on the Cost of Medical Care, Medical Care for the American People. Washington, D.C.: 1936.
3. C.T. Pihlbland, "National Health Survey: Preliminary Reports." Washington, D.C.: National Institute of Health and U.S. Public Health Service, 1937.
4. J.W. Mountin, E.H. Pennell, and V. Nicolay, "Location and Movement of Physicians." Public Health Reports, Volume 57, 1942.
5. Commission on Hospital Care, Hospital Care in the United States. New York, New York: The Commonwealth Fund, 1947.
6. P. Stan, The Social Transformation of American Medicine. New York: Basic Books, Inc., 1982.
7. E.E. Flook and P.J. Sanazaro, Editors, Health Services Research and R & D in Perspective. Ann Arbor, Michigan: Health Administration Press, 1973.
8. J. Newhouse, The Economics of Medical Care. Menlo Park, California: Addison-Wesley Publishing Company, 1978.
9. J. Newhouse, The Economics of Medical Care.
10. P.A. Corning, The Evaluation of Medicare... From Idea to Law. Washington D.C.: Social Security Administration, Office of Research and Statistics, 1969.
11. P. Starr, The Social Transformation of American Medicine.
12. D.P. Rice, "Coverage of the Aged and Their Hospital Utilization in 1962: Findings of the 1963 Survey of the Aged." Social Security Bulletin, Vol. 27, No. 7, July, 1964.
13. S. Davidson and T. Marmor, The Cost of Living Longer. Lexington, Massachusetts: Lexington Books, 1980.
14. M.S. Freeland and C.E. Schendler, "National Health Expenditure Growth in the 1980s: An Aging Population, New Technologies, and Increasing Competition." Health Care Financing Review, Vol. 4, No. 3, March, 1983.
15. J.R. Lave and L.B. Lave, "Hospital Cost Functions." American Economic Review, Volume 60, 1970.
16. Health Care Financing Administration, "Implementing the End-Stage Renal Disease Program of Medicare." Office of Research, Demonstrations, and Statistics, March, 1981.
17. S. Davidson and T. Marmor, The Cost of Living Longer.
18. R.M. Gibson and D.R. Waldo, "National Health Expenditures, 1981." Health Care Financing Review, Vol. 4, No. 1, September, 1982.
19. Department of Health and Human Services, "Report to Congress, Hospital Prospective Payment for Medicare." December 1982.
20. Social Security Administration, unpublished preliminary results of State rate-setting studies. Office of Research and Statistics, 1975.
21. F.A. Sloan, "Government and the Regulation of Hospital Care." American Economic Review, Vol. 72, No. 2, May, 1982.
22. R.M. Gibson and D.R. Waldo, "National Health Expenditures, 1981."
23. Health Care Financing Administration, Professional Standards Review Organization 1973 Program Evaluation. January, 1979.
24. K. Davis, "Public-Sector Cost Containment Programs." Lecture at Johns Hopkins University, April 22, 1983.
25. K. Davis, "Public-Sector Cost Containment Programs."
26. Health Care Financing Administration, "National Hospital Rate-Setting Study." Office of Research, Demonstrations, and Statistics, April, 1980.
27. Office of Technology Assessment, "Diagnosis Related Groups (DRGs) and the Medicare Program: Implications for Medical Technology." A Technical Memorandum, July, 1983.
28. The Board of Trustees, Federal Hospital Insurance Trust Fund, "1983 Annual Report, Federal Hospital Insurance Trust Fund." June, 1983.
29. Department of Health and Human Services, unpublished 1982 hospital expenditures data.

Data Use and Analysis

Session M

ON THE USE OF STATE TAX FILERS DATA FOR ASSESSING THE NATURE OF HEALTH CARE EXPENDITURES: THE MINNESOTA EXPERIENCE FROM 1978-1980

M. Nagi Salem and Paul D. Gunderson
Minnesota Center for Health Statistics

The topic we have chosen to present in this session, the use of state tax filers for assessing the nature of health care expenditure in Minnesota, is timely, yet difficult. Timely in that health care cost and the use of secondary data are the forefront concern of this conference as well as the public, yet difficult since state capacity to monitor cost over time is so new. In fact, prior to 1976 no one had information about the total scope of health care cost in Minnesota.

As we are all aware, routine sources of health care expenditure data in recent years have consisted of randomized on-site household surveys, telephone interviews, medicare information systems and third party insurance carriers. Each of these data sources has limitations including data collection costs and comprehensive coverage of an area's populations.

The Minnesota Center for Health Statistics recently explored the use of samples of Minnesota income tax data for the years 1978 and 1980 as a secondary data source for assessing the distribution of itemized medical deduction by types of health care expenditures. The descriptive study reported herein focuses on:

(1) The scope of health care expenditure in Minnesota and the proportion of costs accounted for by medical tax deductions,

(2) A description of sampling procedures used to select itemized tax deduction records,

(3) A description of the distribution of medical deductions by type of expenditures, i.e., insurance, medicine and drugs, hospitals, physicians, etc.,

(4) An examination of the change in the medical components of Consumer Price Index (C.P.I.) between 1978 and 1980 relative to the change in medical tax deductions and average Minnesotan's income for the same period of time.

The Scope of Health Care Expenditures in Minnesota

The Minnesota Coalition on Health Care Costs estimated the total sum of expenditures by source of funds in 1980 to be 3.960 billion dollars. Forty two percent was paid by federal, state and local governments, and 27 percent paid by private insurance carriers and health maintenance organizations. The remaining 31 percent consists of out-of-pocket health care expenditures.

The question that emerges is "what proportion of these expenditures could be accounted for by reported medical tax deductions?" Minnesota ranked six in the nation in personal income tax, which places the state in the highest tax quartile. It's safe to assume that this distinction is associated with a liberal tax code that encourages recording of health care expenditures by those who itemize deductions. Minnesota residents who choose to itemize deductions could deduct full payments for: medical care or hospitalization insurance costs; medicare insurance deducted from social security checks; medicine, drugs, vaccines and vitamins prescribed by a doctor; hospitalization costs; fees paid to physicians, dentists and other health care professionals; institutions to provide care for mental illnesses or physical handicaps; examinations, x-ray and insulin treatment; ambulance services; vision, dental care, and medical and surgical aides; and lodging to receive medical care away from home. However, some health related payments could not be used as medical deductions. These items include travel to obtain medical and dental care; travel for rest or change; payments for personal hygiene items; payments for life insurance; payments for medicare included in FICA tax; and funerals, burial expenses and cemetery lots.

The proportion of medical tax deductions relative to total health care costs and out-of-pocket health care expenditures is an indication of the importance of using tax data to begin to understand the nature and the distribution of health care costs. Medical deductions accounted for approximately 17 percent of the total expenditures in 1980 in Minnesota. Also it accounted for about 62 percent or about 750 million of out-of-pocket health expenditures excluding insurance.

Sampling Selection of Itemized Tax Deduction Records

A stratified random sampling technique was utilized to select samples of itemized tax records. The Minnesota Gross Income variable was the basis for stratification. This variable is basically the federal adjusted gross income after certain modifications, either additions and/or subtractions being made. However, these modifications do not apply to all taxpayers. Hence, it is possible that Minnesota Gross Income could be negative. For each of the two tax year records used in this study, the selection process was initiated with the generation of twenty subfiles each representing one of the 20 defined income strata. The records within each stratum were then sequenced and a random number generator was used to select the required number of records based upon selection rates which were previously defined. The selection rates varied from one percent for records in the income range $10,000-$19,999 to 100 percent for records presenting the income range of $200,000 and over. It should be noted that greater weighting at the high and low end of the income strata was thought to be more likely to produce an accurate reflection of the population of tax filers.

The result was a selection of two samples with approximately 20,000 records for each year. The samples have a proven track record when compared to actual population statistics. For example, the difference between the sample and actual

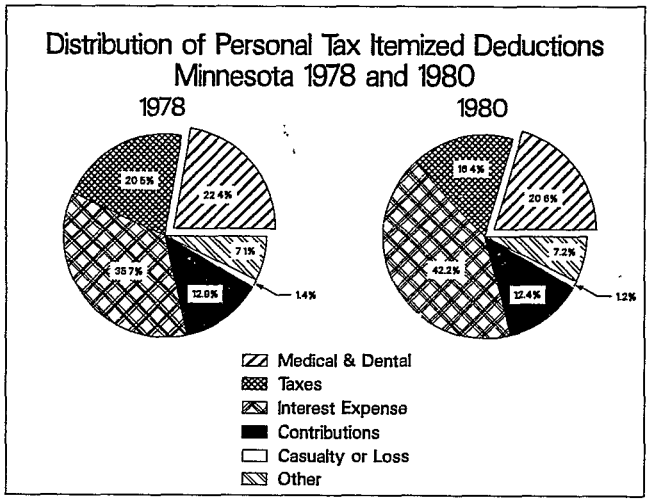
population statistics in the amount of personal deduction did not exceed .2 percent.

Distribution of Medical Deductions by Type of Expenditures

About 62 percent of all Minnesota tax filers had an incentive to file itemized deductions. However, the incentive to itemize varied by income. About 60 percent of those with reported Minnesota gross income between $1 and $10,000 itemized deductions. The percentage of those with income of over $30,000 was approximately 96 percent.

Personal tax deductions were grouped into six categories, namely: medical and dental, taxes, interest expense, contributions, casualty or loss, and other. The distribution of personal tax itemized deductions by type of deductions (Figure 1) showed small variation between 1978 and 1980 which may be an indication of the stability of the system. Figure 1 shows that medical and dental deductions accounted for about 21 percent of total deductions. The difference in interest expense between 1980 and 1978 reflects the beginning of the increase in interest rate.

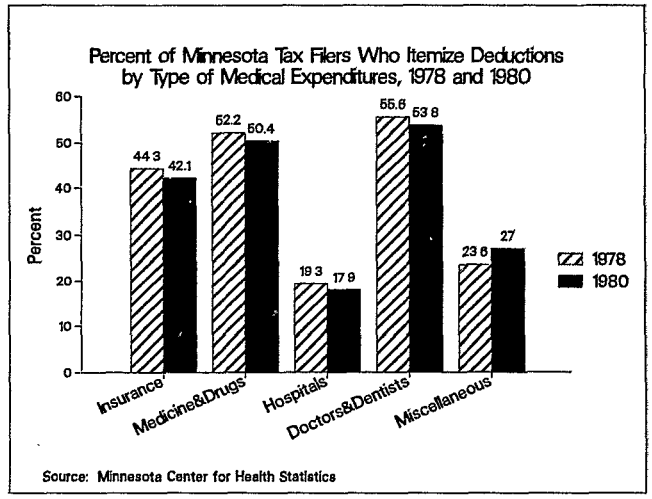
Figure 1



The Minnesota tax code requires the recording of medical tax deductions by tax filers into the following five categories: insurance; medicine and drugs; hospitals, physicians and dentists; and miscellaneous medical deductions. It was expected that all Minnesota tax payers who itemize deductions and claim medical deductions will not necessarily claim all types of medical deductions (Figure 2). In fact, in 1980 only about 18 percent claimed hospital costs, 43 percent deducted insurance expenses, and approximately 50 percent claimed both medicine and drugs as well as physician and dentist fees.

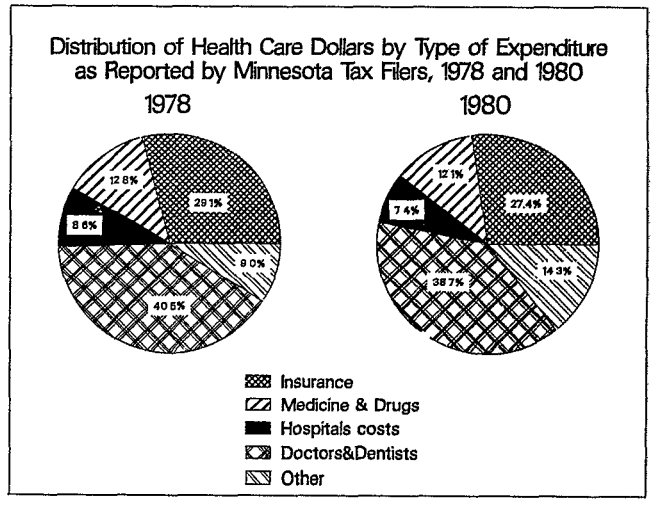
Note the difference in the percentage of those who claimed hospital expenses between 1978 and 1980 was -1.4 percent, while for those who deducted miscellaneous medical costs it was 3.4 percent. These percentages represent the range of differences between 1978 and 1980. The proximity of 1978 percentages to those of 1980 provide for additional demonstration of the reliability of the use of tax records for measuring health care costs.

Figure 2



Data related to the distribution of medical tax deductions by type of expenditures showed that despite the increase in expenditures between 1978 and 1980 the distribution of expenditures by type varied little. Figure 3 shows that relative to total medical tax deductions hospital costs were the least (7.4 percent in 1980) while physician and dentist fees were the highest (38.7 percent in 1980). Health insurance consumed about 27 percent, medicine and drugs cost slightly more than 12 percent. Other medical cost such as eyeglasses, hearing aids, and ambulance services amounted to about 14 percent.

Figure 3



In terms of dollars the average cost of medical services paid by Minnesota taxpayers varied by type of services and over time (see Table 1). Physician and dentist fees ranked the highest while cost of medicine and drugs ranked the lowest. The difference between average cost in 1978 and 1980 may be interpreted as a measure of inflation assuming that taxpayers had purchased the same quality and quantity of services.

Table 1 indicates that income has an effect on the purchasing power of medical services. Hence the reported deductions for health care expenditures varied by level of income. Insurance cost was the lowest for those whose gross income ranged between $20,000 and $29,999, then it showed a steady increase with the increase in income.

Table 1
Average Deductions for Health Care Expenditures
Minnesota Tax Filers, 1978-1980

Minnesota Gross Income Range	Insurance		Medicine and Drugs		Hospital		Physician and Dentist		Other Deductions	
	1978	1980	1978	1980	1978	1980	1978	1980	1978	1980
$-20,000 to 0	576	686	171	162	349	331	531	621	191	252
$1 to 9,999	401	438	142	158	306	388	324	371	316	469
$10,000 to 19,999	375	454	125	147	267	282	371	442	178	363
$20,000 to 29,999	322	359	137	146	175	230	408	412	182	221
$30,000 to 49,999	354	392	147	154	218	178	503	510	180	306
$50,000 to 99,999	473	482	181	183	385	311	806	734	242	669
$100,000 to 199,999	571	590	221	224	606	623	998	1260	916	583
$200,000 and Over	456	536	296	287	788	579	1964	1500	1246	1214
Average of All Filers	366	413	137	152	247	264	406	460	211	338

Cost of medicine and drugs varied also by income level. While the costs remained fairly constant between $-20,000 and $29,999 in 1980, it showed a gradual increase beyond the $30,000 income level where it reached about $300 for those whose reported gross income were $200,000 and over.

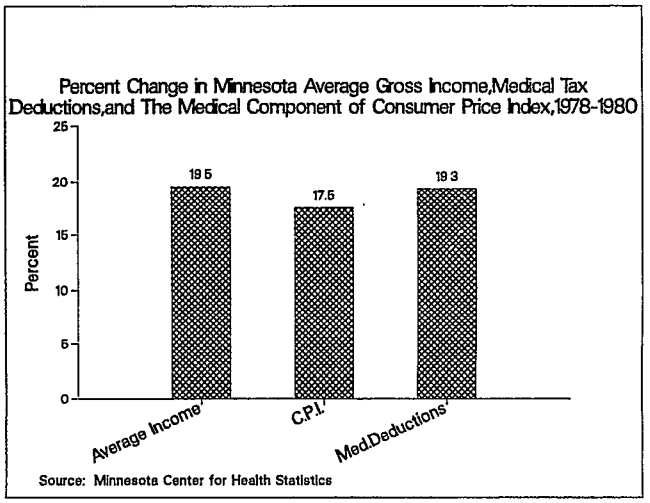
The cost associated with fees paid to physicians and dentists exhibited the same trend as that for medicine and drugs. However, the average cost of fees paid to health professionals in 1980 was approximately $420 for those whose reported gross income ranged from $20,000 to $29,999; and $1,500 for those with income over $200,000. The smallest deductions for physician and dental services were recorded by those with gross income between $1 and $9,999.

Costs deducted for hospital services and miscellaneous medical expenses show basically the same trend as other types of medical expenditures. In general, those whose reported gross income was over $100,000 in 1980 reported the largest deductions for medicine and drugs, hospital expenditures, payment for physician and dental services and expenditures for other medical services. Those who reported negative gross income in 1980 recorded the largest deductible for insurance expense.

Percent Change in Average Gross Income, Medical Tax Deduction and C.P.I.

The reliability of medical tax deductions data and hence its usefulness in assessing the nature of health care expenditures can be checked by comparing the increase in medical deductions with the changes in the medical component of the Consumer Price Index (CPI) over time. The average medical deductions have increased from $912 in 1978 to $1,088 in 1980, resulting in a percent change of 19.3 over the two year period (Figure 4). The medical care CPI for Minneapolis and St. Paul increased by 17.5 percent in the same period of time. It is interesting to note that Minnesota average gross income showed the same magnitude of increase as that for medical deductions. The average Minnesota taxpayer's gross income changed from $10,500 in 1978 to $12,550 in 1980. The percent change was 19.5.

Figure 4



Summary

The descriptive study reported here provided some evidence for the usefulness of secondary data such as tax filers data in assessing the

nature of health care expenditures. Despite the increase in expenditures between 1978 and 1980, the distribution of expenditures by type varied at a minimum. This finding, in addition to the similarity between the increase in the average medical tax deduction and the medical component of the Consumer Price Index, may add confidence in using tax data to study the nature of health care cost.

Unlike data reported elsewhere in Minnesota, health insurance is factored into this analysis for the first time and it accounted for 27 percent of total medical tax deduction. Costs for other health services were 12 percent for medicine and drugs, 40 percent for physicians and other health professional fees, 8 percent for hospitals, and 12 percent for other miscellaneous services. Medical tax deduction excluding insurance accounted for approximately 62 percent of out-of-pocket health services cost in Minnesota.

INITIAL FINDINGS FROM THE STUDY OF WISCONSIN MEDICAID UTILIZATION AND EXPENDITURE PATTERNS

Pam Schnagl, M.S., Ira Kaufman, M.S., and Carol Weidel
Wisconsin Bureau of Health Statistics

BACKGROUND

The Federal Social Security Act was amended to create Title XIX in 1965. The Wisconsin Medical Assistance Program was authorized by state statutes effective July 1, 1966. The program has been jointly financed by state and federal funds. The federal share of the costs has increased since the enactment. Currently, the split is approximately 58 percent federal dollars and 42 percent state general purpose revenue.

Governing state administration of the program is a set of federal guidelines outlining eligibility requirements and benefit levels. The federal government requires that states provide services to residents who are categorically needy. Individual states may set different criteria (demographic, income, asset) in determining their categorically needy population. Wisconsin, along with over 30 other states, also extends its coverage to their medically needy residents. The criteria for determining eligibility as medically needy also varies by state. The federal regulations allow states the latitude to structure their own programs by stating only the minimum required services and a long list of optional services. Although there are guidelines, their lack of specificity creates a dynamic program within and between states.

Because of its relatively liberal eligibility requirements and benefit package, Wisconsin ranked tenth in total Medicaid payments and eleventh with respect to recipient population in the fiscal year, 1979 (The Medicare and Medicaid Data Book, 1981). The general purpose revenue (GPR) budgeted for the Medical Assistance Program in 1980-81 represented approximately 13 percent of the entire state budget. Currently, Medicaid ranks as the fourth largest state program.

Since its creation, Medicaid has grown rapidly. The increasingly high costs of the program have led to various strategies to control expenditures. Policy makers have been interested in data analysis that would provide input into decisions aimed at controlling these costs.

In the past, the available data has primarily been comprised of two, separate fixed reporting subsystems that are part of the Wisconsin Medicaid Management Information System (MMIS). The Management and Administrative Reporting Subsystem provides aggregate information on providers, recipients, operations and finances for use in directing the program, and the Surveillance and Utilization Review Subsystem provides an exception reporting capability designed for postpayment utilization review of providers and recipients. Neither reporting subsystem has the capability of fulfilling the need of Medicaid policy makers for detailed analysis of individual definable segments of the Medical Assistance population using date-of-service information. The Wisconsin Bureau of Health Statistics developed a data system to respond to this need. Unlike the other two MMIS reporting subsystems, this system answers questions requiring detailed study and allows the user to design analyses.

This data base is an eight percent rotating sample of eligibles and their claims and services and is updated semiannually. Two percent of the sample drawn is replaced so that the same individuals are not continually studied. The significant number of individuals carried into the next year's sample allows for studies over time. The data base contains demographic, eligibility and service information. The sample offers policy makers the ability to base decisions on detailed information of utilization and costs corresponding to when services were actually received. There is flexibility in using such data, so that individual analyses can be requested without encumbering the huge expenses.

DATA

Data Sources

The Medicaid data presented here were drawn from the eight percent extract of the eligibility and paid claims tapes for 1980. Because the sample was rotated, the analysis was actually conducted on a six percent sample. The eligibility files provided recipient number, sex, birth and death dates, medical status and number of eligible days. The claims files provided information on recipient number, enrollee characteristics, services received, days billed for long-term care, and payment amount for each claim. The link between individual enrollment and paid claims files was established using the unique recipient identifiers on each file. Appendix 1 describes the variables used in this study.

Data Limitations

There are limitations inherent in the use of such a sample. The data is not as timely as would have been desired. Initially, in order to get the data fully adjusted, the sample extract could not be obtained earlier than one and a half years following the date of service. This problem was due to administrative rules allowing Medicaid

providers this period of time to submit their billings and adjustments. Therefore, at the start of this study only 1980 data was available.

There is a certain amount of error associated with sample estimates. By comparing aggregated estimated cost figures with data maintained by the Bureau of Health Care Financing, it was evident that this error is minimal. Estimates averaged within three percent of comparable actual expenditures.

Other problems were encountered in preparing the data for analysis. Since the extract was obtained from a billing system, certain items appear not to have been thoroughly edited. There were cases that had medical status assignments that did not coincide with their age group. For example, an individual could be 15 or younger and have a medical status classification of aged. These cases (491) were included in the analysis but corrected by giving them a medical status classification that more appropriately corresponded with their age category. These reclassifications affected less than 2 percent of the cases in the AFDC, aged and institutionalized aged categories.

In developing an analysis that associated costs and payments with certain enrollee characteristics, it was assumed that these characteristics would remain fairly constant over the course of one year. This assumption was tested. It was found that less than 3 percent of the eligibles changed medical status.

Because the data is from a sample, there may be instances when the data is broken down to a level where there are too few cases to make reliable estimates. As the level of aggregation becomes finer the estimates will be less stable.

In constructing the person-based records, it was discovered that there were 273 cases of the 33,906 cases where there were paid claims data but no eligibility information. Demographic information for these cases was taken from the paid claims file and the cases were assigned a person year equivalent of one.

FINDINGS

Determination of Population at Risk

Policy analysis of the Medicaid program tends to center around the extent to which specific subpopulations need services. Studies have shown that sex, age and medical status do affect use (Blaugh et al, 1982; Cromwell et al, 1982).

Similar information on relevant populations has served as a basis for setting up private insurance plans. The findings of this study will illustrate the potential of using MMIS data to develop utilization and cost schedules for sex-age-medical status groups that can be used in developing an insurance (reimbursement) plan.

For Medicaid policy makers to address these analytical concerns in setting up plans for their state programs, it is necessary to know the likelihood of a person to use a specific service and the associated costs. Due to changes in eligibility status, enrollees enter and leave the program continuously. Rate of turnover has implications in constructing the use and payment measures that permit this policy analysis.

An estimate of the Medicaid population at risk of using services can be constructed by counting all persons enrolled during the year and also by calculating a person year equivalent (PYE) which is based on the actual number of days the person is enrolled. A comparison of the number of eligibles calculated each way indicates the importance of using the PYE count in assessing Wisconsin's utilization and expenditure patterns.

The total number ever enrolled in 1980 was 1.4 times larger than the PYE count (see Table 1). If this difference is constant for all subpopulations of the Medicaid population, the units of service and cost per person will be uniformly underestimated by using the ever-enrolled count and can be easily adjusted for. However, this is not the case. Although the ever-enrolled count was always larger than the PYE count, the magnitude of this difference varied by sex, age and medical status. For the AFDC males aged 45-64, the ever-enrolled count was 90 percent higher than the PYE count. For the institutionalized disabled-AFDC males, however, the ever-enrolled count was only two percent higher. This shows a large range of turnover rates among Medicaid enrollees.

Examination of the relative proportions in each category of the PYE count and the ever-enrolled count also showed the differences in turnover rates by sex, age and medical status. Hence because of the high differential turnover rates of the eligible population in Wisconsin, PYE counts were used to define the enrolled population.

Composition of the Enrolled Population

Assumptions can be made regarding utilization and costs based on the composition of the population. In developing a plan, it is important to know the composition of a population and how the composition relates to the need for services.

In 1980, the size of the Wisconsin enrolled population was 408,246 (see Table 2). Females comprised approximately 61.4 percent of the Medicaid population. Roughly 38 percent were 15 years of age or younger. As in other state programs, the AFDC were the largest single

eligibility group, constituting 71.6 percent of all enrollees.

Males tended to be younger than females. In every age group except under 15, females predominated males by close to two times the male count. In the 75 + age category, the female figure was almost three times the male count.

Regardless of medical status, females comprised the largest faction of those eligible. The differences in numbers varied by medical status group. For example, the female count was over two times the male count in the institutionalized aged category, but only 1.25 times greater in the disabled category. This particular difference may be due to differential mortality.

By comparing percentages of medical status groups comprising particular age groups, it was found that AFDC eligibles constituted 97.1 percent of those under the age of 15 and 87.2 percent of those between the ages of 15 and 44. Approximately 17.5 percent of the population were over the age of 65, while 15.8 percent of this population were classified as aged, regardless of institutionalization. Although there is an association between age and medical status, it was hypothesized they would have differing effects on utilization and expenditures.

Examination of the combination of sex, age and medical status showed that regardless of age and medical status females still outnumbered males. There were only four sex, age, medical status combinations where the male count was greater than the female count. The association between age and medical status did not disappear. Hence, the enrolled population is found to be primarily female, AFDC, and concentrated in the younger age groups.

Utilization

Comparisons across services could not be made because of the problem in measuring units of differing procedures within services. This analysis restricts itself to examining the variability in utilization within a particular service by characteristics of the population.

Although utilization should vary by the compositional characteristics, use should not be proportional to the percentage the group comprises of the total population. For example, young, AFDC females are a majority of the population, but their use of Medicaid services would not be average. Therefore, examination of use patterns are an important component in deciding insurance plans.

Analyses of the rates for each service by enrollee characteristics indicated certain patterns. For all services, use was greater for females than males (see Table 2). The largest difference between the sexes occurred in the use of drugs (1.63 times higher for females); then physician use; and thirdly, hospital admissions.

The rates varied by age. There were similarities in the relationship between age and utilization by type of care. Use of drugs and long-term care increased with age. Rates of physician and hospital outpatient utilization and rates of hospital admission exhibited similar age patterns - an increase in use until age 65 and then a decrease in the use of services. The size of this increase or decrease varied by service. This age pattern may have been due to the fact that Medicare covers some of these services. Although those in younger age categories comprised the largest percentages of the enrolled population, the rates in the older age groups tended to be higher than in the younger age groups.

The rates varied by medical status. The relationship between medical status and utilization was not the same for each service. The institutionalized and the disabled used more ambulatory services (physician, hospital outpatient) and drugs than the other groups. The AFDC utilized more dental services than any other groups and the disabled and aged (noninstitutionalized) had the highest hospitalization rates. Note that though those classified AFDC were the largest group within the total population, they used less of all services except dental. The institutionalized made up approximately 10 percent of the population, but ranked either one or two in the use of most services.

After controlling for sex, the relationship between age and the use of a particular service remained for most services. With the introduction of age, females no longer had consistently higher rates than males. Only in the utilization of drugs did females have consistently higher rates than males regardless of age. Males 75+ had higher use for all services excepting drugs and other. Males, 65-74, had higher rates of hospital and long-term care admissions as well as utilization of "other" services. The reason female utilization is greater in certain age groups for some services may be due to childbearing.

Although the amount of the difference in utilization between medical status groups changes, the relationship between medical status and utilization remained apparent regardless of sex. The institutionalized continued to rank one or two in use of the services in which they had similar rankings before the introduction of the sex variable. The AFDC ranked fourth or fifth in the utilization of most services. Females still tended to have higher rates than males, although the difference between the rates varied.

The differential in rates by age is still evident after controlling for medical status. For most medical status groups, drug

utilization increased with age. For almost every service, the relationship between age and utilization was observed among the AFDC. The age patterns were also evident for the aged and institutionalized aged. There was more variability in utilization by age in the disabled group. Regardless of age, the institutionalized tended to have higher drug and physician utilization than the other medical status groupings. The disabled still tended to have higher rates of hospital admissions. The AFDC continued to use the least amount of most services. Therefore, the effects of medical status cannot be completely accounted for by age and vice versa.

Examination of the combination of all three variables showed that although some patterns remained similar to those exhibited when examining the effect of the variables jointly or individually, additional information was acquired. For example, the disabled males had higher rates of dental utilization than females. However, when examining the rates by age, only in one age group did males display higher use rates. The AFDC population still had the lowest average payments for most services regardless of age and sex. Although dental payments still tended to be higher, the under 15 disabled females had higher payments than the under 15 AFDC females. This was not true for males. The additional information serves to underscore the importance of examining in as much detail as possible relationships between population characteristics and use in developing insurance plans.

Payments

The preceding analysis clearly demonstrated that both the amount and type of service a Medicaid enrollee will use in a year varies depending on his/her sex, age and medical status. If the type of procedures provided in each service category are the same for all sub-groups of enrollees (equal intensity), the cost differential will be proportional to subgroup service utilization. Therefore estimating the Medicaid program liability or setting the rates associated with the insurance coverage would be straight forward. This analysis will proceed to demonstrate that intensity differences do exist. Although use indicates different needs for covered services, it tends to show only the direction not the magnitude of the costs of the services for the average enrollee in the subgroup.

Average payments per enrollee (Table 3) controls for differences in size of the population and changes over time, allowing for meaningful comparisons and assessment of the effects of the population characteristics on costs. Relationships as well as the magnitude of the differences for the types of services would be obscured by merely noting the average for all services. So the discussion of the findings will concentrate on average payments associated with the types of services rather than merely examining the average payment across service types.

The average payment for females was larger than males in all service categories. The largest difference occurred in drug utilization followed by physician utilization. This relationship was found also between sex and the utilization measures. Although average payment per enrollee was higher for females, the percent difference did not correspond to the differences found in utilization.

Changes in the average payment by age mirrored the changes in utilization rates by age for each service except for dental services. For example, the average payment increased with age for drugs just as did the utilization rates. The magnitude of the differences between the younger age groupings was greater than the differences found in drug utilization.

Average payments also varied by medical status in a manner similar to the utilization rates. The disabled had the highest average payments for hospital inpatient services. However, the AFDC had higher average payments for physician services, hospital inpatient and outpatient services than would be expected. The institutionalized aged had lower average payments than would be expected for physician services given the ranking of utilization levels. The average dental payment for the AFDC was higher than any other medical status group which would be expected given their higher utilization rate. The relationships between medical status and average payment again appeared similar to that exhibited between medical status and utilization. The magnitude of the difference between medical status groups was not similar. For example, the institutionalized aged utilized 2.24 times more drugs than the aged (noninstitutionalized) but their average payment was 2.82 times that of those aged that were not institutionalized.

When controlling for sex, the relationship between age and average payment was still present. The age patterns were similar to those found in the utilization and admission rates. Females no longer had consistently higher average payments. The patterns evident in average payments by medical status varied by sex for some of the service categories. For drugs, dental services, long-term care and other, sex made no difference in the ranking of average payments. In all other categories, however, sex did affect the ranking. In most instances, the medical status group with the highest average payments for a particular service remained the highest regardless of sex. Males did have higher average payments than females by medical status grouping in each service except drugs. The medical status grouping that this occurred in varied by service type.

Medical status interacted with age changing some of the patterns evident when examining the individual effects. The age effects on average service payment varied by medical status group. The age patterns were evident for the AFDC and all aged, regardless of institutionalization. For most services, there was increased variability in the age pattern for the disabled. The data also showed the effects of medical status on average payment given age for those age groups in two or more medical status groups. Except in the age category 65-74, the disabled still had higher average payments for hospital inpatient care. The AFDC had dramatically higher average payments than the disabled for dental services. By age, however, the difference between those under the age 15 was relatively small.

In some instances, sex, age and medical status interacted to add information regarding the relationship and the relative size of the difference in average payments between males and females, age and medical status groups. When looking at all three variables, some of the relationships paralleled those found in utilization such as the sex patterns for the AFDC. But others, like the age pattern for the AFDC, did not.

Though there were similarities, the changes in relationships of the population characteristics to average payments and the utilization measures is illustrative of the major impact that intensity has on the cost of services that cannot be accounted for by use (need). Changes in magnitude of the differences also indicate the effect of an intensity variable but to a lesser degree.

Discussion

Although there is no test for the significance of these effects or assessment of relative impacts, this study makes it evident that utilization and expenditures vary by the composition of the population. No attempt has been made to decompose the effects of all factors on payments. It is clear though that utilization is not the only factor determining expenditures. An intensity factor may also affect costs.

More specifically, this study shows the following:

1) In Wisconsin, when evaluating the characteristics of the enrolled population, calculating rates of use, and average payment figures, person year equivalent counts should be used rather than the count of the ever-enrolled.

2) The percentage a subpopulation is of the total Medicaid population is not indicative of what their utilization and expenditures might be.

3) The individual effects of sex, age, and medical status on utilization and expenditures are evidenced. When these variables are combined jointly, variability is introduced to the patterns established independently. When all three variables are introduced, however, the patterns established by the variables individually are even more disrupted.

4) Even though age and medical status are most likely associated, there is variability in utilization and expenditures within medical status associated with age.

5) Relationships between a compositional characteristic and a utilization or expenditure measure may appear to be the same when taken in combination with another characteristic. However, the magnitude of the variability does change.

6) The composition of the population has similar affects on both utilization measures and expenditure measures.

These results have implications in controlling Medicaid costs. The differential utilization and expenditure patterns by the enrollee characteristics must be taken into consideration when determining adequate reimbursement levels for enrollees in HMO and primary provider programs. As has been demonstrated through this study, reimbursement that does not take into account the characteristics of a particular population in a plan may be setting rates that are over-reimbursing or under-reimbursing the provider. For example, the AFDC population use more dental services than any other medical status grouping and has higher average payments for such services. From this analysis, it is also known that within the category AFDC, the 15-44 year olds have higher utilization and costs associated with their dental care. Furthermore, within this grouping, females have higher utilization and costs than males. A detailed reimbursement mechanism could take into account these nuances in use and payments. In essence, the Medicaid program could be set up as an insurance program. The impact on reimbursement of serving certain populations would be known.

These analyses would be useful in making budgetary projections. If the use and costs associated with different groupings are known, the effect of changes in enrollment patterns will be more readily understood. By calculating an inflation factor for each service the program will be better able to project costs in future years.

This study demonstrates that MMIS data can be used to support analysis that can serve as input into policies and programs aimed at controlling costs while attempting to provide equity in care. Though the specific findings in this report are not applicable elsewhere than Wisconsin, certain methodologies, possibilities for analysis and their policy implications will be extremely useful to other research and planning analysts attempting to manage their state Medicaid programs.

Appendix 1

DESCRIPTION OF VARIABLE MEASURES

A. Compositional Variables

1. Sex: male, female.

2. Age: under 15, 15-44, 45-64, 65-74, 75+. Age was calculated using date of birth on the eligibility file.

3. Medical Status: AFDC, Aged, Disabled, Institutionalized Aged, Institutionalized Disabled-AFDC. Institutional status was determined by using the number of days billed for long-term care on the paid claim file. If the number of days billed was over 30 days, then the case was considered institutionalized.

B. Enrollment

1. Ever enrolled: a count of all persons on the eligibility file.

2. Person year equivalent (PYE): a sum of fractions of years into yearly subtotals as follows:

$$PYE_y = \frac{\sum_{i=1}^{n} d_i}{366}$$

where PYE_y is the person year equivalent enrollment for year y
n is the number of enrollees
d_i is the number of days that person i was enrolled during year y.

C. Service Types

All claims were adjusted prior to the formation of these tapes. Services were formed using a flow chart approach. Cases dropped out of the number of available for selection as soon as they were chosen as defining a particular service. Services were determined in the following order:

1. Drugs: Coded by presence of a drug claim type. Primary provider type not taken into consideration.

2. Long-Term Care: Included those with nursing home claim type regardless of primary provider type.

3. Hospital Inpatient: All claims denoted as institutional inpatient regardless of primary provider type.

4. Hospital Outpatient: Included those with claim type of institutional outpatient and those, regardless of claim type, with a primary provider type of general and psychiatric hospital.

5. Dental: Those with a professional claim type and a primary provider type of professional group (dental/oral surgery) or dentist.

6. Physician: Those claims with primary provider types of physician (M.D.), osteopath (D.O.), professional group (clinic), osteopath group and physician group.

7. Other: All the claim/provider type combinations that were not included in the specified service types (for example, EPSDT, chiropractic services, optometric services, social service/personal care, home health agencies, specialized clinics, etc.).

D. Utilization Variables

1. Hospital inpatient admissions per 1,000 PYE:

sum of the number of inpatient claims for a group divided by the PYE count for that group. This number is then divided by 1,000.

2. Long-term care admissions per 1,000 PYE:

sum of the number of people with an accomodation code present in a group divided by the PYE count for that group. This number is then divided by 1,000.

3. Utilization per 1,000 PYE:

(for hospital outpatient, dental, physician/clinic, drugs and other)

sum of the number of claims for a service in a group divided by the PYE count for that group. This number is then divided by 1,000.

E. Payment Variables

1. Total payments:

sum of the total paid for service(s) in a group.

2. Average payment per enrollee:

sum of the total paid for a service in a group divided by the PYE count for that group.

Acknowledgements

This study was done at the request of the Wisconsin Bureau of Health Care Financing. The authors wish to thank William Jollie, Consultant with John O'Keefe and Associates, who was instrumental in conceiving the research question and the applicability of the results. In addition Ray Nashold and Hilde Neujahr, director and deputy director, allocated the resources within our Bureau so that the research could be completed within its timeframe. Russell Kirby, Henry Krebs and Douglas Murray, research analysts with the Bureau of Health Statistics, gave helpful comments and suggestions. The authors also appreciate the secretarial support of Virginia Erickson, also of the Bureau of Health Statistics, and the staff of Internal Services, Division of Health.

Due to space considerations, the tables could not be presented. The tables are available upon request.

Address: Pam Schnagl
Wisconsin Bureau of Health Statistics
P. O. Box 309
1 West Wilson Street
Madison, WI 53701

References

1. Baugh, David K. et al, "Cost Containment: What Goes Wrong," Paper presented at the 1982 APHA meetings in Montreal.

2. Cromwell, Jerry L. et al, "Trends in Tennessee Medicaid Acute Care: Use and Expenditures 1974-1978," Health Care Financing Review 3(4), June 1982.

3. Harder, Diane, "The Wisconsin Medical Assistance Program: Eligibility Requirements, Program Benefits, Funding and Expenditure Levels," Wisconsin Legislative Fiscal Bureau, unpublished manuscript, May 1982.

4. Kaufman, Ira, "Inclusion in the Fiscal Intermediary Contract of a Policy Evaluation Review and Utilization (PERU) File," Internal report, January, 1982.

5. The Medicare and Medicaid Data Book, 1981, ORD, HCFA.

6. Schnagl, Pam, "Determination of Physician Services," Internal report, August, 1982.

COST SAVINGS OF HOSPICE HOME CARE TO THE MEDICARE HOSPITAL INSURANCE PROGRAM

Charles H. Brooks and Kathleen Smyth-Staruch,
Case Western Reserve University

INTRODUCTION

National Medicare expenditures have been increasing very rapidly. In 1981 $43.5 billion was spent on personal health care services provided to beneficiaries. This amount was $7.8 billion higher than the expenditures in 1980 and almost $28.0 billion higher than the expenditures in 1975 [Refs 1,2]. Thus, the average annual rate of increase from 1975 to 1981 was 18.7%.

It was in partial recognition of this fact that Congress passed the Tax Equity and Fiscal Responsibility Act (P.L. 97-248) of 1982. Two major provisions of this legislation are directed at containing the upward cost spiral of the Medicare Program. The first provision will change the method of reimbursing participating hospitals from a cost-related, per diem basis to a flat-rate system of payment according to a hospital's case mix (i.e., Diagnosis Related Groups, or DRGs). A second provision will expand the coverage of the Medicare Hospital Insurance Program (Part A) to include hospice care for terminally-ill beneficiaries. It is this latter provision that interests us here.

The rationale for expecting hospice cost savings is derived from the concept of palliative care. According to Craven and Wald [Ref 3]:

> "Hospice care is most appropriate for people who have fatal diagnoses, for whom curative therapies have failed and hence active treatment is no longer warranted. The challenge in dealing with such patients thus shifts from curing to caring -- to maintaining patients as symptom free as possible, allowing them to maintain control over their remaining lives, and helping them and their families deal with the impending death with dignity and meaning."

Hospice services, therefore, are meant: (1) to provide medical care for the continuing control of symptoms such as pain, nausea, diarrhea, constipation, anorexia, and the like; (2) to concentrate on bedside nursing to provide comfort and close attention to easing physical distress and providing emotional support; (3) to focus on the family unit by teaching the dying patient and family members to cope with the situation and including them in the caring process; and (4) to provide spiritual care through ecumenical services, group discussions, and through an atmosphere of love and concern [Ref 4].

A critical dimension of hospice care is that palliative care services can be provided to dying patients both in and out of a hospital setting. It recognizes that during the course of a terminal illness there can often be an interplay between institutional and non-institutional care, particularly during the final months of life. Thus one of the principal ways in which hospice care can be cost-effective is that it can change the "locus of care" away from the use of relatively expensive hospital acute care beds toward the use of the patient's own bed [Ref 5].

Whereas previous studies have estimated the potential cost savings of hospice care to range from hundreds to thousands of dollars per patient [Ref 6], no study has clearly demonstrated that these savings stem from the substitution of less costly home care services for more expensive hospital inpatient care. The present investigation, which is a population-based, retrospective analysis of insurance claims data, was designed specifically to test this hypothesis.

METHODS

This is a study of all residents of Cuyahoga County, Ohio (metropolitan Cleveland) who were 65 years old or older when they died of cancer between April 1 and December 31, 1981, and for whom Medicare-Part A paid out benefits for services used during the last 24 weeks of life. The cancer deaths to Cuyahoga County residents were identified by the State of Ohio Department of Health which prepared a computer tape containing the following information: (1) the deceased person's name and social security number, if known; (2) age and birthdate; (3) sex and race; (4) date of death; and (5) cause of death, coded according to the 9th Revision of the International Classification of Diseases [Ref 7]. This Cancer Death File was then matched, using the deceased person's social security number, name, age and sex, with a Medicare-Part A disbursement file.

For each decedent who was matched by this procedure, an analytic data file was constructed in which Medicare hospital insurance payments for hospital inpatient care, hospital-based physician services, nursing home care, and home care visits were tabulated separately for the last 24 weeks of life, the last 12 weeks of life, the last 8 weeks of life, the last 4 weeks of life, and the last 2 weeks of life. The data were arranged in this format of decreasing time intervals so that the differences in the use of medical care and the amount of Medicare liability between the hospice and non-hospice cancer decedents could be studied as death approached.

Four hospice programs in Cuyahoga County, Ohio participated in the study. Three programs were administered by community-based home health care agencies and the fourth was a hospital-based home care program. None of the hospices managed inpatient beds. Each hospice agency provided the names, admission and discharge dates, and when possible, social security numbers of Cuyahoga County residents served between April and December 1981. This information was critical because it facilitated the comparative analysis of the Medicare payments for the hospice and non-hospice decedents and made possible an analysis of the payments for hospice care which could be compared to those payments occurring before the Medicare beneficiary entered a hospice care program.

The results of the matching process are shown in Table 1. A deceased person was identified as

a "hospice cancer decedent" if hospice care was used at anytime during the last 24 weeks of life. Of the 1,693 subjects in the baseline population, 1,148 were matched with the Medicare-Part A disbursement file, yielding an overall match rate of 68%. One hundred and eight of the matched decedents had received hospice care at some time during the last six months of life, and 1,040 had not. The proportions of the hospice and non-hospice cancer decedents in the study and baseline populations were identical: 9% and 91%, respectively.

*** Table 1 ***

Because this cost-effectiveness analysis was a population-based, retrospective study with a relatively large number of cases, results of tests for statistical significance were not reported. Rather, the focus of the analysis was on detecting meaningful differences in Medicare payments between the hospice and non-hospice cancer decedents. A relative difference of at least 20% and greater than $100 per cancer decedent was employed to indicate a substantial cost difference between the comparison groups.

THE MEDICARE HOME CARE BENEFIT

The Medicare Hospital Insurance Program did not officially recognize hospice care as a reimbursable service during the study period. It did not recognize, for example, a physician's prognosis of six months or less of life as a condition of eligibility for receiving home care benefits. Nonetheless, it did reimburse home health care agencies for limited amounts of services which partially reflected the hospice concept of palliative care, specifically: part-time skilled nursing care; physical, speech and occupational therapy; part-time services of home health aides; medical social services; and medical supplies and equipment provided by the home health agency.

Medicare-Part A paid the reasonable cost of up to 100 home care visits in a twelve-month period provided that: (1) the beneficiary had been hospitalized for at least 3 consecutive days; (2) the home health care was for a condition which had been treated in a hospital or skilled nursing facility; (3) the care needed was part-time skilled nursing care, physical therapy or speech therapy; (4) the beneficiary was confined to home; (5) a doctor determined that home health care was needed and established a plan of care within 14 days after discharge from a hospital or skilled nursing facility; and (6) the home health agency participated in the Medicare insurance program.

Despite the non-recognition of hospice care as a reimbursable service and the restrictions on home care services outlined above, what was the cost savings of hospice home care for the Medicare-Part A Program? Were these savings the result of a substitution effect?

FINDINGS

Table 2 shows the length of service on hospice care for the 108 cancer decedents in the study population who chose this mode of palliative care. Half of the decedents entered a hospice program within the last four weeks of life, 28% were hospice patients between four and twelve weeks, while just 22% of these decedents received hospice home care services longer than twelve weeks. The mean length of service on hospice care was 48.1 days, or approximately 7 weeks.

*** Table 2 ***

To isolate the relative cost savings of hospice home care, Medicare-Part A payments for the hospice and non-hospice cancer decedents were compared for time intervals before death in which only those decedents who were on hospice care for the entire duration of a given time interval were included for analysis. By this strategy, any use of medical care and associated Medicare payment that was incurred before hospice care was chosen did not affect the group comparison and, as a result, permitted a more accurate assessment of the cost-effectiveness of hospice home care.

Preliminary analysis showed that there was very little difference between the comparison groups regarding the use of hospital-based physician services and nursing home care. Instead, the major group differences applied solely to the use of hospital days and home care visits. That is, the hospice study subjects, while on hospice care, used substantially fewer hospital inpatient days than the non-hospice cancer decedents (Table 3). The differences in hospital use ranged from 4.3 fewer days during the last two weeks of life to 11.9 fewer days during the last twelve weeks of life. For example, the non-hospice subjects spent virtually half of the last two weeks of life (6.9 days) in a hospital bed, whereas the hospice beneficiaries were hospitalized only about one-fifth of this time (2.6 days). On the other hand, the hospice study subjects used appreciably more home care visits. The group differences ranged from an additional 5.8 visits during the last two weeks of life to an additional 23.2 visits during the last twelve weeks of life.

*** Table 3 ***

The apparent substitution of home care visits for hospital days was also evidenced by the shift in utilization which occurred for the cancer decedents who decided to receive hospice home care within the last 24 weeks of life. The data are presented in Table 4. Before beginning hospice care, these persons were hospitalized one day for every 5.9 days of conventional cancer care and received one home care visit for every 32.4 days of care. While on hospice care, this use shifted perceptibly: to one hospital day for every 8.3 days of care and one home care visit for every 3.7 days of care. Stated otherwise, the relative use of hospital care decreased 41%, whereas the use of home care services increased more than eightfold.

*** Table 4 ***

The marked differences in medical care use heavily influenced the group differences in Med-

icare liability for hospital care and home care (Table 5). The average expenditures for hospital care were much lower for the hospice beneficiaries, the group differences ranged from $1,245 per decedent during the last two weeks of life to as much as $3,232 per decedent during the last twelve weeks of life. The home care expenditures for the hospice patients, however, were considerably higher. The average cost differences to Medicare-Part A ranged from an additional $218 per person during the final two weeks of life to an additional $987 per person during the last twelve weeks of life.

*** Table 5 ***

The substitution of home care visits for hospital days was also indicated by the change in how the Medicare-Part A dollar was spent before and during hospice care (Table 6). The average Medicare expenditure before hospice care was selected was $6,291 per cancer decedent. For every dollar spent, 95 cents went for hospital care, 3 cents went for hospital-based physician services, and 2 cents went for home care. Interestingly, this distribution of the dollar closely matched that of the non-hospice Medicare beneficiaries, suggesting that the cancer illness of the hospice decedents was being treated conventionally in a hospital setting before hospice home care was chosen. Once on hospice care, however, the portion of the Medicare dollar spent on hospital care decreased to 74 cents, whereas the portion spent on home care increased to 22 cents.

*** Table 6 ***

The overall cost savings of hospice care for the Medicare Hospital Insurance Program, including the expenditures for hospital-based physician services and nursing home care, is presented in Table 7. These data demonstrate unmistakably that the average Medicare payment for the hospice study subjects was much lower than that for the non-hospice subjects. The group differences, which ranged from $1,045 per cancer decedent during the last two weeks of life to $2,278 per decedent during the last twelve weeks of life, represented a relative cost savings of from 34% to 50%.

*** Table 7 ***

CONCLUSION

Although Medicare-Part A did not recognize hospice care per se as a reimbursable service, but rather paid hospices only for the allowable cost of a standard home care visit, this retrospective analysis of insurance payments has shown that hospice home care was cost-effective for the group of beneficiaries in Cuyahoga County, Ohio who died of cancer between April and December 1981. Further, the data clearly demonstrated that the lower cost of hospice care was due primarily to the substitution of less costly home care visits for more expensive hospital inpatient days.

REFERENCES

1. Gibson, R. and D. Waldo, National health expenditures, 1981. Health Care Financing Review 4:1-35, 1982.

2. Gibson, R., National health expenditures, 1979. Health Care Financing Review 2:1-36, 1980.

3. Craven, J. and F.S. Wald, Hospice care for dying patients. American Journal of Nursing 75:1816-1822, 1975.

4. Nelson, R.J., Hospice: An alternative solution to the problem of caring for the dying patient. Colloquy, pp. 22-23, 1974.

5. Greer, D.S., V. Mor, S. Sherwood, J. Morris, H. Birnbaum and M. Pozen, The National Hospice Study Design. Unpublished paper presented at the 109th Annual Meeting of the American Public Health Association, Los Angeles, CA, 1981.

6. Brooks, C.H., The potential cost savings of hospice care: A review of the literature. Health Matrix 1:(forthcoming), 1983.

7. International Classification of Diseases, Ninth Revision (ICD-9). Geneva, Switzerland: World Health Organization, 1977.

ACKNOWLEDGEMENTS

This study was funded by the Hospice Council for Northern Ohio, Blue Cross of Northeast Ohio, and the W.M. Keck Foundation Research Scholars Program of the Health Systems Management Center, Case Western Reserve University.

Table 1. Matching results for cancer decedents 65 & over according to hospice care status

	All Cancer Decedents	Hospice Cancer Decedents	Non-Hospice Cancer Decedents
Baseline Population	1,693 (100%)	159 (9%)	1,534 (91%)
Cuyahoga County Residents 65 & Over Who Died of Cancer Between April 1 and December 31, 1981			
Study Population	1,148 (100%)	108 (9%)	1,040 (91%)
Cancer Decedents 65 & Over Matched on Medicare-Part A Disbursement Files			
Match Rate	68%	68%	68%

Table 2. Length of service on hospice care among Medicare cancer decedents 65 & over

	Medicare Cancer Decedents		
	N	%	
Less than 2 Weeks	25	23	1 to 13 Days
2 to 4 Weeks	29	27	14 to 27 Days
4 to 8 Weeks	20	19	28 to 55 Days
8 to 12 Weeks	10	9	56 to 83 Days
12 to 24 Weeks	21	19	84 to 168 Days
More than 24 Weeks	3	3	More than 168 Days
Total	108	100	Total
6.9 Weeks =	Mean	=	48.1 Days
3.9 Weeks =	Median	=	27.5 Days

Table 3. Medicare hospital days and home care visits among hospice and non-hospice cancer decedents 65 & over (per cancer decedent)

Time Interval Before Death	Hospice Decedents	Non-Hospice Decedents	Difference
Hospital Days			
Last 2 Weeks	2.6 (83)	6.9 (1040)	- 4.3
Last 4 Weeks	5.6 (54)	11.6 (1040)	- 6.0
Last 8 Weeks	8.5 (34)	17.7 (1040)	- 9.2
Last 12 Weeks	9.9 (24)	21.8 (1040)	- 11.9
Home Care Visits			
Last 2 Weeks	6.2 (83)	0.4 (1040)	+ 5.8
Last 4 Weeks	11.2 (54)	0.7 (1040)	+ 10.5
Last 8 Weeks	18.4 (34)	1.0 (1040)	+ 17.4
Last 12 Weeks	24.4 (24)	1.2 (1040)	+ 23.2

Table 4. Hospital days and home care visits per cancer decedent according to the pre-hospice and hospice care status of Medicare cancer decedents 65 & over

	Pre-Hospice Care	Hospice Care
Days of Care	119.9	48.1
Hospital Days	20.4	5.8
Home Care Visits	3.7	12.9
Ratio of Days of Care to Hospital Days	5.9:1	8.3:1
Ratio of Days of Care to Home Care Visits	32.4:1	3.7:1

Table 5. Medicare-Part A expenditures for hospital care and home care among hospice and non-hospice cancer decedents 65 & over (per cancer decedent)

Time Interval Before Death	Hospice Decedents	Non-Hospice Decedents	Difference
Hospital Care			
Last 2 Weeks	$ 781 (83)	$2,026 (1040)	- $1,245
Last 4 Weeks	$1,496 (54)	$3,429 (1040)	- $1,933
Last 8 Weeks	$2,479 (34)	$5,200 (1040)	- $2,721
Last 12 Weeks	$3,160 (24)	$6,392 (1040)	- $3,232
Home Care			
Last 2 Weeks	$ 237 (83)	$ 19 (1040)	+ $ 218
Last 4 Weeks	$ 439 (54)	$ 29 (1040)	+ $ 410
Last 8 Weeks	$ 778 (34)	$ 41 (1040)	+ $ 737
Last 12 Weeks	$1,036 (24)	$ 49 (1040)	+ $ 987

Table 6. How the Medicare-Part A dollar was spent during the last 24 weeks of life according to hospice care status

	Hospice Cancer Decedents		Non-Hospice Cancer Decedents
	Pre-Hospice Care	Hospice Care	
Hospital Care	95¢	74¢	95¢
Hospital-Based Physician Services	3¢	1¢	2¢
Nursing Home Care	0¢	2¢	2¢
Home Care	2¢	22¢	1¢
Total	100¢	100¢	100¢
Average Medicare Expenditure	$6,291	$2,307	$9,013
N	108	108	1,040

Table 7. Medicare-Part A expenditures for all medical services among hospice and non-hospice cancer decedents 65 & over (per cancer decedent)

Time Interval Before Death	Hospice Decedents	Non-Hospice Decedents	Difference	Relative Savings
Last 2 Weeks	$1,053 (83)	$2,098 (1040)	- $1,045	50%
Last 4 Weeks	$1,972 (54)	$3,557 (1040)	- $1,585	45%
Last 8 Weeks	$3,328 (34)	$5,414 (1040)	- $2,086	39%
Last 12 Weeks	$4,403 (24)	$6,681 (1040)	- $2,278	34%

Data Collection and Processing Methods in the National Medical Care Utilization and Expenditure Survey (NMCUES)

Session N

INNOVATIVE ASPECTS OF THE NATIONAL MEDICAL CARE UTILIZATION AND EXPENDITURE SURVEYS

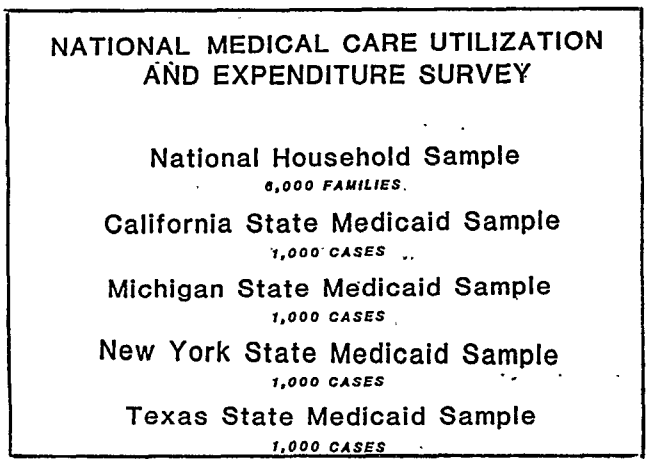
Gordon Scott Bonham, University of Louisville

The National Medical Care Utilization and Expenditure Survey (NMCUES) was a major data collection effort by the Department of Health and Human Services. It was principally built upon earlier surveys and had few untested procedures. As such, however, the survey could develop more fully a number of techniques, and employ them in a coordinated data collection effort. The innovations in the NMCUES are not so much in the parts, but in the way the parts have been integrated to produce a large dataset which can be used to address a great variety of analytical needs. This paper describes many of the NMCUES procedures and how they contribute to the purposes of the survey and the usefulness of its data [1].

BACKGROUND

The NMCUES was jointly sponsored by the Health Care Financing Administration and the National Center for Health Statistics. The data were collected by the Research Triangle Institute, the National Opinion Research Center, and Syste-Metrics, Inc. under contract to the Department of Health and Human Services. The interest of the Health Care Financing Administration centered on the Medicare and Medicaid populations, and their health and health care. Data were needed to evaluate potential effects of changes in the programs. The interest of the National Center for Health Statistics focused upon the total population. The data needed to produce national estimates of health, health care, and health care expenditures.

The interests of the two sponsoring agencies were reflected in the two types of samples. The first was a national area probability sample of about 6,000 households representing the civilian noninstitutional population of the United States. The second were stratified samples of about 1,000 cases each from the Medicaid eligibility rolls in California, Michigan, New York, and Texas. These four states contained a large proportion of the National Medicaid population, and yet the large sample within each State allowed analysis of different State programs. (Since Medicare is a uniform program throughout the nation and sufficient numbers of elderly would be included in the National sample, a special sample was not needed.)



Although there were a number of samples, the data collection procedures were identical in each. This allows the data from one sample to complement data from another, with differences attributable to differences in the sampled populations. The collected data relate to five substantive areas:

- Health.
- Health care.
- Cost of health care.
- Payment for health care cost.
- Health insurance.

Much analysis of the NMCUES data will focus upon people and their health-related experiences. However, social and program policy often involves the family or an aggregate of related people defined as a "case." The NMCUES data were collected in such a way as to permit analysis using the family or the Medicaid case as the unit as well as the person. In addition, analysis can be made of charges and other characteristics of health care events or of health conditions.

The NMCUES was a panel study, with the same households visited 4-5 times at approximately 3-month intervals. The data can be viewed as either a cross-sectional aggregate or as longitudinal data over a 12-month period of time. It all depends on the nature or purpose of the analysis.

HOUSEHOLD SURVEY

Information through household interviews formed the basis of the NMCUES. A number of procedures were instituted to make this information as accurate as possible. These procedures ranged from identifying and labeling each unit of information to having the respondent confirm or update the information in the data files.

Identification and Linkage

About 1,400 data items were collected for each of about 36,000 people in the five rounds of interviewing. One of the most important procedures, therefore, was to establish and maintain unique identification and linkage for each analytic unit. In addition, procedures were needed to assure that data did not get lost or duplicated.

The reporting unit was the basic data collection unit, although it might not always be the basis analytic unit. A reporting unit was defined as all related people living together at the time of interview. A reporting unit record was initially developed for every selected

address in the National sample and every selected case in the State Medicaid samples. The record contained all the sampling information, including the Medicaid identification numbers in the State Medicaid samples. A reporting unit identification number was computer generated for each record. This number contained a check digit designed to catch any transcription or keying errors. The identification number and sampling information were computer printed onto the Control Card assigned to the interviewers. Printed bar codes and light pens were used for half of the Control Cards to further reduce errors.

The reporting unit identification remained constant throughout the panel survey, following all or some of the members. Since a reporting unit was defined as related people living in the same residence, any people who left the residence required a new reporting unit with its own identification. However, any new reporting unit could be linked to the original one. Students living away at school were considered separate reporting units but linked to their parental family for analytical purposes and in case they were living at home during subsequent interviews. Each member of the original case in the State Medicaid samples was followed if case members did not live together. The reporting unit linkage allows the case, as it appeared in the original Medicaid eligibility file, to be reconstructed for analysis.

The Control Card prepared for each sample unit contained six preprinted person identification numbers. Extra Control Cards contained additional preprinted person identification numbers to insure two people could not be assigned the same number. Person numbers had a check digit similar to the reporting unit numbers to guard against errors. Records of linkage between persons and reporting units were maintained, since people did not always stay in the same reporting unit throughout the fieldwork.

Health problems or conditions can exist over a period of time and be associated with many events of health care. Each condition for a person was therefore assigned a unique number by the interviewer when it was first mentioned. The condition number was recorded in the questionnaire whenever the condition was associated with health care, illness, or permanent limitations. A computer printed list of previously reported conditions was part of the information given the interviewer about the person before each follow-up interview.

The Summary of Responses, discussed later, required a special set of linkage numbers. The Summary was designed to allow information collected during one interview to be updated or corrected during a subsequent interview. Unique linkage was established between the data item in the original computer record subject to change and the printed Summary on which the change was recorded. This insured that when the change was later entered into the data base, it would supercede the appropriate original information.

Finally, to guard against loss of data or duplicate entries of data, each separate document used to collect data had a unique form number. Page numbers, interview round numbers and keying dates were also keyed as part of the data records. While identification numbers and linkages may not seem central to the purpose of a survey, they often determine how useful the data will be for analysis. Identification numbers and linkages received a great deal of attention in the NMCUES, and the result is an accurate, useful and flexible data base.

Time Frame

One of the first decisions made on NMCUES was the time period to be covered. A twelve month period was chosen for a number of reasons: 1) it is a common period of time, 2) it is long enough to measure many sequential events and be free from seasonality, and 3) it is short enough to have reasonable costs and permit analysis in a reasonable time after the beginning of data collection. The specific twelve months of calendar year 1980 were chosen for additional reasons: 1) a calendar year is commonly used for income tax and health insurance purposes, 2) a calendar year has the significant Christmas and New Years holidays around its beginning and end to help define the period of data collection, and 3) 1980 was the year of the Census which could provide data on population characteristics to calibrate and amplify survey data. Data collection during 1980 did require special permission, however, and a concerted effort was required by the contracting firms to get the survey in the field at the beginning of 1980.

Interviewing Frequency and Boundaires

The NMCUES was a panel survey with five rounds of interviews approximately three months apart. In general, the accuracy of respondent recall decreases as the length of time from the event increases. Two-week recall periods like those used in the National Health Interview Survey minimize recall problems, but would require 26 interviews to produce yearly histories for people. Too frequent interviewing also creates problems with panel attrition.

In a feasibility study in 1975-76, the effectiveness of a two month interval between interviews was about the same as a one month interval [2]. It was less expensive, however, and less burdensome on the respondents. In the 1977 National Medical Care Expenditure Survey, even a 2-month interval was found to be too short because of the time needed to get data keyed, summarized, and back to the field for the next interview.

A three month interval was chosen for the NMCUES as the best solution to minimize recall problems, respondent burden, and fieldwork problems. The interval between the first and second round of interviewing was slightly more than three months due to problems in data processing and creating the Summary of Responses. Subsequent intervals were slightly less than three months so the final round of interviewing occurred on schedule.

The time period for which data were collected in the first interview began with January 1 and ended with the date of interview. It was anticipated that a January 1 reference data would minimize the problem of telescoping events either into or out of the time period. For intermediate rounds, the reference period was from the previous interview up to the day of interview. During the last interview, data were collected from the previous one up through December 31. Respondents originally interviewed in April 1980 recieved their fourth and final interview in January 1981. Those originally interviewed in February or March 1980 received their fifth and final interview in February or March 1981. Illness, travel, or other reasons prevented some respondents from being interviewed at the desired three month intervals. The overriding objective, however, was to collect data for the full calendar year of 1980 with as close to three month intervals between interviews as possible.

Questionnaire

The basic set of data in NMCUES related to illness and medical care. This set was collected through an unchanging core questionnaire during each round of interviewing. The constant core questionnaire resulted in efficient interviewer training and computer programming. The unchanging core was an important factor in reducing the time required to create the Summary of Responses for the next interview from 13 to 7 weeks. The unchanging core also permitted the third and fourth rounds of interviewing to be conducted largely by telephone as respondents were familiar by then with the questionnaire and the desired data.

Data on characteristics of people that were unlikely to change during the year were collected through questionnaire supplements. The major supplements were in the first and last rounds. As these interviews were conducted face-to-face, show cards could be used to aid data collection. Data collected using the supplements did not become part of the Summary of Responses and were not updated or corrected by the respondents.

Summary of Responses

A Summary of Responses (see illustration below), was computer generated for each reporting unit from core information reported in previous interviews. The Summary was a solution to a major concern in collecting data on medical care expenses. Data should be collected on the occurrance of doctor or other medical care as soon as possible, since underreporting increases with time since the event. However, the cost of the care is often unknown until a bill or statement arrives, and this may be some time after the care. The Summary of Responses was used to allow charge and payment data for medical care to be entered into the database when they became known. The occurrance of the medical care itself, however, was collected and entered the data base as soon as possible.

The Summary aided data collection in additional ways. It helped define the reference date for the interview, and whether a particular medical visit or medicine had already been reported. It was a valuable tool to prevent events from being telescoped into the interview reference period. The Summary allowed respondents to make corrections to medical care or insurance coverage erroroneously

NATIONAL MEDICAL CARE UTILIZATION AND EXPENDITURE SURVEY
SUMMARY OF RESPONSES

HEALTH CARE SERVICES FOR JOHN SMITH FOR THE PERIOD 01/01/80 - 03/12/80

RU ID # 7654321
PID 1234567

DATE OF CARE	TYPE OF VISIT OR SERVICE	MEDICAL PERSON OR PLACE AND ADDRESS	SERVICES RECEIVED	--CHARGE INFORMATION-- SOURCE OF PAYMENT	AMOUNT
01/05/80	DENTAL VISIT	DR. SAMUEL JONES	FILLINGS (02)	FAMILY	$35.00
			FLOURIDE TREATMENT	----------	------
			----------	----------	------
			----------	----------	------
				TOTAL CHARGE	$35.00
02/18/80	HOSPITAL OUTPATIENT	EAR, NOSE, & THROAT CLINIC	DIAGNOSIS/TREATMENT	FAMILY	20%
			X-RAYS	BC/BS OF NC	80%
		WAKE MEDICAL CENTER	LABORATORY TESTS	----------	------
		RALEIGH, NC	----------	----------	------
				TOTAL CHARGE	NOT KNOWN
02/21/80	MEDICAL VISIT	DR. JANE GREENE	GENERAL CHECK-UP	FAMILY	$45.00
		RALEIGH, NC	----------	----------	------
			----------	----------	------
			----------	----------	------
				TOTAL CHARGE	$45.00
02/18/80	PRESCRIPTION	AMPICILLIN	1 TIME	FAMILY	$6.82
				BC/BS OF NC	NOT KNOWN
				----------	-----
				----------	-----
				TOTAL CHARGE	$6.82
03/12/80	HEALTH INSURANCE	BC/BS OF NC	Q7 - PRIVATE PLAN		

		----------			

assigned to or omitted for a person. Finally it communicated legitimacy and importance of the survey to the respondents.

Copies of the Summary of Responses were sent to both the respondent and the interviewer just before an interview. The Summary for a reporting unit contained separate pages for each member. The Summary was cumulative with an entry printed for each medical care event. An entry included the type of care, the date, the provider, the purpose, the total charge, and the sources and amounts of payment. Additionally, health insurance or program coverage at the previous interview was printed, including Medicaid and Medicare identification numbers when appropriate.

Interviewing Aids

The NMCUES employed a calendar as a memory aid and a convenient way to maintain information for the next interview. The complexity of the NMCUES suggested against the respondent recording information in a formal diary. However, there are definite advantages to respondents recording information about sickness or medical care at the time it occurs. A calendar with spaces for noting health care was a constant reminder to the respondent of the study and encouraged information to be recorded as an aid to future reporting. A pocket attached to the calendar provided a convenient place to store medical bills or statements. Interviewers reported that the calendar was well worth the expense. Respondents frequently considered it a gift.

Payments were made to respondents upon the completion of the first, second, and last interview. Totalling $20, these incentive payments engendered goodwill and probably contributed to the very high panel retention during the five rounds of interviewing.

Letters of introduction helped in the initial household contact, especially among the State Medicaid samples. The letters, combined with an integrated computer system for locating current addresses and great cooperation from its social services, helped achieve a 97 percent completion rate in the Texas Medicaid sample.

ADMINISTRATIVE RECORDS

Perhaps the unique feature of the NMCUES was the planned linkage of household reported information and administrative record information to produce a more extensive and accurate database. Survey data by itself is only as good as the reporting of people. People may not want to report the information, may not remember, or may not have known the information in the first place. Good survey design can reduce the first two problems, but cannot affect the third. Since Medicaid pays all the cost for covered care, it was expected that few Medicaid beneficiaries would know the cost of their care. Further, they might know little about the provider or the exact nature of their medical problems.

Administrative data also have limitations. They are organized in response to the requirements of a program which is often different than research needs. Administrative records can provide information on health and health care covered by the program, but they cannot provide information about what was not covered by the program. There is also speculation on how accurately claims reflect actual problems, care, and charges. Finally, administrative data may be incomplete or contain unintential error. With their different strengths and weaknesses, combining the two types of data produces a much more valuable dataset for analysis.

Medicaid Eligibility

Medicaid eligibility during each month of 1980 was requested from States for people reported to be covered by Medicaid in all the household surveys. Medicaid identification numbers were collected from respondents during each interview. In the State Medicaid samples, Medicaid identification numbers obtained in drawing the samples were also available. The numbers were sent to the state in which the person lived. Names and addresses of people reported to be covered by Medicaid but from whom an identification number was not obtained were also sent, as some states had the capability of retrieving eligibility by name.

The monthly eligibility for Medicaid was linked to person-level information from the household survey. The linked data are important for understanding the program and how it affects people. Health needs and care can also be studied during periods of non-eligibility.

Medicaid Claims

Medicaid claims data in California, Michigan, New York, and Texas were extracted for people in the respective State Medicaid household samples. (Claims data for Medicaid beneficiaries in the National household sample were not obtained due to the time, effort and expense involved for the relatively few beneficiaries in each state.) Claims records were in different forms and availability in each state, since Medicaid is a state administered program. Data had to be obtained from the counties in New York, since there was no centralized record system. Claims data were first converted to a standard format, and then combined when necessary to represent a single event of health care.

These combined claims, except for prescribed medicines, were matched and linked to household reported medical events. Prescribed medicine claims were aggregated for the year for a person. The matched claims data provide the charge and payment data generally missing from the household interviews, and insight into the accuracy of each source of data. The unmatched data provide information on completeness of reporting.

Medicare Claims

Medicare Part A and Part B claims data were extracted from the Federal Medicare files for people from all samples with a reported Medicare identification number. Part A claims for in-

patient hospital care were aggregated for a single hospital stay and matched with household reported stays. In the Part B ambulatory claims, all care received by a person during a certain span of time from a single provider and billed through a single fiscal intermediary is represented by a single claim record. Therefore, Part B Medicare claims could not be matched with single events of medical care reported by the household. They were therefore aggregated to the person level. Aggregated claims data still provide a great amount of additional and validating information.

TYPES OF ANALYSIS

The NMCUES database is rich in potential uses. It has a wealth of information available to address specific topics, has flexibility on the units of analysis, and has high quality and coverage. While all potential uses cannot be described, some indication of the range of analysis can be mentioned. These stem from combining different subject areas, different analysis units, and the different sources of data.

SUBJECT AREAS
Health
Health Care
Cost of Health Care
Payment for Health Care
Health Insurance

ANALYSIS UNITS
Family
Medicaid Case
Person
Health Care Event
Health Condition

National estimates of the amount and cost of personal health care can be made. These National estimates can be detailed by characteristics of families, people, type of care, or type of condition requiring the care. Estimates can also be made on the source of payment for the health care charges. In addition, the burden of health care costs on individual families or people can be analyzed. This burden may be for all health problems, or for specific health conditions. Both the total yearly expense and the pattern of care and expense during the year can be considered.

Analysis of the Medicare program can take many different forms. NMCUES household and record data can be used separately or in combination to provide national estimates of the medical care covered by the program. The data also address the extent of health problems and care not covered by the program. Possible effects of changing program coverage can also be explored.

The Medicaid programs within four large states can be reviewed with NMCUES data. The data can show the amount of movement into and out of the program during a year, the amount of the health needs and health care covered by the programs and the amount of needs and care not covered by the programs. Both comparisons between states and comparisons between a state and the nation can give insights into the Medicaid program and to population differences. The effects of changing program coverage within any of the four states or at the national level can be explored.

Beyond substantive issues, the NMCUES database has potential for many methodological studies. They center around the amount of agreement between household reports and medical provider reports. While the NMCUES data do not define which source is accurate (or more accurate), a number of insights can be gained.

CONCLUSION

The National Medical Care Utilization and Expenditure Survey did not use any innovative techniques, yet incorporated or further developed many procedures which resulted in a high quality, flexible, and useful data base. One major thrust was the collection and combining of data from both households and administrative records. Special state household samples selected from Medicaid eligibility rolls complement a national area probability sample which permits study of the Medicaid program within and between states.

Household data were collected identically in the four State Medicaid samples and the National samples, and form the basis for analysis of health, health care, and health care cost. Many different procedures were used in NMCUES to produce as accurate data as possible. Panel interviewing at approximate 3-month intervals collected data for calendar year 1980. The short reference period, defined boundary points and the use of calendars were designed to reduce error. A Summary of Responses permitted households to correct or update data in a planned and efficient way.

The level of detail at which the data were collected, combined with careful attention to linkage, produced a dataset that can be used to address many types of questions. Health, health care, health care cost, and sources of payment can be analyzed using the family, the Medicaid case, a person, a health care event, or a health condition as the unit. Total yearly estimates or patterns within a year can be analyzed. Finally, estimates can be made for the total National population, any of its component groups, and for the Medicaid populations within four large states.

REFERENCES

[1] For further description of the survey, see

National Center for Health Statistics, G. S. Bonham: Procedures and questionnaires of the National Medical Care Utilization and Expenditure Survey. National Medical Care Utilization and Expenditure Survey. Series A, Methodological Report No. 1. DHHS Pub. No. 82-20001. Public Health Service. Washington. U.S. Government Printing Office, Mar. 1983.

[2] Health Services Research and Development Center: Medical Economics Survey--Methods Study, Final Report. Contract 106-74-150-HRA. Baltimore, Md. Johns Hopkins Medical Institutions, 1977.

ACCESSING AND USING ADMINISTRATIVE RECORDS TO AUGMENT THE HOUSEHOLD SURVEY DATA IN NMCUES

Embry Howell, SysteMetrics, Inc.

A unique aspect of the National Medical Care Utilization and Expenditures Survey was the acquisition of administrative records in the form of automated and manual claims, enrollment, and provider records from state and county agencies for Medicaid households and from the federal government for Medicare households. This data collection activity was performed by SysteMetrics, as a subcontractor to the Research Triangle Institute.

The acquisition of administrative records served several purposes for the survey. The primary ones were:

- Sampling: 4 State Medicaid Household Survey (SMHS) states (California, Michigan, New York and Texas)
- Verifying eligibility: All Medicaid persons in the Household Survey (HHS)
- Identifying hospital stays and medical visits: Medicare (all HHS states) and Medicaid (4 SMHS states)
- Identifying cost of care: Medicare (all HHS states) and Medicaid (4 SMHS states)

Several types of administrative records were accessed, including:

- Medicare Enrollment Records
- Medicare Claims Records
- Medicaid Enrollment Records
  - Manual Records (38 HHS states and upstate New York)
  - Automated Records (4 SMHS states: California, Michigan, New York City, Texas)
- Medicaid Claims
  - Manual Records (Upstate New York)
  - Automated Records (California, Michigan, New York City, Texas)
- Medicaid Provider Files (4 SMHS states)

In order to acquire and use these records several steps were necessary. Agreements were negotiated with the Health Care Financing Administration for obtaining all Medicare records, with 38 state agencies for obtaining manual Medicaid enrollment data, with the four large Medicaid states for obtaining automated enrollment and claims files, and with upstate New York counties for obtaining manual enrollment and claims data. As part of those agreements, we obtained complete file documentation, established a liaison with staff members in each organization, and determined a schedule for data acquisition.

File contents were reviewed and a uniform set of variables to be extracted was determined. Coding forms and abstraction instructions were developed for collecting manual data. Code maps and computer programs were developed for extracting automated data. A close working relationship with state and federal staff was quite important during this phase of the project, in order to assure that the content of files was clearly understood.

Data collection proceeded in several parallel phases.

1. Enrollment and claims data from upstate New York counties was abstracted directly from manual county files.
2. A coding form was sent to each of the 38 states in the Household Survey, requesting a monthly verification of enrollment for each person who reported Medicaid eligibility.
3. Medicaid enrollment files were obtained from the four large Medicaid states early in the project for sampling purposes and were updated later to obtain information on persons who became enrolled in 1980.
4. Automated claims files from the four State Medicaid Household Survey states were obtained in the fall of 1981, in order to assure that claims for all 1980 health care events had been obtained and processed by the states.
5. Medicare claims and enrollment records were also obtained from the Health Care Financing Administration in late 1981.

The major output files were:

- A uniform Medicaid enrollment file containing identifying information and months of enrollment for all Medicaid individuals in both the Household Survey and the State Medicaid Household Survey.
- A uniform Medicaid claims file, containing detailed information on diagnoses, services, and payments in 1981 for all persons in the State Medicaid Household Survey.
- Cleaned and edited Medicare claims and enrollment records for all persons reporting Medicare coverage in the Household Survey and State Medicaid Household Survey.

These files were forwarded to RTI for merger with the survey data. This "merge and match" process is still underway.

Tables 1-3 present an initial comparison of Medicare and Medicaid enrollment counts from the survey only, and from the "best estimate" obtained by merging survey and administrative records data. They show that Medicare enrollment went up slightly (by about 2%) and Medicaid enrollment decreased (by 3.5%) when administrative records were merged with survey data. Enrollment counts for persons who had dual coverage increased by more than 10%.

Table 1

Preliminary Comparison of Survey and Best Estimate Data
U.S. Medicare Enrollment in Thousands
NMCUES Household Survey, 1980

	Survey	Best Estimate	%Increase
Total	26,521	27,024	1.9
Age			
<65	3,817	4,051	6.1
65-74	14,622	14,791	1.2
75-84	6,813	6,867	.8
85+	1,269	1,315	3.6
Sex			
Male	11,428	11,701	2.4
Female	15,093	15,323	1.5
Race			
White	23,774	24,023	1.0
Black	2,399	2,599	8.3
Other	384	402	4.7

Table 2

Preliminary Comparison of Survey and Best Estimate Data
U.S. Medicaid Enrollment in Thousands
NMCUES Household Survey, 1980

	Survey	Best Estimate	%Increase
Total	21,710	20,956	-3.5
Age			
0-6	4,350	4,176	-4.0
7-20	7,546	6,694	-11.3
21-64	6,382	6,458	1.2
65+	3,432	3,628	5.7
Sex			
Male	8,619	8,420	-2.3
Female	13,091	12,536	-4.2
Race			
White	13,826	13,446	-2.8
Black	7,188	6,779	-5.7
Other	696	731	5.0

Table 3

Preliminary Comparison of Survey and Best Estimate Data
U.S. Joint Medicare/Medicaid Enrollment in Thousands
NMCUES Household Survey, 1980

	Survey	Best Estimate	%Increase
Total	3,737	4,207	12.6
Age			
<65	802	1,002	24.9
65-74	1,602	1,792	11.9
75-84	1,009	1,032	2.3
85+	324	381	17.6
Sex			
Male	1,333	1,501	12.6
Female	2,404	2,706	12.6
Race			
White	2,916	3,195	9.6
Black	709	880	24.1
Other	112	132	17.9

SPECIAL FEATURES OF THE DATA PROCESSING FOR THE NATIONAL MEDICAL CARE UTILIZATION AND EXPENDITURE SURVEY

Barbara A. Moser, Research Triangle Institute

I. INTRODUCTION

Data processing for the National Medical Care Utilization and Expenditure Survey (NMCUES) encompassed three distinct components: data processing for field operations support, data processing for construction of the NMCUES database, and data processing for the NMCUES Analysis and data management system. Certain unique aspects of NMCUES required that special features be designed and implemented as integral parts of the total data processing system.

First of all, NMCUES was three surveys in one: the National Household Survey (HHS), the State Medicaid Household Survey conducted in four states--California, Michigan, New York, and Texas, and the Administrative Records Survey (ARS). The Administrative Records Survey was a separate survey, however, the National Household Survey and the State Medicaid Household Survey were conducted concurrently and required that the interviews be processed together.

In addition to the fact that NMCUES was three surveys in one, it was accomplished by three contractors. RTI was the prime contractor, with NORC (the National Opinion Research Center in New York) and SMI (SysteMetrics in Bethesda, Maryland) as subcontractors. For the HHS and SMHS household surveys, RTI and NORC divided the interviewing and data processing responsibilities. Each firm interviewed half of the National Household sample and the samples in two of the four states, and completed the data processing of their own interviews. RTI was primarily responsible for all the development of the software to do the NMCUES data processing. However, RTI and NORC were responsible for their share of the production data processing necessary for field operations support. This required that identical control systems and data processing software be set up for production running between the two companies.

II. FIELD OPERATIONS SUPPORT

Control of the field operations required knowing the day-to-day status of every case in the sample, interviewer assignments and their status, the data receipt status, i.e., the cases that had been received for processing and their status for the next round of interviewer assignments. This type of control was based on the Reporting Unit status. Therefore, a special Field Operations Receipt Control Enumeration (FORCE) system was developed and put in place for each firm to operate. This system was small and cost-efficient enough that it was run daily with all receipt events for a day's work, and produced field status reports to assist the supervisors in management of approximately 400 field interviewers.

However, to fully support the field operations through five rounds of interviewing the same respondents, a very large participant based control system had to be implemented. This system, called SUMISS (Survey Monitoring System), maintained the status on all persons sampled and interviewed in the study. Exactly the same system was implemented for both RTI and NORC. Each of the survey tasks was monitored by SUMISS which maintained data on each participant and tracked these participants throughout the five rounds of the panel survey. This meant that the control system not only monitored a participant's movement and status from round to round, but it also monitored the data processing steps. As each case was pre-machine edited, coded, and keyed, an event was sent to the control system indicating the status of each interview document for the Reporting Unit. These events determined when all parts of the interview package were ready for the next phase of processing. After data entry, special "forms integrity" processing software was run on the data to determine that all components of an interview package had passed through these steps. In this way, RTI assured that all data required for a particular Reporting Unit assignment had been processed.

SUMISS controlled the production of the field operations assignment forms -- the Control Card and the Summary. The Control Card was a computer generated form which previded the interviewer with complete information necessary to locate and interview the NMCUES Reporting Unit. It included the composition of the Reporting Unit and all details about the individuals to be included. Changes in the Reporting Unit composition and location were recorded on this instrument for the next

The Summary was a computer generated summary of reported visits from prior rounds of interviewing. It was first generated in Round 2 from Round 1 reported data and was generated for all subsequent rounds with cumulative data reported in prior rounds. It allowed for the updating of all information about the reported visits including charges, payments, and sources of payment. The Summary also showed the previous round's reported health insurance coverage and the reference period date.

The next three steps processed the data for computer generating the Control Cards and Summaries to be sent to the field for each subsequent round of interviewing. Special software was required to reorganize the data from the questionnaire format into person and provider visit segments of information in a form that could be processed for Control Card and Summary production and building of the initial database files. Particular information about individuals needed for the next round of interviewing was extracted from the data entry files to update the control system. Specifically, this information included change of address, change of participant status, updates of demographic data (age, sex, birthdate, etc.), the participant's relationship to the head of household, marital status, and the interview status for the preceding interview.

Reported conditions were also pulled from the data entry files to build a unique condition file. The Control Card contained a cumulative list of unique conditions from all prior interviews. Information that was required for the

Summary was split off into files to be processed and added to the Summary files. This included updates to Summary information from previous rounds and newly reported visits or insurance information from the current round.

Finally, the Control Card and Summary themselves presented special requirements for field operations support. The design of the NMCUES survey required that participants be tracked regardless of where they moved, either from one particular physical location to another location, or from one Reporting Unit to another. Therefore, the control system and the appropriate production of the Control Card and Summary required development of special SUMISS software which would allow for this tracking process. Information which indicated that a person had moved from one Reporting Unit into another Reporting Unit or had created a new Reporting Unit was extracted from each interview package to update the control system. The address of the person moving was then used to produce a new field interview assignment for Reporting Units generated in this manner.

When the control system received events indicating that the data to produce the Control Card and the Summary had been processed, a special request to the Summary and Control Card systems was automatically generated to produce both of these assignment documents at the same time. Even though the Summary and Control Card systems were parallel processing systems, they were coordinated so that a field interviewer's assignment was generated at the same time. That is, the request for Summaries and request for Control Cards were sorted by field interviewer ID, and when the print jobs were run, the field staff could simply separate Summaries and Control Cards and sort them into packages to mail to each individual field interviewer. A duplicate of the Summary was generated and mailed to the respondents at the same time. The Control Card and Summary was the assignment for the next round of interviewing, and when that assignment was mailed, the control system automatically generated an event that was sent to the Field Operations Receipt Control Enumeration (FORCE) system indicating that the next round's assignment for the Reporting Unit was in the field.

III. DATABASE CONSTRUCTION

Requirements and special features of the database construction task were probably not so different from most other database constructions of survey data. Consistency edits and recodes and imputations for missing data were all a part of the database construction of the NMCUES data. There was a special requirement that all NMCUES Summary updates and corrections had to be linked back to the initial questionnaire reported event, and updates to that event record were made to the Summary data. Since unique record IDs were constructed from the original questionnaire form and maintained in the raw data files, the linkage to Summary data events was not a particular problem. The exception to this was in the area of health insurance coverage and events and corrections on the Summary form.

The Summary allowed for a respondent to update, change, delete or add information to his or her record. The Summary carried health insurance coverage information reported in the previous round by each Reporting Unit. For a particular Reporting Unit, one or more persons could be covered by multiple insurance plans, with the same plan covering one or more persons in that Reporting Unit. The linkage of this health insurance coverage information to individuals within the Reporting Unit was primarily based on the order in which the insurance coverage was reported. This specific order was repeated and printed on the Summary for review. Therefore, if a respondent changed the insurance information on the Summary by deleting, adding, or changing the order, the linkage back to the coverage and insurance plan information reported in the questionnaire was virtually destroyed. Since the insurance information as reported on the Summary was expected to be the most accurate and the most up-to-date, preparing this information for the final database delivery required some special data processing efforts.

Other requirements for database construction dealt with the complexity of the forms and the many variable aspects of the data. For example, each Reporting Unit could have multiple people (the range went from 1 person Reporting Units to 24 person Reporting Units); each person could report no medical events or multiple medical events, i.e., no medical visits, 1 medical visit, or many medical visits. Each visit could have a maximum of three conditions associated with it.

Due to the complexity of the data format, collection, editing, and coding procedures, quality control checks were implemented at each step in the data collection, processing, and construction phases. First, the field operations staff sampled and verified 10 percent of the interviews assigned. Second, the pre-machine edit and coding staff selected a ten percent sample for re-edit and re-coding. Unacceptable error rates resulted in retraining or release of the editors and coders. In addition to the routine quality control on these two steps, RTI and NORC exchanged a sample for independent coding and resolution.

At data entry, a 10 percent sample, by form, by keyer was selected and double re-keyed for three-way comparison. Error rates were computed by variable and by character for each operator. Retraining was done for any operator who had greater than 1 percent error rate per form. An overall error rate of less than .50% was maintained.

The next step following data entry was the transmission and reorganization of the data into fixed length segments corresponding to parts of the questionnaire. After this step a 5% sample was selected for manual verification against the questionnaire. An error rate was computed from this item by item comparison. The overall error rate was less than 3/4 of 1 percent. As a final check, variables in the deliverable 12 month data files were verified against the 5 percent sample.

Similar verification procedures were conducted on the Summaries and Control Cards produced in subsequent steps. These quality control

procedures made it possible to produce these highly accurate assignment forms within six weeks of the previous assignment.

The database construction for NMCUES has been a three-step process. The first database delivered was the Twelve-Month Database of the household information. This database was constructed by taking the initial respondent information, linking it to the Summary data updates, preparing imputations for a small subset of items, and adding imputed records for persons who were eligible for periods during the survey but did not respond.

The weights computation component required maintaining the person history which could then be verified against that person's survey eligibility days and the number of days the person responded in the survey. The computation also involved the status of a person at any given time, with respect to his family and Reporting Unit.

The second step in the construction of the NMCUES database was the construction of a set of Analytic files according to client specifications. This database included additional editing and cleaning, extensive sets of recodes for certain variables, further imputation for missing data, and some special variable constructions. A third step has been the development of a set of Public Use files. The Public Use files, a database developed from the Analytic files, contain only HHS respondent data which have been further cleaned, edited, recoded, and imputed for missing items.

Finally, work currently underway will provide a final NMCUES database which includes all of the State Medicaid Household Survey respondent information, the linked claims information from the Administrative Records Survey, and a set of charge variables imputed and derived from the survey and claims data. The task of creating these charge variables is called the "Best Estimation" and will provide the client with variables which represent the best estimate of total charge, sources of payment, and payment amounts. The Best Estimation task has included detailed study of state program characteristics in order to achieve a process which would provide appropriate best estimate data for future use. A major requirement of this Best Estimation task and ultimate database construction was the linkage of claims reported by each of the four state Medicaid agencies to the household reported medical care events. The event level linkage of this data was so complex that after several matching attempts and a review of various computer programs to match data, it was determined that the event level matching could only be done by an individual who was extremely knowledgeable about the database, the procedures used to collect the data, the program characteristics of each state, and the requirements for matching. Therefore, the entire matching process of all of the event level claims to household reported responses was manually done by one person, and was quality controlled and verified by an NCHS representative who was very knowledgeable about the task. Special data processing programs which fit the characteristics of the Medicaid programs in the individual states were prepared to assist with the matching, and to ultimately produce a matched survey and claims record for use in the Best Estimation task.

IV. DATA MANAGEMENT SYSTEM FOR THE ANALYSIS DATABASE

Currently, the total NMCUES database contains approximately 5,000 unique variables which are being used for several analysis efforts and are being used continuously to construct new variables and recodes. Currently this use is primarily by the NCHS and HCFA clients and the analysts who are either under contract to NCHS and HCFA or are NCHS/HCFA staff. In order that the inefficiencies associated with duplication of effort in recreating variables and redoing tabulations be reduced to a minimum, RTI developed a monitoring system which maintains information on each variable in the NMCUES database. This Variable Monitoring System is also associated with a file log and Table Monitoring System which maintains information about analysis files which have been created and tabulations which have been completed, respectively. The Variable Monitoring System is unique to the NMCUES database, and was designed specifically to carry sufficient information about each variable to identify that variable's source, decision logic tables specified for variable creation, and the file in which it is maintained. The Variable Monitoring System also provides the user with a data dictionary for any file he/she wants to build, given a specific set of variables which are in this system. These data dictionaries also can be utilized to create SAS input for creation of SAS files which contain those variables. At this point, the Variable Monitoring System simply monitors the variable and provides its descriptions and its location on other files. It may be used to create data dictionaries describing a file to be created. However, the ultimate system would be development to the point that a specified set of variables could be input and the resultant output would be a file with those variables on it, in the specified order, with documentation by the Variable Monitoring System. This capability is not outside the realm of possibility and could be developed if budget and schedule permitted.

Special tabulations software, specifically tailored to the NMCUES data sample design and type of variables created, has also been developed. This software is being utilized in analysis to properly estimate the variance of the complex NMCUES sample and present the analysis results in meaningful, readable form.

When interpreting NMCUES data it is important to account for the sampling variability inherent in the population estimates. This is usually done by providing an estimate of the standard error of the quantity of interest. The standard error can be used to construct a confidence interval for the pouplation value being estimated. This provides a range of values which can reasonably be expected to include the population value. RTI's survey data analysis software efficiently and routinely calculates standard error estimates accounting for the complex NMCUES sample design. In addition, the analysis software directly interfaces with RTI's table manage-

ment and production software discussed above so that the requisite reliability tables can be produced with a minimum of effort.

The production of detailed tables is greatly facilitated by RTI's table management and production software. This software produces report ready tables from output files created by RTI's survey data analysis programs. Noteworthy among the capabilities of the table production system are:

- concatenation of several independent tables
- rearrangement of a table's dimensions
- calculation of row or column percentages
- term-by-term operations on data from several tables
- tabular printing of results
- extensive title, footnotes, and row and column labeling

The user has a high degree of control over the format of the printed tables (title, labels, field widths, number of decimal places, etc.). In addition, the computer generated tables can be directly transferred to the IBM System 6 word processing system without any manual keying.

In summary, the NMCUES data processing system had a variety of unique requirements in support of field operations, database construction, and data analysis. These requirements were resolved with special software features of the system in those areas. The major special software features were in the control system, Control Card and Summary production, the linkage of Summary updates to household reported data, the linkage of claims and computation of "Best Estimate" data, the Variable Monitoring and the tabulations analysis software. Quality control procedures at each data processing step was required to meet the schedule and maintain the quality of data specified for the NMCUES project.

Techniques to Measure and Improve the Completeness and Quality of Vital Records

Session O

BIRTH CERTIFICATE COMPLETION PROCEDURES AND THE ACCURACY OF MISSOURI BIRTH CERTIFICATE DATA

Garland Land and Bill Vaughan
Missouri Center for Health Statistics

INTRODUCTION

The birth certificate serves a dual legal and statistical function. The legal items such as name, sex, date of birth, county of birth, etc., must be accurately recorded to meet the standards and acceptance of adjudicating agencies. Any errors on legal items detected by the hospital, local registrar, state registration office, parents, or person of record are usually corrected.

However, the medical statistical items found in the lower half of the certificate many times are not given as close a scrutiny. Studies have shown that there is considerable difference in the completeness and accuracy of the items on the birth certificate specifically when the legal items are compared to the medical items.[1,2,3,4,5]

The National Center for Health Statistics provides registration officials a handbook to assist in completing the birth record.[6] Included in the instructions are procedural suggestions to insure proper recording. However, the actual procedures which are used in hospitals for completing birth certificates are generally unknown. The relative value of one procedure over another is also unknown.

A survey was sent to the 119 Missouri hospitals who delivered infants in 1980. All hospitals responded to the survey.

The survey was designed to answer the following general questions:

1. What is the source of information for each of the medical items?
2. Who and by what procedure are the medical items entered on the certificate?
3. What quality control procedures are used?

Using the survey results, we compared the birth data for hospitals using different procedures. From this comparison we can evaluate the effect of one procedure over another.

FINDINGS

The Missouri birth certificate has all the items recommended by the National Center for Health Statistics. The medical and pregnancy items on the standard certificate plus two additional items on the Missouri certificate were studied. The two unique Missouri items studied were indication if Cesarean and mother's prepregnancy weight.

For the eleven medical and pregnancy history items on the Missouri birth certificate, the hospital personnel were asked to indicate the source which usually furnishes data for the birth certificate. Six different sources were indicated on the survey--physicians who delivered the infant, obstetrical personnel, prenatal care record, physician caring for the infant after delivery, mother while in the hospital, and medical record personnel. The infant's record was indicated by several hospitals as an additional source. Presumably this record is compiled from several sources including the physician delivering the infant, physician taking care of the infant and obstetrical personnel. Several hospitals also indicated that more than one source is used. This may occur when physicians use different procedures in completing the record. Or several sources may be consulted to insure accuracy or completeness.

As shown in Table 1, the physician delivering the infant is the usual source of information for the various types of complications including those that occurred during pregnancy, during labor, and indications for C-sections. The physician delivering the infant is also the primary source of information for the malformation section. The physician caring for the infant provides the malformation information in less than 20 percent of the hospitals.

There is no apparent preference in obtaining the Apgar score from either the physician or the obstetrical personnel. Table 2 shows that the pregnancy history and mother's weight before pregnancy is obtained in most hospitals from the mother while in the hospital. The prenatal care information is secured from either the mother or the prenatal care record on about an equal basis. In over one-third of the hospitals, the mother is asked to recall while in the hospital the date of her last menses, the month prenatal care began, and the number of prenatal visits. In two-thirds of the hospitals, the mother is asked to remember her weight before pregnancy began.

These results indicate that in several instances the primary source is not being used to complete the birth certificate. Namely, the physician delivering the infant usually completes the malformation section rather than the physician caring for the infant. In several hospitals, the prenatal care, menses, and prepregnancy weight information is obtained from the mother while in the hospital rather than from the prenatal care record. This study did not try to determine if the mother's recall was different from the prenatal record. We did, however, try to determine if there were detectable differences in the aggregate data according to the source. Hospitals were grouped according to the source of information and who wrote the information on the birth record. The rates or means for the groups of hospitals using similar procedures were computed for each appropriate data item.

The 1976-80 malformation rates for groups of hospitals according to the source of the malformation data is shown in Table 3.. The differences in the rates are a function of both the true malformation rates for the hospitals and the completeness of reporting. Because hospital size is correlated with malformation rates[7] and may be correlated with the birth certificate procedures, the malformation rates were adjusted by hospital size. However, the adjusted rates did not effect the ordinal relationship of the nonadjusted rates so the nonadjusted rates are used.

The malformation rates are lowest when either the obstetrician or pediatrician is the sole source of the data. There is no discernable difference between the rates for the two physicians.

The rates are about 20 percent higher when the malformation data is provided by hospital staff or a combination effort of the physician and hospital staff.

The 1980 birth certificate data on prenatal care and prenatal visits for hospitals grouped by source of information was analyzed for both completeness and quality.

Missouri has an extensive computer edit and query system so the number of unknowns is very small regardless of reporting source. Table 4 shows those hospitals that reported using both the prenatal record and the mother to obtain the prenatal care and prenatal visits information had the lowest incompleteness rate. Presumably these hospitals use one source primarily and the other source as a backup. This naturally leads to more complete information than relying on just one source. Because the incomplete percentages are so small and the records have already passed through a quality control system, it is difficult to claim that either the mother or prenatal record provides more complete prenatal information.

However, Table 5 shows that there is a small bias when the mother reports the prenatal care information. Hospitals using the mother as the source of information report earlier prenatal care and slightly more prenatal visits on the average. One possible explanation for the discrepancies is that the mother might be including the visit to test for pregnancy as a prenatal visit. Another possibility is that she includes the last prenatal visit which is not recorded on the prenatal care record. This may happen when the prenatal record is sent to the hospital in anticipation of the delivery but before the last prenatal visit occurs.

To determine if this is a problem, the hospitals were asked if the last prenatal visits are included on the birth certificate if the prenatal record is sent to the hospital before the last visits occur. Of the 42 hospitals which use the prenatal record, ten do not include the last visits. As shown in Table 5, these ten hospitals had an average of 8.9 visits as compared with 10.7 visits for the hospitals which attempted to report the unrecorded visits. When the last prenatal visit is included, the prenatal record and the mother's information are very similar.

The month prenatal care began and prenatal visits were also checked to determine if there was any heaping on even numbers or proclivity to a particular value. There was no even number heaping for either of the variables. However, the distributions were distinctive. Hospitals using the prenatal record reported 44 percent of the mothers received care in the first two months whereas hospitals using the mother as the source reported 56 percent received early care. The same pattern was shown for prenatal visits. Hospitals using the prenatal records reported 52 percent of the mothers with ten or fewer visits compared with 39 percent when the mother is the source. There also appears to be some heaping on 12 visits when the mother is the source. The prenatal record showed 18 percent of the mothers had 12 visits whereas 26 percent of the mothers reported 12 visits.

The hospitals were also grouped according to source of information providing the date of last normal menses and mother's weight before pregnancy. The different sources had comparable means for length of pregnancy and mother's weight.

The day of last normal menses was more incomplete when the mother provided the information in the hospital as opposed to obtaining it from the prenatal record. However, there was a higher incompleteness rate for mother's weight for those hospitals using the prenatal record.

The mother's prepregnancy weight was checked for heaping on 0's and 5's for the different sources. Heaping did not vary much between sources. However, the overall heaping was substantial. Over 60 percent of the weights regardless of source of the information was reported to end in 0 or 5. The reason for this is that a scale cannot be used to determine this information for the vast majority of women. Only 15 percent of the women receive care during the first month of pregnancy. Thus, at least 85 percent of the women must recall their weight before pregnancy after the first month of pregnancy. Over 70 percent of the women are providing the information while in the hospital.

The actual recording procedures were also surveyed along with the source of information. Several procedures are used to record the malformation and complication sections.

In about 60 percent of the hospitals, the physician writes the malformations and complications on the birth certificate. 23 percent of the hospitals use either the obstetrical personnel or medical records staff to complete these items. The physician writes the malformations and complications on work sheets for hospital personnel to transcribe on the birth certificate in 15 percent of the hospitals. While the work sheet is not used very frequently for the malformation and complication items, it is heavily used for the other medical items. 98 percent of the hospitals use a work sheet for all or some of the medical items.

As with the sources of information, the varying procedures used to complete the medical items were compared with the quality of the data.

The hospitals that have poorest malformation reporting are those that rely upon the physician to write the malformation on the birth record (Table 3). Substantial improvement in reporting is shown when the hospital staff completes the birth certificates from information provided by the physician on a work sheet or when the hospital personnel complete the malformation section from information in the medical record and/or personal observation. Use of a work sheet rather than direct entry on the birth certificate by the physician possibly gives hospital staff a better opportunity to study the records and make changes if the physician's information is incomplete. It should be noted that while there are some differences in the rates according to the procedure used, all rates are still substantially lower than expected. The expected malformation rate for Missouri is approximately 27 malformations per 1,000 live births.[8] Using the work sheet gave the highest rate which was only 11.9. So while who provides the information and completes the records does make a difference, the underreporting of malformations is a problem regardless of the procedure used to complete the record.

The hospitals were asked if information provided by the physician was cross-checked with hospital records. Hospitals with fewer deliveries

were more likely to cross-check the information--67 percent of the hospitals delivering less than 200 births per year cross-check whereas 39 percent of the larger hospitals cross-check the information. To see if cross-checking had a demonstrable effect on reporting, the malformation and complication rates for the two groups of hospitals were compared. The rates were about 20 percent higher for hospitals that used a cross-checking procedure. The observation that smaller hospitals tend to cross-check information and have better reporting is consistent with the Missouri malformation study.[9]

One of the recommended procedures in the hospital handbook is that the informant who is usually the mother sign the work sheet and the birth record. The surveyed hospitals were asked if the mother sees the completed birth certificate before it is sent to the local registrar. 13 percent of the hospitals do not show the completed record to the mother. 22 percent only show the mother the work sheet and not the completed certificate. Even hospitals that show the birth certificate to the mother follow different procedures. 5 percent of the hospitals show the mother only the completed upper or legal portion of the record. 22 percent of the hospitals show the mother the completed information down to but not including the complications section. 35 percent of the hospitals show the entire completed record to the mother.

CONCLUSION

There is a wide variety of procedures and practices being utilized by Missouri hospitals in completing birth certificates. Some hospitals noted the use of more than one procedure. This is probably due to the differing preferences of the physicians on staff. The prenatal record would be considered the primary source document for the prenatal information. However, hospitals are as likely to obtain the prenatal information from the mother while in the hospital as to obtain it from the prenatal record.

In two-thirds of the hospitals, the congenital malformations section is completed from information provided by the physician who delivered the infant. This is in deference to the fact that it is the pediatrician and obstetrical staff who care for the infant and have first-hand knowledge of its health.

Other good record management procedures such as cross-checking different sources are done by 50-60 percent of the hospitals depending upon the data item. Allowing the mother to view the entire completed certificate only occurs in 35 percent of the hospitals.

These different procedures have a significant effect upon the quality of the data. Some hospitals have what could be described as a team approach to completing the record. The physician, obstetrical staff, prenatal record, and mother are all consulted and utilized where appropriate. These hospitals have lower item incompleteness rates and better quality data. Hospitals that use only one source which may be the secondary source have the poorest quality data.

This study demonstrates birth certificate completion procedures do effect the quality of the medical information. With proper information and training, hospitals should be able to improve their procedures which can enhance the quality of the birth data.

REFERENCES

1. Querec, LJ: Comparability of Reporting Between the Birth Certificate and the National Natality Survey. US DHEW Series 2, No. 83, PHS 80-1357, April 1980.
2. Carucci, P: Reliability of Statistical and Medical Information Reported on Birth and Death Certificates. New York State Department of Health Monograph No. 15. New York State Department of Health, Albany, N.Y., May 1979.
3. Floyd, L, Lavoie, M and Terry, JS: The Status of Birth Certificate Information in Georgia. Journal of the Medical Association of Georgia 70:871-873, December 1981.
4. Montgomery, TA, et al.: Live Birth Certificates - Evaluation of Medical and Health Data in California. California Med. 96(3):190-195, Mar. 1962.
5. Lilienfeld, AM, et al.: Accuracy of Supplemental Medical Information on Birth Certificates. Pub. Health Rep. 66(7): 191-198, Feb. 16, 1951.
6. Hospital Handbook on Birth Registration and Fetal Death Reporting, US DHEW Publication No. (PHS)78-1107, February 1978.
7. Bakewell, JM and Land, G: Completeness of Congenital Anomalies Reported on the Missouri Birth Certificate: Comparison with the Hospital Discharge Abstract. Missouri Center for Health Statistics Publication No. 6.3, Jefferson City, Missouri, June 1981.
8. Ibid., p. 11.
9. Ibid., p. 11.

TABLE 1

PERCENT OF MISSOURI HOSPITALS
USING THE PHYSICIANS, HOSPITAL PERSONNEL,
OR THE PRENATAL RECORD AS THE SOURCE OF INFORMATION
FOR SELECTED BIRTH CERTIFICATE ITEMS

BIRTH CERTIFICATE ITEM	PHYSICIAN DELIVERING THE INFANT	PHYSICIAN CARING FOR THE INFANT	HOSPITAL[1] PERSONNEL	PRENATAL RECORD	OTHER
Apgar Score	36.1 (5.9)[2]	5.0	35.3		17.7
Complication of Pregnancy	77.3		3.3	5.9	13.5
Other Illnesses or Conditions Affecting the Pregnancy	73.1		4.2	9.2	13.5
Complication of Labor and/or Delivery	82.4 (6.7)[2]		5.9		5.0
Congenital Malformation	65.6 (9.2)[3]	10.9	5.1		9.2
Indication if Cesarean	79.0		8.4	0.8	11.8

[1]Obstetrical or Medical Records Personnel

[2]Combination of Physician Delivering Infant and Obstetrical Personnel

[3]Combination of Physician Delivering Infant and Physician Caring for Infant

TABLE 2

PERCENT OF MISSOURI HOSPITALS
USING THE MOTHER AND/OR THE PRENATAL RECORD
AS THE SOURCE OF INFORMATION
FOR SELECTED BIRTH CERTIFICATE ITEMS

BIRTH CERTIFICATE ITEM	PRENATAL RECORD	PRENATAL RECORD AND MOTHER WHILE IN THE HOSPITAL	MOTHER WHILE IN THE HOSPITAL	OTHER
Month Prenatal Care Began	34.4%	21.8%	34.4%	9.4%
Number of Prenatal Visits	37.8%	15.1%	34.4%	12.7%
Date Last Normal Menses Began	28.6%	21.0%	41.1%	9.3%
Mother's Weight Before Pregnancy	16.0%	15.1%	66.4%	2.5%
Pregnancy History	8.4%	17.7%	67.2%	6.7%

TABLE 3

1976-80 MALFORMATION RATES FOR HOSPITALS
GROUPED ACCORDING TO WHO PROVIDED AND WHO ENTERED
THE MALFORMATION DATA ON THE BIRTH CERTIFICATE

SOURCE OF INFORMATION	NUMBER OF HOSPITALS	MALFORMATION RATE PER 1,000 LIVE BIRTHS
Physician Delivering Infant	78	8.8
Physician Caring for Infant	13	8.6
Hospital Personnel	6	10.2
Other[1]	22	10.6
PERSON ENTERING INFORMATION		
Physician Delivering Infant	52	8.4
Physician Caring for Infant	2	6.0
Obstetrical Personnel	8	11.9
Medical Records Personnel	27	9.6
Other[1]	30	9.5

[1]More than one person provides or writes the information.

TABLE 4

PERCENT OF UNKNOWNS ON SELECTED DATA ITEMS
ON THE BIRTH CERTIFICATE ACCORDING TO THE
SOURCE OF INFORMATION, MISSOURI 1980

SOURCE OF INFORMATION	MONTH PRENATAL CARE BEGAN	NUMBER OF PRENATAL VISITS	DATE LAST NORMAL MENSES BEGAN		MOTHER'S PREPREGNANCY WEIGHT
			MONTH	DAY	
Prenatal Record	0.9%	0.7%	1.0%	14.5%	3.1%
Prenatal Record and Mother While in Hospital	0.3%	0.4%	0.6%	14.9%	0.4%
Mother While in Hospital	0.4%	0.9%	0.9%	22.3%	0.3%
All Sources	0.5%	0.8%	1.0%	19.6%	0.6%

TABLE 5

MONTH PRENATAL CARE BEGAN
AND AVERAGE NUMBER OF PRENATAL VISITS
FOR HOSPITALS GROUPED BY SOURCE OF INFORMATION

SOURCE OF INFORMATION	AVERAGE MONTH PRENATAL CARE BEGAN	AVERAGE NUMBER OF PRENATAL VISITS		
		TOTAL	LAST PRENATAL VISITS INCLUDED*	
			YES	NO
Prenatal Record	3.0	10.1	10.7	8.9
Prenatal Record and Mother	2.7	10.4	10.3	N.A.
Mother	2.4	11.1	11.1	11.0
All Sources	2.6	10.7	10.8	9.7

*If the prenatal record is sent to the hospital prior to the delivery of the infant, are the last prenatal visits unreported on this prenatal record included in the number reported on the birth certificate?

IMPACT OF FIELD REPRESENTATIVE ON QUALITY OF DATA

Linda A. Bowman Nancy K. Briette PENNSYLVANIA DEPARTMENT OF HEALTH

I. INTRODUCTION

I would like to discuss with you today the field program activities in Pennsylvania and their effect on quality and completeness of data reported on birth and death certificates. Also, I will touch briefly on some of the activities of the quality control, statistical and querying units and how they have identified problem areas and worked with the field program to try and resolve some of the reporting difficulties.

II. BACKGROUND OF THE PENNSYLVANIA VITAL REGISTRATION SYSTEM

To gain an understanding of Pennsylvania's experience, problems and successes with the Vital Records field program, it is necessary to know some of the background of the system. In 1974, the Division of Vital Records was moved from Harrisburg, to New Castle, Pennsylvania. The central office of vital records is located in New Castle which is where all birth, death, fetal death, marriage and divorce certificates are eventually processed, microfilmed and filed. The field program operates out of New Castle, but the quality control, statistical and data processing units are in Harrisburg, which is approximately 250 miles from New Castle. You can readily understand that this could cause some difficulty. Regular contact is maintained between the two units through daily telephone calls, special weekly conference calls and onsite visits between the two offices.

The field program to improve quality and completeness of reporting began in 1979. There is and always has been only one field representative who maintains contact with approximately 300 hospitals, 3000 funeral directors, 330 local registrars and hundreds of other personnel who may be involved in some part of the vital registration system. Last year there were over 160,000 births, nearly 120,000 deaths, and 3,500 fetal deaths.

Field Representative Activities

The field representative is responsible for a wide variety of activities, both in the field and administrative tasks at the vital records central office in New Castle. His job duties are split 50/50 between field and administrative activities. It is anticipated that field activities will increase this year due to additional review of the data concerning problem reporting areas.

Basically, the field program consists of training, problem solving and public relations. Each one of these aspects is necessary to the effectiveness of the field program.

Training in correct completion and submittal of certificates is very important and the field program initiates this training even before health personnel become actively involved in the vital registration program. Training sessions are held three times per year for graduating funeral directors. These training sessions involve instruction by the field representative through lecture and a slide presentation. A training session similar to this is also held five to six times per year for graduating medical record administrators to instruct them in the correct completion of birth and fetal death certificates. These training sessions not only instruct individuals in correct completion and submittal of forms, but also emphasize the importance of the data and outline the many uses.

Training sessions are also held for newly licensed physicians, coroners, medical examiners and local registrars. There are approximately 330 local registrars in Pennsylvania who are responsible for 1) examining each birth, death and fetal death certificate for appropriate completion; 2) querying for missing or unclear information; 3) replacement of torn or defaced forms, reproduced or carbon copied forms, those forms not signed by the certifier, etc.; 4) issuance of burial, cremation or removal permits; 5) issuance of certified copies of death certificates while the original is in their possession; and 6) after review and completion of certificates, submittal of all completed forms to the Division of Vital Records in New Castle. You can see that local registrars have quite a variety of duties and it is the field representative's responsibility to ensure that each one is following the same procedures. Each registrar receives a detailed manual which outlines their responsibility with regard to the system, querying procedures, instructions for forms submittal, permit and certified copy issuance, etc. Training sessions are held as needed, usually when a problem area is encountered and additional training is deemed necessary. Since there is not a high turnover among local registrars, ongoing training sessions are not as necessary as they may be with other health professionals in the Vital Registration System.

Even though training is provided to those individuals responsible for certificate completion, problems do arise. The vital registration system is complex and many people are involved which makes a high potential for error. With regard to the birth certificate, there is usually a high turnover with those involved in completing the form, especially hospital personnel. Although a problem may be corrected one month, if someone else is hired or becomes involved, the same problem or a different one could arise. The physician or midwife who attends the birth is legally responsible for the completion of the birth certificate but usually the record is completed by other hospital personnel. Also, it has been discovered that in some hospitals the birth certificate is completed at various locations by different individuals which can cause considerable difficulty. In one specific case, we discovered that the medical records department was two blocks away from the hospital. Needless to say, this hospital had a high number of reporting errors or missing data.

Completion of death certificates does not seem to pose the same kinds or as many problems. There is little turnover among funeral directors and they are required by law to complete the certificate and obtain the medical certification. The only part of the certificate requiring medical information is cause of death which must be certified by the physician. Whereas, on the birth certificate there are at least 15 medical/statistical items which may be completed by just as many different persons in the hospital. It seems that the reporting problems encountered with birth registration are due primarily to the nature of the system.

The field representative spends about half his time at the Division of Vital Records in New Castle. When the certificates are received there, they are batched in county order, alphabetized, sequentially numbered, and data entered. Forms that must be queried for invalid data or blanks are 'kicked out' in a computer edit and returned to the query unit for appropriate action. The query unit and field representative work closely with regard to problem reporting areas. For instance, if a particular hospital continues to make the same error or consistently leaves a certain item blank and the query unit is unable to resolve the problem, the field representative will send a form letter or special letter to that hospital outlining the reporting problem and how to correct it. From this point the field representative must wait one to two months until the certificates are processed from the month the letter was written to see if the correct reporting procedures have been implemented. If errors still persist, a phone call is made referring to the problem and correspondence. If the problem still is not resolved a site visit is made to the facility. Unfortunately, this entire process could take many months and in the meantime the data continue to be reported incorrectly. A way we are trying to resolve this will be discussed later. On the average, 20-30 site visits are made per year, but this number is likely to increase due to additional knowledge of where problem areas are occurring.

Another important part of the field representative's duties is the public relations activities which are ongoing throughout the year.

The field representative regularly attends conventions involving data providers such as funeral directors and local registrars. Usually an exhibit booth is set up where the slide presentation is shown, handbooks are distributed, form supplies can be ordered and publications showing the many data uses can be reviewed. Questions from data providers can also be answered. The public relations part of the job increases the rapport between data providers and the central office. It is necessary that they know the importance of their part in the Vital Registration System.

III. DETERMINING PROBLEM AREAS

Determining problem reporting areas which may have slipped through the query process was not done to any large extent in the early years of the Vital Statistics Cooperative Program contract. Emphasis was placed on converting the existing system to comparability with federal guidelines. This involved redesign of all birth, death and fetal death certificates, implementation of revised data entry procedures and compilation of new coding and editing instructions. A considerable amount of time was also spent just becoming familiar with the vital statistics data collection and reporting system, since many of the staff members were new to the program. Not until 1981 were the data closely examined as to the quality and completeness of reporting on birth and death certificates.

Based on statistical reports completed by the Health Data Center, it was evident that there were higher amounts of unknown data reported on the birth certificates than on the death certificates. As I stated earlier this is partly due to the nature of the two

systems; specifically more statistical/medical items are included on the birth certificate which may be completed in different areas of the hospital, by inappropriate personnel, or the hospital may not keep the required records for completion of certain items on the certificate. In this last case, a hospital is permitted to enter "unknown" and the item will not be queried. Unfortunately, an entry of "unknown" or "not available" is considered a valid entry. This is one of the major problems with incomplete reporting on birth certificates and we are trying to decrease the amount of "unknown" entries reported by data providers.

Since the medical portion of the birth certificate was identified as the major problem area, we decided to concentrate on data items from that section. Seven of these items are used extensively in statistical reports and by other units in the Department of Health, such as the WIC Program and Maternal and Child Health, so reducing the amount of unknown entries is very important. Often these statistics are broken down to the county and minor civil division level and if the amount of missing information is high for one particular area, the resulting statistics can be very misleading. We have noticed that certain counties or regions do have more reporting problems than others.

Specifically, the seven items examined from the medical portion of the birth certificate were:

1. Education of Mother
2. Number of Previous Live Births Now Living
3. Number of Previous Live Births Now Dead
4. Number of Previous Terminations < 16 weeks
5. Number of Previous Terminations >= 16 weeks
6. Month of First Prenatal Visit
7. Birthweight

There are additional items from this section such as Apgar Score and Length of Pregnancy which also cause some reporting difficulties, but it was decided to concentrate on those items which were reported in statistical publications or were requested most often by data users. The remaining items in the medical section will be reviewed sometime in the future.

The seven selected items were examined for each hospital and at the state level for amounts of unknown data. This was done for data reported from 1979 thru 1981. Since the field representative started in 1979, we could discover what impact was made during the first three years of the field program. We did not compare data prior to 1979 because the Vital Statistics Cooperative Program contract was implemented at that time and revisions were made to the existing systems and certificates. We do know, based on statistical reports, that amounts of missing information or poor quality data were high prior to 1979. In addition to checking for unknown data, the types of entries for two of the items, month in pregnancy of first prenatal visit and birthweight, were reviewed. For these two items, there are two types of valid entries. For prenatal care a valid entry could be the number of months in the pregnancy when prenatal care began or the actual named month of the year. The number of months in the pregnancy is the preferred type of entry and the way it is asked for on the certificate since it can easily be converted to trimester of first prenatal visit which is reported in statistical publications and needed by many data users. If the named month of the year is entered, a complicated conversion to month in pregnancy could be done so that trimester of first prenatal visit can be generated. This conversion program has not yet been built into the system, so entries reporting actual month of the year are counted as unknown data for statistical reporting purposes. Another problem encountered with this item is that some hospitals are entering the number of weeks in the pregnancy of the first prenatal visit. We have not yet determined how many hospitals are doing this because it is not evident unless you examine actual certificates. It is unfortunate that the quality and completeness of this item is poor when so many data users require trimester of first prenatal visit statistics.

Valid entries for the birthweight item are grams or pounds and ounces. For statistical reporting, birthweight in grams is the preferred method. The conversion from pounds and ounces to grams is not a problem, except that some exactness in the data is lost in the conversion process. Unfortunately, only 8% of the births have their birthweight reported in grams on their certificate. It seems that reporting in grams should not be a burden to the hospitals and we will be encouraging them to do so in our next revision of the certificate.

As I said, the analysis of these seven items was done for three years, 1979-1981. There were some encouraging results as well as some areas which will require some special attention. First, the encouraging results:

1. The overall amount of unknown or not available entries decreased substantially for all seven items.

2. The items with the most improvement were the pregnancy history items including previous children now living, those now dead, and pregnancy terminations before 16 weeks and 16 weeks or longer.

3. The prenatal care item still contains large amounts of unknown or unusable information, but the reporting has substantially improved since 1979. At that time there were almost 11,500 unknown or incorrect entries which was over 7% of all resident births. This figure dropped to approximately 7,600 in 1981, less than 5% of all resident births. Preliminary statistics for 1982 show that this figure will drop again.

4. Improvement has also been shown in the number of cases where the birthweight was reported in grams on the certificate. In 1979 only 9,400 cases had the birthweight recorded in grams. This figure increased to over 12,000 in 1981. The amount of unknown responses have substantially decreased.

Although some success has been realized, we still have a long way to go. Based on our evaluation it is apparent that a small percentage of the hospitals are responsible for the majority of the reporting problems. In one respect this is good because we can concentrate on them and try to resolve their reporting difficulties. On the other hand, this poses some additional problems. These few hospitals may have special reasons why they are having reporting difficulties which cannot be readily solved. For instance, if a hospital does not normally collect prenatal care information, how do you encourage them to do so? If that particular hospital is responsible for a large percentage of births in a specific county, this could distort the prenatal care statistics for that county. In fact, we have run into this problem for several counties in Pennsylvania because a few large hospitals are not collecting the prenatal care data at all or are not reporting it correctly. Although completion of the birth certificate is mandatory, the medical portion of the certificate has always had large amounts of unknown data. We can only encourage the data providers to report accurately and completely in this section and hope that additional contact with the hospital will improve the data reported.

We are pleased with the improvements so far and feel these improvements are due to the following reasons:

1. A good rapport has been established by the field program with data providers.

2. Health professionals are being trained in correct completion of certificates prior to their participation in the vital registration system.

3. Data providers are becoming more familiar with the new certificates and instructions since the system has been revised.

4. Statistical and quality control units have increased their participation in determining reporting problem areas.

5. With the experience gained in the first four years of the Vital Statistics Cooperative Program, staff can now spend more time resolving reporting problems.

6. Data users are becoming more involved after showing them the results of the analysis.

IV. FUTURE PROBLEM SOLVING ACTIVITIES-BIRTH REGISTRATION SYSTEM

Due to the problems discovered in the analysis of selected birth certificate data items, it was decided to have the field representative spend a larger percentage of his time on improving the completeness and/or quality of the reporting of those items. In the past, the field representative concentrated more effort on the death registration system. Since that system is encountering fewer problems, he will now be able to spend more time on the birth system. A computer generated "report card" has been prepared for each hospital showing the percent of unknown or inaccurate responses for selected data items. These reports can be organized by county, region, etc., and onsite visits scheduled for those hospitals experiencing reporting difficulties. Organizing the reports by region can be beneficial for the hospitals so they can easily see how they compare to other hospitals in their area. Also, if the data providers realize that their reporting status is being closely examined and are given feedback through the field program, they may report more conscientiously in the future. We are fortunate in one respect because most hospitals do report complete data, so onsite visits will happen in only a small percentage of the cases. For those hospitals which will require an onsite visit, their data will be re-examined three months after the visit to determine if the onsite training had any beneficial effect. For those hospitals who have a history of accurate and complete reporting, a

letter will be sent thanking them for their good work, encouraging them to continue, and letting them know that an onsite visit is not necessary at this time.

Another task the field representative is currently working on is the development of a directory of hospital personnel involved in completion of birth and fetal death certificates. This directory would be set up for each hospital and will outline which person to contact for specific items on the certificate. In those cases where the certificate is completed by more than one individual or department, this manual will be especially helpful for querying by local registrars or vital records personnel in New Castle and should help if site visits are necessary.

Additional or revised instruction manuals are being developed for the query unit, local registrars, funeral directors and other individuals involved in the program.

In addition to checking for amounts of unknown responses, the statistical unit is also determining problem areas with regard to the quality of data by preparing special computer validations to discover reporting errors which may slip through the regular querying procedures and standard manual and machine validations done by the vital statistics program. It is not feasible for the query unit to check certificates for types of responses, for instance, whether the response makes sense in relation to other items, due to the large volume of certificates and nature of the system. Also, special validations cannot be built into the system because not enough certificates are affected to make it worthwhile. Therefore, special validations are done by the statistical unit after the full year tape file is created. For instance, certificates were examined for those births where a high number of previous pregnancy terminations were reported. These certificates were examined and it was discovered that some hospitals were reporting the number of weeks in the pregnancy when the termination occurred instead of the actual number of terminations. Several hospitals have made this same type of error and the quality control unit is trying to clear up the confusion that some hospitals have with the pregnancy history data items.

Special validations are run against other items, too, such as length of pregnancy, birthweight, mother's age, and prenatal care. These special validations and examination of certificates with questionable responses have provided valuable information concerning problem areas with regard to the quality of data reported. These computer validations have also determined areas which may need special attention since these kinds of errors are not checked for by the query unit or discovered in the regular monthly validations.

To improve the quality of the prenatal care and birthweight data, we have met with the Division of Maternal and Child Health which is part of the Pennsylvania Department of Health. Since they make extensive use of these two items, they were very interested in the reporting problems we have had in the past three years. This division has offered us their support in trying to resolve these problems and will be contacting those hospitals which are reporting incorrectly. Since the Division of Maternal and Child Health has some influence with the hospitals due to resource allocations, we are hopeful that the reporting problems of certain hospitals will be resolved soon.

V. FUTURE PROBLEM SOLVING ACTIVITIES-DEATH REGISTRATION SYSTEM

Although there are fewer problems with regard to complete data on death certificates than on birth certificates, this does not mean there are no problems with the death registration system. There are two areas which will be receiving special attention in the next few months. These are geographic residence coding for infant deaths and improvement of cause of death specificity.

The major difficulty with residence coding for infant deaths is that infants who die soon after birth and do not leave the hospital often have their residence entered as the municipality where the hospital was located rather than the mother's residence. This happens especially in cases where the mother has been discharged from the hospital and the infant has not been discharged. This can distort infant death rates for particular areas. To try and solve the problem, we are going to match the death certificate with the corresponding birth certificate and determine whether the residence coding on the death certificate is the same as the mother's residence on the birth certificate. We will match only neonatal deaths with their matching birth certificate since it is usually the very young infants whose residence may be recorded incorrectly.

After it has been determined which hospitals are reporting the residence incorrectly, either a site visit will be made or special letter will be written

to the hospital, outlining the problem, how to resolve it, and the importance of accurate reporting for this item.

Another area of special interest is the quality of cause of death reporting in the medical certification section of the death certificate. Until now items in the medical certification section had not been queried because these were mainly the responsibility of physicians, rather than funeral directors, and there was a lack of resources available for initiating a query program comprehensive enough to make an impact. With the implementation and transfer of the online coding direct system to the Division of Vital Records in New Castle in late 1982 and increased computerization, the cause of death query program became more of a possibility. We feel that better data concerning cause of death can be realized with the implementation of a special cause of death query program. As an example, unspecified malignant neoplasm is coded as 199.1 using the ICD-9 coding structure. If we were to query the physician and find out that the primary sight of the neoplasm was the lung, the code would become 162.9. Both responses are acceptable, but the second one is much more complete.

Health statisticians have recognized that one of the essentials of reliable mortality statistics is complete and accurate medical certification of causes of death. Since the underlying cause of death is used for death tabulations, greater attention has been focused on this problem. We must be concerned with how physicians report causes of death as well as what they report.

While there are various ways of approaching the problem of improving medical certification, the most direct method is to write letters to physicians who are inadequately completing medical certification sections of death certificates. Those responsible for cause of death coding in the Division of Vital Records will determine which certificate should be queried since they are more familiar with the medical certification section.

This query program is currently in the planning stages. To test this program, only those certificates where unspecified malignant neoplasm is entered will be queried. In this way, we can discover if the querying has any effect on the quality of the data, work out any problems that are encountered, and decide if the continuation and expansion to other causes is feasible. We hope to test this new querying system later in the year.

VII. SUMMARY

I hope I have given you an idea of how Pennsylvania is utilizing our field representative and quality control and statistical units to improve the quality and and completeness of data reported on certificates. One important fact we have discovered is that it takes the statistical, quality control and query units working closely with the field program to make the system more effective. There are problems the field representative may not be aware of without input from those other three units. Although the field program has only been in existence since 1979, it has had a definite beneficial impact on the quality and completeness of the data reported since 1979. We are very pleased with the rapport that has been established between the field representative and the data providers.

Although it sometimes seems that we find new problems each day, we feel this is because we are paying more attention to what is actually being reported, not just whether it will pass through the computer as a valid response even though it may be incorrect.

The Vital Statistics Cooperative Program is still somewhat new in Pennsylvania.

With the implementation of additional field and quality control activities and the encouragement of other units in the Department of Health, we feel that the problems we currently have will be resolved or drastically reduced.

We welcome any ideas or comments you may have and hope we have given you some ideas which may be useful for your program.

Thank you.

A TWO-PART BIRTH CERTIFICATE FOR MAINE

Ellen M. Naor, Marian B. Perkins, Carl S. Gerry; Maine Department of Human Services

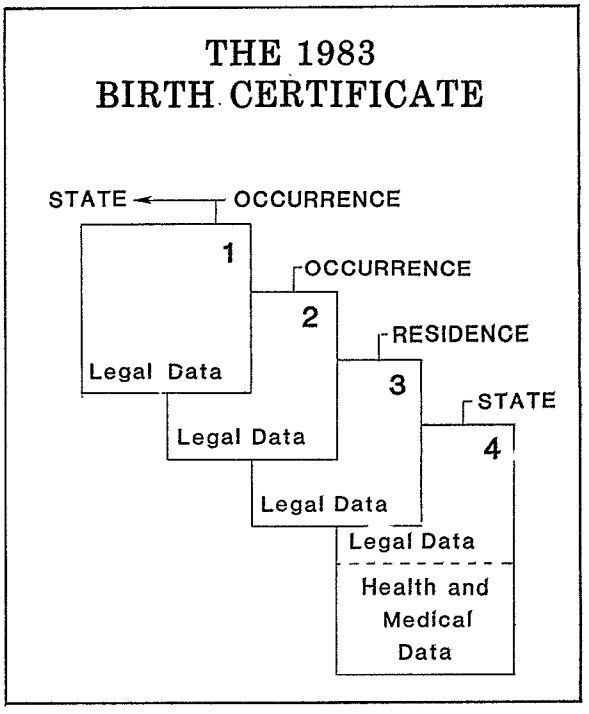
Introduction

Over the last several years, concerns have been voiced by various groups in Maine about the quality of the statistical data obtained via the birth certificate. Areas of inaccurate and/or incomplete reporting were identified as were various processing problems. There was also some concern about the confidentiality of the health and medical information reported on the birth certificate. This stemmed from the statutory requirement that birth certificates be signed by the local registrar in the town of birth before being forwarded to the central state repository. The local registrars in Maine are city or town clerks. The majority of town clerks work for small towns with populations of 5,000 or less and many work out of their homes. As long as the Maine birth certificate copied the U.S. Standard Certificate in having the health and medical information on the original (first) copy of the certificate, this sensitive information was necessarily routed to the local town clerk together with the legal information certifying the fact of birth. Many in the medical community in Maine felt this to be an improper disclosure of confidential medical information, with potential for misuse.

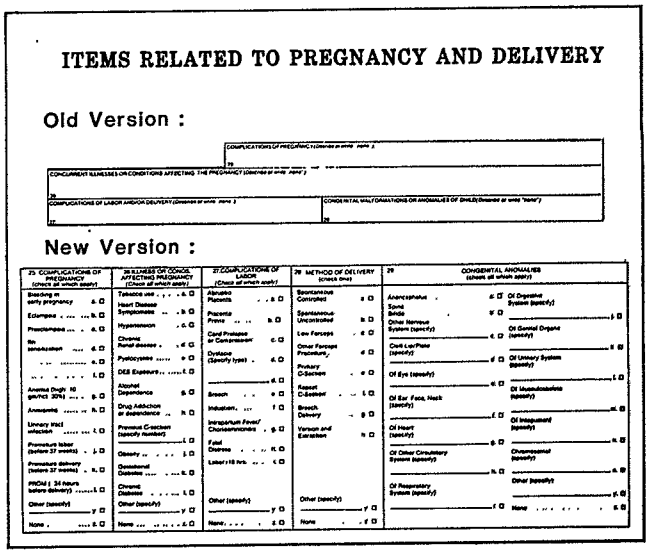
The dual problems of data quality and confidentiality protections were attacked by modifying both the certificate itself and the procedures and systems for handling it. Redesign of the certificate physically separated the legal and medical data, allowing the medical data to be sent directly to the Department. The revision incorporated several features designed to increase the accuracy and completeness with which information is recorded and reported. In addition, extensive changes were made in coding and processing, and the computer system was redesigned to reduce agency errors, correct problems in updating the statistical data base, and preserve the integrity of historical tapes. This paper will describe the key features of these changes and outline the resources required for implementation.

The Certificate

The physical form of the new certificate is the key to protecting the confidentiality of the health and medical information it contains. The "two part certificate" has, in fact, four pages, separated by carbon paper. The first chart illustrates the distribution of the various component parts. The top three pages are identical copies of the legal portion of the official State of Maine Certificate of Live Birth, containing the information provided on a certified copy of the record. The top copy is forwarded to the State Office of Vital Records after signature by the local registrar in the town of birth. The second and third copies are filed by the registrars at the place of birth and the place of residence, respectively. The fourth page of the certificate was added this year. It contains both the legal portion of the record (the carbon accounts for this) and the confidential information reported for medical and health use. The fourth page is sent by the hospital directly to the State Office of Vital Records, thus promoting the privacy of this information by restricting its travels. The fourth page has no legal status since it does not bear the signature of the local registrar; it is used only for establishing the statistical record containing both demographic and medical data. A preprinted number on all pages permits linkage of the forms received by the State Office.



A checklist format was adopted for a number of medical items related to pregnancy and delivery. These items had previously appeared on



the certificate as blank spaces to be filled in by the physician, following the format of the model U.S. Standard Certificate. We found, however, that the spaces were all too frequently left blank, whether or not they applied. In addition, responses were often in the wrong place, requiring the coder to read each entry carefully and determine to which item it belonged - a time consuming judgement prone to error unless the coder was highly trained. A checklist of common complications and illnesses was developed to serve as both a convenience and a reminder to the physician filling out the certificate. A specific item was added to identify the method of delivery. The format for congenital anomalies, too, was revised to encourage more complete reporting. The checklist format for these items is used with variations by several other states and appears to offer great promise for increased reporting.

The System-Clerical Operations and Computer Processing

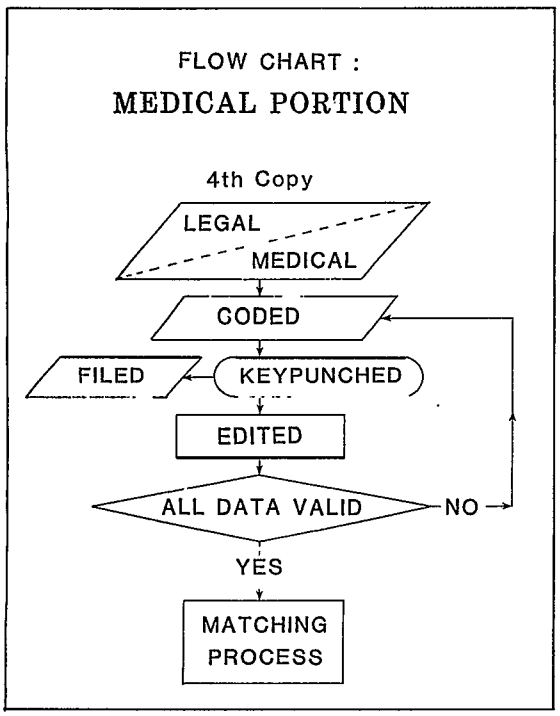
The system for handling birth certificates was completely redesigned to make maximum use of available computer facilities to speed the processing of incoming records, facilitate their storage and retrieval, maintain the privacy of health and medical information, and promote the accuracy of the statistical information contained on the certificate. Conversion to an on-line system was considered but had to be rejected due to lack of time and funds for the needed developmental work. The system which finally evolved is based in part on that developed by the Connecticut Bureau of Health Statistics in 1978 and 1979.

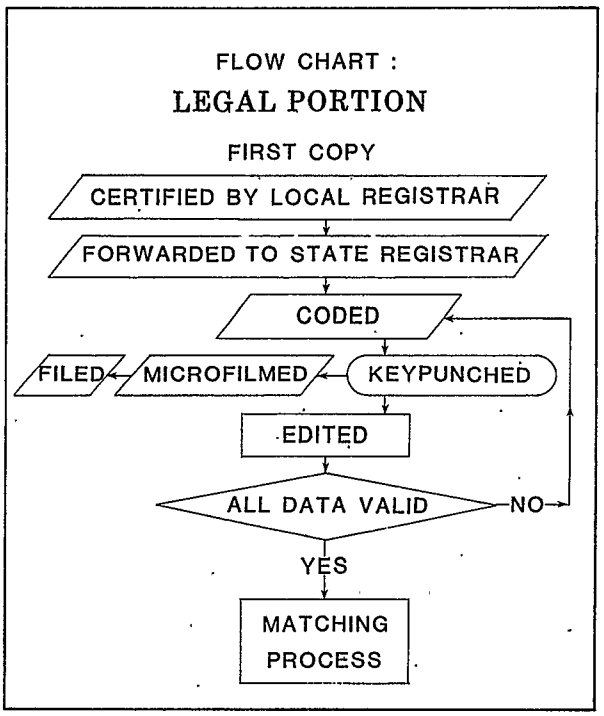
The computer record for each birth in the state is initiated by the arrival of the medical form in the Office of Vital Records. The form is used as a coding worksheet since it has no legal status. The major items coded are the places of residence and occurrence. An alphabetic geocode was developed to speed this process, using a ten-position code to uniquely identify each town and unorganized territory in Maine. Self-coding and checklist formats are used for most of the medical items, so little additional coding is needed although the entire form is reviewed for completeness and consistency. We require that a response be supplied for every item. Medical forms on which items are left blank are returned to the hospital for completion.

The medical forms are batched weekly for keyplex entry. The batch number serves as the key for retrieving medical forms as needed for error correction and verification, quality control, and research projects. All information on the form is keyed, including the name, address, birthdate, and other data from the "legal portion" of the medical form. Input of this form creates the computer record for each birth and supplies all data elements it will contain except for the official certificate number.

In order to improve the quality of the statistical/medical information, stringent edit procedures were developed as part of the load program. Records containing errors identified by the edit program are not accepted. The weekly error listing is returned to the birth records coder for correction, thus providing feedback on coder performance. Corrected records are resubmitted with a subsequent weekly batch.

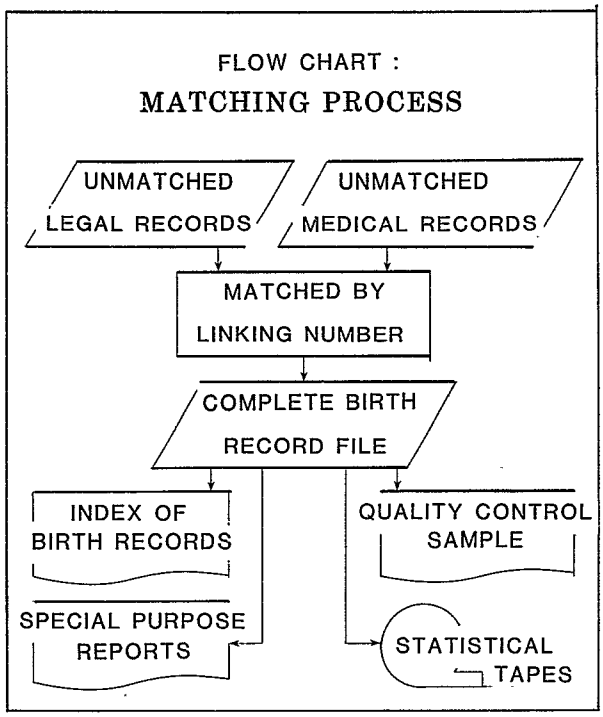
The legal form for each birth constitutes the official birth certificate. Upon signature by the local registrar in the town of birth, these records are sent to the State Office of Vital Records in monthly batches. After review for completeness, the records are sorted and held by county; new forms are interfiled as they





are received. At the end of the month, a sequential file number is assigned using a numbering machine - this is the official birth certificate number. Selected items are then keyplexed: linking number, name, birthdate, and certificate number. The legal forms are microfilmed and booked.

A key conceptual feature in making this new certificate and system consistent with the historical data series is the reliance on the legal form as the basis for the permanent computer record. Although it is the medical form which creates the initial computer record for each birth, it is linkage with the legal form containing the offical certificate number which triggers release from the interim file and entry into the final, permanent computer file. The final record is thus comparable with those of earlier years in which both legal and medical data were entered from a single form.



A computerized merge/match procedure is implemented each month after the legal certificates are keyed. In brief, this consists of the following sequence of operations:

(a) Medical and legal files are matched on linking number, infant's name, and birthdate.
(b) Matching records are consolidated into a single record.
(c) The official certificate number (carried on the legal record) is assigned to the combined record.
(d) Non-matches (approximately 1%) are held over for 2-3 months and eventually researched clerically.

A monthly file is thus created of complete, edited birth records. The records contain both legal and medical information and use the official birth certificate number as the record identifier.

A dilemma we struggled with early on was the potential for piling up enormous amounts of paper - two sheets, in fact, for each birth in the state instead of only one. As we analyzed our objectives for the new system, however, we found we had no need to retain the medical forms indefinitely since all of the material they contain is now on the corresponding computer record. There is one exception to this - those records containing a "write-in" for an "other-specify" response (approximately 2%) will be kept on file for future special studies.

Several types of outputs are produced by the system to meet the needs of various health and statistical agencies. Most of these will be familiar to vital and health statisticians. First, to meet the immediate needs of the Office of Vital Records to retrieve certificates efficiently, indexes of birth records are produced monthly, semi-annually, and annually. Secondly, a sample of records is selected for use by the quality control program of the National Center for Health Statistics (NCHS) as well as for use at the state level to monitor the level and type of errors which persist in the final file. The corresponding medical forms are located using the weekly batch numbers and copied for comparison with the computer record and transmittal to NCHS. Thirdly, a monthly statistical tape is created and a copy sent to NCHS as part of the Vital Statistics Cooperative Program. Aggregated on an annual basis, the statistical tapes constitute the major resource for statistical analyses of every description. Finally, a number of special monthly reports and listings are produced for use by public health programs within the state, such as immunization, newborn services for home deliveries, and high risk infant services. Our previous practice was to provide these programs with copies of the birth certificates. Now, computer reports are tailored for each user to provide only the specific data elements required by each.

Implementation and Resource Requirements

There were a number of phases and milestones in the process of revising the birth certificate and the associated systems work. Key features will be reviewed briefly here.

IMPLEMENTATION

	1982												1983	
PHASES	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F
PLANNING	X	X	X	X	X	X	X	X	X	X										
FORMS DESIGN							X	X	X	X	X									
FIELD WORK													X	X	X	X	X	X		
SYSTEMS DESIGN											X	X	X	X	X	X				
SYSTEMS DEVELOPMENT														X	X	X	X	X	X	X
EVALUATION OF CLERICAL TASKS												X	X	X	X	X	X	X	X	X
DRESS REHEARSAL																		X	X	
IMPLEMENTATION W/LIVE DATA																			X	X

Overall planning and design work for the new system stretched over more than a year. Initially, the focus was on the certificate itself and on the flow of forms from the time they were completed by the hospital until they arrived in the central State Office of Vital Records. The initial planning group included individuals from the agencies and organizations which had expressed dissatisfaction with the then current certificate. Physicians, epidemiologists, the Child Health Program Manager, and a researcher participated in this process together with the State Registrar of Vital Records and administrators of the statistical unit. This group developed a prototype certificate and proposed modifications in the handling procedures which would correct the identified problems. The planning group was subsequently expanded to include the Bureau's systems analyst, and the areas in which substantial systems modifications would be needed were identified.

Developmental activities in the central Office of Vital Records focused on intensive field work which preceded the introduction of the new form and facilitated a smooth transition. The education programs, scheduled over the last quarter of 1982, were tailored for specific audiences. In addition, the clerical operations in the office and the forms flow between the Vital Records Office and the keyplex and microfilming units were thoroughly evaluated. Procedures were worked out for selecting the quality control sample, assembling the relevant medical and legal forms, and transferring the official certificate number to the sample records. Needed new equipment was ordered and extra clerical help was arranged for January and February of this year, when 1982 records would be coming in on the old forms and 1983 records, on the new forms. In December, 1982, and January, 1983, dress rehearsals identified problems in handling or processing the records which we were able to correct before the new certificate went into effect.

Systems work-analysis, design, and programming - was started about six months before the scheduled implementation date of January 1, 1983. This phase included the preparation of new coding and keying instructions, procedures and programs for alpha-input geocodes, preparation of edit programs and new correction/update procedures and programs, development of the programs for matching legal and medical forms, sample selection, and preparation of special purpose listings.

The total developmental, or "front-end", costs were moderate, Personnel costs were the largest component: a total of one annual full-time equivalent was invested on the project. This included seven person-weeks by administrative personnel, 30 by systems staff, and 15 by the vital records staff. Additional identifiable costs were $3,500 for printing new manuals, travel costs and computer time. The overall increase in running costs due to the introduction of the revised certificate is estimated to be approximately $3,500 per year. It appears that the extra clerical time required for researching nonmatched forms and batching and handling medical forms, etc., will be balanced by the increased efficiencies of the new system.

Evaluation

The new certificate appears to be a satisfactory approach to the identified problems of data quality and confidentiality. Initial evaluation, based on the first six months of 1983, suggests that there has been significant improvement in the accuracy and completeness of reporting.

Completeness of reporting, considered in terms of item completeness, has improved, as shown in the accompanying chart. The frequency of omitted information has dropped noticably since introduction of the new form on problem items such as education, prenatal care, and complications. Overall completeness of reporting will be documented through analysis of late-filed records over the next five years.

PERCENT OMISSIONS BY ITEM
1978-82: 5% SAMPLE
1983 : 100% COUNT

ITEM	YEAR					
	1978	79	80	81	82	83*
EDUCATION	0	1.1	1.1	2.0	0.9	0.0
DATE LMP	1.2	1.1	0.6	1.4	1.3	0.3
PRENATAL CARE	1.8	1.7	0.6	2.3	0.9	0.5
BIRTHWEIGHT	0	0	0	0	0.1	0.3
APGAR SCORE	0	2.9	1.6	1.8	1.2	0.6
COMPLICATIONS, ETC.	5.1	3.5	4.2	1.4	1.5	0.9

*FIRST SIX MONTHS

Several approaches are being used to assess the accuracy of information which is reported on the birth certificate. Data for the first six months of use yield preliminary insight into its performance. Problem indicator rates were calculated for items for which the checklist format

INCIDENCE OF POSITIVE FINDINGS

ITEM	1978	79	80	81	82	83*
ITEMS WITH FORMAT CHANGES						
CONGENITAL ANOMALIES (%)	1.9	1.5	1.4	1.2	1.2	2.1
COMPLICATIONS OF PREGNANCY (%)	7.7	9.2	8.8	10.5	11.4	19.4
CONCURRENT ILLNESSES (%)	3.3	4.1	4.4	5.3	5.1	22.5
COMPLICATIONS OF LABOR (%)	18.0	21.7	22.5	23.6	24.6	22.2
ITEMS WITHOUT FORMAT CHANGE						
EDUCATION OF MOTHER (% <H.S.)	21.3	19.9	19.4	18.3	17.3	17.1
MONTH PRENATAL CARE BEGAN (% 2ND & 3RD TRIMESTER)	20.7	18.9	18.3	16.9	17.1	15.8
# PRENATAL VISITS (% <10)	34.6	32.6	33.4	29.3	26.4	26.7
LOW BIRTHWEIGHT (% <2501 G)	5.3	5.3	5.1	5.5	5.1	5.5

*FIRST SIX MONTHS

was introduced on the new certificate and those for which no changes were made. There was a startling increase in the frequency of positive findings reported on the new certificate for the checklist items. For example, congenital anomalies increased from 1.2 per 100 live births in 1981 and 1982 to 2.9 per 100 in 1983. Complications of pregnancy increased from 11.4 per 100 live births in 1982 to 19.4 in 1983. The reporting of concurrent illnesses increased most dramatically, from approximately 5 per 100 live births in 1978-1982 to 22.5 per 100 in 1983. Those problem indicators for which no format changes were made, e.g., low education level, month prenatal care began, number of prenatal visits, and low birthweight, did not vary in frequency after introduction of the new certificate. Although these data are based on a short time period, we believe they are a reflection of the success of the checklist format in promoting better reporting and more accurate coding and processing.

The regular quality control sampling program provides an indication of the faithfulness of the computer record in reproducing the information reported on the certificate. The overall rate of errors which persisted in the final statistical file during the period 1978-1982 ranged from 4-15%. In the first six months of 1983, the edit program identified 4.4 erroneous records per 100. Since all of these errors are being corrected, we expect that the error rate on the final file will be substantially lower than in previous years. This improvement reflects the success of the edit programs, improved coding quality due to checklist formats, and more effective feedback to coders.

In conclusion then, we believe that we are well on the way to accomplishing the purposes of the revision. Better protection of confidential information - a major objective - has been achieved. Mailing the medical information directly from the hospital to the State Office safeguards against access by unauthorized individuals. Similarly, the preparation of specialized output reports for various health programs restricts access to confidential information. Specific items of information are provided only to those with a need for the item and the authority to have such information. The experience of these first six months is optimistic, suggesting that the revised format and associated systems changes are producing significant improvements in the quality of data obtained from Maine birth certificates.

Using Data to Estimate or Reduce Costs

Session P

MODERATING HEALTH COSTS USING ANALYSIS OF INDIVIDUAL PHYSICIAN PRACTICE PATTERNS

William M. Burke, M.D., Barbara Ladon, H. Williams Wisotzkey
HealthPro, Inc., Worcester, Massachusetts

Abstract

The total national health expenditure for health services is the result of the interaction of a number of complex factors. Since 1973, Professional Standards Review Organizations have been charged with the task of assuring the appropriate utilization of hospital services for Medicare and Medicaid participants with the dual objectives of cost and quality control. A variety of methods have been attempted with mixed results. During the last 18 months, one PSRO has had success with a method which incorporates several approaches: Diagnosis Related Groups (DRGs) based casemix analysis; physician-specific review; and centralized (non-delegated) retrospective peer review. Results have been assessed for both utilization and quality measures.

Introduction

In her excellent presentation this morning at the second plenary session, Dorothy Rice reviewed some of the complex factors resulting in the increase of health services expenditures since the mid 1960's.

National health expenditures were 4.4% of the gross national product in 1950. By 1982, they had risen to 10.5%. The national health expenditure was $12.7 billion in 1950, $26.9 billion in 1960, $247.4 billion in 1980, and by 1982 had risen to $322 billion. The amount of disposable income Americans spent on health insurance premiums more than quadrupled between 1950 and 1980.

On July 1, 1965, the implementation of Titles XVIII and XIX of the Social Security Act introduced a major change in the demand factors for health care in the United States. The rapid increase in health care costs, which followed, has focused attention on the utilization of health care services, especially hospital care as the single largest item in the Medicare budget. It is, therefore, not surprising that a major effort has been directed at assessing the appropriate use of hospital care.

In response to these pressures, in the 1972 amendments to the Social Security Act, Senator Bennett of Utah sponsored a section which created a series of peer review organizations called Professional Standards Review Organizations (PSRO). PSROs were brought into existence to help control costs and assure the quality of health services purchased by the federal government. The peer review organization originally established as a PSRO to serve the central Massachusetts area is HealthPro, Inc. The service area of HealthPro is roughly the middle third of the state of Massachusetts, approximating the borders of Worcester county. There are 15 acute care hospitals in the area with approximately 33,000 Medicare discharges annually. This report concerns a new and successful review program HealthPro has implemented to address appropriate utilization of hospital services.

HealthPro was established in 1974 as a non-profit physician organization involved in utilization review and quality assurance programs. Its mission has been to insure that quality health care is provided in the community it serves and that health services are medically necessary, meet professional recognized standards of quality and are appropriately utilized at the most economical level consistent with professional standards.

The organization collects discharge data on each case under review, over 350,000 cases since its inception. The discharge hospital abstract includes the Uniform Hospital Discharge Data Set (UHDDS) and selected other variables which describe demographic and health characteristics of the population. It has utilized the DRG approach to analysis since 1976, as an effective mechanism to assess health care utilization for a large population. HealthPro was one of the first PSROs to work on DRGs with Yale University. The HealthPro database was part of the early tests of DRGs in a utilization process.

HealthPro, using DRGs, has developed a set of physician-specific profiles which are casemix adjusted. The use of this discharge data by areawide peer review committees, provides an epidemiological approach to identifying groups of physicians who are high utilizers of hospital services.

When it began review programs as a PSRO in 1976, HealthPro, as required by law, initiated a plan of 100% concurrent review. In 1979, a system of focused review was introduced which consisted of concurrent review of a randomly selected sample of cases. This system did not provide an adequate mechanism to focus on problem areas. In 1981, a physician-based focused review system was initiated. Physician-specific concurrent review continued. It was soon replaced by a centralized retrospective review system. The primary factor leading to this change was the realization that concurrent review was not effecting admissions (as opposed to length of stay). National data has indicated that Medicare patients have had a higher adjusted admission rate in the central Massachusetts area than in both the state of Massachusetts and the rest of the country. Although an excess of 20% of the admissions reviewed did not meet acute care criteria at

the time of admission, less than 1% were denied for payment as being "not medically necessary" under the concurrent review program.

It is the results of this physician- specific retrospective review program that will be reported on today. The data below deals only with patients whose care was covered under the Medicare program in central Massachusetts from January 1, 1982 - June 30, 1983.

Methods

Data are collected from the medical records of each Medicare patient discharged from a hospital in the central Massachusetts area. The data collection instrument (abstract) used in 1982 included 57 distinct data elements. There are 11 elements that are used to develop the screen against which physicians' practice patterns are compared to identify "high utilizing" individual physicians. The particular items selected (Table 1) are significant in that they exemplify the ability to devise an effective screening program from a small group of data elements.

TABLE 1

Data Elements on Abstract Necessary to Construct HealthPro Review Screens

1. Hospital Number
2. Admission Date
3. Admission Status
4. Attending Physician
5. Admission Certification
6. Surgeon
7. Discharge Date
8. Disposition
9. Discharge Diagnoses
10. Procedures and Dates
11. Number Consults

This report concerns the results of the initial phase of the retrospective review process. During this period, all Medicare cases at the 15 area hospitals had an initial review. It was originally anticipated that the initial review and subsequent re-reviews could be completed within a 12 month period of time; unforeseen logistical problems intervened and the first complete review/re-review cycle took 18 months to complete.

The review program consists of a 20% sample of discharges across the 15 hospitals in the area. For scheduling reasons, no fewer than 125 and not more than 600 cases were reviewed in any one hospital for one review cycle. Of the cases reviewed, 20% were chosen by random sample. The rest were physician- specific. The 20% random sample was elected to provide an estimate of areawide patterns and to compensate for the loss of concurrent review data collection capability. See below for a description of screening criteria.

The actual process begins with a nurse review-coordinator reviewing the medical records for the selected physicians against a Severity of Illness/Intensity of Service (SI/IS) criteria. The SI/IS criteria set was initially developed by INTERQUAL, a national consulting organization, and then locally modified by central Massachusetts physician review panels. Medical records that "meet criteria" in the review-coordinator audit were not subject to further review. Any cases that did not meet criteria were reviewed by a physician in the traditional peer review fashion. Cases that the physician found questionable were reviewed with another physician. The two (2) physician panel also reviewed the pattern of individual physician's practices. Their assessment was brought to an areawide physician panel of nineteen (19) physicians which had at least one representative from each of the 15 hospitals in the HealthPro area. Each physician whose cases were reviewed was notified of the review and the results. The quality assurance committee and hospital administrator at each hospital was also informed of the results.

The panel choose action steps based on the initial review, as follows:

1. Letter of recommendation for exemplary behavior;
2. Notice of cases not meeting criteria (minor problem);
3. Notice of cases not meeting criteria; the type(s) of problems found (e.g. a pattern of inappropriate admissions; lack of discharge planing or lack of follow-up of abnormal laboratory results); and a second review in 6 months (moderate utilization problems);
4. Notice of cases not meeting criteria; the types of problems (see above); and of a second review in 3 months (major utilization problems and quality care problems).

Personal meetings; corrective action plans including such things as mandatory second opinions; pre-admission review or mandatory consultation; and retrospective denial of payment are the type of actions taken when required after a second (or third) monitoring.

During the initial review cycle, 7 specific screens were devised to select physicians for further review. The 7 specific screens developed were:

1. Severity of Illness/Intensity of Service (SI/IS)

The first screen was developed using the percentage of the area's 1981 caseload that did not meet either SI or IS criteria at initial review. Physicians were selected who had greater than or equal to 25% of their cases not meeting SI or IS criteria at initial review. Because there was not a concurrent review program in 1982, this screen was not used after December 31, 1982.

2. Short Stay Discharges

An assumption was made that short stay medical admissions could potentially be "not medically necessary". The 1981 areawide experience identified that 14% of all cases (excluding maternity and deaths) were discharged from area hospitals in less than or equal to 3 days without having had any procedure performed. Physicians who had 20% or more of their cases discharged in 3 days or less with no procedure performed were selected for review.

3. Expected Bed Days

HealthPro has selected diagnostic categories in which a hospital or a physician's practice pattern differs from what would have been expected after analyzing the community's experience. Discharge data are analyzed by Diagnosis Related Groups to determine the normative use of bed days areawide. The caseload of each physician and hospital are then compared to the areas (by DRG) to identify providers whose use of bed days exceeded the norm.

Based on the results of the analysis, four diagnostic areas were chosen as targets for review. Any physician whose experience in these four diagnostic categories exceeded "the expected" by ten percent (10%) was reviewed. Similarly, the two diagnostic categories accounting for the highest number of days used over expected were identified by hospital. The physicians whose practice patterns caused the excessive use of days were reviewed.

Finally, each physician's practice pattern is casemix adjusted, using DRGs, and compared to the community. Any physician whose use of bed days exceeded the norm for similar patients by 10% was selected for review.

There are parameters for minimum caseload established for each group.

Nursing Home Discharges

Based on area norms, this screen identifies physicians who had 40% or more of their nursing home patients discharged in 10 days or less (excluding deaths) to nursing homes without having had any procedures performed.

5. Pre-operative Length of Stay

HealthPro's existing policy is that a one day pre-operative stay is appropriate for elective procedures. Physicians whose practice pattern differs from this policy by having a pre-operative stay greater than 1.5 days per case was identified for review. Cases included for this analysis were those in which the attending physician is also the operating physician.

6. Consults

The consult screen identifies physicians who ordered two (2) or more consults in 3% or more of their discharges. This screen is based on the norm that 97% of area discharges have less than two (2) consultations.

7. Physician Advisor Referrals

Physicians were identified who had 20% or more of their cases referred to a Physician Advisor during the hospital stay in 1981. Because there was not a concurrent review program in 1982, this screen was not in use after December 31, 1982.

Results

There were approximately 750 physicians who regularly admit Medicare patients. Approximately 25% of the physicians were reviewed because they met the screening criteria (see above). Additional physicians met the screens but were not reviewed due to resource limitations. The physicians who were reviewed clustered into 3 categories. 22.8% had mild or no problems and were not subjected to further review. 48.9% had moderate problems and were apprised of this in a letter stating the areas which presented the problems. These physicians were thought to have severe problems; they were notified of the problems and scheduled for re-review in 3 months.

TABLE 2
Physician-Specific Results of Initial Monitoring

	No.	%
Mild or None	39	22.8
Moderate	92	48.9
Severe	57	30.3
TOTAL	188	100.0

Out of 33,000 Medicare discharges in 1982, 2,302 cases were selected for initial retrospective review. After review, HealthPro was in disagreement over the medical necessity of the admission in 318 (13.8%) of the cases.

Of 27,291 days of care reviewed, HealthPro disagreed with 5,964 (21.9%) days. Medical care practice issues, including such things as inappropriate use of ancillary services, lack of follow-up of laboratory tests, and inappropriate management quality of care were found in a somewhat surprising 6.7% of cases.

In both the moderate and severe groups, some physicians were not included in the follow-up study due to retirement, moving out of the area, or death. As a result, 141 physicians of the original 149 were re-reviewed.

Physicians with moderate problems demonstrated a marked change in their patterns of care at the 6 month re-review. The admission disagreement rate was halved to 4.9%. Days in disagreement were also dramatically reduced from approximately 19% to 8%. There was a similar change in disagreement rate with medical practice issues. Eighty-seven (87) of the original 92 physicians with moderate problems were caring for Medicare patients at the time of the 6 month re-review (Table 3).

TABLE 3

Disagreement Statistics for 87 Physicians with Moderate Problems

		Initial Review	6 Month Review
Admission	No.	116	37
	%	10.0	4.9
Days of Care	No.	2634	719
	%	18.8	7.9
Medical Practice	No.	34	8
	%	3.0	1.1

Physicians with more severe problems were re-reviewed after 3 months. There were 54 physicians who were subjects of the 3 month re-review (Table 4).

TABLE 4

Disagreement Statistics for 54 Physicians with Severe Problems

		Initial Review	3 Month Review
Admissions	No.	195	48
	%	24.6	12.8
Days of Care	No.	3381	642
	%	34.2	19.6
Medical Practice	No.	116	45
	%	14.6	12.0

As was the case with physicians who had moderate problems, the physicians with severe problems showed some change of behavior. The pattern of change, however, is different. The physicians with severe problems halved their admission disagreement rate but even their improved rate was higher than the original rate noted for physicians with moderate problems. Days in disagreement were not as dramatically reduced as in the instance of physicians with moderate problems. There was essentially no change in the rate of disagreement concerning medical practice issues.

Discussion

A striking and unexpected finding of this study is the strong inverse relationship between a high inappropriate utilization of the acute hospital and the quality of medical care provided. Physicians with the most severe utilization problems also accounted for the vast majority of the medical practice problems. Physicians identified having utilization problems were able to modify their utilization behavior to some degree after the rather mild intervention of simple notification. The same was not true for medical practice issues.

It is the current practice of HealthPro to meet personally with physicians with severe problems in a peer review educational session to resolve medical practice issues. If this physician to physician discussion is not effective in modifying disagreement rates, a meeting between the physician, HealthPro and the quality assurance committee of the hospital is requested. When such meetings are held, their goal is to develop a specific corrective action plan for the physician in question. It has been hoped that this intense intervention would result in more gratifying behavior changes in the area of medical practice issues for those physicians with the most severe problems.

Based on the results of the initial monitoring cycle, those few physicians with serious problems have not improved to the point that their practice patterns fall within the area norm. Intense monitoring and interventions are being undertaken in an attempt to modify this behavior.

The retrospective review process described served as a cost-effective mechanism for screening large volumes of data to identify physicians with aberrant practice patterns. The great majority of the problems were either minor or were rectified by the physicians (and hospitals) within a short time after being notified of their existence by the external review organization. What remains is the

difficult task of behavior change in the small number of residual problems.

In addition to identifying utilization problems, the review system has been an effective strategic method in changing utilization behavior. The management information system and use of DRG casemix adjusted hospital discharge data proved to be an efficient screening tool and made it possible for HealthPro to carry out an effective program.

IMPROVING CERTIFICATE OF NEED DECISIONS THROUGH BETTER DATA USE AND ANALYSIS

James L. Dallas, State of Washington

The conclusion of this paper is that the decisions made on Certificate of Need applications in the state of Washington over the past few years have been improved by the availability, better use, and analysis of data. The paper explains how this occurred -- how decisions were made previously, what additional types of data became available, how data was analyzed, what difference this made in terms of final outcome, and what limitations still exist.

How Decisions Were Made Previously

The process of arriving at a decision through the Certificate of Need process in this country has been characterized by some to be far from predictable. David Porter, in a recent publication on hospital architecture, spoke of the CN process in this way, "Many biases will be tried and tested and politics will tend to enter in." [1]

The act of deciding whether to approve, reject, or modify a Certificate of Need application in Washington State was and still is partly a matter of professional judgement rather than entirely being a methodical process of applying uniform criteria to reach a conclusion. Over the course of the program's history, however, the decision-making process has become increasingly uniform and predictable. It has been an evolutionary process. When Washington State began its CN program in 1971, there were very limited guidelines or criteria for reviewing applications. The quality of staff analysis depended greatly on the experience of the particular analyst.

In the absence of detailed policy and guidelines, analysts were inappropriately thrust into the combined role of planner and regulator. The constricted time frame of a CN review (90-120 days) did not generally allow adequate time to resolve major policy issues. CN decision-making in these early years (1971-1974) was based heavily on how persuasively an application was written, the reputation of the applicant, and how forcefully a case for approval was presented.

Staff found objective criteria and used them whenever possible. Since the criteria were not established in regulation or as policy in a state health plan, however, staff relied largely on precedent to achieve some degree of consistency. Moreover, criteria could be applied differently because of the wide variation in the circumstances surrounding individual projects.

Health data contained in these early applications focused primarily on the applicant; little, if any, information related to the impact on the overall health system. Information provided in the application was often non-specific. The depth and quality of information varied from application to application. Comparisons between applications offering similar services were difficult.

Arguments and issues developed over basic data instead of substantive issues. For example, debate centered over such a basic question as what population projection method should be used. Many applicants presented historical data for their institution only, without paying attention to changes in the community. Staff did the best they could at a time when there were essentially few decision rules.

What Types of Additional Data Became Available?

Staff at the state level and in local planning agencies over time began to build a common data base and to work out agreements on how this information would be used to make decisions. In 1975 the Board of Directors of the Puget Sound Planning Council adopted a hospital development guide. [2] The Washington State Health Coordinating Council adopted a State Hospital Bed Projection methodology in 1979 [3] and in 1980, Washington's first state health plan. This marked a turning point. Although other criteria are used in CN, the existence of these documents proved significant.

The state hospital bed projection method proved to be especially useful. Providers, regulatory agencies, and HSAs all had previously worked many months and agreed upon a step-by-step methodology for determining planning area bed needs. A single population projection method was agreed upon, areawide occupancy rates were set, and criteria and standards were established that defined how the methodology was to be used. The methodology not only provided for baseline bed projection of hospital patient days, it also set out a method of adjustments and outlined the documentation necessary in order for the state and other review agencies to accept adjustments.

The importance of the state health plan was that it contained the standards and methods that were necessary to apply criteria specified for CN review in state law. While CN criteria had grown in number from 1971-1981, a common set of decision rules for applying them had been lacking until development of the state health plan and its accompanying bed projection methodology.

While having better planning and the information to support it is an important step to improved CN decisions, there must be capable individuals to analyze the applicability of plan documents to individual applications. Fortunately, Washington State had both. Using these

new sources of information, staff began to apply them on a system and community basis to compare hospitals with one another. The artificial and, at times, stifling barriers between the planning staff and the CN staff were removed and they began to work together on the review of CN applications. While it was appropriate to have the individuals setting CN criteria be different from those applying them to an application, close communication is nevertheless essential if the criteria are to be mutually understood and applied as intended by policy makers.

In addition to the hospital bed projection methodology being developed, a patient origin study was conducted in 1980. An earlier study was done in 1977, but now with the bed methodology in place, the 1980 study was of far greater value in CN decision-making.

The state also collects hospital utilization data through a monthly (HUR) report. The HUR predates the bed methodology and is illustrative of earlier attempts to improve the hospital data base. When used in tandem the HUR and patient origin study can be used to assess the impact hospitals are having on each other and can be used to analyze a hospital's CN request to determine if, in fact, they can attract new patients.

An array of information sources are used by CN analysis in developing a staff recommendation. The continuum shown in Figure 1 illustrates the relative usefulness of these information sources.

Assisted by this expanded scope of information, and newly-developed analytical tools to better use existing information, the nature of CN decision-making changed. It became more orderly, consistent, and methodical. Professional judgments are still necessary, but the choices for those making decisions are clearer and the process for those submitting applications is more predictable. The changes that took place both in CN data and in decision-making are shown in Figure 2.

How has data improved CN decisions? What difference has it made in final outcomes?

CN decision making has improved in several ways because of better data use and analysis. It is now possible to compare projects with one another because a common data base is generally accepted by all participants in the process.

With agreement by all providers and regulators on a single bed projection method, coupled with the availability of a common data base in the form of a patient origin study (regularly updated) and an improved hospital utilization report, CN decision-making is more consistent. Applicants know in advance what information to submit. By making all involved in a CN decision aware of the same information, extraneous arguments over what numbers to use are minimized. Issues and arguments can now be focused on major policy issues such as, "Which tertiary services, if any, should be performed outside major urban areas in Washington State?"

With a common data base, applications can be compared with one another. Batching of applications and concurrent reviews are now feasible and appropriate for some services.

As an example of how improved information had a direct impact on a project, an applicant requested a multi-million dollar addition to an existing hospital. Part of the documentation for the increase included assumptions about market share increases. Using state health plan criteria, the bed methodology adjustment process, and factoring in patient days which had already been spoken for in the prior approval of other hospital projects, staff was able to independently project market share. They concluded that the hospital would not increase its market share for certain services. The compromise eventually worked out was based on the analysis staff developed. The result was a major reduction in the cost and size of the project.

The dollar impact of improved data is not easily quantified. The usual statistics presented to show the impact of CN speak only to denials and withdrawals.

Washington State has never prided itself on its denial rate, preferring instead to work with an applicant to modify a proposal to bring it in line with appropriate standards and recommendations of review agencies.

Table 1 illustrates CN program activity during the three most recent fiscal years, 1981-1983. The combined total of denials, withdrawals, and reductions has increased over time when viewed as a percentage of total expenditures received. In FY '81, the total was 16%, in FY '82 - 18% and in FY '83 - 19%.

What Limitations Still Exist?

Despite improved CN decisions caused by better use of data and analysis, there are certain limitations to the information currently available. The process has not reached the point where many, other than staff and applicants, understand the full significance of data generated as a result of the CN process. An HSA Board of Directors, state officials who have many other responsibilities besides CN, and an appeals hearing officer with little or no background in health services can still sometimes be swayed more by persuasive emotional appeals than by forecasts and other data. There are very few "hard" numbers (e.g., number of staff and patient mix); most information is "soft" (e.g., projections of market share).

Among the options being developed to improve the process are the following:

1. An adjustment process outside of CN where market share analysis and adjustments are completed prior to applications being submitted. The CN program would then be responsible for allocation based on the agreements already reached through the planning process. Under this process, the state would be responsible for determining

planning area limits and local planning agencies would be responsible for allocation within planning areas.

2. Greater use of batching/concurrent review to compare similar projects using common criteria and standards.

Conclusion

The Certificate of Need process does not suffer from a lack of data. What is needed is development and use of selected information that results in attention being focused on major policy issues rather than on extraneous, non-essential issues. Washington State, through the development of common methods to analyze selected data, has made CN decision-making more consistent and meaningful.

Figure 1

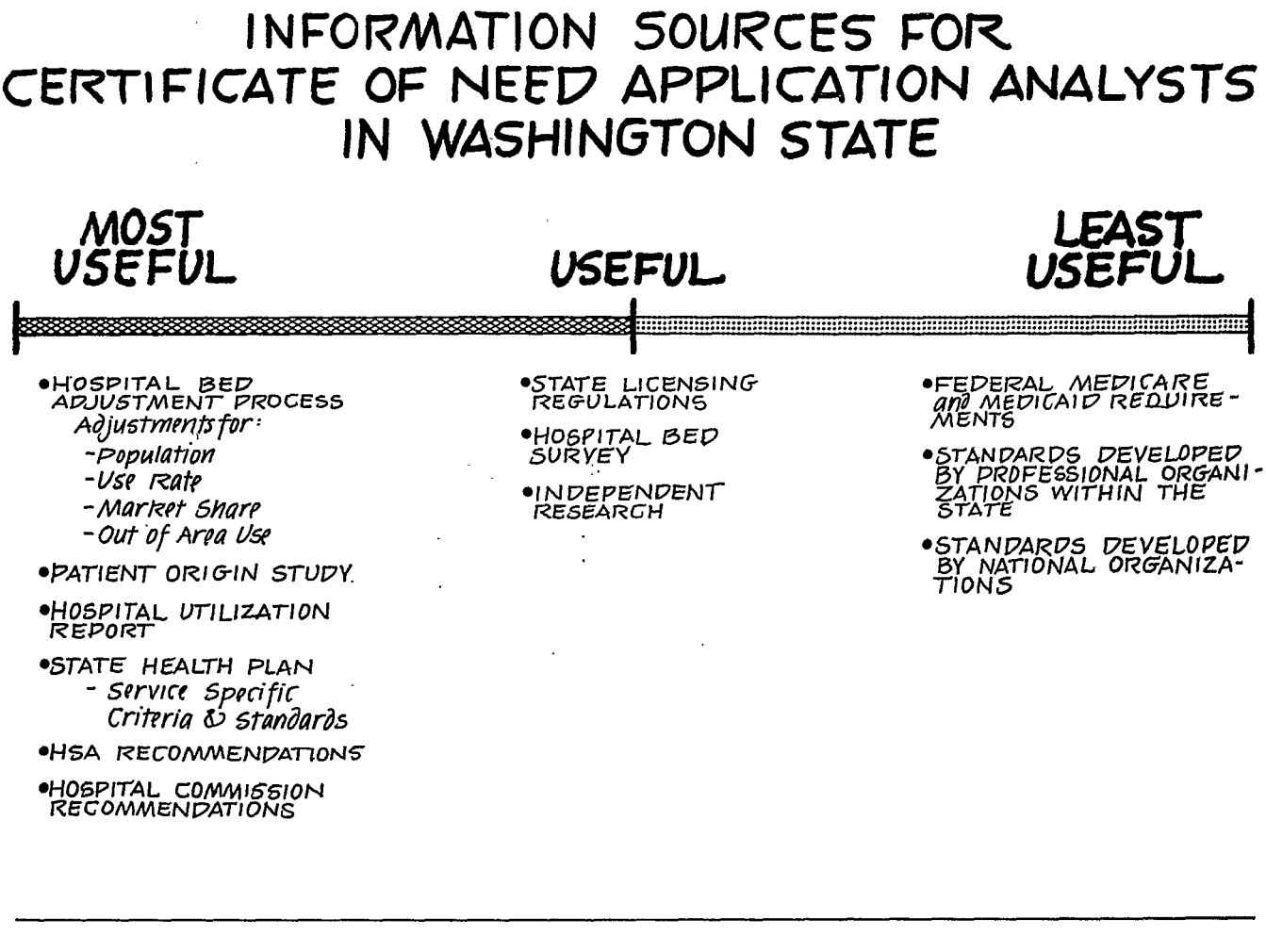


Figure 2

MODIFICATIONS TO WASHINGTON STATE'S CERTIFICATE OF NEED PROGRAM 1971 – 1983

Changing Characteristics of Data Used for CN Data

FROM:	TO:
• BASED ON HISTORICAL TRENDS	• POPULATION - BASED
• APPLICANT - CENTERED	• COMMUNITY OR PLANNING AREA - BASED
• GENERAL	• SPECIFIC

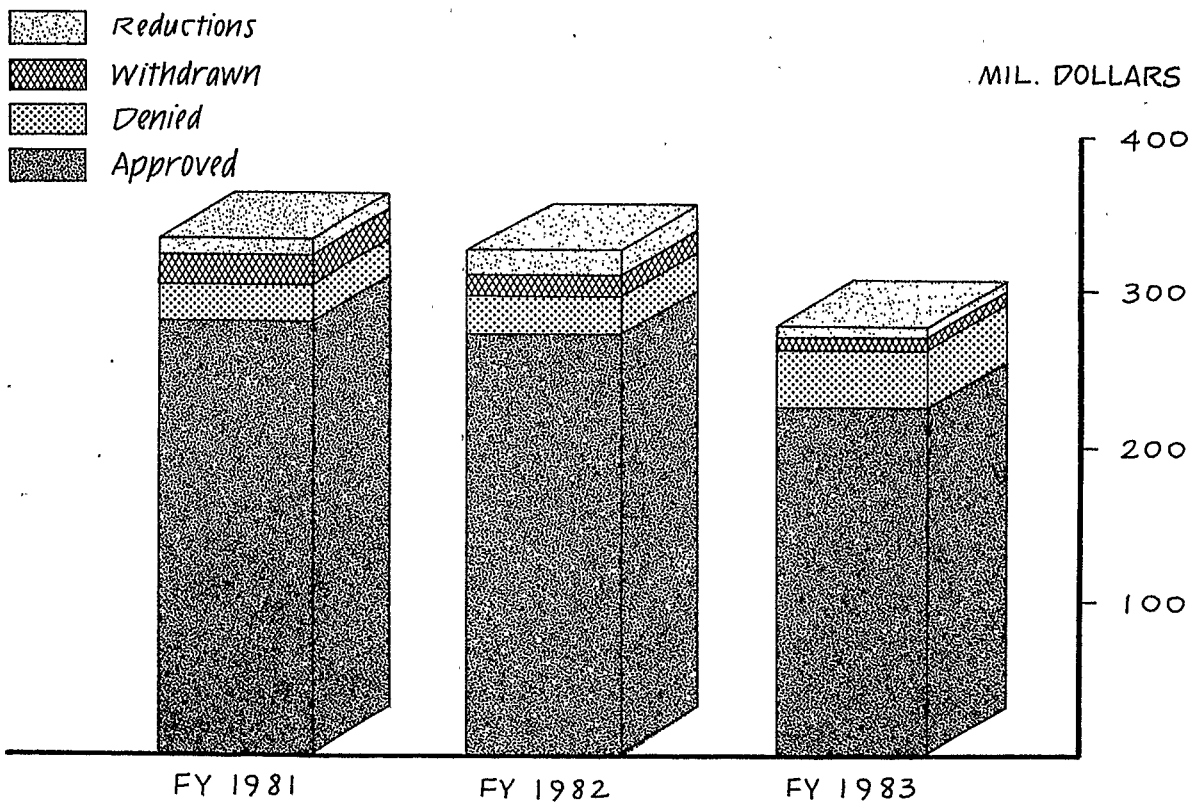
Changing Characteristics of CN Decision Making

FROM:	TO:
• PROJECT REVIEWED ON ITS OWN MERITS	• PROJECT REVIEWED IN RELATION TO SIMILAR PROJECTS
• CN APPLICANTS NOT CHALLENGED BY THEIR PEERS	• CN APPLICANTS CHALLENGED BY THEIR PEERS
	• ASSUMPTIONS BY AN APPLICANT CHALLENGED
• ABSOLUTE NEED	• RELATIVE NEED

Table 1

CERTIFICATE OF NEED PROGRAM ACTIVITY

FY 1981, 1982, 1983



References

[1] Porter, David, Hospital Architecture: Guidelines for Design and Renovation. AUPHA (Association of University Programs in Health Administration) Press, Ann Arbor, 1982, p. 13.

[2] Puget Sound Health Planning Council, Hospital Development Guide for the Central Puget Sound Region. October, 1975.

[3] Washington State Health Coordinating Council, Washington State Hospital Bed Projection Methodology. August 7, 1979.

A SYSTEM FOR MONITORING COST CONTAINMENT EFFORTS

Nancy J. Hurwitz, Independent Consultant

While well-known and accepted methods for reducing health care costs are being widely implemented by those interested in containing health care costs, it appears that these methods are being applied without sufficient knowledge of their ultimate effect or effectiveness. In too many cases, it is not known whether efforts applied are, in fact, needed, or if agreed that a particular intervention is appropriate, that the approach is effective in achieving its goals once instituted.

This paper will demonstrate that a more informed approach to cost containment is possible. The system for monitoring cost containment efforts described herein was originally designed for a large national advisor to multi-employer pension and welfare plans and is applicable to all employers. It permits the identification of areas in need of improvement, while also allowing for the tracking of efforts already in place. Further, the system permits for the range of data sophistication found in and fiscal resources available to a wide variety of users and potential users (i.e., employers and benefits managers).

Cost containment is built on the combination of successfully lowering costs of the delivery of health care and influencing the use of health services by the various segments of the population. The monitoring system has, therefore, been designed to use and exploit these two factors, namely, costs and utilization. It compares data commonly collected and maintained by the employer's health insurance carriers with existing national, regional or local data concerning health care costs and utilization.

The monitoring system, and it is stressed that this is a monitoring system, not a cost containment system, is built upon the completion of five separate task areas, which are shown below. These are:

1. The identification of data that are appropriate and useful for monitoring cost containment efforts;
2. Identification of data available from the carrier or claims processor/administrator;
3. Identification of appropriate outside data sets and individual data elements for comparison with carrier-provided data;
4. The actual comparison of carrier-compiled data with past data (if available) and with outside data sets; and
5. The determination of desired changes in utilization and costs.

Each of these task areas is now discussed in more detail, beginning with Item 1 - the identification of data that are appropriate and useful for cost containment monitoring.

Knowing what data are appropriate and useful for monitoring cost containment, is the cornerstone to this very simple monitoring system. This area addresses not only what specific data data elements are necessary, but also the form in which they will be most useful.

Data that are of interest should include measures of utilization, both inpatient and ambulatory, such as discharges and physician visits, and also some indicators of costs of services, such as room rates.

Since this system relies on the comparison of two sets of data to identify areas in need of cost containment, and since the comparison of unadjusted numbers alone is often misleading, the monitoring system depends on the use of rates and computed values. Therefore, exact data needs may be broken down into two types: numerator or observed data, and denominator or total at-risk data. Numerator data consist of those on subjects directly impacted by cost containment - costs and utilization. For inpatient programs, the area of health care where it is known that the most resources are expended, data of interest should include at a minimum, numbers of discharges (not admissions) by diagnosis, lengths of stay by diagnosis, and numbers of operative procedures by specific procedure. It should be noted that for more sophisticated users, DRGs could be substituted for diagnostic data, although, at present, little comparison data are available in that form. Ambulatory utilization data should at least include numbers of physician visits by diagnosis. Costs should include charges for semi-private rooms and charges for specific surgeries.

These raw numbers describing utilization are, however, of little value to an analyst without knowing something about the population who used the services, and also about the population who would be affected by any cost containment measures instituted. Therefore, knowledge of the age and sex composition of the users and potential users of health care services, as related to the raw numbers, is of vital importance. This forms the basis for the denominator data. These data should follow the same form as the numerator data, but are concerned with information on totals, such as total numbers of discharges and total population, and at-risk population figures.

Knowing what data are needed for monitoring, the next step is to match these data needs with data that are similar and are readily available from the carrier/administrator/processor, and what data might be obtained if claims or beneficiary files were manipulated differently. While it would be most advantageous for all of the desired data elements identified in the first step to be available immediately for monitoring purposes, it is recognized that the primary purpose of insurance companies is to pay claims, and not to provide data, and therefore, this will usually not be possible. In fact, it may be better and easier to implement the monitoring system gradually, allowing for the readily available data to be observed and used at the start of the monitoring program, while accomodating for sufficient time at the carrier for any changes that may be necessary in computer

programming and forms to produce the desired data.

The identification of available data should be done in conjunction with the health insurance carrier/administrator/processor. Prior to meeting with the carrier, a complete review of all claims and eligibility forms should be undertaken to determine what data elements are routinely collected and what might reasonably be expected. Information gleaned from this exercise will aid in making knowledgeable and reasonable requests of the carrier/administrator/processor.

When requesting data from the carrier, make sure that the person you are requesting data from is familiar with not only the insurance policy and its specific benefits but also with the capabilities of the company's data systems. In most cases this will not be the accounts person. Doing this simple thing will reduce the frustration level and time needed to acquire data for monitoring health care cost containment efforts.

Although rates and computed values are needed and desired as end products for comparisons, the carrier should not be asked to supply information in this form. Instead, they should be asked to supply only raw, well-documented frequency data. In this way, it will be easier to control the quality of data being used for comparisons, and also allow firms or employers with multiple carriers to minimize the differences among those carriers. In addition to asking for raw data only, it should be made clear that data are desired on a date-of-service rather than a date-of payment basis. This will ensure that data are reported for similar time frames, with similar extraneous factors at play, and also permits comparison with outside data bases reporting for the same time period. In addition to these requests of the data processor, a full accounting of all beneficiaries and their dependents, by age and sex groups, should be obtained so that meaningful rates and values can be computed.

The third step in developing a cost containment monitoring system is to identify appropriate complementary outside data sets that might be used as confirmation and comparison with carrier-provided data.

When originally researching this particular area of the system, fifteen separate data bases from seven different governmental and private sources were reviewed in depth. They included data collected and reported by the National Center Health Statistics (i.e., Health Interview Survey, Hospital Discharge Survey, National Ambulatory Medical Care Survey, National Medical Care Utilization and Expenditure Survey, Master Facility Index) and Medicare (20% sample), on the governmental side; and the American Hospital Association (i.e., Annual Survey of Hospitals, National Hospital Panel Study), IMS America (i.e., National Disease and Therapeutic Index, National Prescription Audit, Hospital Record Study), the Health Insurance Association of America (i.e., Prevailing Healthcare Charges System, Mail Survey of Hospital Semi-Private Room Charges), and CPHA, on the private sector side. In addition, four other data bases were reviewed, but immediately rejected from further consideration because they were not on target, or were only in the very early stages of their development. While local comparison data are of interest and in some areas surely available, this system does not address them since its original intent was to design a system for mass application across the country at a reasonable cost to the user.

In considering which of the fifteen sources would be best employed for use with a health cost containment monitoring system, a number of factors were taken into account, namely:

1. Reliability of data provided;
2. Timeliness of data;
3. Specificity of data;
4. Ease of using data; and
5. Cost of obtaining data.

Of these factors, the most important is reliability. This factor, therefore, was used as a first-cut criteria for reducing the number of available data sources.

Reliability of data was judged on the basis of statistical confidence of estimates provided, and one thing that is integrally tied to confidence: the manner in which the data were collected. In reviewing the authorities, it is acknowledged that surveys, with sufficient sample size and representation, and adequate response rates can provide statistically reliable estimates of utilization and costs. Further, surveys based on records, such as inpatient records, even with their recognized problems, are still more reliable than those based on recall, such as data that might be gotten through personal or telephone interviews of past events. With this in mind, eight candidate data sources can be removed from consideration as providers of comparison data. Therefore, the best sources when considering reliability are, for inpatient care utilization: The National Hospital Discharge Survey, the HCFA/Medicare 20 percent sample survey, the Hospital Record Study, and the CPHA/PAS 800,000 and 2,000,000 record research files. For ambulatory data, the National Ambulatory Medical Care Survey and the National Disease and Therapeutic Index are considered. And for costs, only the surveys of the Health Insurance Association of America are available. The National Medical Care Utilization and Expenditure Survey may prove to be a valuable resource for cost data in the future, but at the present time, too little is known of its data capabilities or precision.

The seven sources discussed above were then examined in light of the four other criteria - timeliness, specificity, ease of use and costs - to determine first choice comparison data sources, and where possible, alternatives. This examination looked at each of the data sources in terms of the most recent complete calendar year for which data were available (timeliness), the specific geographic level at which data were available and the populations surveyed (specificity), how the data are usually reported (ease of use), and the annual charge for published data (costs).

Sources of inpatient data were looked at first. Data from the HCFA/Medicare 20 percent sample were immediately dismissed from further

consideration due to the restrictive nature of the population under study - disabled and over age 65, only - and because of the lack of relatively timely data. Data from the Hospital Record Study were also eliminated. Data estimates presented in the HRS reports are actually from a small CPHA/PAS data file (400,000 records). Further, costs of procurement of data from the HRS are prohibitively high.

The Hospital Discharge Survey and the Professional Activity Study research files are, therefore, the logical choices for obtaining inpatient utilization data. CPHA/PAS has more recent data available currently, but data from the 1981 HDS should be available shortly, if they are not already available (HDS data generally run one year behind data from the PAS). Data for 1982 shoud be available from CPHA/PAS now or in the very near future. PAS provides more specificity in terms of geographic location than the HDS, offering nine breaks as oppossed to four general regions; however, data are reported as frequencies rather than rates, which would necessitate an additional statistical calculation in order to determine differences with beneficiary population experiences later.

Perhaps the most decisive determinant in choosing the HDS over the PAS data base, or visa versa, could be the cost of the data. Charges for published data from the National Center for Health Statistics for HDS data are minimal, ranging from no charge to $9.00. Charges for PAS data, while not nearly as high as the Hospital Record Study, are nonetheless, high enough to give second thought to ($4,900 to $5,900). If the cost of the PAS is not deterring, and despite the additional calculation required, PAS may be better because of its ability to respond relatively quickly, and because they are a more timely source of comparison data. If the cost of data is a factor, data from the HDS are more than adequate, and also have the advantage of providing estimated rates rather than frequencies.

In terms of ambulatory health care service utilization, only two choices existed: the National Ambulatory Medical Care Survey and the National Disease and Therapeutic Index. These two surveys are conducted using similar study methodologies - panels of physicians in private practice recording characteristics of patient encounters on specified days over the weeks during the year. However, response rates for the two surveys differ greatly. NAMCS has a much higher response rate. Surprisingly though, the estimates made by the two systems are remarkably alike. The differences in response rates, in combination with the high cost of procuring data from the NDTI ($22,000), make the NAMCS, despite its being one year behind the NDTI, the logical first choice for obtaining ambulatory utilization data.

No alternatives currently exist for cost data that are available from the Health Insurance Association of America surveys. The issue that should be examined here is whether the expenditure required to acquire the comparison surgery charges ($1,500) and average room rates ($35) justifies the limited amount of of data that would be returned.

No comparison data sources could be found for certain types of data that should be of interest to those monitoring health care costs and utilization. These include data on ancillary services utilization, specifically laboratory and radiology, costs per prescription by diagnosis, and charges incurred for specific diagnoses. In researching the available comparison data sources, it was found that while most people agree that these types of data are important for getting a complete picture of costs and utilization, no one was aware of a source of these data.

Now that two sets of complementary data have been compiled, that is the carrier and outside data, the actual comparison of data elements can be undertaken. Depending on what outisde data sets have been chosen, rates, averages, or expected values will have to be computed first for carrier data. Data in these forms can then easily be compared with data provided by the outside data sources through the use of simple analytical techniques such as odds ratios, t-tests, and Chi-square.

As in all comparisons, the trick to using the data is in knowing how to interpret the results. In other words, if differences are seen, are they real differences? Using this type of monitoring system, i.e., one employing both carrier data specific to an employer and appropriate comparison data, the employer or benefits manager is able to determine whether his measures of utilization and costs are radically different from those of his specific geographic region, and thereby identify areas that might benefit from some type of cost containment intervention. Further the employer can ascertain if ongoing cost containment strategies have been effective before complete baseline data are available. By comparing his few points of data with the corresponding retrospective comparison data elements, the user may determine if his strategy or something else occurring in the environment is responsible for observed changes in costs or utilization, thus indicating where additional efforts may be needed.

Further, aside from a strictly statistical identification, the same techniques may be used along with a standard reporting format for the actual monitoring of strategies already in effect, that is, for areas with more than one year of data. This report should show the employer's data for the current year and past year, along side appropriate national, regional, or local data. It should also document which cost containment strategies have been put into effect to impact on what measure or utilization or costs. The last section of the form should show the goals for the strategies, what their actual or achieved levels of performance were, and the variances from the stated goals. This can show, at a glance, how effective different cost containment interventions have been, which ones should be retained and continued, and which ones would be best abandoned.

In conclusion, while the system described above does not itself directly reduce costs or change utilization, some health care cost savings may be realized by pointing out where

and when health care cost containment strategies can be employed most effectively and in which directions limited program resources should be marshalled.

Methodology

Session Q

A METHODOLOGY FOR ESTIMATING COST SAVINGS RESULTING FROM REDUCTIONS IN HOSPITAL INPATIENT USE

Kirk Phillips and Sal Bognanni
Health Policy Corporation of Iowa

Introduction

Health care costs continue to spiral well beyond the rate of inflation for the rest of the economy. The largest single component of health care costs is hospital inpatient use. Targeting cost containment efforts on reducing hospital inpatient use could produce considerable savings. This paper presents a methodology for placing a "price tag" on the potential savings by reducing inpatient use. Patient discharge records and total inpatient costs of each hospital were used to calculate hospital inpatient costs for residents of a defined geographic area. The analysis is population-based and provides an important conceptual link between hospital inpatient use and hospital per capita costs.

The methodology builds on the work of John Griffith and Dr. John Wennberg and can be applied to any defined geographic area by using a few critical data elements.

Methodology

Data Acquisition

To apply the population-based approach, data from the patient records of Iowa residents should be obtained from all hospitals that serve Iowans. The hospitals could be located in Iowa or in other states, or even could be located outside the United States. The impossibility of a 100% survey is apparent. Thus, an attempt was made to obtain data from those hospitals which serve a number of Iowas. Patient abstracts were purchased from Servi-Share of Iowa to obtain individual, but not identifiable, patient data on all patients discharged during September 1980 from Iowa hospitals and Iowa residents discharged from as many hospitals in coterminous states as were willing to participate.[1] All Iowa hospitals participated, as did 15 hospitals from neighboring states. This report used the following information from the patient records:

- hospital identifier
- residence of patient (zip code, county when possible)
- age of patient
- sex of patient
- date of admission
- date of discharge

The above were converted into the required data items. A zip code-to-county match program was used to assign a county code to any patient record missing the county code. A special code was given to all non-Iowa residents. Also the length of stay was calculated using the date of admission and date of discharge.

The 1980 population counts used to calculate rates were obtained from the Iowa Office of Planning and Programming, August 1980 series of population projections.

Iowans Served By Out-of-State Hospitals

The data obtained from hospitals outside the periphery of Iowa poorly represented the care of Iowans traveling out of the state for care. Complete patient abstracts were not readily available from many of the hospitals queried. There, a decision was made to use the 1980 patient origin data only for information on Iowans who received care in Iowa hospitals. Information on Iowans who recieved care in out-of-state hospitals would be estimated. The best source available to us was a printout, "Distribution of Discharges of Medicare Beneficiaries Age 65 and Over by County of Residence and Hospital Where Discharge Occurred 1977" from the Health Care Financing Administration. This information was from a 20% sample of all hospitalizations under Medicare.

The feasibility of using this Medicare sample for patient origin was tested against 100% patient origin data from Maine, Massachusetts, and Vermont. Results in the Vermont study typified the findings. Only 4.4% of the population were <u>not</u> assigned the same way for the ziptown aggregations needed for Vermont analysis.[2] With over 95% of the population being assigned the same way, it would seem that the Medicare A sample can be a good proxy for complete patient origin studies.

Age-Sex Adjustments

Age is an excellent predictor of hospital use. Elders have the highest rate of hospital use, followed next by the 45-64 age group. Females use hospital services at a rate higher than males. Because different age-sex groups have different hospi-

talization and hospital utilization rates, county populations with dissimilar age-sex structures would be unequally weighted with respect to their aggregate rates. Age-sex standardization or age-sex adjustment is a method for compensating for dissimilar age-sex structures in county populations. The following steps were taken to age-sex adjust discharges and patient day rates for each county:

1. The proportion of the State of Iowa population for each of the 8 age-sex groups (females age 0-14, 15-44, 45-64, 65 and over and males age 0-14,15-44, 45-64, 65 and over) was calculated.
2. The proportion of population for each county for each of the 8 age-sex groups was calculated.
3. For each age-sex group, the proportion of the Iowa population was divided by the proportion of each county population to find the adjusting factor.
4. The number of discharges/patient days for each county for each age-sex group was calculated and the result was then multiplied by the corresponding adjusting factor. (Counties that have higher than the state average proportion for particular age-sex groups had adjustment factors less that 1.0 which reduced the number of discharges or patient days for those age-sex groups).
5. The eight resulting calculations for each county were summed and the total was divided by the County population. The result was multiplied by 1,000 to obtain discharge/patient day rates per 1,000 population.

Steps in Calculating Discharge and Patient Day Rates

For each Iowa county, several age-sex adjusted discharge and patient day rates were calculated. A list of these rates with some notes about how they were calculated follow:

1. Estimated rates per 1,000 population for county residents using Iowa hospitals during the year.
   All Iowa hospital patient records from the sample survey were aggregated by hospital to obtain the number of discharges and patient days for each hospital. The total number of discharges and patient days during the whole year for each hospital was obtained from the Annual Report for Hospitals and Related Health Facilities. The yearly total of hospital discharges (and hospital patient days) was divided by the sample total of hospital discharges (and hospital patient days) to obtain an annualizing factor. In 1980, the annualizing factors for discharges ranged from 9.8 to 14.0 and for patient days 9.9 to 17.0. The appropriate hospital discharge annualizing factor was added to each patient record, as was a calculation of the patient days on each record times the corresponding hospital patient day annualizing factor.

   All non-Iowa resident patient records were omitted from the records from Iowa hospitals. The remaining records were aggregated by county of patient residence to obtain the number of annualized discharges and annualized patient days for each county for the eight age-sex groups. After multiplying by the age-sex adjusting factors and summing by county, the resulting county numbers times 1,000 were divided by the corresponding county population to obtain county rates per 1,000 population.

2. Estimated rates per 1,000 population for county residents using Iowa and non-Iowa hospitals during the year.
   The estimated rates calculated above were for county residents using only Iowa hospitals during the year. To obtain estimated rates for county residents using both Iowa hospitals and out-of-state hospitals, the rates calculated above were increased. To do this, each discharge rate calculated above times 100 was divided by corresponding estimated percent of county residents discharged from Iowa hospitals. The estimated percent of county residents discharged from Iowa hospitals was calculated by subtracting the estimated percent out-of-state discharges (described earlier) from 100.0. Similar calculations were made to estimate the patient day rates.

STEPS IN CALCULATING ACUTE CARE COSTS

In this report, county resident costs for hospitalization are shown in gross and per capita figures, based on the methodology shown below. These costs were obtained from audited Medicare financial reports made available by Blue Cross of Iowa and the Iowa Health Department. After obtaining the total 1980 inpatient cost for each Iowa hospital, the following statistics were calculated:

1. Adjusted Total Hospital Expenditure

The total hospital cost figure was reduced to exclude the cost of non-Iowans receiving care. This was accomplished by estimating the percent of each hospitals' patients traveling from out-of-state, and reducing the total hospital expense accordingly.

2. Proportion of patient origin for each hospital

The percent of patients served in each county was calculated for every hospital. This percent figure was used as a proportion to show the degree to which certain counties are served by each hospital.

3. Allocation of costs

The total inpatient costs for each hospital were divided among its counties served according to the proportion of patients served. Where, for example, 65% of a hospital's patients lived in a single county, 65% of the hospital's inpatient costs were assumed to be used by patients in that county. The resulting cost figures which were brought to counties from the various hospitals serving them, were summed to represent the cost of hospitalizing its residents in Iowa hospitals. These cost figures were inflated by factors described earlier, which reflect Iowans' use of non-Iowa hospitals. This method assumes that the cost of non-Iowa hospitals is similar to those of Iowa hospitals attended by each county's residents.

4. Per capita costs

The cost of inpatient care per person in each county was developed by dividing the total allocated hospital costs by the 1980 population for each respective county. These costs were age/sex standardized to account for differences in county population structure. This adjustment was made by applying the ratio of adjusted patient days over non-adjusted patient days to the per capita costs for each county.

5. Cost savings with reduced use rates

In this paper, two variables were applied to the county hospitalization costs, which demonstrated cost savings; they are: (a) normative or expected use rates, and (b), variable costs in reducing hospital expenditures. These variables were used in the following manner:

a) Percent reduction in patient day use -

Where county patient day use rates were greater than the norm, e.g., 1287 days per 1,000, the percent of that county's actual patient day use rate above the norm was calculated as follows:

$$\frac{\text{Act. Rate - Exp. Rate}}{\text{Exp. Rate}} \times 100 = \text{\% Reduction}$$

Expected use rates at 1287 and 1046 were used in this study.

b) Cost savings

The above percent was applied to the total county cost for inpatient care, resulting with the gross cost savings, due to the proportional reduction in hospital patient day use.

$$\frac{\text{Total County Cost} \times \text{\% Reduction}}{100} = \text{Gross Cost Savings.}$$

c) Adjusted cost savings

The variable cost portion of the above figure was calculated at a 50% level. Accordingly, the following shows a method for calculating cost savings, accounting for variable costs.

Gross Cost Savings x .50 = Adjusted Cost Savings @ 50% variable cost

d) Net Cost Savings

The adjusted cost savings from the above was subtracted from each total county cost for counties having patient day rates above the norm. Resulting county costs were summed to represent the statewide cost savings as a result of reducing hospital use to the expected patient day use rate.

[1]All civilian acute care hospitals participated. Information was not solicited from the mental health institutes, the state schools for the mentally retarded, prison hospitals, nor veteran hospitals.

[2]John E. Wennberg, M.D., and Alan M. Gittelsohn, Ph.D., A Small Area Approach to the Analysis of Health System Performance, DHHS Publication N.(HRA)80-14012, August 1980, pp. 48-51.

SELECTED BIBLIOGRAPHY

American Hospital Association. Hospital Statistics, 1981 Edition. Data from the American Hospital Association 1980 Annual Survey.

Department of Health and Human Services, 1979. Fixed and Variable Costs of Hospital Services, Final Report.

Department of Health and Human Services, 1980. National HMO Census, 1980.

Department of Medical Care Organization, School of Public Health, The University of Michigan. Medical Care Chartbook, Seventh Edition. AUPHA Press, 1980, Ann Arbor.

Gittelsohn, Alan. Department of Health and Human Services, 1979. "Uses of Health Data in Health Planning." The Public Health Conference on Records and Statistics.

Governor's Commission on Health Care Costs, 1981. Background Statement.

Griffith, John R. Measuring Hospital Performance. An Inquiry Book, 1978, Blue Cross Association, Chicago.

A COMPUTER SIMULATION APPROACH TO PROJECTIVE COST ANALYSIS IN CORPORATE BENEFIT PLANNING

David R. Anderson, Control Data Corporation

Most large companies in this country spend millions of dollars every year on employee health benefits, and these costs are increasing rapidly. With such large and growing expenditures, many companies are becoming interested in programs designed to contain health costs. However, given the substantial cost of implementing some of these cost containment programs, as well as practical limitations on the number which can be implemented during each budget period, decision makers must attempt to select those programs which maximize their return on investment.

This paper first summarizes the development of cost containment as a critical issue in employee benefit programs. It then describes our use of computer simulation to estimate return on investment of proposed and existing health benefit programs, and discusses some of the advantages of this technique over more traditional approaches.[1] Finally, it briefly outlines some of our plans for future applications of computer simulation to employee benefit planning.

Growth of Employee Benefits

Employee benefits have grown steadily as a percentage of total compensation, increasing from about 25 percent of payroll in 1965 to more than 37 percent in 1981. During this same period, health related benefits have grown more rapidly, more than doubling from 4.4 percent of payroll in 1965 to nearly 10 percent in 1981.[2]

Several factors account for this general growth in benefit costs to employers and for the growth of health related benefit costs in particular. Benefits have been influenced directly or indirectly by legislation, such as social security cost increases and tax laws providing favorable treatment of benefits compared to wage increases. They have also been influenced by labor negotiations and the desire of employers to attract and retain capable employees (Rosenbloom and Hallman, 1981).

Health related benefits have been subjected to additional cost pressures, leading to their particularly rapid rise. First, and foremost among these additional pressures are incentive structures which encourage or fail to discourage high benefit usage levels. For example, health care insurance which pays a high percentage of charges by fee-for-service providers after a low deductable, and provides more complete coverage for inpatient care, has produced a cost spiral of crisis proportions for employers paying for this insurance, as well as for society in general. Another example of such incentive structures is the typical sick leave program which encourages employee abuse by providing no compensation for accrued sick leave to terminating employees (Harvey et al., 1983). Second, cost shifting by health care providers to those most able and historically "willing" to pay has also increased the health cost burden on employers, and this problem will intensify with widespread implementation of new federal Medicare patient reimbursement policies. Finally, an aging work force gradually requiring more medical care is also influencing health benefit costs for many employers.

Health Cost Containment

While benefit adequacy and satisfying employee needs are key considerations in employee benefit planning, cost is also an important concern. Employers must constantly balance desired benefits against available funds, and the growing cost crisis in health care has shifted emphasis to the cost side of the equation.

Few employers desire to reduce benefits to halt cost increases, nor do most believe that they can maintain current levels of benefits while also reducing costs. Rather, the more typical objective is to control, or contain cost increases through more effective design and management of benefit programs. Employers are developing a variety of approaches to achieve this objective of health cost containment. For example, some employers are self-insuring their health care plan to obtain better data on where their health care dollars are being spent, as well as to eliminate insurance vendor administrative costs. Some are also placing a greater emphasis on cost sharing in their health insurance plan to reduce unnecessary utilization by employees in the plan (Newhouse et al., 1982).

In addition to modifying existing benefits to more effectively manage costs, employers are also introducing new benefits specifically designed to control or reduce health costs. Examples of such benefits include employee education and health promotion programs, disability management, utilization review, and health care provider relationships such as HMOs and preferred provider organizations. Table 1 lists some of the benefit program changes which have been implemented or are being considered by Control Data in our efforts to contain health related costs.

Need for Cost Analysis

Some benefit program changes designed to contain health costs require minimal internal development or external investment, with the major expense of implementing the change being to communicate it effectively to employees. An example of such a low cost change is modifying deductable or co-insurance rates in the health care plan. However, other benefit program changes demand major investments long before any return can be expected. For example, a health promotion program focusing on lifestyle change (e.g., fitness, smoking cessation, nutrition, weight control) can be very costly to implement, but the major pay-back in reduced health related costs cannot be anticipated until five to ten years following implementation.

Regardless of the cost of the benefit program, it is important to estimate its likely

Table 1. Health Benefit Cost Containment Programs Implemented or Being Considered by Control Data

HEALTH INSURANCE PLAN DESIGN

*Self-insurance
*Preventive care
*Second medical opinion
*Outpatient surgery
*Employee cost sharing

UTILIZATION REVIEW

Pre-admission certification
*Concurrent review

PROVIDER RELATIONSHIPS

*HMO
Preferred provider org. (PPOs)
*Coalitions
Negotiated rates

DISABILITY PLAN DESIGN

*Return to work program
Flexible time off

HEALTH PROMOTION

*Diabetes/blood pressure screening
*Health risk appraisal
*STAYWELL lifestyle change

*Employee assistance
Financial incentives

EMPLOYEE EDUCATION

*WellTimes health news
*Cost awareness
Health consumerism

*Program(s) implemented

return on investment as early in its history as possible. Ideally, cost impact estimations should be made when potential changes are initially considered as part of the more general policy analysis process. Such planning phase estimations reduce the likelihood of costly mistakes and permit the allocation of limited resources to those programs likely to generate the greatest net return. Cost analysis should also continue throughout the implementation phase as part of the evaluation of the program, to assure that the anticipated return on investment is being realized.

Of course, some sort of cost analysis is almost always performed in benefit program planning and implementation. However, cost analysis methods too often include a major element of "gut feel" and a dearth of empirical data. Systematic, empirically oriented cost analysis still tends to be the exception, often because of the difficulty and expense of organizing existing data or performing special data collection and analysis. However, a systematic data based approach to cost analysis is crucial to the overall success of benefit planning and, given the potential costs of poor decisions, should be formally integrated into the benefit planning process.

Computer Simulation in Cost Analysis

Assuming adequate support for data collection and analysis, a different kind of obstacle to effective cost analysis for management decision making is encountered. That obstacle is the difficulty of integrating the great volume and diversity of analysis results often generated into a single bottom-line statement about cost impact. Faced with a plethora of charts and graphs representing individual cost components, decision makers are often overwhelmed by detail and revert to a "gut feel" approach to cost analysis.

A tool we have found very useful in overcoming this obstacle is computer simulation. The formal requirements of designing and programming a computer simulation model inherently demand integration of results and bottom-line orientation for a number of reasons. First, it is necessary to develop a detailed model of the process being analyzed and to specify the relationships among elements of the model. Second, an optimal simulation model incorporates all results of individual cost analyses considered relevant to determining overall cost impact. Finally, it is necessary to specify a priori the desired output of the model, not only in terms of content but also how it is to be formatted.

In addition to being a valuable integrative tool, computer simulation has other significant advantages as a general approach to cost analysis in benefit planning. Once the basic model is programmed, it can readily be modified to perform sensitivity and "what if" analyses. Sensitivity analyses test the effects on estimated cost impact of changing potentially invalid assumptions in the model. "What if" analyses test the effects on estimated cost impact of various changes in the program being modeled.

Given these general statements about the usefulness of computer simulation in cost analysis of benefit programs, I will outline two specific computer simulation models we have developed to assess cost impact of Control Data benefit programs. Both models project cost impact over a ten year period following program implementation. The first projects the cost impact of STAYWELL, a health promotion program already implemented at Control Data; the second

projects the cost impact of a flexible time off program being considered to replace our current sick leave and vacation programs.

STAYWELL Program Cost Impact Model

STAYWELL is a voluntary health promotion program developed by Control Data for all full-time employees and their spouses. STAYWELL is also being marketed to other employers. The program consists of a health screening/risk appraisal, health education and lifestyle change courses, and structured and unstructured group activities. STAYWELL was initially piloted at two Control Data work sites in 1979 and is currently being offered to more than 25,000 employees in 109 facilities. Specific aspects of the program are described more thoroughly elsewhere (Naditch, 1981).

We are conducting a very comprehensive evaluation of STAYWELL. The objectives of this evaluation are to assess the efficacy of STAYWELL in producing and maintaining risk factor change and to determine the economic consequences of the program for Control Data. To achieve these objectives detailed data bases are being compiled on participation in the program and on health risk factor trends. These data bases are being linked with company-wide data bases containing employee demographic information and health insurance claims to determine relationships among participation, health risks, health status and health claim costs.

Although the STAYWELL program evaluation will not be completed for at least five more years, we wanted to be able to make preliminary estimates of return on investment to assist in marketing the program. Our strategy was to develop a computer simulation model combining available STAYWELL program evaluation data with relevant published data.

While STAYWELL addresses numerous lifestyle risk factors, acceptable cost data were not available for several of them. Consequently, risk factors presently included in the model are smoking, hypertension, lack of exercise, and failure to use seat belts (e.g., Kristein, 1982; Hortunian, Smart and Thompson, 1981). Given this lack of data in some program areas (e.g., stress, nutrition), we believe that estimates generated by the model are probably less than actual STAYWELL cost impact.

While the computer simulation model itself is computationally complex, the underlying conceptual model of program impact upon which it is based is quite simple. Briefly, in order for savings to be realized by the organization implementing the program, an eligible individual must (1) possess a cost related risk factor, (2) participate in program activities relevant to that risk factor, (3) reduce the risk factor to a low-risk level, and (4) be employed by the organization for the program year being modeled.

Computationally, the model deals with the entire eligible population as a group, rather than estimating cost impact at an individual level. This group or macro-simulation approach requires specifying risk factors, participation, risk reduction, and turnover as probabilities rather than as dichotomous yes/no variables. These probabilities are specified based on STAYWELL evaluation results and the demographic characteristics of the population being modeled. For example, risk factor probabilities for an eligible population are estimated from its demographics by using tables created by applying discrete multivariate modeling techniques to health risk profile data collected by the STAYWELL evaluation (Bishop et al., 1975).

In order to estimate return on investment accruing to the organization from the STAYWELL program, modeled results for risk reduction in the eligible population are applied to data on costs associated with risk factors. These cost data combine STAYWELL evaluation results relating claim costs to risk factors with cost data from the scientific literative. In the risk areas of smoking, lack of exercise, and hypertension it is assumed that the full cost impact of risk reduction is realized only gradually over a multiple-year period following change in the relevant behavior (cf., Paffenbarger, 1979). On the other hand, the cost impact of seat belt use is assumed to be realized instantaneously following change in that risk factor.

The STAYWELL cost impact model outputs the following bottom-line estimates for an organization:

* Excess costs for each of the next ten years due to risk factors present in the eligible population, assuming no change in risk factor prevalence;

* Reduction in excess costs for each of the ten years following STAYWELL implementation;

* Net cost impact to the organization in each of the ten years following STAYWELL implementation, if yearly STAYWELL program costs are entered as data in the simulation.

To simplify interpretation of the output for decision makers, all estimates are provided in January 1983 dollars. Also, a range of potential outcomes is provided by computing high and low, as well as expected value estimates for each set of output.

The STAYWELL cost impact model is very successful in achieving the objective of reducing a large volume of data to a readily understandable summary of program cost impact. However, the current version is a very simple model compared to the underlying reality it represents. Much of this simplicity derives from the macro-simulation approach of aggregating across individuals and performing calculations using group probabilities. Such aggregation makes it impossible to take into account recognized covariation among predictors of cost impact. For example, our evaluation has shown that people tend to engage in multiple risk behaviors, and we suspect that successful intervention with such multiple risk individuals will yield greater cost reduction than estimated by the model.

The next model to be outlined, which simulates the cost impact of a flexible time off program being considered by Control Data, is much more complex than the STAYWELL model relative to the underlying reality it represents. This flexible time off model uses a micro-simulation approach, calculating cost impact estimates for

individual employees rather than for aggregates.

Flexible Time Off Cost Impact Model

Control Data currently has a sick leave program typical of many other employers. Paid sick leave accrues every pay period; it is to be used when employees are absent from work due to brief periods of illness; and upon termination of employment no compensation is given for accrued sick leave. We also share with many other employers a concern with increasing sick leave costs and a belief that our traditional program leads to sick leave abuse. With no incentives for not using sick leave and a "use it or lose it" contingency in the policy, such abuse is almost certain to occur (Harvey, 1983).

Because of these problems with the current policy, our disability management department began searching for alternatives which would reduce or eliminate sick leave abuse and contain costs at or below current program costs. A particularly attractive alternative was a flexible time off (FTO) program, which vests sick leave days and combines them with vacation time into an undifferentiated block of days that employees can use at their own discretion. Such a program has been successfully implemented by Hewlett-Packard and a number of other large employers. The appeal of a flexible time off program was that, as well as eliminating the incentive for abuse, it would give employees the responsibility to manage their own time and would put an absolute and predictable ceiling on costs.

While the general concept of flexible time off was very appealing to management, a lot of questions remained. Possibly the most difficult was what to do about sick leave employees have accrued under the current policy, which in some cases amounts to hundreds of hours. Other questions had to do with how many FTO days should be provided to employees, how the number of FTO days provided should be related to tenure, whether the short-term disability waiting period should be shortened to assure that employees would have FTO days available for vacations, and so on. In order to develop a specific flexible time off design proposal, decision makers required three types of information concerning solutions to each such question:

* Will this solution meet employee needs?

* Will this solution be accepted by employees?

* What will this solution cost?

Given the large number of alternative options being considered, and the desire on the part of decision makers for the cost implications of each, a computer based cost analysis was essential. The desire to project costs far into the future given a dynamic employee population specifically suggested that we develop a simulation model of future employee paid time off, based on available information on past sick leave and vacation use.

Building the simulation model began by obtaining a 1982 year-end extract from our computer-based employee master file, containing data for each employee on job category, tenure, salary, sick leave use and balance, vacation use and balance, short-term disability use and certain other data potentially related to the flexible time off program design. Next, this data set was analyzed to determine the relationship between employee characteristics and time off use. For example, sick leave use was found to be related to job, tenure, and sick leave balance.

Once we understood the relationships in this paid time off data set, we were able to develop our basic FTO computer simulation model. Figure 1 presents employee flow through the model, which is the same for both the current sick leave/vacation policy and the proposed flexible time off policy. Summarizing this flow, a decision is first made concerning whether to retire or terminate the employee during the year being modeled. If the employee is retired or terminated, the various accounts (e.g., vacation, FTO) are settled and the employee is replaced. If the employee is not retired or terminated, then (1) age, tenure and salary are incremented; (2) time off accruals are awarded; (3) time off use is modeled; and (4) time off use and balance accounts are adjusted.

The FTO model is unlike the STAYWELL model in that it models individual rather than aggregate behavior and cost impact. Time off use is assigned based on each specific employee's characteristics, with a random error term added to more effectively simulate the underlying system (Lehman, 1977). While significantly more complex to program than an aggregate model, this micro-simulation approach was essential for the FTO model to make acceptable estimates, due to substantial covariance among the variables. For example, turnover probabilities and time off use were found to be closely related to tenure, and

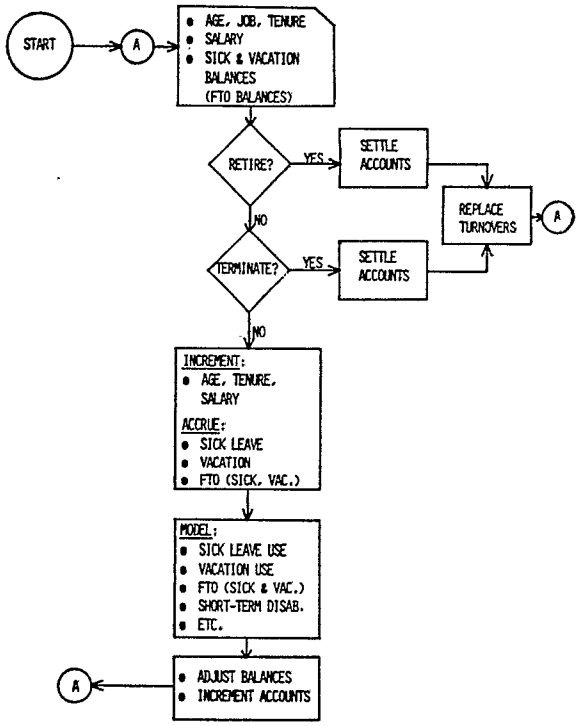


FIGURE 1. EMPLOYEE FLOW THROUGH FTO COST MODEL.

the nature of this relationship varied substantially across job categories.

The FTO simulation model processes a sample of employees or the entire employee population through 10 yearly cycles. Output is generated for our current policy and multiple specific FTO policy options. For each year, the simulation outputs background detail on balances, accruals, usage, turnover, and counts of employees exhausting accruals or balances. This latter item is considered a "human" cost impact measure of the policy being modeled. The simulation also outputs a ten-year cash flow summary, a ten-year accrued cost summary, and present value cost comparisons for the entire ten-year period at several different discount rates.

The basic flexible time off model was programmed assuming no change in employee's tendency to use sick leave unless they terminated. (We had found from our analyses that under the current policy, employees used significantly more sick leave in the year they terminated than similar employees who did not terminate that year.) Thus, changes in cost under flexible time off in the basic model would result solely from constraints of the policy itself. More than a dozen variations of this basic model have also been programmed to estimate the cost impact of:

* accelerating the FTO accrual rate,
* allowing employees to "cash out" excess FTO days once a certain balance had been accrued,
* shortening the short-term disability waiting period,
* adding the balance of the "grandfathered" sick leave bank to an employee's retirement plan benefit,
* vesting a portion of the "grandfathered" sick leave bank, and
* assuming changes in employee behavior under the new policy.

Wherever possible, costs estimated by the model were compared with projections of experts from inside and outside of the company. We found that the simulated results closely matched expert opinion and went well beyond what these experts were able to project using more traditional and less empirically based methods.

Conclusion

Developing computer simulation models can be a very complex and time consuming undertaking. However, they are proving to be a valuable tool in Control Data's health cost containment efforts for a number of reasons. First, they permit integration and efficient analysis of large and diverse data sets relevant to the benefit program being analyzed. Second, parameters and design details in a computer simulation model can easily and rapidly be modified to perform sensitivity analyses which set bounds on expected cost impact of a program, or "what if" analyses which test the effects of altering program design. Finally, the formal process of model development often leads to identification of issues and cost implications overlooked in simpler, less systematic approaches to cost analysis.

Management at Control Data has recognized the advantages of computer simulation in benefit planning and actively supports further applications of the technique. We are currently planning to develop a computer simulation model to project the cost impact of health insurance plan design changes, such as modifying deductables and co-insurance rates. We also plan to model the cost impact of employee behavior change in response to such design modifications (Newhouse et al., 1981). In the longer term, we visualize simulation models which integrate a broad range of benefit programs, permitting us to estimate the cost impact of multiple program changes and flexible benefit approaches. We also believe that this sophisticated benefit planning capability can be of tremendous value to other employers in their cost containment efforts. Therefore, we plan to introduce benefit planning simulation models to the marketplace in the very near future.

FOOTNOTES

[1] I would like to thank Robert S. Maier, also of Control Data Corporation, for his assistance in designing and programming the computer simulation models described in this paper.

[2] Information on historical employee benefit costs was obtained from U.S. Chamber of Commerce reports cited in the references section of this paper. The definition of health related benefits used here includes the benefit payment categories of: workers' compensation, life and health insurance, short-term disability, salary continuation or long-term disability, dental insurance, and paid sick leave.

REFERENCES

Bishop, Y. M. M., Feinberg, S. E., & Holland, P. W. Discrete multivariate analysis: Theory and practice. Cambridge, Mass.: MIT Press, 1975.

Employee benefits 1981. Washington, D.C.: U.S. Chamber of Commerce, 1982.

Employee benefit historical data: 1951 - 1979. Washington, D.C.: U.S. Chamber of Commerce, 1981.

Harvey, B. H., Schultze, J. A., & Rogers, J. F. Rewarding employees for not using sick leave. Personnel Administrator, May 1983, 55-59.

Harvey, B. H. Two alternatives to traditional sick leave programs. Personnel Journal, May 1983, 374-378.

Hortunian, N. S., Smart, C. N., & Thompson, M. S. The incidence and economic costs of major health impairments. Lexington, Mass.: D. C. Heath, 1981.

Kristein, M. M. The economics of health promotion at the work site. Health Education Quarterly, 1982, 9 (Special Supplement), 27-36.

Lehman, R. S. Computer simulation and modeling. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

Naditch, M. P. Wellness program reaps healthy benefits for sponsoring employer. Risk Management, October 1981, 21-24.

Newhouse, J. P., Manning, W. G., Morris, C. N., et al. Some interim results from a controlled trial of cost sharing in health insurance. New England Journal of Medicine, 1981, 305, 1501-1507.

Paffenbarger, R. S. Countercurrents of physical activity and heart attack trends. In Havlik, R. J., & Feinleib, M. (eds.) Proceedings of the conference on the decline in coronary heart mortality. NIH Pub. No. 79-1610, May 1979.

Rosenbloom, J. S., & Hallman, G. V. Employee benefit planning. Englewood Cliffs, N. J.: Prentice-Hall, 1981.

POLICY ISSUES AND DATA NEEDS: COST CONTAINMENT

Charles E. Bennett, Illinois Department of Public Health

I. INTRODUCTION

My presentation today is based upon excerpts from a larger, internal working document[1] and on-going analyses being undertaken in the Office of Health Planning in the Illinois Department of Public Health (which is the State Health Planning and Development Agency). One of my assignments has been to propose and write a justification for "a reasonable health data system." One response to that assignment is a paper focusing on the data needs related to the major policy area of concern: cost containment. It is from that paper and subsequent analyses that this presentation is derived.

In addressing the data needs related to cost containment, my first step was to attempt to identify and organize--that is, to show the interrelations among--numerous factors which affect health care costs and expenditures. Figure 1 graphically illustrates my efforts to show those relationships among factors and variables. While this is not a comprehensive or exhaustive model, it does present within a limited space a rather large number of inputs which contribute to health-related costs and expenditures and, at the same time, to health outcomes or the products of the health care system. While it is not a "true" path analysis diagram depicting comprehensively empirically ascertained correlations between or among the various elements, it does indicate by the connecting lines and arrows a logical relationship among the elements and, where available, some preliminary "numbers" or relative weights among inputs as derived from the literature of this field. The absence of "hard numbers or data", I hope, need not be seen as a weakness of the analysis to date but, rather as an aide to increased awareness of what we have yet to learn or discover. The "model", then, is an admixture of what has been observed, quantified and analyzed with what is still theoretical or hypothetical; in the larger paper some pains have been taken to document sources of knowledge and theory. The figure depicts relationships which presently exist, some which may have been hidden or out of mind, along with some relationships which do not yet exist or which exist only in limited geographical and organizational settings.

At least seven (7) purposes are served, I believe, by Figure 1 and accompanying analyses:

(1) First, this figure should help to alert policy and decision-makers of the complexities involved in the operation of the health care system. It would appear, for example, that there are not a few, easy solutions to the containment of increases in health care expenditures. Rather, there are a lot of small, partial solutions, often to be regionally or locally applied. Awareness of that, it is hoped, should help to avoid frustrations based upon simplistic, unrealistic expectations of the total effects of particular changes upon the health care system and health care expenditures.

(2) It begins to illustrate, graphically, what is known and what is, as yet, merely theoretical.

(3) While it indicates some factors currently or potentially amenable to change by health planners, it also reminds us of factors which lie beyond our control but which, if planning is to be successful, must be accounted for or responded to.

(4) It creates a framework in which to model or analyze the possible results of alternative assumptions or arrangements of elements of the health care system.

(5) It helps to identify areas for further investigation, data collection and analyses.

(6) It helps to highlight what is being "covered" or currently ignored by cost-containment planning efforts.

(7) By directing the attention of policy/decision-makers to what might otherwise be overlooked, it is hoped it will help to avoid or reduce unintended and harmful side-effects of changes elsewhere.

In the time available, I shall attempt to present illustrative examples of each of these several purposes.

II. EXAMPLES OF PURPOSES SERVED

Purpose 1- To illustrate the complexities of the system.

Purpose 2- To illustrate what is known and what is theoretical.

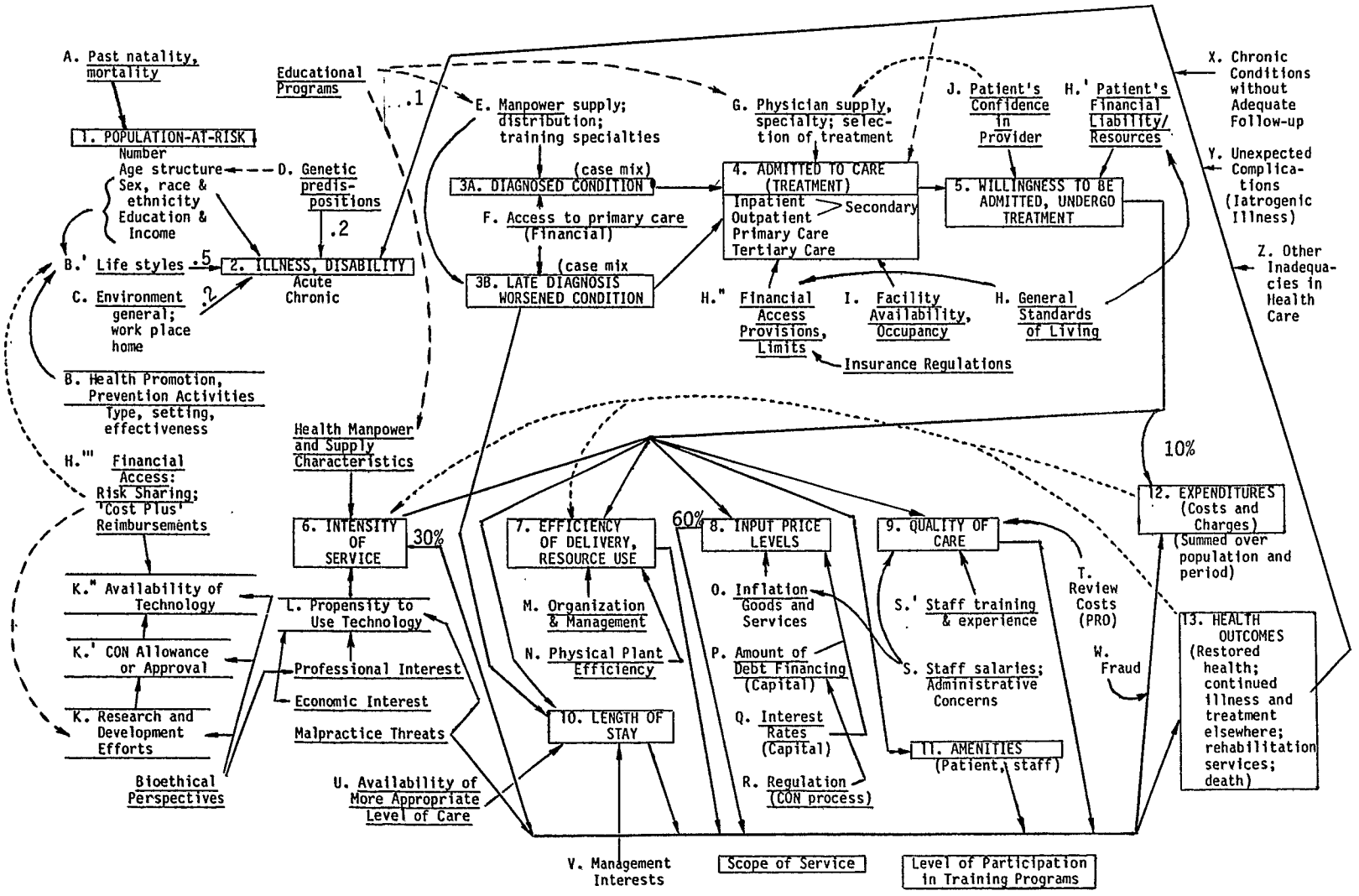
Example 1. Three Basic Components of Increases in Hospital Expenditures

Paul L. Joskow, Professor of Economics at the Massachusetts Institute of Technology, in

---

[1] While this report was developed as a part of planning efforts funded under Federal Health Planning Grant OP000285, deliberations as to the relative merits of options under consideration are still underway. Consequently, the views expressed to the Conference, in support of its objectives, must be understood to be the author's personal opinions and not, necessarily, the official position of the Illinois Department of Public Health (State Health Planning and Development Agency).

FIGURE 1. HEALTH CARE EXPENDITURES MODEL INPUT VARIABLES AND PATH ANALYSIS



an article contained in A New Approach to the Economics of Health Care (edited by Mancur Olson, 1981), writes:

> It is convenient to think of increases in hospital expenditures as being composed of three basic components:
>
> - increases in the cost of inputs (labor, capital, materials)
> - increases in the scope and intensity of hospital services resulting from technological change and from increases in demand
> - increases in the quantity of services reflecting population growth and changes in demographic characteristics, and increases in individual demands for care.[2]

Jaskow concludes on the basis of available data and analysis that, in the period of the past decade or two (it is unclear whether he speaks of the period 1960-1979 or only of the period 1970-1979),

> About 60 percent of the increase in expenditures has resulted from increases in input prices, about 10 percent from additional admissions and outpatient visits, and about 30 percent from an increase in the "intensity" of care.[2]

To this he adds the comments:

> (It should be noted, however, that input price changes are not likely to be completely independent of the demand for factor inputs.) The increase in the "intensity" of care appears to be the result of a high rate of technological change.[2]

Those findings are reflected in Figure 1 in several places. Considering the figure, for the moment, to be a model of hospital expenditures, the arrow and number (10%) just above box 12 (health expenditures) reflects the "additional admissions and outpatient visits" reflecting population growth, changes in demographic characteristics, and increases in individual demands for care. The 60% beside box 8 reflects expenditure increases due to increases in the costs of labor, capital, and materials. And the 30% beside box 6 reflects the expenditure effect of increases in the scope and intensity of hospital services resulting from technological change and from increases in demand. To illustrate one aspect of that increase in intensity: Nationally, in 1960, 10 percent of community hospitals had intensive care units; by 1969 the figure had risen to 44 percent.[3] In Illinois, by 1982, 80 percent of the community hospitals reported having an intensive care unit.[4] The contributing factors below and to the left of box 6 (including K, K', K", L, H ''' , and others) both reflect Joskow's view that "the increase in the 'intensity' of care appears to be the result of a high rate of technological change" and add to it. Technological change, itself, does not occur in a vacuum or as a simple response to population created demand, but reflects, a care providers to use the technology and the ease with which its costs can be covered. Moreover, the propensity of professionals to use new technology itself arises from numerous contributing factors. More on this point later. (See Purpose 7, below.)

Example 2. Factors Affecting Illness

In his 1979 report on health promotion and disease prevention, the U.S. Surgeon General made reference to an earlier work published by the government of Canada.[5] That work introduced a concept which views all causes of death and disease as resulting from four contributing factors:

- inadequacies of the existing health care system;
- behavioral factors or unhealthy lifestyles;
- environmental hazards; and
- human biological factors.

The Surgeon General went on to report:

> Using that framework, a group of American experts developed a method for assessing the relative contributions of each of the elements to many health problems. Analysis in which the method was applied to the 10 leading causes of death in 1976 suggests that perhaps as much as half of U.S. mortality in 1976 was due to unhealthy behavior or lifestyle; 20 percent to environmental factors; 20 percent to human biological factors; and only 10 percent to inadequacies in health care.[6]

And, he observed:

> Even though these data are approximations, the implications are important....[6]

As before, the numbers or percentages obtained by the analyses of others are presented in the figure, associated with several factors leading to illness and disability (box 2), together with further elaboration of the contributing factors. Please note, the percentages reflect estimated impact of these determinants upon mortality; I am presently unaware of similar estimates pertaining to acute or chronic morbidity, or to hospitalizations (which would be of special interest).

Clearly the number of ill and disabled persons in any area is a function of the size of the population at risk. The effects of "human biological factors", which I have called "genetic dispositions" (including both species and individual characteristics) vary with the age and sex of populations. "Behavioral factors or unhealthly lifestyles" are a reflection, at least in part, of numerous cultural and subcultural influences including ethnicity, education (formal and informal), income, and the presence or absence of vigorous local efforts in health promotion/disease prevention. It is useful to recall that people spend time in several distinct environments - the home - which may be safe or hazardous, - work places and/or several faceted milieu, including a willingness of physicians and other health

schools, in addition to the "general environment" of air, water, traffic hazards, roadways, and so forth.

"Inadequacies in the existing health care system" can mean many different things. Shortages or maldistribution of health manpower and facilities and/or limited financial access are potential inadequacies which, in the figure, I have depicted as relating more to diagnosis and treatment than to illness per se. Absence of certain prevention activities - e.g., immunization - could 'contribute' indirectly to disease. But of particular concern in an analysis of factors which contribute significantly to health costs and expenditures are iatrogenic illnesses and chronic conditions without adequate follow-up. The last of these two is frequently interconnected, also, with lifestyles, so the weight (10%) given to the connection between these factors and illness cases (box 2) may not be strictly accurate. (At the least, some additional interaction term is needed.) The impacts of iatrogenic and chronic illness to repeated hospital admissions has been at least partly documented, however, and should not be overlooked.[7]

Researchers Zook and Moore recommend among other suggestions:

> Hospital economy measures should be targeted more precisely on those small groups of patients who require much longitudinal care or demonstrate a high probablility of readmission.[8]

Purpose 3- To remind us of factors beyond our control which must be accounted for or responded to in planning.

Example 3. Population Size and Composition

The influence of population size and structure has already been mentioned. Together they represent only a relatively small part of the increase in hospital expenditures in the past decade. But if one assumed for a moment, either that nothing could be done about cost inflation or that inflation was no longer a factor, then population size and structure would be seen to play a more significant role in whatever increases in expenditure will yet arise. Even less than inflation, population size and structure are not amenable to influence by health planners. For that reason, and because their ramifications are contrary to cost containment efforts, it is important that both planners and policy-makers understand those ramifications and prepare for them. To do so should help to prevent unrealistic expectations and promote planning to help cover unavoidable health care costs.

In Illinois, for example, if all other factors in the hospital care system remained constant, the changes in the size and age structure of the state's population between 1980 and the year 2000, would bring about a 10.2 percent increase in hospital admissions. A seven point five percent (7.5%) increase would be due to simple increase in the size of the population, a 2.5 percent increase would be due to the aging of the population, and 0.2 percent increase would be due to the interaction of these factors. Because lengths of stay and general costs of treatment, like admission rates, also increase with age, the increases in days of care and hospital costs would be well in excess of 10 percent. In other states or local areas, the particular effects might be quite different, but need to be anticipated even if uncontrollable.

The magnitude of changes due to age structure and population size to be anticipated in the long-term care system is even greater than that for hospitals. Awareness of the magnitude of those changes increases the pressure to extend alternatives to current arrangements, wherever possible.[9]

Deliberations about such a policy bring us to examples of the next purpose.

Purpose 4- Create a framework to model or analyze the possible results of alternative assumptions or arrangements of elements in the health care system.

Example 4. Financial Arrangements for Hospital and Long-Term Care Services.

Illinois studies are showing that in the coming decade and a half large amounts of capital construction and debt financing costs might be avoided in Illinois if conversion of unused hospital space to long-term care services can be promoted.

Such conversion might, also, enhance the operating efficiency of the hospital plant (variable N). Among the points yet to be examined, however, are the reimbursement formulas for hospital and long-term care services. Anecdotal evidence presently suggests that under current arrangements hospitals make more money by retaining empty hospital beds than by filling a nursing care bed. What appears to be needed is analyses of a series of simultaneous equations in which adjustments to both sets of reimbursement formulas can be examined to ascertain the likely overall effects in both systems together and, by the way, help to anticipate the costs to the State of Medicaid supports within both systems.

Example 5. Effects of Reductions in Length of Stay.

It has long been observed that the South and the West have shorter average lengths of stay than do the North Central and Northeastern regions of the country. Parts of those differences are due to case mix, and to age structure of the populations; part of those differences are due to styles of practice of the providers, both physicians and hospital managers, involved in delivering care. Those styles, in turn, are influenced in part by local financial arrangements, participation in HMOs, etc. While it is clearly necessary that modifications of care practices not threaten quality of care or

healthy outcomes, within those limits efforts seem desirable to promote reductions in lengths of stay.

To obtain some idea of the magnitudes of reductions that might be possible, while controlling for case mix and age structure, we took a look at lengths of stay for eighteen major diagnostic categories reported for the South by the National Hospital Discharge Survey.[10]

The South was chosen because its age structure is quite similar to that of Illinois, thus minimizing the effect of age as a confounding variable. We presently lack diagnostic data on hospitalizations in Illinois, so the rates of patient discharges by condition as reported for the North Central region was used as a surrogate for Illinois. Applying, then, the South's average length of stay (ALOS) by diagnosis to the North Central region's discharges by diagnosis we find an "expected" average length of stay over all categories of 6.9 days. That is higher than the ALOS in the South (6.8), and considerably lower than the ALOS in the North Central states (7.5). It diverges even farther from the average length of stay reported for Illinois in 1981 by the American Hospital Association (8.0)[11] or the Illinois Department of Public Health Annual Hospital (Utilization) Survey (8.4)[4].

We cannot ascertain with presently available data how much of the difference between Illinois and the North Central region is due to case mix and how much is due to practice styles. But if the ALOS in Illinois could be brought down to equal the ALOS by diagnosis reported in the South without sacrificing quality of care/health outcomes, -that is, if a 17.9 percent reduction in days could be reasonably accomplished - that would represent a reduction of 2.89 million patient days in Illinois. Added costs of physician office and outpatient care would still have to be figured into calculations of net savings in expenditures to the general public and the State, but significant savings seem quite possible and deserving of further exploration. The effect of such a shift in ALOS upon bed need calculations is also significant. Using 76 percent occupancy rate (U.S. average) as a standard, the inpatient needs of Illinois patients (1981) could be served by 47,875 beds or 4.18 beds per 1000 population. The current ratio of acute care beds to population is 5.57 per 1000.[12]

The magnitude of the potential savings which appear to be involved, and our present ignorance of Illinois-specific conditions illustrates the fifth purpose of the figure, namely:

Purpose 5- To help identify areas for further investigation, data collection and analysis.

Perhaps analyses such as these, conjectural as they are, will help to secure the political and financial support needed for data so essential to effective planning.

Purpose 6- To highlight what is "covered" and what is ignored by current cost-containment planning efforts.

The above example also illustrates the sixth purpose of this figure, namely, to highlight areas which are under active consideration and implementation efforts, and those which are not being pursued. (A) Length of stay is a topic minimally addressed at present in Illinois. Furthermore, (B) Manpower training programs, supply and distribution, while related to service delivery[13], are planned with little analysis as to their cost-implications for the health care delivery system as a whole. (C) Health promotion/disease prevention activities, though not totally ignored in the past, seem likely to obtain even more attention in the future because of their bearing upon lifestyles and health. (Efforts to assess the effectiveness of various approaches, to improve their impact were possible and to reduce wasted effort where effects cannot be demonstrated, appear likely to be of increasing importance in coming years.) (D) Creation of "feedback" mechanisms within individual hospitals, to appraise physicians of the charges for services provided to their patients and to compare one's practice patterns with those of his peers within a specialty, -indicated on the figure by dotted lines between boxes 12 and 13, and 6 and 7- is still such an innovative practice as to deserve mention in the July, 1983 issue of Hospital Progress. (E) One pro-competition theorist has suggested that if a patient shared a larger portion of his health care costs, he would, sooner or later, reflect on the service provided and decide whether or not the physician treated him appropriately, ordered only the truly necessary tests, and so on. An adverse judgment is presumed to result in lack of recommendations of the provider to the patient's relatives and friends, and a consequent loss of business by a high-cost provider. While this theory is open to much question (and a virtual absence of any data to test it), it is clear that over the past twenty years there has been a marked improvement in the ability of prospective patients to bear health costs through cost-sharing of another sort, namely, insurance.

Example 6. Insurance Coverage.

The rapid increase of insurance coverage over the past 20 years has had the effect of spreading the costs of medical care and the financial risks of illness around much more broadly than before. For that reason, among others, people are willing to undergo treatments which before wide insurance coverage and sharing of costs among the insured (not co-payments) they would have opposed. Such cost-sharing/risk-sharing has several dimensions. One obvious dimension is to offer protection against "indiscriminate hazards" - the unavoidable accident when the

other guy is at fault, "the luck of the draw" as to genetic make-up and so on. Another is to share the risks or distribute the costs between different age cohorts. I personally favor that - and am willing to help support older co-workers and senior citizens during my productive years on the expectation that some day I will also benefit from such an arrangement. But current insurance systems frequently share the costs/spread the financial risks in two other ways also. They put into the same group persons with quite different actuarial expectations due to self-imposed risks (smoking, drinking, excessive eating, ignoring of high-blood pressure) and, unless you are buying health insurance privately, make no distinction as to premiums paid by or on behalf of such high-risk individuals. A potential constructive factor influencing lifestyles and total health expenditures is, thus, noted mostly by its absence from the scene. And insurance programs spread the costs among geographic regions of a state, between high utilization and low utilization areas, so that regions which make little use of health care services, subsidize regions which have high use rates.[13] Not only does this further penalize people in areas wherein special efforts may be undertaken to adopt low-intensity or lower frequency use patterns; it further hides the real costs of care from the people in high use areas.

Purpose 7- Direct attention to what might be overlooked, to help avoid or reduce unintended and harmful side-effects of changes in the health care system.

It appears plausible, at least, that "financial access" has been one of the factors contributing to the rapid expansion of medical technologies in the past decade or two insofar as costs and new technologies were frequently, if not invariably, covered by an insurance system of cost plus reimbursements and ever-upward adjusted premiums, raised year after year with little question. DRG or other reimbursement capping arrangements, while curtailing costs, may have as an unintended side effect the discouragement of technological research and development. Technologies of an experimental nature may have to go through much more extensive testing before they are adopted on a wide scale. And while that may not be bad, the uncertainties and reduced chances of recovering costs early may, in the absence of special funding, curtail research and development efforts and slacken the pace of new breakthroughs.

III. CLOSING REMARKS

I hope this brief presentation has been informative and stimulating. Many of you, I am sure, could add to this model. Tables and more complete documentation are contained in the larger paper from which this is taken, along with a number of ideas for cost-containment actions to be undertaken in both the public and the private sectors. I hope that the larger paper will be published in the late fall or this coming winter. Meanwhile, your comments, suggestions, additions and questions are all welcome.

Thank you for your kind attention.

REFERENCES

[1]Illinois Department of Public Health, Office of Health Planning, "Health Policy Issues and Health Data Needs: Cost Containment", (A policy background and discussion paper, forthcoming, 1983.)

[2]Paul L. Joskow, "Alternative Regulatory Mechanisms for Controlling Hospital Costs." Pp. 219-257 in A New Approach to the Economics of Health Care, Mancur Olson, ed. (Washington D.C.: American Enterprise Institute for Public Policy Research, 1981).

[3]Morgan, Daly and Murawski, "Dollar and Human Costs of Intensive Care." Journal of Surgical Research, 14 (May, 1973): 441-448.

[4]Illinois Department of Public Health, Office of Health Planning, Division of Health Information and Evaluation, unpublished tabulation from the Annual Hospital Survey for 1981.

[5]Marc Lalonde, A New Perspective on the Health of Canadians, (Ottawa: Information Canada, 1974).

[6]U.S. Surgeon General, Healthy People: The Surgeon General's Report on Health Promotion and Disease Prevention, 1979. (Washington D.C., Government Printing Office. DHEW (PHS) Publication No. 79-55071), pp. 8-9.

[7]Schroeder, Showstock and Roberts, "Frequency and Clinical Description of High-Cost Patients in 17 Acute-Care Hospitals", New England Journal of Medicine, 300 (June 7, 1979): 1306-1309.

[8]Christopher J. Zook and Francis D. Moore, "High-Cost Users of Medical Care." The New England Journal of Medicine, 302 (May 1, 1980): 996-1002.

[9]Illinois Department of Public Health, Office of Health Planning, Division of Health Information and Evaluation, Briefing Papers on the Health Needs of the Elderly. 1983.

[10]Department of Health and Human Services, National Center for Health Statistics, Utilization of Short-Stay Hospitals: Annual Summary for the United States, 1980. (Data from the National Health Survey, Series 13, No. 64), March, 1982, Table 15.

[11]American Hospital Association, Hospital Statistics, 1982 edition.

[12]Illinois Health Facilities Planning Board, Inventory of Health Care Facilities and Need Determination by Planning Area, Parts 1-4, 5th Edition (June 1983); and rough estimate of Illinois mid-year 1983 population at 11,452,000 (up from 1981 and 1982 census estimates of 11,444,000 and 11,448,000 respectively). (1980 Census figure for Illinois: 11,426,000.)

[13]John Wennberg and Alan Gittlesohn, "Variations in Medical Care Among Small Areas", Scientific American, 246 (April 1982), Number 4, pp. 120-134.

National Medical Care Utilization and Expenditure Survey (NMCUES) Data Use and Analysis

Session R

NATIONAL MEDICAL CARE UTILIZATION & EXPENDITURE SURVEY: LINKING ADMINISTRATIVE & SURVEY DATA TO IMPROVE A DATA SET's POLICY RELEVANCE

Larry Corder, Health Care Financing Administration

The Health Care Financing Administration (HCFA) and the National Center for Health Statistics (NCHS) co-sponsored the 1980 National Medical Care Utilization and Expenditure survey (NMCUES) As each agency was responsible for planning and funding this ambitious enterprise, a survey structure and design was arrived at that met several requirements of each party. HCFA required substantial emphasis on Medicare and Medicaid beneficiaries in the survey design for the following reasons: (1) a paucity of person based information concerning Medicaid enrollees (2) the cancellation of the current Medicare Survey (CMS) which eliminated HCFA's only Medicare data source which combined survey and administrative data at the person level. NCHS required a national survey of medical care utilization and expenditures with a sample of sufficient size to address general health care issues traditionally in their domain of interest.

These interests, both convergent and conflicting gave rise to a survey design with three major elements: (1) a national probability sample of 18,000 persons called the household sample (HHS) (2) a sample of 1,000 Medicaid enrollee cases in each of four States (New York, California, Michigan and Texas) drawn from Medicaid enrollment files at the close of 1979. These surveys were jointly called the State Medicaid Household Survey (SMHS). Both SMHS and HHS employed data collection instruments that were essentially identical. Data was collected at approximately three month intervals concerning 1980. Waves 1, 3, and 5 were conducted in person and 2 and 4 by telephone. The last major survey element was the Administrative Records Survey (ARS). While the label may be misleading, efforts in this survey component were aimed at collecting eligibility and claims data from Medicaid and Medicare records for persons in the HHS and SMHS who were deemed to be enrolled in those programs. Data concerning Medicare and Medicaid covered individuals was collected under strict confidentiality rules in the three survey components and none may be used for fraud or abuse investigation purposes.

The ARS consisted of several components which were designed to optomize the collection of administrative data for a given cost. First, as the Medicare system maintains a central record depository and it is a uniform, national bill paying system; all persons in the HHS and SMHS deemed to be covered had their records of claims extracted and their basis of eligibility recorded Second, all persons deemed to be covered by Medicaid based on data from the survey had their basis of eligibility (AFDC, etc) checked in the appropriate State Medicaid system enrollment file. The Medicaid program, unlike Medicare, does not maintain a national claim record system at a central location. Rather, the Medicaid program is a State run activity, operating under broad Federal guidelines. For this reason it was necessary to individually query each State which contained an HHS primary sampling unit concerning survey persons who resided in that State. This truly costly and cumbersome procedure required by the nature of the Medicaid program limited the amount of data that could be collected on "Medicaid" individuals in the HHS to type of eligibility. The third component of the ARS collected detailed claims information for persons in the four State SMHS.

The ARS was conceived and executed to minimize the type and amount of missing or incorrect data in the two household survey components of NMCUES. HCFA is primarily concerned with the behavior of its beneficiaries, so within fiscal constraints, the ARS was designed to eliminate missing use and expenditure data attributable to the programs (Medicare and Medicaid) and ensure that a verified count of enrollees was available along with their basis of eligibility. This strategy to minimize missing or incorrect data concerning Medicare and Medicaid programs is superior to logical and statistical imputation of coverage and events in that it relies on a bill paying system whose reason for existence is the orderly identification of enrollees and the appropriate payment to providers for covered services. However, the use of an administrative system designed to verify eligibility and pay bill to supplement information collected from survey respondents about program enrollment and use of covered services is neither low cost or operationally elegent. Its value, from the perspective of the NMCUES, lies in the ability to more completely and accurately characterize the medical care use of HCFA beneficiaries than any other approach once the survey data and administrative data for individuals have been joined together to form "best estimates" of program enrollment, medical events, costs of those events, and sources of payment for services.

It is unfortunately true that persons enrolled in the Medicare and Medicaid programs do not often provide accurate information concerning their eligibility and their medical care utilization and expenditures. It is not entirely their fault. For example, persons enrolled in the Medicaid program usually don't know how much the care which they receive costs because the information is not made available unless given by an accomodating medical provider. Indeed, it is often difficult to elicit an affirmative response to questions concerning enrollment in the Medicaid program. While the Medicare program does require that cost reports be sent to all beneficiaries, nearly all persons enrolled in the Medicare program are aged. The aged, as a group, do not report events as accurately as the remainder of the population. Indeed, partial payment by Medicare for certain services adds to a certain amount or respondent confusion.

HCFA embarked on the NMCUES project to generate an analytically flexible person based data file which would serve to fill the gap in

our understanding of Medicaid enrollee health related behavior and continue the time series begun by the Current Medicare survey. Within fiscal constraints, the ARS made those two goals attainable. Without the ARS component the covered medical care utilization and expenditure patterns of SMHS respondents would be unknowable and the same experiences in the Medicare population would be subject to extensive underreporting. With the ARS in place data which can only be collected from surveys (i.e. out of plan use, coverage under public and private plans and programs, health status measures, employment, income, and sociodemographics) is combined with "best estimates" of enrollment, use, and expenditures based on survey and ARS reports.

NMCUES is thus unique in that it is often the sole source which can address program policy issues/which neither household surveys such as the Health Interview Survey or Administrative data systems such as the Medicare Statistical System could address separately. Consider an analysis of any set of policy options where total income and total expenditures for medical care are required. These two variables are common enough in any consideration of altering in eligibility requirements for a program, content establishing sliding scales for the forgiveness of copayments, or determining the cost to the Treasury for an option to cover a certain group above an arbitrarily selected medical expenditure amount.

Yet neither a household survey or an administrative system could provide information to address these fundamental questions above. The survey and administrative data would both suffer from lack of completeness for the expenditure variable; the survey from underreporting and the administrative data due to lack of inclusion of non-covered services, of processing error, and late claim filings. For the other variable, income; the survey would contain it and the administrative data would not. HCFA, through a contractual agreement, plans to continue to make such information available to inform the policy process with verified counts of program enrollees under the Medicare and Medicaid programs, best estimates of total expenditures and sources of payment and data which may only be collected from surveys such as health status and total income. Further, the Agency has embarked on an ambitious publication series program to rapidly disseminate such information to the interested public as well as program managers.

It is precisely the above logic, the requirement to collect certain necessary policy relevant data from surveys, that leads to the collection of complementary enrollment, use and expenditure data in both household and administrative surveys. Both data sets are incomplete but where they overlap in the area of covered services, it is necessary to develop a means of determining which one is correct, that is, which one best approximates the truth.

This operationally difficult procedure of arriving at an approximation of the truth concerning use and expenditure for covered services has traditionally been called "best estimation." Enrollment and claim data for persons in the household survey were derived from the Medicare Statistical System and the four States' Medicaid Management Information Systems. Enrollment data was collected for HHS persons from Medicaid Enrollment files in 38 States. Claims file construction and content limited attempts at event level matching to the following types of events: Medicare hospital claims and Medicaid hospital, doctor visits, and other medical expense reports. All other events were grouped and matched at the person level.

Three independent judges established the criteria for a match between house-hold and claims system reported, events employing a sample of individuals reported experience from both systems. These persons implicitly derived matching criteria from all the data available on each record ie. date of service, source of payment, place of service, name of respondent, provider name and address, and total charge. Thereafter, matching was conducted by a coding staff employing rules developed by the judges. Matched events then used data from the administrative record for the total charge and amount Medicare or Medicaid paid.

Of all the events from the survey and administrative data, some are matched. Others do not form survey administrative pairs. Certain events represented in the survey do not appear in the claim file. These events will not usually appear in the best estimate file if it is clearly a covered service used during a period when the person was covered by the program. These unmatched survey events may occur because of retroactive denial of a claim, use of a noncovered service, lack of program enrollment failure to file a claim, misreporting or vague respondent reporting, or lack of completeness in the administrative file. Administrative events unmatched to survey events were counted as visits in the best estimate files. Actual payment for a service is sufficient evidence for inclusion.

Thus, the matching procedure at the event level produces 3 classes or groups of reported services: matches, unmatched survey, and unmatched administrative events. Only one cell in Figure one remains unaccounted; those covered events which were porportedly represented in neither the survey nor the administrative data files containing survey records, administrative,

Figure One: Allocation of Survey and Administrative Events

Survey Events	Administrative Events: Yes	Administrative Events: No
Yes	Match	Unmatched Survey
No	Unmatched Adm.	Unaccounted

and best estimates records will shortly be available to the public. This approach was chosen to allow interested investigators the opportunity to choose between and among data sets as well as the option of creating their own best estimate strategy.

Special Application

The NMCUES survey design was embarked upon to meet certain specific programmatic needs. Other strategics for supplementing survey data with medical provider records and/or insurance company files have served the same general purpose in other projects. All such multiple data source project may later be evaluated to determine the degree of correspondence between the data sources for certain items and the pattern of relationship between items in the different data sources. For example, it might generally be acknowledged that there will be some difficulty associated with conducting a NMCUES cycle II in the future due to fiscal austerity. Should that be true, a series of studies could establish the relation between Medicare/Medicaid claims in the ARS and best estimate reports of covered and uncovered services and expenditures. Once established, this relation could be applied to the current and constantly updated Medicare Statistical System data and whatever State Medicaid data is available to estimate expenditure information for the desired period. Should methods of this sort along with other data source "Aging" techniques prove useful, then the cost of conducting surveys of program participants should be reduced by increasing the necessary period between surveys to check the relation between known events from current claims and data which may only be collected from surveys which include household and administrative reports combined in best estimates.

Conclusion

Best estimates techniques clearly make the range of policy activities which may be informed by reliable and accurate data much larger than those available from survey or administrative sources separately. This outcome strengthens the case for the combination of disparate data sources from administrative functions with household interview data in future data collection activities.

DEVELOPMENT OF A MACROECONOMIC-DEMOGRAPHIC MODEL OF HEALTH CARE EXPENDITURES

Joseph M. Anderson, ICF Incorporated

This paper describes a comprehensive, long-term macroeconomic-demographic model that is being developed for the National Institute on Aging (NIA) to investigate the effects of demographic and economic changes on the future level and composition of health care expenditures. The project is using the National Medical Care Utilization and Expenditure Survey (NMCUES), combined with other cross-section and time series data on the population, economy, and health care expenditures. This paper first describes the purpose of the project and the general approach. Second, it describes the existing NIA Macroeconomic-Demographic Model (MDM), with which the new health care expenditures model is being integrated. Third, it provides an overview of the health expenditures model.

PURPOSE AND APPROACH

The purpose of the project is to develop a long-term simulation model to study trends in health care expenditures, including the effects of:

- o demographic change,
- o long-term economic change,
- o public policy.

Because the level and types of health care services demanded differ greatly by age, sex, and other demographic characteristics, change in the size and structure of the population are associated with major changes in the level and composition of national health care expenditures. NIA is particularly interested in the effects of "population aging"--the absolute and relative increase in the numbers of older persons that will occur as the "baby boom" cohorts, formed during the period from the late 1940s through the mid-1960s, reach older ages.

Evolution of the economy over the long-term affects health care expenditures. The model is designed to study the effects of changes in relative prices, in household and national income levels, and in technology.

Public policy can significantly influence the level and composition of health care expenditures and the distribution of the costs of providing health care services. The model is designed to investigate the effects of alternative policies, such as alternative approaches to the financing of health care, alternative tax policy measures, and income maintenance programs, especially those affecting the income of the elderly.

To accomplish these objectives a comprehensive model of health care expenditures is being developed and integrated with the existing NIA Macroeconomic-Demographic Model. The focus of the project is the development of a structural model of the supply of and demand for various types of health care that takes into account input costs, technology, household characteristics, preferences, and incomes. The health expenditures model incorporates both cross-section data on individual households--including demographic and economic attributes and health care expenditures--and time series data on aggregate expenditures for various types of health care from various sources, other categories of consumer expenditures, and other economic variables. This health expenditures model is then integrated into the existing NIA Macroeconomic-Demographic Model, which includes a population projection model and a comprehensive representation of the operation of the labor market and the process of economic growth in a general equilibrium framework. The integrated health expenditures model takes advantage of the demographic detail and the general equilibrium framework of the existing Macroeconomic-Demographic Model.

THE NIA MACROECONOMIC-DEMOGRAPHIC MODEL

The existing Macroeconomic-Demographic Model is composed of a core macroeconomic and demographic modeling system and a set of five peripheral models that depict the operation and behavior of the major components of the retirement income system. The core model has three major parts: a population projection system, a macroeconomic growth model, and a labor market model. The five major elements of the retirement income system that are modeled are the Old-Age, Survivors and Disability Insurance System (OASDI), the private pension system, the public employee retirement system, the Supplemental Security Income (SSI) system, and the Medicare system.

The Population Model replicates the U.S. Census Bureau population projection methodology. It projects the total U.S. population by age and sex for each year from 1970 through 2055. Fertility rates, mortality rates, and net immigration are determined exogenously. The user specifies an ultimate completed cohort fertility rate. An appropriate set of age-specific fertility rates is then calculated and the corresponding population is projected. Mortality and immigration can also be varied exogenously by the user. The base population estimate is from the 1980 Census.

The Macroeconomic Growth Model is an adaptation of the Hudson-Jorgenson four sector long-term econometric forecasting model. It depicts the formulation of working, spending, and savings plans by households; and production, investment, and employment plans by businesses. It projects the demand for and supply of goods and services and depicts the equilibration of demand and supply by price adjustments and changes in consumption and production decisions. This long-term economic growth model is characterized by a more careful depiction of the

determinants of supply than most other econometric forecasting models, which focus on the determinants of aggregate demand. The data for the development of the macroeconomic growth model were derived primarily from the U.S. National Income and Product Accounts.

The demographically disaggregated Labor Market Model depicts three basic aspects of the labor market: the demand for labor; the supply of labor; and the simultaneous determination of labor and capital services input along with compensation, output, and employment. The derived demand for labor inputs is investigated by modeling the aggregate production technology of the private U.S. economy, focusing on the substitutability among age groups in the production process. Labor supply is measured in total annual manhours worked by each of twenty-two age-sex groups. Total manhours worked by each group is the product of the group's population, labor force participation rate, employment rate, and hours worked per year. The labor supply-demand system is fully integrated into and solved simultaneously with the Macroeconomic Growth Model for the input levels and prices of capital services and of labor, the unemployment and participation rates of each age-sex group, the level of output, consumption and investment, and other economic variables. This model is based on both establishment level data on employment, hours worked, and compensation of workers, and household level data on labor force participation, employment, and unemployment of individuals of various ages and sexes, collected by the Bureau of the Census and the Bureau of Labor Statistics.

Three pension system models and two Federal transfer models are integrated with the core Macroeconomic-Labor Market Model. In many instances, these models currently use fixed actuarial assumptions rather than a system of behavioral equations.

The Social Security Model depicts the determination of contributions into and benefit payments from the retirement (OASI) and disability insurance (DI) systems. Annual contributions are derived from the estimates of total compensation by age and sex generated by the Labor Market Model by estimating covered earnings and the taxable earnings base and applying statutory and projected tax rates. Total annual benefit payments are calculated by estimating the average benefit level and number of beneficiaries for each of fourteen benefit categories. A primary insurance amount (PIA) is estimated for individuals classified by year of birth, sex, and initial year of eligibility by applying the statutory provisions for the calculation of average indexed monthly earnings (AIME) to the hypothetical earnings records of typical individuals in each age-sex cohort and using the statutory benefit formula. Average payments for the fourteen types of benefits are keyed to the estimated primary insurance amounts. The model then projects balances for each of the trust funds each year from 1970 through 2055 and can be used to estimate the level of tax collections that would be required to finance projected benefits and the implications of alternative social security policies.

The Private Pension Model depicts the aggregate behavior of three types of pension plans--defined benefit plans, defined contribution plans, and individual retirement plans (IRAs, Keoghs, TSAs, etc.). Private pension coverage, participation, and vesting rates, by age and sex, were estimated using the Pension Supplement to the May 1979 Current Population Survey (CPS). Contributions to each type of plan are estimated by applying appropriate contribution rates to the earnings of each age-sex group. The model applies age-sex specific retirement benefit acceptance rates to estimate the population of beneficiaries and calculates average pension benefits by applying prototypical pension benefit formulae to the estimated earnings records (for defined benefit plans) or contributions (for defined contribution and individual plans) of the individuals of each age-sex cohort. Pension fund assets for each type of plan are derived from total contributions and benefit payments and the rate of return projected by the Macroeconomic Growth Model.

The Public Employee Pension Model specifies that all public employees are in defined benefit plans (in fact, 98 percent of public employee participants are so covered). Public employment is divided into seven sectors: Federal Civil Service, military officers, military enlistees, state and local hazardous duty, state and local general administrative, state educators, and local educators. The seven sectors are distinguished because the characteristics of the work forces and pension plans differ significantly among these groups. Given these distinctions, coverage, participation, and vesting rates and benefit acceptance probabilities are estimated using the same techniques as used by the Private Pension Model, drawing on actuarial data developed by the Federal Civil Service and the Department of Defense as well as data from the May 1979 Current Population Survey. Contributions and benefit payments are calculated separately for each sector of employment.

The Supplemental Security Income (SSI) Model depicts separately the operation of the programs for the blind, for the disabled, and for the aged. The eligible aged population is projected by applying an income distribution model to the population and earnings levels forecast by the Population and Labor Market Models to estimate the earnings, social security benefits, and other income of the elderly population. Age-specific SSI participation rates calculated from Social Security Administration (SSA) data are applied to the eligible population to estimate the number of beneficiaries. Average benefits are projected using data on average Federal benefit payments and average state supplements. Blind and disabled beneficiaries are estimated by age and sex applying historic incidence rates to the projected population.

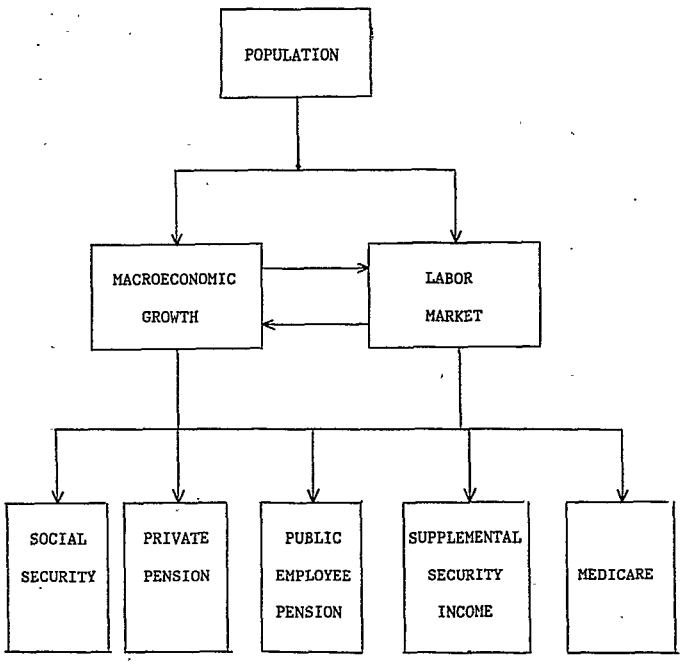
The existing Medicare Model forecasts revenues using the same approach as the Social

Security Model. Expenditures are forecast by disaggregating total beneficiaries and expenditures by age, sex and category of service. There are up to 26 age-sex groups for each of six services (inpatient hospital care, home health care, physician services, etc.). For each age-sex-service category expenditures are projected by applying estimated ratios of recipients per capita and expenditures per recipient to the projected population of the age-sex group. With the development of the health expenditures model the Medicare Model is being reformulated and a Medicaid Model is being added.

Figure 1 depicts the operational linkages between these models within the existing Macroeconomic-Demographic Model. At the start of any simulation year, the Population Model initially forecasts the new size and composition of the population. These population figures are principal inputs into the Macroeconomic Growth Model and the Labor Market Model, which operate simultaneously to project levels of aggregate economic activity and the labor market outcomes for twenty-two different age-sex groups. These projections of the state of the economy and the disaggregated labor market are inputs into the simulation of each of the three pension system Models and the two transfer income models currently included in the MDM.

FIGURE 1

STRUCTURE OF THE EXISTING
MACROECONOMIC-DEMOGRAPHIC MODEL



The Macroeconomic Growth Model and the Labor Market Model are the only models in the existing Macroeconomic-Demographic Model which actually simulate market processes. Each has one or more demand and supply relationships that jointly determine an equilibrium set of market outcomes. In the Macroeconomic Growth Model, Newton's method for solving a set of simultaneous equations is used to find the market equilibrium. In the Labor Market Model, the Gauss-Seidel method is used to solve the equations of the model. The major macro models--Population, Macroeconomic Growth, and Labor Market--employ many lagged variables in their equations, ensuring that one year's results play an important part in determining the next year's results.

Consistency is maintained between models because the outputs of the core macroeconomic modeling system serve as inputs to the pension models. In other words, the entire model operates from a consistent set of accounting relationships. All of the economic variables--compensation, employment, GNP, etc.--are defined identically throughout the model.

THE HEALTH EXPENDITURES MODEL

Development of the health expenditures model involves six tasks:

(1) Preparation of cross-section and time series data on health care expenditures that are mutually consistent and are consistent with the National Income and Product Accounts and with other data on consumer expenditures;

(2) Projection of providers' costs as functions of input prices;

(3) Estimation of consumers' prices from providers' prices;

(4) Development of the demand system for health care services;

(5) Linkage of the demand for health care expenditures to total personal consumption expenditures; and

(6) Translation of total health expenditures into payments from specific sources.

(1) The first step in the development of the health expenditures model is the preparation of comprehensive and consistent cross-section and time series data on health expenditures by type of service and by source of payment. We use as the basis for this system the National Health Accounts developed by the Office of Research, Demonstrations, and Statistics of the Health Care Financing Administration (HCFA). Each year from 1965 through 1981, HCFA has estimated health expenditures cross-classified into the types of services shown in Table 1 and the sources of payment shown in Table 2. We adjust these data to make them consistent with the National Income and Product Accounts (NIPA), published by the Department of Commerce, which

TABLE 1

NATIONAL HEALTH ACCOUNTS EXPENDITURE CATEGORIES BY TYPE OF EXPENDITURES

Health Services and Supplies
- Personal Health Care
  - Hospital Care
  - Physicians' Services
  - Dentists' Services
  - Other Professional Services
  - Drugs and Medical Sundries
  - Eyeglasses and Appliances
  - Nursing Home Care
  - Other Health Services
- Prepayment and Administration
- Government Public Health Activities

Research and Construction of Medical Facilities
- Research
- Construction

TABLE 2

NATIONAL HEALTH ACCOUNTS EXPENDITURE CATEGORIES BY SOURCE OF PAYMENT

Direct Payments

Third-Party Payments
- Private Health Insurance
- Philanthropy and Industrial In-Plant
- Government
  - Medicare (Federal)
  - Medicaid
    - Federal Expenditures
    - State and Local Expenditures
  - Other Public Assistance Payments for Medicare Care
    - Federal
    - State and Local
  - Veterans' Medical Care
  - Defense Department Medical Care
  - Workers Compensation
    - Federal Employees
    - State and Local Programs
  - State and Local Hospitals
  - Other Public Expenditures for Personal Health Care
    - Federal
    - State and Local
  - Government Public Health Activities
    - Federal
    - State and Local

provide the basic data for the macroeconomic growth model with which the health expenditures model is integrated.

HCFA now has underway an effort to disaggregate the expenditure type categories over several different age groups. HCFA has also developed a set of price indexes for the various types of health expenditures that are used to analyze past changes in expenditures and to forecast future changes. We use these indexes for the development of supply and cost equations for the health expenditures model.

The primary source of cross-section data for the development of the health expenditures model is the National Medical Care Utilization and Expenditure Survey (NMCUES) developed by the National Center for Health Statistics (NCHS). NMCUES was designed to collect data on the U.S. civilian non-institutional population during 1980. Information was obtained on health, access to and use of medical services, associated charges and sources of payment, and health insurance coverage. NMCUES consisted of three survey components. The national household survey comprised about 6,000 randomly selected households that were interviewed five times over 14 months beginning in early 1980. The State Medicaid household survey consisted of about 4,000 households selected from the Medicaid eligibility files in California, Michigan, New York, and Texas. Each household was interviewed five times over the 14 months during 1980-81. The administrative records survey was used to obtain information on program eligibility and payments for medical care for persons receiving Medicare and Medicaid. Data were obtained for approximately 31,000 persons in NMCUES--17,600 in the national household survey and 13,400 in the four State Medicaid samples. Both samples excluded people living in institutions, members of the Armed Forces, and people residing outside the United States.

In order to produce a data base suitable for estimation of the health expenditures model, the information in the NMCUES is aggregated into the types of expenditure and sources of payment shown in Tables 1 and 2. Because NMCUES includes only the civilian non-institutionalized population, the data must be augmented to include health care provided to members of the armed services and to institutionalized persons to make it consistent with the National Health Accounts.

(2) The second step in the development of the health expenditures model is to generate health care providers' prices as functions of prices of capital and labor services generated by the macroeconomic growth/labor market model. These are "reduced form" supply equations, since the prices of all goods and services can be expressed as functions of factor prices and the level of technology. These supply functions are specified to have a translog form. We use the price indexes in the National Health Accounts as proxies for provider prices, relating them to the prices of capital and of the services of labor of different age groups.

Subsidies of various types of health care are reflected in adjustments to factor prices in the equations for some services or adjustments to the providers' prices directly. Departures of the prices of some types of services from trends in average input prices are reflected in time trend variables or as biased technological change.

(3) The third step in construction of the health expenditures model is to translate providers' prices into prices paid by consumers. This requires estimation of a set of factors for each type of health service and for each demographic group giving the proportions of expenditures from private sources in each time period. This set of factors is first constructed for 1980, the year for which the NMCUES data are available. We then extrapolate the set of factors backward based on totals of public and private sources of funding from the National Health Accounts. These data are supplemented by the National Medical Care Expenditure Survey (NMCES), conducted in 1977, and the National Health Interview Surveys. Extrapolations of the price factors for future projections can reflect alternative policies for funding health care expenditures.

(4) The fourth step in construction of a health expenditures model is to develop an appropriate representation of the demand for health care services. We develop a two stage model, disaggregating total consumer expenditures into health care and other expenditures and then disaggregating health expenditures over the various types of services shown in Table 1. We estimate the share of health expenditures allocated to each type of service by each type of household as a function of the prices of the various types of health care services and the demographic characteristics of the households. The parameters of the set of allocation equations are first estimated from cross-section data on the allocation of expenditures among individual types of health care from the NMCUES.

The cross-section analysis provides estimates of the effects of demographic characteristics on health care expenditures. To estimate the effects of changes in prices on health care expenditures over time requires analysis of time series data. To develop a model appropriate for analysis of time series data, we aggregate the equations for each household to develop expressions for aggregate expenditure shares of each type of health care as functions of aggregate prices for each type of health care and the proportions of aggregate health care expenditures by each of the demographic groups. The estimation of aggregate prices takes into account the fact that different households pay different prices for the same health care service, depending on their demographic characteristics and income. We estimate this model by combining time series and cross-section data. The time series data are developed using the National Health Accounts data and using average utilization rates for demographic groups derived from the Health Interview Survey.

(5) The fifth step in constructing a health expenditures model is to link the demand for health expenditures to total consumption expenditures. First, a price index for health care as a whole for each household is constructed as a function of the prices of the types of health services and the quantities consumed by households with different demographic characteristics. Then, we allocate total consumption expenditures for each household between health care and other expenditure categories as a function of the aggregate prices of health care and of the other categories of goods and services. Estimates of the share of health expenditures in total household consumption expenditures are developed by using data on personal consumption expenditures on health care and on other goods and services from the 1972-73 Consumer Expenditure Survey (CES) and the National Income and Product Accounts.

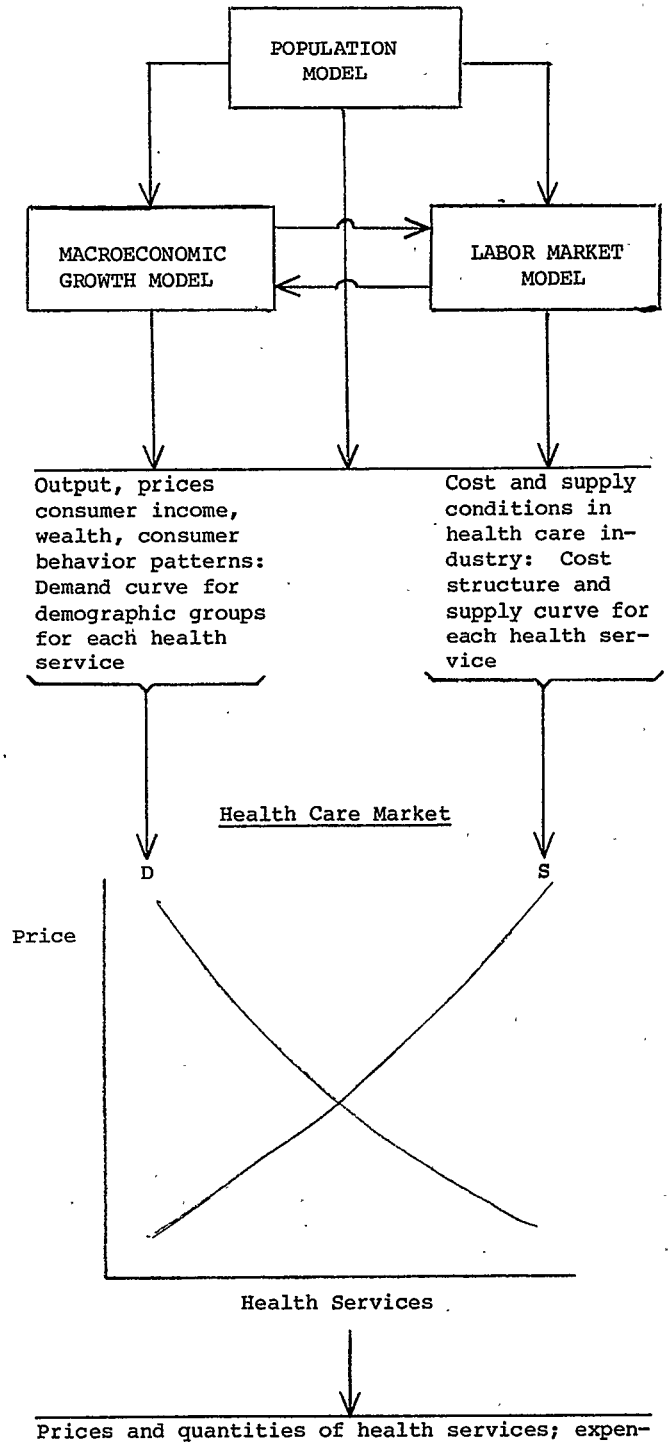
(6) The final step in constructing a model of health expenditures is to translate the total costs for each health service into payments from the various sources per unit of health care utilization. For this purpose, we estimate a set of factors for the proportion of expenditure on each type of health care that is provided by each source for each demographic group in each time period. This set of factors is similar to and is estimated in the same way as the factors described earlier in step (3) that are used to translate providers' prices into consumers' prices. The basic data for estimation of these factors are the estimates of expenditures by source of payment from the National Health Accounts. We use NMCUES to establish the distribution by sources of payment in the base year. This is supplemented by NMCES. NMCES included a survey of providers and a survey of employers and insurers, so it includes more information about sources of payment than NMCUES, which included only a survey of consumers.

To simulate the health expenditures model, we generate prices of capital and labor services and the price and quantity of aggregate personal consumption goods and services from the macroeconomic growth and labor market models. We generate providers prices from the supply equations described in step (2). We translate providers' prices into consumers' prices and project utilization of health services of the various types by each household using the methodology described in steps (3), (4), and (5). We then translate the providers' costs for health services into payments from each source to each type of household for each type of service. We then sum over services to estimate the total benefits paid to each group from each source and sum over households to obtain the total payments from each source for each type of service to obtain a projection of data in the form presented in the National Health Accounts. Finally, we sum payments over all services for each source to estimate payments from that source in each future year.

Figure 2 displays the structure of the proposed model schematically.

FIGURE 2

STRUCTURE OF THE NEW
MACROECONOMIC-DEMOGRAHPIC MODEL
OF HEALTH CARE EXPENDITURES



NATIONAL ESTIMATES OF THE SOCIAL AND DEMOGRAPHIC CHARACTERISTICS AND MEDICAL EXPENDITURES OF THE POPULATION WHO REPORT A MENTAL HEALTH CONDITION IN 1980

Marvin A. Feuerberg, Larry G. Kessler, Carl A. Taube, National Institute of Mental Health

INTRODUCTION

This paper presents national estimates of the population characteristics which differentiate those persons who report a mental health problem during 1980, and how that reporting is related to the volume of and expenditures for medical services. This preliminary report is based on data from a national household survey of about 6,600 families accounting for 17,123 people, the 1980 National Medical Care Utilization and Expenditure Survey (NMCUES). These preliminary findings represent a first analysis of the NMCUES data and illustrate, at least by implication, how this data set is useful for examining a number of questions with respect to mental health service system use.

This paper will be divided into four sections: 1) the research context that frames this research and the research to follow; 2) a description of the research design, data source, and limitations; 3) a presentation of findings; and 4) an outline of future research with the NMCUES data base.

RESEARCH CONTEXT

Consistent and significant findings from the health services research literature can be summarized as follows:

1. A substantial portion of the U.S. population - estimated to be at least 15% (Regier, 1978; Dohrenwend, 1980) - is affected by a mental disorder in a given year. It is estimated that, if evaluated, an additional 13% of the population would show severe psychological distress not accompanied by clinical psychiatric disorder, what has been characterized as "demoralization" (Dohrenwend, 1980).

2. A number of clinical and epidemiological studies have found not only a substantial portion of mental disorder among the U.S. population in general, but that rate varies by race, ethnicity, age, social class, and sex (Ilfeld, 1978; Dohrenwend, 1980). It should be noted that the variation in rates has been obtained generally from community studies and not from a national probability sample as available from the NMCUES.

3. About 60% of those estimated to have a mental disorder are seen in the outpatient medical sector (Regier, 1978). The prevalence of mental disorder in primary care practice has been estimated to be between 15% and 50% (Hoeper, 1979). Clearly, the prevalence of mental disorder among primary care users is higher than the general population.[1]

4. A substantial research literature points to a consistent association between reports of psychologic and physical distress as well as an association between both of these and physician utilization (Eastwood, 1972; Mechanic, 1976).

5. Patients with a mental disorder consume a disproportionately large share of general medical services (Liptzin, 1980). For example, a recent report of four medical programs in three different organized health care settings found that patients with a mental disorder diagnosis visited general medical departments from 1-1/2 to 2 times as frequently as patients without such a diagnosis (Hankin, 1982).

6. The causal determinants for this association between mental disorder (or distress) and medical utilization are open to interpretation, but the fact of the association is not in doubt. There are at least five major competing explanations for the association: that the somatic symptoms or physical illness causes or is a contributing cause in mental disorder and psychologic distress; that mental disorder (or distress) causes or is a contributing cause of somatic symptoms or physical illness; that both physical and mental disorder are concomitants found in people who have a generalized vulnerability to environmental stress; (in the case of those who actually seek mental health treatment) that use of mental health and general health services reflect "illness behavior", a generalized tendency to seek help; (in the case of those who actually seek mental health treatment) that the use of both types of services reflects opportunity or access to both types of care. Of course, none of the above are mutually exclusive; there is some research that has tried to examine a number of these competing explanations (Tessler, 1976; Mechanic, 1982).[2]

This higher utilization of medical services by patients with mental disorder is reflected, of course, in higher associated medical expenditures. The spiraling costs of medical care, exceeding the general inflation rate, and pressure from third party payers and policy makers to contain these costs, add a special urgency to research on these issues. This analysis provides an initial step in this process. We ask in this initial analysis of the NMCUES data: 1) What population characteristics differentiate those persons who report and those who do not report a mental health condition in 1980? 2) How does having a mental health condition relate to the volume of and expenditures for medical services? In sum, we hope to be able to describe high risk populations and their associated volume of and expenditures for medical care.

RESEARCH DESIGN

Data Source

The 1980 National Medical Care Utilization and Expenditure Survey (NMCUES) is the data

source for this paper. NMCUES is a household sample survey of the civilian, noninstitutionalized population. This preliminary report is based on data from about 6,600 families accounting for 17,123 people. Households were interviewed four or five times during 1980 and early 1981, a 14-month period, at approximately 3-month intervals. The survey provides detailed information on health conditions, insurance coverage, utilization and expenditures for health services for calendar year 1980. This survey has the advantages of providing data across all service settings, of including all medical utilization and not just covered services, and of including both utilization and expenditures.[3]

Mental Health Condition, a Key Variable

For this analysis we have used the analytic person file, with the addition of a key variable created out of the condition file - a variable which we call MHC, mental health condition - which has been appended to the person file for us by NCHS. Since this is a key variable in this initial analysis, it is important to note how MHC is defined and what we think it represents and what it does not represent.

There are a number of places in the NMCUES interview where "conditions", departures from well-being, can be reported: disability days, medical provider visits, hospital stays, and prescribed medicines. The interviewers record these "conditions" and these, in turn, are coded to the Ninth Revision, International Classification of Diseases (World Health Organization, 1977), as adapted for use with household surveys by the National Health Interview Survey (National Center for Health Statistics, 1979). If an individual has received any of the mental health codes, ICDA categories #290-316, they have what we have labeled as a mental health condition (MHC). This variable, then, attaches to an individual a mental health condition if that individual receives from the survey any of the ICDA codes #290-316 at any time during 1980. MHC is a self-report measure, a self-report that is not solicited from a list of pre-determined categories. It is not a clinical screening question, much less a clinical assessment. We believe it to be a measure of the individual's association of mental health with a visit or event (for example, disability). Using this measure, 5.6% of the population report a mental health condition in 1980. Given the estimates of mental disorder in the general population discussed above, MHC is clearly an underestimation. This is not surprising; we think a variety of reasons could account for this underestimation: the stigma attached to reporting a mental health condition; some mental disorders are not directly linked to a visit or event (for example, a personality disorder); some people have a tendency to somaticize their mental problems, and what is reported - fatigue, listlessness, insomnia - would probably not be recorded for these people as a mental health condition.

In sum, MHC is a self-report measure, a measure of the association of mental health with a visit or ill-health event, and the NMCUES provides the most recent national probability sample to address the question of subpopulation differences in reporting MHC as well as reliable estimates of the volume of and expenditures for associated medical services.

Limitations

A number of limitations with the NMCUES for the analysis presented here should be noted. First, the survey is limited with respect to the institutionalized aged. If someone over 65 is institutionalized for the entire year, or if they are institutionalized for part of the year but established a separate household after the date of the first interview, then they were excluded from the sample. Given the large number of the elderly in nursing homes, estimates for the noninstitutionalized elderly may seriously misrepresent the elderly population as a whole. Second, in addition to the important qualifications noted above with respect to MHC, different population groups are likely to differ with respect to their tendency to report symptoms, both physical and psychological, and their tendency to go to the doctor when they feel ill. For example, these differences are found when one compares women to men (Ingham, 1982). These differences will not only affect subpopulation estimates of MHC and the associated medical expenditures, but the total estimates are similarly affected.

FINDINGS

The findings are reported in Tables I-IV.[4-5] Table I reports (in column percents) the demographic, economic, regional and other population characteristics of persons who do (MHC) and do not report any mental health condition for 1980 (No MHC). With respect to the demographic characteristics of age, sex, race, education, and marital status, there is nothing particularly surprising. Women and the not married are overrepresented among those with MHC. Those who report a mental health condition are no more likely to be over 65 than those persons who do not report MHC.

With respect to economic characteristics of income, employment in 1980, and insurance coverage, we find that income does not seem to be related to MHC. Employment status, particularly whether the individual worked full-time, year-round was strongly related to MHC. There is somewhat greater concentration of Medicaid and Medicare recipients among those persons who report a mental health condition, a finding which suggests that a closer look at this population may be warranted.[6]

The other population characteristics which appear highly related to MHC are perceived health status and the number of ICDA codes reported during 1980.[7] Again, this is not surprising: people who report mental health problems are likely to report physical

conditions as well.

Table II controls simultaneously for age and sex which have been found to be important in a number of studies. (Note here that the column percents do not add to 100% because it is the age groups which are percented.) It should be noted that with the exception of the under 19 age group, in every age category the percent reporting MHC for females exceeds that of males; indeed with that one exception, any of the female age categories exceeds all of the male categories in reporting MHC.

Table III reports on the volume or utilization of services by mean annual visits. It should be noted that "other medical provider" is a catchall category that includes persons such as chiropractors, speech therapists, faith healers, psychologists, nurses, as well as medical and osteopathic doctors. We find that those who report MHC have from 1.3 to 3.9 times as many visits as those who do not report MHC.[8] The differences between the means are consistent with previous research. The distribution of MHC and No MHC around discrete visit categories is as indicated by the means: No MHC concentrated among the low visit categories and MHC more concentrated among the high visit categories. For example, 13.4% of those persons with MHC had zero visits, in contrast to 34.4% of those with No MHC. At the other extreme, the visit category of 20 and over, 5.6% of those with MHC were represented here, in contrast to .8% of the No MHC category.

Although not presented here because of space limitations, we also controlled for age and sex and found that for both males and females, the visit ratio of MHC/No MHC declines for the various visit categories as age categories increase. For example, with respect to physician visits for females, the ratio is 2.8 for ages 0-18, and that ratio declines uniformly to 1.6 for those over 65.

Table IV reports on the expenditures for medical services over various cost categories.[9] We find that 1.7 to 4.5 times as much was spent by those reporting MHC compared to persons not reporting MHC. As in the visit data, we controlled for both age and sex and found that for females, but not for males, the ratio declined for each of the cost categories as age increases. For example, for females ages 0-18 with respect to total costs, the ratio of MHC/No MHC is 3.0. The ratio declines to 1.9 for females ages 19-34, to 1.8 ages 35-49, to 1.4 ages 50-64, and to 1.2 ages 65 and over.

In sum, we have found a number of population characteristics which identify a high risk population, at least in the sense of reporting a mental health condition. In addition, this mental health condition is substantially related to the volume and expenditures for medical services. Most important, these findings have been found for a national probability sample of the U.S. population.

FUTURE RESEARCH

As we stated earlier, the findings reported above are preliminary, highly descriptive, and we plan more extensive analysis in the future. Some of the important research issues are listed below.

First, with respect to the analysis presented in this paper, the key variable, mental health condition (MHC), may be modified for further research so that some of the problems with this measure might be remedied. For example, MHC might include persons who do not report MHC but who have received a psychotropic drug or treatment from a mental health specialist.

Second, an analysis will be directed to the characteristics of mental health use or treatment. The focus here, in contrast to the analysis presented in this paper, will be on the rate and volume of use rather than users. In addition to overall prevalence rates, differences between the various population groups that we addressed above could be directed to inpatient vs. outpatient use, general vs. specialty sector use, and psychiatrist vs. other MD vs. use of other medical providers.

Third, another analysis might be addressed to the relation of use to need for the various population groups. The key variable here, of course, would be some measure of need. Possible definitions might be "mental health condition" or "mental health condition together with disability associated with it." Hence, one focus of analysis might differentiate groups by the number of mental health visits per 100 bed-disability days.

In addition to the highly descriptive studies noted above, a number of additional analytic efforts are clearly possible with the NMCUES data. For example, a study might be directed to the costs of an episode of mental health treatment. The determinants, magnitude, and variability of such costs may be of particular interest to policy makers who have been considering the inclusion of mental health treatment in various proposals of reimbursement by Diagnostic Related Groups (DRGs).

The above list of studies is clearly not exhaustive. However, it should provide some indication of research questions with respect to the mental health service system that can be profitably addressed with the NMCUES data.

Table I. Population Characteristics of Persons[1] Who Do (MHC) and Do Not Report Any Mental Health Condition (No MHC)

Population Characteristics	MHC % (Column Percent)	No MHC % (Column Percent)	Population Characteristics	MHC % (Column Percent)	No MHC % (Column Percent)
Age in Years			Insurance coverage[4]		
0 - 18	16.5	31.0	No insurance	6.6	9.1
19 - 34	28.2	26.9	Medicaid	8.4	5.7
35 - 49	22.3	15.7	Medicare	13.9	9.1
50 - 64	19.2	14.3	Private	63.1	71.5
65 - and older	13.8	12.1	Other	8.1	4.6
Sex					
% Female	62.9	51.0	Region		
			Northeast	25.4	20.7
Race/Ethnic			North Central	22.2	26.8
% Black	6.7	12.0	South	29.2	31.3
% Hispanic	5.5	6.9	West	23.2	21.2
Education[2]			Residence		
0 - 8	19.5	14.8	SMSA - Central City	29.4	29.2
9 - 12	51.2	55.0	SMSA - Not Central City	43.3	39.8
13+	29.3	30.2	Non-SMSA, Urban	13.6	13.9
			Non-SMSA, Rural	13.7	17.1
Marital Status					
Under 17 years	14.0	29.0	Number of ICDA Codes		
Married	46.0	43.8	Reported During 1980:		
Not married or unknown	40.0	27.2	0 -	0	15.7
			1 - 3	27.0	53.4
Income[3]			4 - 9	60.6	28.6
Poor	15.4	12.0	10 - 12	8.5	1.7
Other low income	17.5	16.0	13 and over	3.9	.6
Middle income	24.4	29.0			
High income	42.6	43.0	Perceived Health Status[5]		
			Excellent or good	70.7	87.2
Employment in 1980			Fair or poor	29.2	12.0
Worked full-time, year round	25.7	37.7			
Worked some of the time	37.2	34.3			
Did not work in labor force	4.8	2.7			
Not in labor force	32.4	25.3			

Notes: 1. Weighted
2. Under 17 years of age excluded
3. Income has been adjusted for family size
4. Insurance categories are mutually exclusive
5. Does not add to 100% because unknown category is not shown

Table II. Percent Reporting Any Mental Health Condition (MHC) by Age and Sex

		(%[1] MHC who report)
Males		
Age in years:	0 - 18	3.4
	19 - 34	4.5
	35 - 49	5.1
	50 - 64	5.4
	65 and older	4.1
Females		
Age in years:	0 - 18	2.8
	19 - 34	7.2
	35 - 49	10.3
	50 - 64	9.2
	65 and older	8.0

Table III. Mean Annual Visits of Persons[1] Who Do (MHC) and Do Not Report Any Mental Health Condition (No MHC)

	Mean Visits			
	Physician	Physician, (Conditional)[2]	Other Medical Provider	Emergency Room
MHC	5.9	6.9	3.5	.4
No MHC	2.6	4.0	.9	.3
Ratio: MHC / No MHC	2.3	1.7	3.9	1.3

Table IV. Mean Annual Medical Expenditures of Persons[1] Who Do (MHC) and Do Not Report Any Mental Health Condition (No MHC)

	Mean Costs in Dollars[1]			
	Total Costs	Physician Visits	Other Medical Provider	Hospital Stay
MHC	$ 1336	$ 185	$ 95	$ 587
No MHC	$ 682	$ 77	$ 21	$ 343
Ratio: MHC / No MHC	2.0	2.4	4.5	1.7

1. Weighted
2. Mean physician visits of persons who have had at least one visit.

FOOTNOTES

1. For a review of the research in this area, see: National Institute of Mental Health, Series D, No. 5, Mental Disorder and Primary Medical Care: An Analytical Review of the Literature, DHEW Publication No. (ADM) 78-661. Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402, 1979.

2. In addition, there has been a considerable research effort on the "offset" effect, that is, the question of whether alcohol, drug abuse or mental health (ADM) treatment reduces or "offsets" subsequent medical care utilization. For a review of research in this area, see Jones, K. and Vischi, T.: Impact of Alcohol, Drug Abuse and Mental Health Treatment on Medical Care Utilization: A Review of the Literature. Med Care, 17 (Suppl.): entire issue, 1979.

3. For detailed information on the NMCUES collection and reporting procedures, see: National Center for Health Statistics, G.S. Bonham: Procedures and Questionnaires of the National Medical Care Utilization and Expenditure Survey. National Medical Care Utilization and Expenditure Survey. Series A, Methodological Report No. 1, DHHS Pub. No. 83-2001. Public Health Service. Washington, U.S. Government Printing Office, March 1983.

4. The reader will note that we have not reported significance levels. At this point, we have not worked out the variances and estimation procedures for the weighted data reported in the Tables because of the complex sampling design of NMCUES. That will be our next step. We expect, however, most of the differences to be significant, given the large N's. What we are more concerned about are differences that are substantial and that the patterns make sense.

5. See the last section of this paper for an outline of future research. Estimates presented in the Tables are for weighted data rounded to the nearest .1%. A weight has been applied to each individual in the NMCUES in order to transform the data into a sample representative of the 1980 civilian noninstitutionalized population.

6. This may be particularly useful, given that in addition to household interview data, an administrative records component obtained data for individuals in the household survey who were on Medicaid and Medicare. Data was also obtained from providers who served these individuals.

7. Of course, the relationship between MHC and number of ICDA codes is contaminated because to receive MHC, an individual must receive at least one ICDA code. Nevertheless, even if we have to subtract an average of one code from each MHC individual, the association would remain strong.

8. One must interpret these findings with caution, however: the physician visit category is somewhat contaminated because some individuals may have received a MHC in conjunction with a visit. This problem is somewhat remedied by comparing the MHC/No MHC for those persons who have had at least one visit (physician, conditional). Here we find the ratio declines from 2.3 to 1.7. This adjustment is not entirely satisfactory, however, because some number of visits for the No MHC category is not for any condition. Here it might be useful to compare MHC/No MHC for those who report at least one condition. The above problem results in a possible bias in the direction of indicating a stronger association between MHC and visits. There are factors, however, that possibly bias these findings in the opposite direction. For example, we noted the underreporting of MHC because of possible stigma. In addition, it is likely that a large number of the No MHC category consists of individuals who are receiving mental health treatment from a general physician in the form of psychotropic drugs.

9. Due to space limitations and what is of theoretical relevance to MHC, we have not reported on all cost categories. Cost categories such as dental visits and prescribed medicines are excluded. Hence, the cost items presented do not total to an amount equal to "total costs."

REFERENCES

1. Dohrenwend, Bruce P., et al: Mental Illness in the United States: Epidemiological Estimates. New York, Praeger Publishers, 1980.

2. Eastwood, M.R., Trevelyan, M.H.: Relationship Between Physical and Psychiatric Disorder. Psychological Medicine 2:363-372, 1972.

3. Hankin, Janet, et al: Use of General Medical Care Services by Persons with Mental Disorders. Arch Gen Psychiatry 39:225-231, 1982.

4. Ilfeld, Frederic W.: Psychologic Status of Community Residents Along Major Demographic Dimensions. Arch Gen Psychiatry 35:716-724, 1978.

5. Ingham, J., Miller, P.: Consulting with Mild Symptoms in General Practice. Soc Psychiatry 17:77-88, 1982.

6. Liptain, Benjamin, et al: Utilization of Health and Mental Health Services in a Large Insured Population. Am J Psychiatry 137:5, 1980.

7. Mechanic, David, et al: Distress Syndromes, Illness Behavior, Access to Care and Medical Utilization in a Defined Population. Med Care 20:361-372, 1982.

8. Regier, Darrel A., et al: The DeFacto US Mental Health Services System. Arch Gen Psychiatry 35:685-693, 1978.

9. Tessler, Richard, et al: The Effect of Psychological Distress on Physician Utilization: A Prospective Study. J Health Soc Behav 17:353-364, 1976.

Recent Uses of Vital Statistics

Session S

SOCIOECONOMIC INDICATORS ON DEATH CERTIFICATES

Gene D. Therriault, New York State Department of Health

INTRODUCTION

As noted by Rosenberg and McMillen earlier this year at an international meeting on socioeconomic differentials in mortality (1), there has emerged an increasing interest in the socioeconomic factors associated with mortality. Part of the dramatic reduction in mortality rates experienced in the industrialized countries has been attributed to improvements in the standard of living as well as advances in medical care and the development of more meaningful and far reaching public health programs and policies. However, attempting to isolate the contribution of socioeconomic factors to this improvement has proved to be a difficult task.

The first decision those of us interested in this type of activity must make is what data base to utilize to study socioeconomic differentials in mortality. The most complete data bases of deaths in the United States are the state administered Vital Records Registration Systems. Reporting is believed to be nearly complete and the quality of information collected has proved to be high. The decision on how best to segment the population according to socioeconomic status for analytic purposes is made difficult because of two factors. First, a definition of what is meant by socioeconomic status and how it can best be measured has certainly not been universally agreed upon. Secondly, even if such a definition existed, the measurements necessary to quantify such status would probably not be available on death certificates where the number of potential indicators is limited.

Kitagawa (2) summarizes the first of these problems when, referring to such characteristics as marital status, race, ethnic group, education, income, occupation and housing unit of residence, she states "Although many of these characteristics are related to a person's socioeconomic status - however it may be defined - no one of them is directly equated with socioeconomic status in the sense that it is accepted as the sole determinant of such status or that it is a fully satisfactory index of socioeconomic status for all research purposes." Stockwell (3) warns of the extent of the problem when he states, "The way in which socioeconomic status is defined (for example, in terms of income as opposed to education or occupation or some combination of these, or even other variables) will largely determine the nature and extent of any resulting relations."

AVAILABLE SOCIOECONOMIC MEASURES

The second of these problems, the sparsity of socioeconomic indicators on death certificates, can be fully appreciated by reviewing the United States Standard Certificate of Death, as developed and recommended for use by the National Center for Health Statistics. The only possible indicators of socioeconomic status contained on this document are residence data, race and occupation and industry.

Residence data allows for the application of United States Census Data to vital statistics whereby the general socioeconomic characteristics of a small geographic area such as a census tract or minor civil division are used as an estimate of the particular socioeconomic status of all individuals residing in that area. Numerous indices, based upon various data collected on the census, have been utilized by researchers (4,5,6,7). Obviously, the more homogeneous a particular area, the better this method works. The major problem with this approach relates to the timeliness of the data. Given that the census is taken only once in every ten years, one is often forced to characterize geographic areas using outdated data.

Data pertaining to occupation, as collected on vital records, is subject to some criticism as to its potential use in socioeconomic stratification. Two principal objections are the difficulty in classifying various occupations in some sort of heirarchy according to status, and secondly in trying to interpret just what is being collected: usual occupation, most recent occupation, etc. Despite these difficulties occupation is generally recognized as an excellent source of information pertaining to socioeconomic status (8) if collected and classified properly.

Race has been a point of major discussion as to whether any observed differences in health status among the races are due to ethnic differences or socioeconomic differences. Hendricks (9) for example, presents evidence which, is claimed, supports the hypothesis that differences between whites and nonwhites in reproductive efficiency are "...predominantly socioeconomic rather than ethnic."

The death certificate used by the New York State Department of Health contains one additional variable for use in socioeconomic analysis which is not contained on the U.S. Standard Certificate: education of the deceased. Education level, perhaps more than any

one indicator, has been relied upon as a measure of socioeconomic status (2,10,11,12). It is easily classifiable, has a natural order to it, and when asked for is fairly accurately reported. One potential problem with the reporting of educational attainment on the death certificate relates to the ability of the informant to correctly report the educational level achieved by the deceased. This problem would appear to be most prevalent in the case of very elderly decedents with no surviving spouse.

MATERIALS

The New York State Department of Health, Bureau of Health Statistics, administers the vital records registration system for New York State, exclusive of New York City, a geographic area referred to as upstate New York. The data base utilized for the present study includes all 1980 and 1981 deaths recorded in upstate New York to residents of upstate New York between the ages of 18 and 64 years. There are a total of 46,053 deaths available for analysis. Occupation is coded according to the three digit 1980 Census Bureau classification. Our Department currently is engaged in a Cooperative Agreement with the National Institute for Occupational Safety and Health to establish an occupational health surveillance system based upon vital records. The data being developed is being used to investigate relationship between occupational exposure and health status but it can also be utilized to investigate relationships between socioeconomic status and health status. Education level is a data item routinely entered onto our computer files.

METHODOLOGY

Recalling the Stockwell statement that emphasizes the critical nature of selecting an appropriate measure of socioeconomic status, two variables available on the New York State Certificate of Death are compared for consistency in analyzing mortality data by socioeconomic status. The education level and the usual occupation of the deceased are independently utilized to establish socioeconomic stratifications. Sex specific, age standardized proportional mortality ratios (13) are calculated for 1980 and 1981 upstate New York deaths between the ages of 18 and 64, segmented into ordered categories according to educational level and occupational status. Such ratios are calculated for specific major causes of deaths.

For educational level, four categories are used: 0-11 years of schooling completed, 12 years, 13-15 years and 16+ years. Occupation is classified according to an adaptation of a scheme proposed by Nam and his co-workers using the 1970 Census Bureau coding system (14). Using a table of weighted averaged occupational status scores summarized for the major occupational categories, by Powers and Holmberg (15), we adapted the Nam system for 1970 to the 1980 coding system used for our data set.

THE CROSS CLASSIFICATION OF EDUCATION AND OCCUPATION

Table 1 presents educational level cross classified by occupational status score for male deaths. If one considers those combinations on the diagonal or within one of the diagonal to be in agreement, approximately 80% of the 25,788 male deaths with both variables stated and occupation codable are similarly categorized by the two methods. An interesting interpretation of this table arises if one considers education level as a measure of "potential" and occupation as a measure of "achievement." Of the 5,155 deaths whose education level and occupation status do not agree, 509 (10%) are below the diagonal and 4,646 (90%) are above the diagonal. Those below the diagonal could be considered those individuals whose education level would indicate a higher socioeconomic status than actually achieved while those above the diagonal could be considered those whose education level would indicate a lower socioeconomic status than actually achieved. Interestingly, there are more than 9 times as many over achievers as under achievers.

Table 1 also points out an advantage to the use of education as an indicator of socioeconomic status as opposed to occupation. While there are 2,148 records with occupation not stated or not codable or indicating unemployed or retired, only 899 records have the education item not completed. Since education is a self-coded item, one does not encounter difficulties in assigning codes to vague entries as one often does in coding occupation.

Table 2 presents similar data for 17,080 female deaths between the ages of 18 and 64. For those records with both variables stated and occupation codable to a socioeconomic class, agreement is slightly better for the females: 86% as opposed to 80% observed for males in Table 1. Of the 1,224 whose education and occupation do not agree, 212 (17%) are below the diagonal and 1,012 (83)% are above the diagonal. While these figures are similar to those observed for males, it should be pointed out that there is a higher proportion of well educated females in lower status jobs than was observed for males.

Table 2 also emphasizes a real difficulty encountered when using

occupation data to characterize socioeconomic status for female deaths - the entry of "housewife" as a usual occupation. Approximately 46% of all the female death records could not be assigned an occupational status score and the main contributor to this problem is the entry of "housewife." This may become less of a problem in the future as more and more women enter the labor force but, for now, it is a serious shortcoming which must be recognized. The completion rate for education is approximately the same for both males and females, over 96%.

STANDARDIZED PROPORTIONAL MORTALITY RATIOS

The standardized proportional mortality ratio (SPMR) is calculated as a measure of excess cause specific mortality among subgroups of the population. Using the 1980 and 1981 age and sex specific total distribution of deaths as a standard, expected number of deaths are calculated. This statistic is a particularly useful device to monitor the changing patterns of current mortality among subgroups of the population for which the populations at risk are not readily available.

Table 3 summarizes the results of the SPMR calculations. For deaths due to motor vehicle accidents for females, both indicators reveal a tendency for higher mortality among higher socioeocnomic groups. This may be reflective of more access to automobiles among the higher social class. It should also be noted that all of the occupation groupings for females have SPMR's in excess of unity. This is due to the exclusion of housewives from any of the occupation groupings. They are included in the noncodable group. Thus, any female with a usual occupation, and therefore more likely to frequently use a motor vehicle, is at higher risk of death due to a motor vehicle accident than those with no usual occupation.

The data for acute myocardial infarction among females reveals a definite inverse relationship between mortality and socioeconomic status, regardless of the SES indicator used. This is not observed among males where there exists only a very slight direct relationship. This could be suggestive of females being better able to physically cope with the stress often associated with higher socioeconomic status. For cerebrovascular deaths, there is excess mortality among the lower socioeconomic group with a more pronounced trend noted for males.

While male deaths due to diabetes do not exhibit any strong patterns, a definite inverse relationship between socioecnomic status and mortality is documented among females. This finding is consistent for both educational subgroups and occupational subgroups. Among the highest socioeconomic class, the SPMR's are both below one-half. The highest prevalence of diabetes has been shown to occur among nonwhite females and this data is supportive of these previous findings.

Education and occupation yield similar results when male homicide deaths are considered. Substantial excess mortality is exhibited among the low socioeconomic group. Yet no such consistency is observed among females. Using education there appears to be a tendency for higher mortality among the lower classes but this is not seen if the population is segmented according to occupation.

Many of the site specific cancers show a direct relationship between socioeconomic staus and mortality. The one exception to this is for neoplasms of the respiratory system where an inverse relationship is exhibited, particularly among males. This may be reflective of more occupational exposure among many of the lower socioeconomic workers or differences in the smoking habits between members of the higher and lower social groups.

The final cause contained in Table 3 is pneumonia. Male mortality is inversely related to socioeconomic status, using either indicator. However, female mortality shows a considerably different pattern for education than for occupation. While the least educated women show considerable excess mortality, all occupation subgroups have SPMR's less than unity. This again may be caused by the exclusion of housewives from any of the occupation subgroups.

SUMMARY

In summary, data available on certificates of death which can be used as measures of socioecnomic status are limited. The U.S. Standard certificates contain only data on residence, race and occupation. The New York State Department of Health's certificate has additionally the education of the decedent.

Analyzing a data set of 46,053 death records indicates a moderate degree of agreement between the results when using either education or occupation as a surrogate for socioeconomic status. There are inconsistencies, however, which indicate that careful consideration must be given in choosing suitable indicators. The analysis of the current data shows more agreement between the two measures when the relationships with cause specific mortality are stronger.

Education has the advantage of being more completely reported and self-coded. Occupation coding is a major task, requiring a large well trained

staff. Many occupation entries cannot be assigned codes in a socioeconomic heirarchy and the entry of "housewife" eliminates close to half of all female deaths from categorization. These arguments for the use of education are contrasted by the argument that education may not be a true measure of achieved socioeconomic status for many individuals and occupation is a more meaningful criteria.

Future advancements in the development of socioeconomic measures in the analysis of mortality data will certainly center on education and occupation. One recommendation that I would have, based upon my work in this area, is the inclusion of education on the U.S. Standard Death Certificate and the adoption of this item by individual states as their death certificates are revised. The ease of collection and the potential value of the variable warrants its inclusion as a collectible item.

REFERENCES

1. Rosenberg H; and McMillen M; New Research Directions on Socioeconomic Differentials In Mortality: United States of America; Presented at the Third Meeting of the UN/WHO/CICRED Network on Socio-Economic Differential Mortality in the Industrialized Societies; Rome, Italy; May 24-27, 1983.
2. Kitagawa, EM: Socioeconomic Differences in Mortality in the United States and Some Implications for Population Policy. In Commission on Population Growth and the American Future, Research Reports, Volume I, Demographic and Social Aspects of Population Growth. Edited by C.F. Westoff, R. Parke, Jr. Washington, D.C., U.S. Government Printing Office: 87-110, 1973.
3. Stockwell, EG: Use of Socioeconomic Status as a Demographic Variable. Public Health Reports, 81: 961-966, 1966.
4. Ilves, A: Socioeconomic Stratification Applied To The Analysis of Health Problems. State of New York Department of Health Quarterly Vital Statistics Review: 11-21, July-September, 1977.
5. Smith, JC; Fry, RE; Pace, ME: A Socioeconomic Classification Technique for Standard Metropolitan Statistical Areas. Atlanta, GA, Center for Disease Control, 1973.
6. Willie, CV: A Research Note on the Changing Association Between Infant Mortality and Socioeconomic Status. Social Forces, 37: 221-227, 1959.
7. Stockwell, EG: A Critical Examination of the Relationship Between Socioeconomic Status and Mortality. American Journal of Public Health, 53: 956-964, 1963.
8. Stockwell, EG: Socioeconomic Status and Mortality in the United States. Public Health Reports, 76: 1081-1086, 1961.
9. Hendricks, CH: Delivery Patterns and Reproductive Efficiency Among Groups of Differing Socioeconomic Status and Ethnic Origins. American Journal of Obst. and Gynec., 97: 608-624, 1967.
10. Socioeconomic Differentials in Morbidity. Metropolitan Life Statistical Bullentin, June 1972.
11. Kitagawa, EM; Hauser, PM: Education Differentials in Mortality by Cause of Death, United States, 1960. Demography, 5: 333-353, 1968.
12. Nam, CB: Socioeconomic Status and Mortality: Needed Research. Presented at the Annual Meeting of the Population Association of America; Seattle, Wash.; April 17, 1975.
13. Kupper, LL; McMichael, AJ; Symons, MJ; Most, BM: On the Utility of Proportional Mortality Analysis. Journal of Chronic Diseases, 31: 15-22, 1978.
14. Nam, CB; LaRocque, J; Powers, MG; Holmberg, J: Occupational Status Scores: Stability and Change. Presented at the Annual Meeting of the American Statistical Association, Atlanta, GA; August, 1975.
15. Powers, MG; Holmberg, JJ: Occupational Status Scores: Changes Introduced by the Inclusion of Women. Demography, 15: 177-182, 1978.

Table 1
Resident Deaths Recorded in Upstate New York
by Education and Occupational Status
Ages 18-64
1980 and 1981
Males

Education in School Years Completed	Occupational Status Score I (Low)	II (Lower middle)	III (Upper middle)	IV (High)	Total Stated	Not* Stated	Total
0-11 (Low)	2,356	2,160	2,259	418	7,193	747	7,940
12 (Lower Middle)	2,723	2,747	5,356	1,969	12,795	885	13,680
13-15 (Upper Middle)	343	221	1,102	933	2,599	326	2,925
16+ (High)	97	69	544	2,491	3,201	190	3,391
Total Stated	5,519	5,197	9,261	5,811	25,788	2,148	27,936
Not Stated	286	176	292	145	899	138	1,037
Total	5,805	5,373	9,553	5,956	26,687	2,286	28,973

* includes occupations not stated, not codable or nonclassifiable into an occupational status group (e.g. housewife, student, retired)

Table 2
Resident Deaths Recorded in Upstate New York
by Education and Occupational Status
Ages 18-64
1980 and 1981
Females

Education in School Years Completed	Occupational Status Score I (Low)	II (Lower middle)	III (Upper middle)	IV (High)	Total Stated	Not* Stated	Total
0-11 (Low)	812	471	359	110	1,752	2,199	3,951
12 (Lower Middle)	1,171	734	2,475	543	4,923	4,328	9,251
13-15 (Upper Middle)	160	34	588	367	1,149	659	1,808
16+ (High)	44	8	193	867	1,112	390	1,502
Total Stated	2,187	1,247	3,615	1,887	8,936	7,576	16,512
Not Stated	86	71	98	25	280	288	568
Total	2,273	1,318	3,713	1,912	9,216	7,864	17,080

* includes occupations not stated, not codable or nonclassifiable into an occupational status group (e.g. housewife, student, retired)

Table 3
Resident Deaths Recorded in Upstate New York
Standardized Proportional Mortality Ratios
Ages 18-64
1980 and 1981

Cause: Motor Vehicle Accidents					
SES Indicator		Low ----------			High
Education	Males	0.82	1.09	1.03	1.01
	Females	0.81	1.02	1.00	1.36
Occupation	Males	1.01	1.14	1.06	1.10
	Females	1.17	1.29	1.20	1.38
Cause: Other Accidents					
SES Indicator		Low ----------			High
Education	Males	1.08	0.98	1.06	0.84
	Females	1.04	1.00	1.04	0.84
Occupation	Males	1.15	0.99	0.97	0.82
	Females	1.04	0.91	0.80	0.88

Cause: Acute Myocardial Infarction

SES Indicator		Low			High
Education	Males	0.97	1.01	1.02	1.03
	Females	1.20	0.98	0.83	0.67
Occupation	Males	0.87	1.03	1.06	1.04
	Females	1.06	1.23	0.90	0.79

Cause: Cerebrovascular Disease

SES Indicator		Low			High
Education	Males	1.12	1.03	0.87	0.70
	Females	1.04	1.01	0.97	0.96
Occupation	Males	1.07	1.07	0.98	0.89
	Females	1.10	1.27	0.98	1.01

Cause: Other Ischemic Heart Disease

SES Indicator		Low			High
Education	Males	0.87	1.03	1.09	1.10
	Females	1.10	1.02	0.83	0.70
Occupation	Males	0.98	0.95	0.99	1.09
	Females	1.00	1.05	0.97	0.84

Cause: Cirrhosis of the Liver

SES Indicator		Low			High
Education	Males	1.10	1.03	0.73	0.83
	Females	0.99	1.01	1.05	0.93
Occupation	Males	1.28	0.89	0.97	0.88
	Females	1.04	0.79	0.81	0.88

Cause: Diabetes

SES Indicator		Low			High
Education	Males	1.10	0.94	1.05	0.94
	Females	1.33	0.94	0.97	0.46
Occupation	Males	0.97	1.21	0.95	0.90
	Females	1.16	0.86	0.70	0.44

Cause: Homicides

SES Indicator		Low			High
Education	Males	1.58	0.84	0.78	0.67
	Females	1.31	1.02	0.78	0.78
Occupation	Males	1.45	0.87	0.83	0.80
	Females	1.17	1.04	0.97	1.10

Cause: Malignant Neoplasms of the Digestive Organs and Peritoneum

SES Indicator		Low			High
Education	Males	0.88	0.98	1.27	1.26
	Females	0.86	1.02	1.09	1.07
Occupation	Males	0.90	0.96	1.00	1.18
	Females	0.96	1.03	1.19	1.13

Cause: Malignant Neoplasms of the Genital Organs

SES Indicator		Low			High
Education	Males	0.65	1.11	0.88	1.52
	Females	1.02	0.98	1.08	1.08
Occupation	Males	0.84	1.04	1.03	1.12
	Females	1.09	1.15	0.97	1.18

Cause: Malignant Neoplasms of the Lymphatic and Hematopoietic Tissues

SES Indicator		Low			High
Education	Males	0.75	0.95	1.36	1.48
	Females	0.71	1.04	1.03	1.34
Occupation	Males	0.84	0.83	0.99	1.28
	Females	0.93	0.99	1.13	1.16

Cause: Malignant Neoplasms of the Respiratory System

SES Indicator		Low			High
Education	Males	1.11	0.99	0.91	0.79
	Females	1.04	1.01	0.92	0.97
Occupation	Males	1.03	1.11	1.05	0.87
	Females	1.04	1.11	1.12	0.95

Cause: Malignant Neoplasms of the Breast

SES Indicator		Low			High
Education	Females	0.70	1.01	1.22	1.49
Occupation	Females	0.74	0.84	1.20	1.31

Cause: Pneumonia

SES Indicator		Low			High
Education	Males	1.40	0.86	0.83	0.47
	Females	1.52	0.90	0.57	0.68
Occupation	Males	1.23	0.70	0.95	0.66
	Females	0.81	0.78	0.80	0.54

CHANGE IN SMOKING AND DRINKING BEHAVIOR AMONG MOTHERS OF LIVE-BORN INFANTS

Kate Prager, National Center for Health Statistics; Henry Malin, Carol Graves, Danielle Spiegler, National Institute on Alcohol Abuse and Alcoholism; Louise Richards, National Institute on Drug Abuse; and Paul Placek, National Center for Health Statistics

INTRODUCTION

The Surgeon General has warned women about the hazards to the infants of mothers who smoke cigarettes or consume alcohol before or during pregnancy:

> Smoking slows fetal growth, doubles the chance of low birth weight, and increases the risk of stillbirth. Recent studies suggest that smoking may be a significant contributing factor in 20 to 40 percent of low weight infants born in the United States and Canada. Studies also indicate that infants of mothers regularly consuming large amounts of alcohol may suffer from low birth weight, birth defects, and/or mental retardation. Clearly, both previously developed habits need careful attention during pregnancy (Office of the Assistant Secretary for Health and the Surgeon General, 1979).

Excessive alcohol use is also associated with a pattern of abnormalities called Fetal Alcohol Syndrome (FAS). The National Institute on Alcohol Abuse and Alcoholism has reported to Congress:

> Some clinical studies of alcoholic mothers and some animal studies give evidence that heavy use of alcohol during pregnancy may result in a pattern of various abnormalities in the offspring...Current estimates conservatively suggest an incidence on the order of magnitude of 1 per 2,000 live births. Some evidence suggests that among 100 mothers drinking more than 1 ounce of absolute alcohol per day one might find one case of FAS. On this basis, FAS would be the third leading cause of birth defects associated with mental retardation, exceeded only by Down's syndrome and spina bifida. Of the three, only FAS is preventable (Malin and Munch, 1980).

Although studies have reported a decrease in drinking (Little, Schultz, and Mandell, 1976) and smoking during pregnancy (Hook, 1976), they have been based on small numbers of women in selected localities.

This paper reports on changes in drinking and smoking behaviors during pregnancy for a national sample of married mothers who had live births in 1980. Variations in smoking and drinking are examined before and during pregnancy according to maternal race, Hispanic origin, age and education.

DATA SOURCE

Data discussed in this paper are from the 1980 National Natality Survey. About 6 months after delivery, married mothers of live-born infants were mailed a questionnaire to assess social and demographic characteristics and prenatal health behavior, including smoking and drinking. This analysis is based on the 56 percent of married mothers who responded to the questionnaire (4,405 respondents).

Several things are important to note. The proportion of births to unmarried mothers varies widely by race, age, and other variables. For example, in the U.S. population nearly 90 percent of all white mothers are married but only 45 percent of black mothers are married. The smoking and drinking behaviors of unmarried mothers may differ from those of married mothers. It is also important important to note that the data in this analysis are unweighted, but selected comparisons with weighted data did not reveal any serious differences.

Mothers were asked which national origin they identified with most. Nearly two-thirds of those who identified with the Hispanic origin reported their race as white. Approximately one-third reported their race as other than black or white, leaving about 1 percent reporting themselves as black. Therefore, in this analysis, race and Hispanic ethnicity are not mutually exclusive categories.

Regarding smoking, mothers were asked, "On the average, how many cigarettes did you smoke per day before you found out that you were pregnant?" and "On the average, how many cigarettes did you smoke per day after you found out that you were pregnant?" Women who reported smoking one or more cigarettes per day were defined as smokers. Regarding drinking, women were asked "Did you drink any alcoholic beverages (beer, wine, or liquor) during the 12 months before your 1980 delivery?" and if yes, "How often did you usually drink alcoholic beverages; that is beer, wine, and/or liquor?" and "On the day or days that you drank, how much did you drink on the average per day?" Identical questions were asked for the periods before pregnancy and during pregnancy. Convenient check boxes with a wide range of possible responses were provided. From responses to the above questions regarding drinking, alcohol consumption, defined as the average daily consumption of absolute alcohol expressed in ounces, is calculated as follows:

> alcohol consumption = ounces of absolute alcohol consumed per day on the average
>
> = 0.5 x quantity factor x frequency factor.

The factor of 0.5 ounces per drink is used to translate the questionnaire response to ounces of alcohol consumed. This is approximately the amount of absolute alcohol contained in one beer, one glass of wine, or one mixed drink. The quantity factor is the average number of drinks per drinking day. The frequency factor translates the answer to "How often do you drink?" into a proportion of the time that is spent drinking. A mother who had one drink three times a week or three drinks once a week would have had the same average daily consumption of alcohol.

FINDINGS

Prepregnancy Smoking and Drinking Behavior

Figure 1 shows the distribution of mothers according to both smoking and drinking. Before pregnancy, nearly one-third of mothers smoked and more than one-half drank. More than one-third of mothers abstained from both smoking and drinking and one-fifth both smoked and drank. However, these behaviors varied substantially among different subgroups of mothers.

Smoking.--The prevalence of smoking was higher among white mothers than among black mothers and other mothers (table A). Almost one-third of white mothers smoked before pregnancy, compared with about one-fourth and one-fifth of black and other mothers, respectively. Less than one-quarter of Hispanic mothers smoked.

Smoking prevalence decreased with age among white mothers. Nearly 1 out of 2 mothers under 20 years of age smoked, but about 1 out of 4 mothers 25 years of age and over smoked. Among all other mothers, the same smoking pattern of decreasing with age prevailed.

Mothers under 20 years of age are excluded from the analysis by educational attainment because some of these mothers would not have had a chance to complete their education. The prevalence of smoking declined steadily with increased education among white mothers. Prevalence ranged from nearly 60 percent of mothers with less than 12 years of education to only 15 percent of those with 16 years or more. Among all other mothers, college graduates were the least likely to smoke, but no other differences by years of education were found.

Drinking.--As with smoking, drinking was more common among white mothers than among black and other mothers. Nearly 58 percent of white mothers drank, compared with 39 percent and 35 percent of black and other mothers, respectively. Forty percent of Hispanic mothers drank.

Although smoking decreased with age and education for all mothers, drinking increased with age and education. Among white mothers, drinking increased from 42 percent for mothers with less than 12 years of education to 67 percent for mothers with 16 or more years of education. Among all other mothers, drinking was less prevalent for those 30 years of age and over but did not vary by education.

Smoking and drinking.--About 22 percent of white mothers both smoked and drank, compared with about 14 percent of black mothers and 13 percent of all other mothers. The prevalence of the combination of behaviors decreased with age and education for both white mothers and all other mothers.

Abstention.--Abstention from smoking and drinking before pregnancy was much higher among black and other mothers than among white mothers. One-half of black mothers, nearly three-fifths of other mothers, but less than one-third of white mothers abstained before pregnancy. More than one-half of Hispanic mothers abstained.

Through 30-34 years of age, abstention increased with age for other mothers and decreased with age for white mothers but varied little by educational attainment.

Typically, smokers were white mothers under 25 years of age with a high school education or less, and drinkers were white mothers 25 years of age and over with more than a high school education. Younger white mothers and white mothers with 12 years of education or less were most likely to consume both alcohol and cigarettes. Hispanic mothers, black mothers, and other mothers were most likely to abstain and least likely to either smoke or drink.

Change in Smoking and Drinking Behavior During Pregnancy

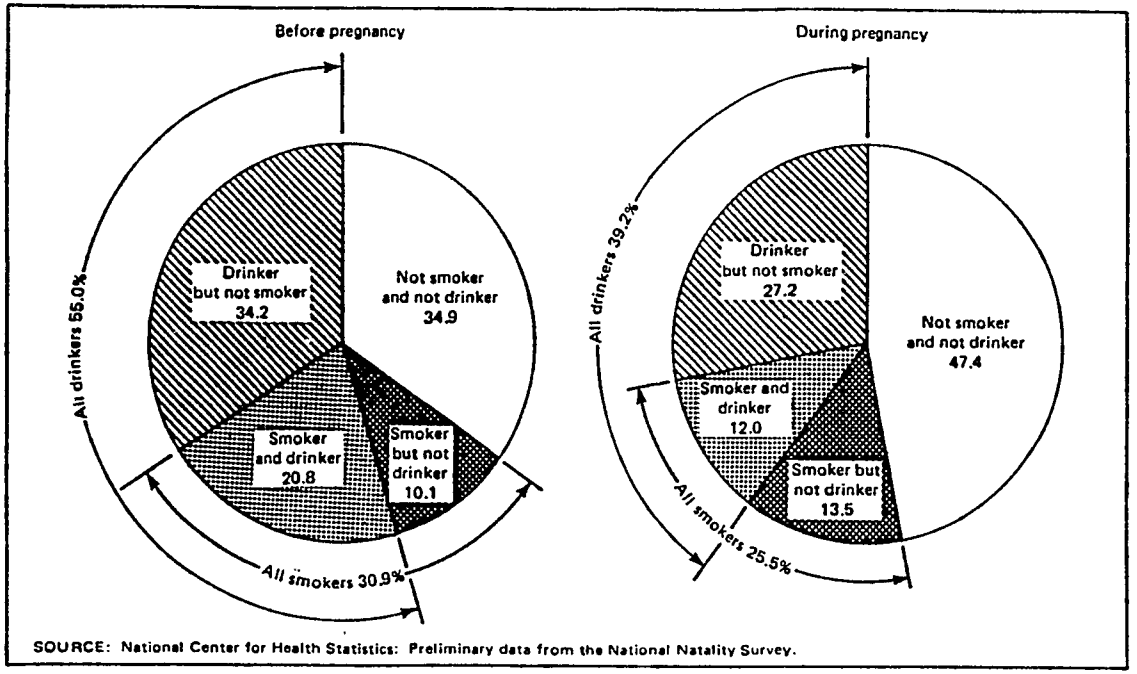
Mothers were much more likely to stop drinking than to stop smoking during pregnancy (table B). Of those who had the habit before pregnancy, 30 percent stopped drinking and 18 percent stopped smoking. Of mothers who both smoked and drank prior to pregnancy, 43 percent gave up either cigarettes or alcohol or both, reducing the prevalence of both habits from 1 out of 5 pregnant women to 1 out of 8 (figure 1). Twenty percent of mothers with either of these habits before pregnancy became abstinent during pregnancy, increasing the prevalence of abstinence from 1 of every 3 to 1 of every 2 mothers.

Change in smoking.--No significant differences by age, race, or Hispanic origin were found in the proportion who stopped smoking. However, educational attainment was directly related to the tendency to stop smoking. Of white mothers who smoked, the proportion who stopped during pregnancy ranged from 10 percent for mothers with the least education to 24 percent for mothers with the most education. Thus, the difference between the highest and lowest educational groups in the prevalence of smoking became even greater during pregnancy than it was before.

Change in drinking.--White mothers who drank were not as likely to stop drinking as were other mothers. For mothers under 35 years, the age of the mothers who drank was inversely related to the tendency to stop drinking. For example, of white mothers under 35 years who drank, the older the mother the smaller the proportion who stopped drinking. However, there were no significant differences by education in the proportion who stopped drinking.

Change in behavior of mothers who smoked and drank.--Mothers who both smoked and drank before pregnancy were more likely to give up alcohol than cigarettes during pregnancy. Of these mothers, 43 percent of white mothers and 35 percent of black mothers gave up one or both behaviors during pregnancy, compared with 56 percent of other mothers. Half the Hispanic mothers stopped one or both habits during pregnancy (table B).

Percent becoming abstinent.-- Of mothers who either smoked or drank or had both habits before pregnancy, about 1 of every 5 became abstinent during pregnancy.



SOURCE: National Center for Health Statistics: Preliminary data from the National Natality Survey.

Figure 1. Percent distribution of married mothers of live-born infants, according to smoking and drinking behaviors before and during pregnancy: United States, 1980

Table A. Smoking[1] and drinking[2] behavior before pregnancy of married mothers of live-born infants, according to selected characteristics: United States, 1980

Characteristic	Number	All smokers	All drinkers	Total	Not smoker and not drinker	Smoker but not drinker	Smoker and drinker	Drinker but not smoker
		Percent of mothers			Percent distribution			
All married mothers...	4,405	30.9	55.0	100.0	34.9	10.1	20.8	34.2
Race								
White.....	3,888	32.0	57.5	100.0	32.3	10.2	21.8	35.6
Black.....	250	24.8	39.2	100.0	49.6	11.2	13.6	25.6
Other[3].....	267	19.9	34.5	100.0	58.4	7.1	12.7	21.7
Hispanic origin								
Hispanic.....	253	23.3	39.9	100.0	51.0	9.1	14.2	25.7
Non-Hispanic.....	4,152	31.4	55.9	100.0	33.9	10.2	21.2	34.8
Age								
All races:								
Under 20 years..	330	47.3	40.3	100.0	39.1	20.6	26.7	13.6
20-24 years.....	1,393	36.8	53.0	100.0	35.1	11.9	24.8	28.1
25-29 years.....	1,608	27.3	58.2	100.0	33.8	8.0	19.3	38.9
30-34 years.....	846	23.2	58.6	100.0	34.2	7.2	16.0	42.7
35 years and over.....	228	25.4	53.1	100.0	37.3	9.6	15.8	37.3
White:								
Under 20 years..	282	48.9	40.4	100.0	38.3	21.3	27.7	12.8
20-24 years.....	1,239	37.9	54.2	100.0	33.5	12.3	25.7	28.6
25-29 years.....	1,442	27.9	60.5	100.0	31.8	7.8	20.2	40.3
30-34 years.....	740	24.9	63.4	100.0	29.2	7.4	17.4	45.9
35 years and over.....	185	27.6	57.8	100.0	31.9	10.3	17.3	40.5
All other:[3]								
Under 20 years..	48	37.5	39.6	100.0	43.8	16.7	20.8	18.8
20-24 years.....	154	27.3	42.9	100.0	48.1	9.1	18.2	24.7
25-29 years.....	166	21.7	38.6	100.0	51.8	9.6	12.0	26.5
30-34 years.....	106	11.3	25.5	100.0	68.9	5.7	5.7	19.8
35 years and over.....	43	16.3	32.6	100.0	60.5	7.0	9.3	23.3
Education[4]		Percent of mothers			Percent distribution			
All races:								
0-11 years.....	432	49.8	40.7	100.0	36.3	22.9	26.9	13.9
12 years.....	1,823	34.4	54.5	100.0	35.1	10.4	24.0	30.5
13-15 years.....	936	25.0	59.4	100.0	33.2	7.4	17.6	41.8
16 years or more.....	884	14.5	63.9	100.0	33.9	2.1	12.3	51.6
White:								
0-11 years.....	332	57.2	41.9	100.0	30.4	27.7	29.5	12.3
12 years.....	1,666	35.7	56.2	100.0	33.1	10.6	25.0	31.2
13-15 years.....	827	25.3	62.6	100.0	30.5	6.9	18.4	44.3
16 years or more.....	781	14.7	67.3	100.0	31.1	1.5	13.2	54.2
All other:[3]								
0-11 years.....	100	25.0	37.0	100.0	56.0	7.0	18.0	19.0
12 years.....	157	21.7	36.3	100.0	55.4	8.3	13.4	22.9
13-15 years.....	109	22.9	34.9	100.0	54.1	11.0	11.9	22.9
16 years or more.....	103	12.6	37.9	100.0	55.3	6.8	5.8	32.0

[1]Smokers are those who smoked at least 1 tobacco cigarette per day.
[2]Drinkers are those who consumed at least 1 drink (½ oz. of absolute alcohol) once a month.
[3]Includes all other races not shown separately.
[4]Includes mothers who are 20 years of age and over.

SOURCE: National Center for Health Statistics: Preliminary data from the National Natality Survey.

Table B. Change in smoking[1] and drinking[2] behavior during pregnancy of married mothers of live-born infants, according to selected characteristics: United States, 1980

Characteristic	With 1 or both habits			With both habits and stopped 1 or both
	Stopped smoking	Stopped drinking	Became abstinent	
	Percent of mothers			
All married mothers...	17.6	29.6	19.6	43.1
Race				
White.....	17.7	29.0	19.1	42.9
Black.....	12.9	37.8	26.2	35.3
Other[3].....	20.8	37.0	23.4	55.9
Hispanic origin				
Hispanic.....	25.4	27.7	20.2	50.0
Non-Hispanic.....	17.3	29.7	19.6	42.8
Age[4]				
All races:				
Under 20 years.....	19.9	39.1	17.4	54.5
20-24 years.....	17.4	33.2	19.1	46.5
25-29 years.....	18.0	29.5	21.3	41.2
30-34 years.....	18.9	22.4	18.5	33.3
35 years and over.....	6.9	28.1	17.5	36.1
White:				
Under 20 years.....	20.3	39.5	17.2	55.1
20-24 years.....	17.0	32.4	18.4	45.9
25-29 years.....	18.1	29.5	21.2	41.6
30-34 years.....	19.6	21.3	17.7	33.3
35 years and over.....	7.8	25.2	15.9	34.4
Education[4,5]				
All races:				
0-11 years.....	9.8	29.5	9.5	40.5
12 years.....	16.9	30.8	19.3	42.0
13-15 years.....	21.8	25.2	20.0	40.0
16 years or more.....	24.2	29.7	25.5	45.9
White:				
0-11 years.....	10.0	30.2	9.5	39.8
12 years.....	16.5	30.2	18.6	41.7
13-15 years.....	23.0	24.5	19.7	40.8
16 years or more.....	24.3	28.5	24.5	44.7

[1]Smokers are those who smoked at least 1 tobacco cigarette per day.
[2]Drinkers are those who consumed at least 1 drink (½ oz. of absolute alcohol) once a month.
[3]Includes all other races not shown separately.
[4]For all other races, number of cases was too few to meet standards for precision or reliability.
[5]Includes mothers who are 20 years of age and over.

SOURCE: National Center for Health Statistics: Preliminary data from the National Natality Survey.

Table C. Smoking and drinking behavior during pregnancy of married mothers of live-born infants, according to type of behavior before pregnancy: United States, 1980.

Behavior before pregnancy	All married mothers		Behavior during pregnancy			
	Number	Percent distribution	Total	None	Level 1	Level 2
Smoking			Percent distribution			
Total.....	4,405	100.0	100.0	74.5	12.0	13.5
None[1]	3,044	69.1	100.0	99.9	0.1	0.0
Level 1 [2].....	412	9.4	100.0	31.3	65.8	2.9
Level 2 [3].....	949	21.5	100.0	11.8	26.8	61.4
Drinking						
Total.....	4,405	100.0	100.0	60.9	35.9	3.2
None[4].....	1,981	45.0	100.0	99.1	0.8	0.1
Level 1 [5].....	1,741	39.5	100.0	34.5	64.6	0.9
Level 2 [6].....	683	15.5	100.0	17.4	64.6	18.0

[1]Did not smoke even 1 cigarette a day.
[2]Smoked 1-10 cigarettes each day on the average.
[3]Smoked 11 cigarettes or more each day on the average.
[4]Consumed less than 1 drink (1/2 oz. of absolute alcohol) less than once a month.
[5]Consumed 0.01-0.19 oz. absolute alcohol each day on the average; equivalent to less than 3 drinks per week.
[6]Consumed 0.19 oz. or more absolute alcohol each day, on the average; equivalent to 3 or more drinks per week.

Source: National Center for Health Statistics: Preliminary data from the National Natality Survey.

Quantity and Frequency of Cigarettes and Alcohol Consumed

The change in amount of drinking and smoking after pregnancy was confirmed among all mothers who drank or who smoked before pregnancy (table C).

Of the 31 percent of mothers who smoked before pregnancy, most (22 percent) smoked more than 10 cigarettes per day. About 12 percent of these stopped smoking during pregnancy, and an additional 27 percent reduced their consumption. In addition, nearly one-third of those who smoked less than 10 cigarettes per day stopped smoking during pregnancy. The prevalence of those who smoked more than 10 cigarettes per day during pregnancy dropped to about 14 percent, and the prevalence of nonsmokers increased from 69 percent to 75 percent. Of the 55 percent of mothers who drank before pregnancy, only 16 percent drank three drinks or more per week. About 17 percent of these stopped their drinking during pregnancy, and another 65 percent reduced their consumption. In addition, one-third of those who consumed less than three drinks per week stopped drinking altogether during pregnancy. The prevalence of women drinking three or more drinks per week dropped to only 3 percent, while the prevalence of nondrinkers increased from 45 percent to 61 percent.

CONCLUSION

Before pregnancy, mothers' smoking and drinking behaviors resemble those of the general population of women in the reproductive ages 15-44 years (National Center for Health Statistics, 1979; National Institute on Alcohol Abuse and Alcoholism, 1982a and 1982b; Malin, Hartsock, and Frank, 1982). When pregnancy is confirmed, reductions in smoking and drinking occur. Although the prevalence of drinking is much higher than the prevalence of smoking among mothers included in the survey, the reduction in drinking is much more pronounced than the reduction in smoking. This is opposite to what might have been expected. Media messages have documented the deleterious effects of maternal smoking for two decades, while messages concerning the hazards of maternal drinking have been more recent and more modest. However, it is likely that giving up nicotine is more difficult than giving up alcohol (Eysenck, 1973). Given the levels of smoking and drinking observed in this study, smoking may have been more deleterious to the fetus than drinking.

Since smoking is usually a daily habit, a strong physiological and psychological dependency can be created. Most of the mothers in this survey who smoked did so daily, and most of the mothers who drank did so less frequently. In fact, the largest number of mothers who drank consumed the weekly equivalent of only one or two drinks. Apparently, drinking three or more alcoholic beverages per week was extremely rare among the women in this survey. Therefore, alcohol addiction probably was not a factor for the majority of drinking mothers and may have accounted in part for the larger reduction in drinking than smoking. Furthermore, one study suggested that moderate to heavy drinkers may decrease their drinking during pregnancy because alcohol becomes distasteful or causes adverse physiological effects (Little, Schultz, and Mandell, 1976).

This article has shown major differences in smoking and drinking by maternal race, Hispanic origin, age, and education. Mothers may have underreported their behavior because of guilt feelings aroused by media campaigns to persuade them to reduce smoking and drinking during pregnancy. If so, the estimates of smoking and drinking could be considered minimum estimates.

Studying the smoking and drinking behavior of pregnant women is the first step in researching the effects of maternal behavior on fetal growth, development, and survival. The next step, already underway, is to explore the effects of smoking and drinking on pregnancy outcome. In these analyses, sociodemographic factors must be carefully controlled to rule out their confounding effects.

REFERENCES

Eysenck, H.J.: "Personality and the Maintenance of the Smoking Habit," in W. L. Dunn, ed., Smoking Behavior, Motives and Incentives. Washington. Winston and Sons, 1973, pp. 113-146.

Hook, E.B.: "Changes in Tobacco Smoking and Ingestion of Alcohol and Caffeinated Beverages During Pregnancy--Are These Consequences, in Part, of Feto-Protective Mechanisms Diminishing Maternal Exposure to Embryotoxins?," in S. Kelly, E. B. B. Hook, and P. Janerich, eds., Birth Defects, Risks and Consequences. New York. Academic, 1976.

Little, R. E., Schultz, F. A., and Mandell, W.: "Drinking During Pregnancy." J. Stud. Alcohol 37 (3): 375-379, 1976.

Malin, H., Hartsock, P., and Frank, J. : "Alcohol Use and Abuse Among U.S. Hispanics and in Latin American Countries," Working Paper No. 61. Alcohol Epidemiologic Data System. Contract No. ADM 281-79-0012. Prepared for the National Institute on Alcohol Abuse and Alcoholism; Alcohol, Drug Abuse, and Mental Health Administration. Rockville, MD., Sept. 1982.

Malin, H. J., and Munch, N. E.: Implicated Alcohol and Disease Entities, Alcohol as a Risk Factor in the Nation's Health. Report of the Alcohol Epidemiologic Data System. Contract No. ADM 281-79-0012. Prepared for the National Institute on Alcohol Abuse and Alcoholism; Alcohol, Drug Abuse, and Mental Health Administration. Rockville, MD., Mar 28, 1980.

National Center for Health Statistics: "Changes in Cigarette Smoking and Current Smoking Practices Among Adults," United States, 1978, by A. J. Moss. Advance Data From Vital and Health Statistics, No. 52. DHEW Pub. No. (PHS) 79-1250. Public Health Service Hyattsville MD., Sept. 19, 1979

National Institute on Alcohol Abuse and Alcoholism: "Alcohol Use and Alcohol Problems Among U.S. Adults, Results of the 1979 National Survey," by W.B. Clark and L. Midanik. Alcohol and Health Monograph No. 1, Alcohol Consumption and Related Problems. DHHS Pub. No. (ADM) 82-1190. Alcohol, Drug Abuse, and Mental Health Administration. Washington. U.S. Government Printing Office, 1982a.

National Institute on Alcohol Abuse and Alcoholism: "An Epidemiologic Perspective on Alcohol Use and Abuse in the United States;" H. Malin, J. Coakley and C. Kaelber. Alcohol and Health Monograph No. 1. Alcohol Consumption and Related Problems. DHHS Pub. No. (ADM) 821190. Alcohol, Drug Abuse, and Mental Health Administration. Washington. U.S. Government Printing Office, 1982b.

Office of the Assistant Secretary for Health and the Surgeon General: Healthy People, The Surgeon General's Report on Health Promotion and Disease Prevention. DHEW Pub. No. (PHS) 79-55071. Public Health Service. Washington. U.S. Government Printing Office, July, 1979.

Third Plenary Session

PROGRAM EVALUATION AS A BASIS FOR HEALTH POLICY FORMULATION

Gwynne R. Winsberg, GRW Associates

Evaluation research is supposed to tell us whether or not a particular program is a success or failure according to the stated goals of the program in question. The Hatch amendment earmarked a 1% set aside of health care appropriations to be used in evaluation of all funded programs. This money, for the most part, has been given to various consulting firms for the purpose of evaluating the many Department of Health and Human Services programs. The consultants have also trained many DHHS employees to carry on evaluation activities in-house as well. Many reams of paper documenting, whether or not a program is capable of being evaluated (evaluability assessment), are to be found on the shelves and in the files of countless bureaus of the various DHHS agencies. The second step after determining the evaulability of a particular program is to let a contract for the actual evaluation of the program. This usually costs two to three times as much as the evaluability assessment and frequently follows even if the evaluability assessment states that the program cannot be evaluated because their was no consensus on either the objectives or the final goals of the program. The final outcome of this process is a report supplied to "policy makers" who may utilize it to either redesign a particular legislatively mandated program or to dismantle the program entirely.

Research done over a period encompassing the last two decades, in a political and economic climate unlike the one existing today, was used to formulate the present competitive philosophy of health care. These competitive strategies will be evaluated eventually in a climate unlike the one in which they were formulated. At the time of evaluation, the original objectives, if any were stated, may be obscure. The political objectives of the moment will guide the evaluation study.

During my time in Washington, I became involved personally in a variety of evaluation studies centered around the successes and failures of the National Health Service Corps (NHSC). The legislative beginnings of this highly controversial program are delightfully laid out in the book "The Dance of Legislation", by Eric Redman (1). The Corps was originally conceived as a plan to place a small cadre of volunteer physicians in private practice in rural areas, in the usual and customary fee-for-service mode of practice - the return of the "Marcus Welby" type of doc to the rural area from which he had departed during the various wars and tribulations of our society. This was certainly a noble and non-threatening objective for any legislator to support. The National Health Service Corps was looked upon, at least in 1974, as probably the only program that was not draining the coffers of the U.S.Treasury - indeed these practices were returning money to Uncle Sam, as the physicians paid off their start-up costs and returned excess fees over and above their salaries. The program looked so good, that legislators thought it would be nice to expand the program to take care of the numerous people in economically depressed areas, may of which were rural - but even more were urban!

Expansion of the NHSC presumed a supply of physicians champing at the bit to serve in undesirable parts of the country and willing to serve the poor of our urban ghettos for considerably less remuneration than their peers would receive serving the affluent in the large suburban areas surrounding the cities where the major medical centers were placed. In addition, the trend since World War II was toward the training of specialists who are dependent on the technology and educational climate of the modern academic medical center environment. This was hardly a fertile field in which to nurture primary care physicians who would be expected to go out on their own with only one year of post-graduate training. Just who would volunteer? Very few, and thus the National Health Service Corps Scholarship Program was born.

In return for tuition and a stipend, a medical student agreed to serve humanity after one year of post-graduate training, year for year of Federal support, after which he or she was free to return for further residency training and the mainstream of well paid medical practice. The placement and scholarship programs were administered in different agencies, which rarely spoke to each other except just before Congressional hearings.

Each year or two, a new rule was added to either the Scholarship or Placement Program as shortcomings were experienced and reports came in from the field, or legislators received feedback from constituents whose practices were perceived to suffer from proximity of a Corps physician. As the Corps grew in size, difficulties in management were perceived within and without the Federal government, but were rarely expressed except in inquiries of Congressional rage. Statements such as - "You are exporting physicians from our State and we have a shortage", "Corps docs are lazy and unmanagable, they are just serving time and lack commitment to this practice", Corps physicians lack medical judgement", "The States can do it better", "The National Health Service Corps is directly competing with the private physician's practice", and the always present cry that "Retention rates are low, so why bother".

The 'first evaluation' attempt was, "An Evaluability Assessment of the Management of the National Health Service Corps" (2). Before this study was completed, it was determined that the Corps management problems could be solved by combining the two arms of the Corps - the Placement and Scholarship Programs into a single Bureau under the aegis of the then Health Services Administration. Certainly, combination and centralization of management functions is desirable under a number of circumstances. However, reorganization without a distinct plan for management does not necessarily guarantee communication or improvement in the function of the program. The evaluability assessment findings were that the management of the Corps Programs, in their present state, was not capable

of being evaluated because of the lack of distinct management objectives. A subsequent contract to evaluate the management of the Corps was not let in this instance, because it was determined that the Corps should have time to formulate objectives, both programmatic and managerial, based on its present day operational mode, rather than on its aquired growth characteristics.

In late 1979, a Senator, who was later to become Secretary of Health and Human Services voiced the opinion that the States could do a better job of training and placing physicians than could the Federal Government. Two independent assessments were undertaken immediately of State programs designed to solve the physician shortage. One of these, a comprehensive study of State aid to health professions education was done by the Congressional Research Service, utilizing a large and expensive consulting firm (3). The second study, under the direction of the Office of Evaluation in the Office of the Secretary of DHHS, which also utilized an expensive consulting firm, directed itself to assessing the role of the States in supporting Corps-like programs (4).

The first study, whose staff followed fast on the heels of the second study's staff, cataloged the large variety of State programs aiding health professions education, but did not describe the management or outcome of those programs. The second study specifically cataloged programs dealing with allopathic and osteopathic medical education and dental education. This study also looked at achievement of objectives (if any), location of management functions, and coordination between scholarship and placement activities. It was concluded that the States could not do it better, and most importantly, that they were having very serious problems in keeping the programs afloat. Legislative objectives were unclear in the majority of the States, and appropriations with the exception of two States, were miniscule. Just as in the Federal Program, those awarding the scholarships had little or no communication with those who controlled any existing placement activities. Moreover, Federal and State programs were frequently at odds with each other. Designation of shortage areas done by the States had little congruence with shortage areas designated under Federal guidelines. States which had a long history of such programs were queried as to why the same shortage areas remained year after year. Uniformly, the answer was that the area was too poor to support a physician without continued subsidy which the State could not supply.

Both studies recommended that there should be attempts to coordinate State and Federal programs. This recommendation was carried out subsequently and came to fruition under a different administration, just as the Federal appropriations for the Corps were cut severely. At the same time the States found themselves in extreme financial distress due to cuts in all Federal health care programs that served the economically disadvantaged.

The question of competition with private physician practice was resolved with another large and costly study, involving an econometric model of ambulatory care production functions. This massive study of the market place, and the effects if any, of the presence of Corps physicians in that market area required the combined direction of Offices of Evaluation of the Health Resources Administration, Health Services Administration and the Office of the Secretary. Several well known economists were engaged as consultants by the consulting firm. The costs continued to escalate as errors in the design and execution were discovered. Concurrently with this study, another study co-sponsored by DHHS and the Robert Wood Johnson Foundation, was launched to look at the ingredients in success and failure of rural primary care projects, most of which utilized NHSC physicians. Both studies involved a survey of the community residents and their attitudes, health status, and utilization of health care.

Neither study was completed in time to inform or influence the policies of the administration initiating the studies. However, preliminary results from both studies showed that the NHSC filled a previously existing void. Early results from the study on market interference showed that for the most part, Corps physicians served a poorer and sicker population than that served by the sometimes nearby fee-for-service physician.

The problem with all evaluation research surrounding the success or failure of the National Health Service Corps was that the evaluators never knew what they were evaluating. One must remember that the original Corps volunteer could be a physician with a completed residency, or even an older physician in mid-career. Depending on what version of the legislation he or she came under, it was possible for the scholarship physician to have only one year of post-graduate training, or be allowed to complete a residency in a primary care speciality. A still later version broadened the definition of primary care. Amount of training and reason for service in a particular practice could have an impact on motivation and consequent productivity. While the Corps was frequently criticized for a seeming lack of stability - the States certainly were shown to have no better a record of retention of physicians than the Corps - again, depending on the stage of training and the area in which the physician practiced. Economically disadvantaged sites were and are unable to retain physicians without some type of subsidy. The solution to this by the present administration was to phase out the Scholarship Program and allow presently obligated physicians to serve by utilizing the Private Practice Option - in essence returning the Corps to its original goal of providing physicians to economically viable environments which just happened to lack a physician because of an overall physician shortage.

Unfortunately, the policy of offering a private practice option comes at just such a time as we are experiencing a glut of physicians who, according to the Rand study, are fanning out under their own steam into every nook and cranny of the country (5). However, we are still allowing Corps physicians to become employees of Community Health Centers (which must pay their

salaries out of steadily diminishing funds when the pressure to supply the health care needs of an increasing population of unemployed individuals is taxing their resources to the utmost).

Each of the evaluation studies commissioned, outlined specific problems and suggested solutions. Very few of these problems were unknown to the program administrators. In many instances, attempts at solution were already underway at the time the evaluation was in the design stage. By the time the request for proposal went out, the proposals evaluated and the projects completed, the findings had little relevance to the political climate then in existence.

A more positive example of evaluation research utilized to formulate health care policy is that surrounding the Community Health Center Program (Sec.330). As usual, programs serving the poor with comparatively few dollars must be made to prove their worth over and over again. In contrast, programs subsidizing fairly substantial segments of our society (e.g. the Medical Schools) are continued well beyond the time when they are proven to be effective). Thus the Community Health Centers (CHC) had to prove that they were cost effective in order to survive. A reworking of the data from a study done on place of seeking care for people who received Medicaid in several West Coast communities, showed that the cost of care decreased after Community Health Centers were in place (6). This decrease in cost is extrapolated from data which show that hospital days per thousand were decreased significantly when the Medicaid recipients left the emergency rooms and hospital out-patient departments to utilize the organized system of care available in the CHC.

In addition, a contract was let for a study to assess CHC cost effectivness, utilizing data from three States that had Medicaid Management Information Systems (7). This study demonstrated that costs decreased when Aid to Families with Dependent Children (AFDC) recipients utilized the CHC. While the present administration worked diligently to divest the CHC program of its funding - the information from these studies was utilized by groups outside the government to inform Congress of the importance of organized systems of care in decreasing Medicaid costs - thus salvaging a program destined for destruction. While one cannot directly attribute the recognition that organized systems of care have the effect of lowering costs to the CHC evaluations alone, it is apparent that they have had some effect on present Health Care Financing Administration policy with regard to encouragement to the States to establish Medicaid pre-payment plans.

The act of evaluating a program can in itself lend political credence to unpopular findings. The establishment and actions of the Graduate Medical Education National Advisory Council (GMENAC) may be viewed in this light. The Division of Manpower Analysis of the Health Resources Administration, for many years, worked diligently to amass the facts and figures necessary to predict numbers and deployment of health professionals in the U.S. When it became apparent that the only predictions possible pointed toward future surpluses, it became necessary to have the health professions schools accept these findings. This was especially necessary since government policy, for a number of years, was directed toward increasing the overall supply of health professionals trained in this country in order to meet a perceived need for more health professionals. Little attention was paid to the fact that only specifically trained health professionals (e.g. primary care physicians and dentists who would practice in underserved areas) were actually in short supply.

GMENAC, that august and learned body, was constituted of respected professionals in the various medical specialties and staffed with an army of excellent government researchers. It was charged with the task of predicting the need for physicians in 1990. The methodology utilized defies description in the short time alloted to this brief plenary session. However, the outcome was the reaching of consensus with regard to a surplus in most specialty areas in 1990. A little discussed corollary to the report is that shortages would continue to exist in primary care specialities in economically disadvantaged areas of the country (8). The process of reaching this consensus deeply committed the medical community to acceptance of the GMENAC findings. Although GMENAC's predictions were very similar to those of the Division of Manpower Analysis, they in essence constituted both an evaluation of the Division's work and validation of the concept that we were training too many physicians.

The Office of the Secretary closely monitored the GMENAC proceedings with the intent of launching a full scale evaluation of the methodology. When it became apparent that GMENAC's predictions would validate previous predictions and support the embryonic policies directed toward less financial support for health professions education, a decision was made to abort any full scale evaluation attempts.

One may infer from the examples discussed here that evaluation research may be utilized in many ways to assist in the formulation of policy. It is useful to both catalog programs and to study their operations. But one also must remember that such research is generally costly and rarely specific. The majority of government programs are formulated to meet perceived political objectives, frequently lost sight of as administrations change. It is then unfair to assess the efficiency of these programs in meeting objectives that no longer exist at the time the study is proposed, much less finished. It is even less fair to evaluate a program before it is in operation for a stated period of time, as is frequently done when a delayed appropriation for a new program start comes through. A safe statement to make about evaluation and health policy is that - Evaluation studies are always at least 4.3 years out of date at the time of completion and will be very probably utilized to formulate policies that may come to exist 4.8 years in the future.

1. Redman, Eric, "The Dance of Legislation". Simon and Schuster, New York, NY. 1973

2. unpublished

3. U.S. Senate Committee on Labor and Human Resources. "State Support for Health Professions Education". Washington, D.C., U.S. Government Printing Office. 1981

4. Gibbons, S.L., Williams, J.L. and Winsberg, G.R., "Short-Term Evaluation of State Educational Service Conditional Support Programs for Allopathic, Osteopathic and Dental Students. MACRO Systems, Inc., 1980. (NTIS PB 87-115248)

5. Schwartz, W.B., Newhouse, J.P., Bennett, B.W. and Williams, A.P. The Changing Geographic Distribution of Board-Certified Physicians. N. Engl. J. Med., 303:1032-1037. 1980

6. Freeman, H.E.,Kielcot, K.J. and Allen, H.M. Community Health Centers: an Initiative of Enduring Utility. Milbank Memorial Fund Quarterly, Health and Society. Vol.60,No.2:245-267, Spring 1982

7. Duggar,B.C., Balicki, B., Lazarus, S., etal. Community Health Cost-Effectivness Evaluation. Report submitted to DHEW/OS/ASPE/NHI Analysis, by JRB Associates, McLean, VA, Feb. 1981

8. "Report of the Graduate Medical Education National Advisory Committee" (HRA 84-651). Washington, D.C. Department of Health and Human Services. 1981

CURRENT CONTROVERSY IN PROGRAM EVALUATION METHOLODOLOGY

Donald L. Patrick, University of North Carolina-Chapel Hill

Confusion over method and purpose in program evaluation is a common experience for many involved in the evaluation process. Such confusion is evident in both written and oral discussions of program evaluation methodology that have been dominated by vigorous debate over the relative merit of different methods, philosophies, and criteria used in the evaluation process. In the past these controversies often have been presented as choices between bipolar or irreconcilable approaches to evaluation: between process and outcome, between standardized and non-standardized treatments, between inside and outside evaluations. More recently evaluation researchers, like professionals in other rapidly expanding fields of scientific inquiry, are moving toward agreement on issues about which they contended hotly only a few years back. At present the differences are more in emphasis between methodological approaches rather than in specific practices recommended.

This paper discusses two of the many areas of controversy in evaluation methodology in relation to the purposes of health program evaluation: (1) the relative emphasis to be placed on quantitative and qualitative methods of evaluation, and (2) the extent to which an evaluator emphasizes internal or external validity in the evaluation design. Before explaining these specific areas of controversy, however, it is important to put these methodological considerations into the context of the evaluation process and its different purposes. This is particularly crucial, since the selection of methodological emphases should be based heavily on the purposes of the evaluation as perceived by participants in the evaluation process.

Evaluation, as the word implies, concerns values and it begins and ends as a political process of determining the worth of a program under consideration.[1] Decision-makers, administrators, clinicians, consumers, officials, and citizens are all inescapably in the business of determining whether policies and programs--that is the social arrangements for providing services--are "good" or "bad."

Evaluation research, as distinct from evaluation, is the application of scientific method to the process of judging worth. Research efforts in evaluation can be judged "by the extent to which they help the political community achieve its end."[2] In this sense, the political community is made up of all participants or stakeholders in evaluation, and the major purpose of research is to "facilitate a democratic, pluralistic process by enlightening all the participants."[2]

Different participants in the evaluation process may have different motives, vested interests, or purposes in evaluating a health program. In this sense they are stakeholder groups that either participate directly or become interested in the evaluation process and its results.[3] These stakeholders include the persons responsible for deciding whether a program is to be started, continued, expanded or stopped, the evaluation sponsor, the organization that initiates and funds the program, clients of the program, administrators, staff, evaluators, competing groups, and others in the local environment who are interested in the program.

The interests and expectations of these stakeholders relate to the following six purposes of program evaluation: (i) to improve the program; (ii) to assess the effects of the program; intended and unintended; (iii) to choose between competing programs; (iv) to advocate the program for wider use; (v) to increase knowledge; and/or (vi) to fulfill a funding requirement. Program administrators and staff may wish to improve an existing program because they recognize service problems which need solutions. Citizens, consumers of services, funding agents and sponsors may be more interested in making the program accountable, that is assessing program effects, whether they be the intended effects of the program or unanticipated consequences. The sponsors may wish to advocate wider use of the program or to choose between competing programs, particularly in an era of cost-containment. Evaluators may simply wish to increase knowledge about a particular phenomena, for example, the relationship between diet and mortality.

A common reason why evaluations get done is because they are mandated or made a requirement for funding . In a recent article, Senator Orrin Hatch, Chairman of the Senate Committee on Labor and Human Resources, wrote, "It remains my firm conviction that every program authorized in the public sector should have a built-in evaluation component which is operated by independent, non-service providers."[4] Since the early 60's, federal support for programs such as community mental health centers, family planning programs, neighborhood health centers, drug abuse programs, social service programs, and educational programs has been contingent on evaluation. Some program administrators and staff may be interested in evaluation simply because they have to be.

Since participants in evaluation may have different purposes and these purposes may easily conflict, it is not surprising that conflicts in purpose are carried into the design evaluations. Debates over the choice between evaluation strategies or the emphasis given to different methodologies may well reflect conflicts of interest, purpose and authority.

Quantitative versus Qualitative Methods

One debate which has simmered for many years in evaluation research centers on the distinction drawn between qualitative and quantitative methods.[5] Quantitative approaches have dominated the field, including the prized randomized experiment, quasi-experiments, sample surveys, and other methods producing numerical results. Qualitative approaches, such as case studies, ethnography, participant observation and narrative or journalistic descriptions of programs, grew out of a need to imbue evaluation findings with "real life." Until

recently, each methodological tradition attracted strong advocates who recommended their approach as the method of choice.

This debate has its origins in a fundamental clash between two philosophical perspectives or world views: phenomenology and logical positivism . In the tradition of phenomenology, the qualitative perspective emphasizes "verstehen" or understanding from a person's own point of view. Observations are seen to be naturalistic or uncontrolled, subjective, oriented toward "process," valid, real and sometimes ungeneralizable. The quantitative perspective, arising from the logical-positivist position, emphasizes facts, objectivity, and deduction. Observations are seen to be controlled, outcome-oriented, reliable, hard and thought to be generalizable.

While the characteristics of these two perspectives are not mutually exclusive, adherence to one perspective or method over another is common. As Donald Campbell writes, "Too often quantitative social scientists, under the influence of missionaries from logical positivism, presume that in true science, quantitative knowing replaces qualitative common-sense knowing. The situation is in fact quite different. Rather, science depends upon qualitative, common-sense knowing even though, at best, it goes beyond it. Science in the end contradicts some items of common sense, but it only does so by trusting the great bulk of the rest of common-sense knowledge."[6]

In program evaluation, using qualitative and quantitative methods in combination is a desirable strategy for converging on the "truth" about the intended and unintended effects of a program. Not all evaluations, however, have the resources, time, or expertise available to mount an elaborate, multi-method research study. Choices have to be made between evaluation strategies, and these choices depend upon the purpose of the evaluation. But who determines what the purposes are? In most cases the program and evaluation sponsors have the greatest influence on why and how a particular evaluation is conducted. As stated earlier, however, other stakeholders may have radically different purposes in mind, and the "technical" or methods side of the research study may reflect the tension of the wider evaluation process.

Evaluators sometimes encourage the use of sophisticated research designs and outcome measures, the very tools of their trade, when program administrators and funding agents have a purpose in mind which does not require the randomized experiment or use of comprehensive instrumentation. For example, the evaluation of an exercise program for elementary school children might be designed by an evaluator to measure elaborate outcomes such as blood lipid levels and electrocardiogram readings in experimental and control groups, when the more important determinant of success for the program administrator, staff, and even the funding agent may be how much "fun" the program was for the children. Rigorous evaluation with randomized experimental and control groups is often unwise unless the program being evaluated is clearly the program of choice.[1] Conflict also occurs frequently when program administrators and staff, convinced of the effectiveness of the program they are conducting, resist the advice of sponsors and evaluators that experimental design and outcome measures are necessary to show that the program has particular desirable effects or should be used on a wider basis.

Identifying the stakeholders in an evaluation and finding out their perceptions of purpose represent one approach to this problem. Although it may not resolve the differences which are uncovered, assessing the evaluability of a program or conducting a systematic survey of stakeholders can help choose the most appropriate evaluation methodologies. [7,8] Wide disagreement in purpose, if unresolved, is a good predictor that the evaluation will fail.

Stakeholders in an evaluation, regardless of their perceived purpose, are constrained by the funds available for the evaluation and the political context in which the evaluation occurs. The sine qua non of any successful evaluation is the expertise of the evaluator and such expertise does not come cheaply, regardless of whether qualitative or quantitative methods are used.

Here I am glossing over another controversy in evaluation by implying that evaluations should be conducted by outside contractors rather than in-house personnel in close proximity to the program. This issue is often decided by how much the program or evaluation sponsor are willing to pay for a particular evaluation effort. In my view, mandated outcome evaluations require "built-in" funds to support contracted evaluators who are independent, non-service providers. Few in-house evaluators or program personnel have the skill or opportunity of viewing a program with the necessary impartiality or the luxury of reporting negative findings without some kind of unwelcome retribution.

In regard to the political context, the qualitative findings of an evaluation may be less acceptable to administrators, staff and program sponsors, particularly if they highlight the negative, non-quantified effects of the program. For example, in an evaluation of a team collaborative model of integrated health-mental health care in which I was the principle evaluator, the quantitative findings indicating program success passed easily through review by staff, program and evaluation sponsors to publication in a journal. The more negative qualitative findings, suggesting that the team organization was not supported by administrators, are still in the unpublished manuscripts drawer of my filing cabinet. Seldom are all key stake holders in an evaluation in the market for the whole truth.

Internal versus External Validity

The relative emphasis to be given to the internal or external validity of findings obtained using different evaluation designs is an issue that has been discussed mainly in relation to quantitative evaluations, although the findings of qualitative investigations are subject to the same logic.[1] Internal validity is the extent to which the design and execution of an evaluation allows definitive statements of outcome.

The essential question for internal validity concerns the confidence with which one can assert that exposure to the intervention produced the outcomes that were intended. The logic of internal validity is to examine rival hypotheses of program effects: what factors, other than the effects of a particular program, might influence clients of a program such that these factors would yield the desired effects? External validity is the extent to which the design of a particular evaluation allows the findings to be generalized. The crux of external validity is the extent to which one can assert that program effects will be replicated elsewhere.

Evaluators in the tradition of experimental design have paid attention to both types of validity, while emphasizing the importance of internal validity. This emphasis is not particularly surprising given the interest of most evaluators in establishing the precise cause-effect relation between program and outcome. However, the search for precision in cause-effect relations in one particular evaluation study is elusive. The finding of desired effects is, at best, temporary insurance that the prejudices of the investigation were confirmed with this particular program, with these particular outcomes, and as measured by these particular evaluators. More often, viewing outcome evaluation findings with an undue emphasis on internal validity or the precision of causal thinking leads to grave disappointment.

The MRFIT or Multiple Risk Factor Intervention Trial, is a recent example of such disappointment.[9] Over 360,000 men aged 35-57 years were examined to select nearly 13,000 at high risk who were randomly allocated to test and control groups and followed for an average of seven years. The test group was subjected to a special intervention program including treatment of high blood pressure, counseling to reduce cigarette smoking, and dietary advice aimed at decreasing serum cholesterol concentration. Control subjects were referred to their usual sources of medical care. The evaluation was conducted at 28 institutions throughout the United States. The experimental group had a modestly lower mortality rate for coronary heart disease (17.9 compared to 19.3 per 1000) and a slightly higher all causes mortality rate (41.2 compared to 40.4 per 1000). The failure to discover a significant beneficial effect has disappointed many, including the authors of the MRFIT report, who have found it difficult to accept the negative conclusions gracefully.[10]

If one views the findings of the MFRIT study from the perspective of external validity, a different conclusion can be reached. No single, large trial of a health enhancement strategy, whether it be diet modification, smoking and/or hypertension control, can provide conclusive evidence of cause and effect. Experiments, even at their best, merely "probe" causal theories; they do not "prove" them.[11] To narrow an investigation or tighten a design to increase confidence in the cause-effect relation may well reduce the relevance of an investigation. After all, planned social action is based on broad inferences that particular interventions, such as diet modification, will produce the desired effect, such as decreased disease risk, in a wide variety of settings, populations and times.

Regardless of resources, effort and time, program evaluations are not laboratory experiments within an isolated system with all the sufficient and necessary conditions in view. The MRFIT evaluation represents progress in that a relatively strong design demonstrated that a multifactor intervention had little effect on the outcome of survival in middle-aged men. This trial tested one set of much believed-in treatments, certainly not the only ones, using the outcome measure of survival, an outcome which may not have been the only one of interest to the participants. If the results had been more positive, the authors might well have recommended major behavioral changes which are unwelcomed by many U.S. men. Regardless of the sophistication, size, time, effort, and expense, "definitive" projects are only the stronger members of a collected body of evidence about cause-effect relationships. The results of so-called critical studies should be seen as predictions for future evaluations which are likely to use less rigourous methods, fewer and different participants, and perhaps a different theoretical model of what affects cardiovascular mortality, survival rate, or other outcomes of interest.

This problem of causal inference and the limitation of resources for evaluation suggest that designs which are feasible for use in many different settings and by many different investigators are most useful. Recent development of "meta-analysis," reviews of research in which the data from different studies are statistically combined, encourage comparable designs which will yield estimates of program effects which can be compared.[12] Even if not randomized trials, a series of similarly designed studies that estimate program effects can lead to conclusions which are useful guides to social action.

Comparing the results of evaluations systematically to uncover consistencies or inconsistencies in findings is to identify the limits of a particular intervention and particular theoretical idea. For example, the same worksite health education campaign to encourage weight loss may work for one group of employees in a particular industry or worksite and not for others in a different line or place of work. In this case, the stability of the program effect across different populations, investigators and times, is equally, if not more, important than the rigourousness of the causal interpretation in a single setting.

The choice of emphasis on internal or external validity in designing an outcome evaluation should occur early in the design stage of the research, bearing in mind yet again the objective of the evaluation. External validity is particularly important if one wants to advocate wider use of an intervention. In testing a brand new experimental program the issue of generalization may not be as important as the internal validity of the evaluation, that is

demonstrating that the program leads to the intended effects. On the other hand, the external validity of the new program itself may be an important issue. We all have probably wondered if the effects of a new and widely acclaimed program, such as smoking cessation clinic or a new weight loss program, aren't due as much to the zeal and dedication of the counselors as to the curriculum or program itself.

Conclusion

It is not possible to develop general guidelines that will resolve adequately all the issues which impinge upon selection of the most ideal method of program evaluation or use of evaluation resources. Every evaluation presents a number of different possibilities for the evaluator and the sponsor, and it is indeed a fine art to compose a good overall evaluation plan that can't be "blown out of the water" by one critic or another. Limited resources will always imply choices between methods, designs, activities and evaluators.

It is most important to be aware of what method choices imply in relation to: (1) the many and different purposes of the evaluation; (2) the funds made available for the research evaluation; (2) the many and varied stakeholders in the evaluation; and, (4) the political context of the evaluation, particularly the clash of interests, the prospects for disillusionment, and the tyranny of the old adage, "he who pays the piper calls the tune." Subscribing tightly to one perspective, method, or design to the exclusion of others is like shooting all photographs with the same lens at the same speed. Some photos will be seriously distorted because conditions are not the same for every snapshot.

Truly elegant program evaluation methodology incorporates both the logic of science and the logic of the real world. Taking the position that methodology is a servant to the essentially political process of evaluation, all methods are means of organizing and accumulating imperfect knowledge about what we should do with our energy and resources. Methodological controversy can be helpful in establishing the limits of any single approach and in motivating efforts to combine and integrate approaches. Such controversy is not helpful when it distracts us from finding out all we need to know about what programs accomplish and what we should expect them to accomplish.

Acknowledgment

The author thanks Shirley Beresford and William Beery for their suggestions and comments on this paper.

References:

1. Patrick, D. (1982). "Evaluation of health care." pp. 238-250 in D. Patrick and G. Scrambler (eds), Sociology as Applied to Medicine. London: Hailliere Tindall.
2. Cronbach, L. (1982). Designing Evaluations of Educational and Social Programs. San Francisco: Jossey-Bass.
3. Mitroff, I., Enshoff, J., and Kilmann, R. (1979). "Assumptions analysis: a methodology for strategic problem-solving." pp. 208-225 in L. Datta and R. Perloff (eds), Improving Evaluations. Beverly Hill, CA: Sage.
4. Hatch, O. (1982). "Evaluations of government programs." Evaluations and Program Planning 5: pp. 181-191.
5. Reichardt, C., and Cook, T.D. (1979). "Beyond qualitative versus quantitative methods." pp. 7-32 in T.D. Cook and C. Reichardt (eds), Qualitative and Quantitative Methods in Evaluation Research, Beverly Hills, CA: Sage.
6. Campbell, D. (1979). "Degrees of freedom and the case study." pp. 49-67 in T. Cook and C. Reichardt (eds), Qualitative and Quantitative Methods in Evaluation Research. Beverly Hills, CA: Sage.
7. Lawrence, J. and Cook, T.J.. (1982). "Designing useful evaluations: the stakeholder survey." Evaluation and Program Planning 5: pp. 327-336.
8. Wholey, J. (1977). "Evaluability assessment." pp. 41-56 in L. Rutman (ed), Evaluation Research Methods. Beverly Hills, CA: Sage.
9. MRFIT Research group (1982). "Multiple Risk Factor Intervention Trial.: JAMA 248: pp. 1465-1477.
10. Stallones, R. (1983). "Mortality and the Multiple Risk Factor Intervention Trial." American J. Epid. 117(6): pp. 647-650.
11. Campbell, D. and Stanley, J. (1966). Experimental and Quasi-Experimental Designs for Research. Skokie, Ill.: Rand McNally.
12. Glass, G., McGaw, B., and Smith, M.L. (1981) Meta-Analysis in Social Research. Beverly Hills, CA: Sage.

Planning and Evaluation Using Vital and Health Statistics

Session T

CREATION AND IMPLEMENTATION OF MATERNAL AND CHILD HEALTH NEED INDICES FOR ALLOCATION OF BLOCK GRANT FUNDS IN OHIO

Michael Fleming, Ohio Department of Health

Background

Beginning with Fiscal Year 1981, federal dollars that once went to categorical health programs have been given to the states in the form of block grants. The three main health block grants are for preventive health, primary care, and maternal and child health. Of the three, the maternal and child health (MCH) block grant is the most significant for states. It greatly exceeds the preventive health block grant in size, with $478 million versus $85 million in total national funding for FY 1983. Although the MCH block grant potentially is smaller than the primary care block grant, the latter seldom is accepted by states because of unfavorable administrative restrictions.[1]

Ohio's FY 1983 MCH block grant is approximately $6 million, an amount which is expected to increase by several million dollars over the next few years. As is the case for most states, Ohio has been faced with the question of how to spend its block grant funds in an appropriate and acceptable way. There are essentially three alternative answers. The first alternative is the status quo, i.e., to continue funding the present categorical, relatively uncoordinated programs, located around the state in an historical, idiosyncratic pattern. The second alternative is for the state to divest itself of responsibility for administering block grant funds, most simply by allocating dollars to counties on a strict per capita basis. Such an approach has an important disadvantage in that it implicitly ignores differences of need across the state.

Innovation is the third alternative for managing MCH block grant funds. It is this option that Ohio has chosen. The Ohio Department of Health's Bureau of Maternal and Child Health is channeling most of Ohio's block grant dollars into a new effort to provide comprehensive, coordinated maternal and child health services where needed in the state. The initiative, called the Child and Family Health Services (CFHS) program, combines multiple categorical programs under one funding umbrella. The CFHS program includes five basic services: 1) program planning and development; 2) public health; 3) family planning; 4) perinatal health; and 5) infant, child, and adolescent health. Regional coordination is emphasized by requiring locally developed CFHS plans of at least countywide scope.

Having decided to use MCH block grant funds for the innovative Child and Family Health Services program, the Bureau of Maternal and Child Health confronted a further question: who in the state should receive how many CFHS dollars? The Bureau's answer, consistent with the spirit of the CFHS program itself, has been that funds should be allocated according to "need."

When the CFHS program first was conceived, no acceptable measures of need for comprehensive maternal and child health services were available. The Bureau of Maternal and Child Health has worked for approximately the last year with the staff of the Ohio State Health Planning and Development Agency (SHPDA) in order to develop need indices. The outcome of that joint effort is the topic of this paper.

Development of the Indices

In developing indices of MCH need, a number of explicit and implicit criteria have been employed. Among the criteria are the following:

Need should be related to socioeconomic status. The CFHS program targets the poor and/or medically indigent population.

Need should be related to health status outcomes. The goal of CFHS is to improve the health of mothers and children.

Need should be measured both in terms of large numbers and large proportions of needy mothers and children. Magnitude and intensity are two distinct and equally valid ways of characterizing need. In making allocation decisions, both kinds of information are useful.

Need indices should be sensitive to administrative priorities. The definition of need in any context requires value judgments. The values embodied in administrative priorities should be integrated clearly into indices of need.

Need indices should not presume a certain mix of MCH services. The CFHS program is designed to allow for local autonomy; defining need in terms of specific required levels of services would greatly restrict such autonomy.

Need indices should be conceptually and mathematically consistent. Although an assessment of need should take into account value judgments, the overall framework for determining need should be as rational as possible.

Indices should discriminate need at a county level. CFHS grants are intended to be county-based, so need indices must correspond.

Need indices should be understandable to lay persons. Acceptance of allocation based on need depends, in part, on the public's ability to understand the assessment of need. Excessive technicality may inhibit such understanding.

Need indices should incorporate available and readily updateable data. Because it is anticipated that the CFHS program will extend for a number of years, present indices must be applicable using new data to assess need on a consistent basis in the future.

Need indices should be operational within a limited time. Given that the beginning of the CFHS program originally was planned for the start of FY 1984, the target date for completion of the needs indices was April 30 of this year.

To meet the preceeding criteria, a rigorous epidemi-

ological/statistical model would require: first, tested knowledge of the relationships among economic status variables, health status variables, and the need for comprehensive MCH services; and, second, local data adequate for projecting such relationships to counties in Ohio. Neither requirement can be met at the present time; therefore, another kind of rationale has been developed for determining need.

The alternative rationale may be called, for want of a better term, an "arbitrary deductive" model of need. The model is arbitrary because it includes factors based on administrative and professional judgment rather than on epidemiological or statistical considerations. And it is deductive in the sense that it was formulated by starting with certain ideas about what should be in an MCH needs model and then deducing from those ideas the elements of the model and their relationships.

The Model[2]

The central elements of the model are two general indices of maternal and child health need: total need magnitude and total need intensity. Total need magnitude is intended to be an index -- although not an actual count -- of the number of persons in a given county in need of comprehensive MCH services. Similarly, total need intensity is an index of the proportion of persons in a given county in need of comprehensive MCH services. Because total need intensity is derived from total need magnitude, it is useful to begin describing the model by examining the latter index. In the narrative below, the general structure of the total need magnitude index is discussed, followed by a more detailed treatment of its elements and subelements.

A basic premise of the model is that total need magnitude should be a function of the separate needs of different MCH subpopulations. For purposes of the model, MCH subpopulations are defined as follows: 1) infants -- ages less than one year; 2) children -- ages 1 - 13 years; 3) adolescents -- ages 14 - 17 years; and 4) women of childbearing years -- ages 18 - 44 years. The chosen age-ranges vary somewhat from commonly used definitions in order to conform to the data available and to avoid age overlapping.

A second premise of the model is that total need magnitude should reflect not only the needs of the MCH subpopulations, but also the relative administrative values placed on MCH services associated with each subpopulation. In other words, administrative priorities concerning the needs of each subpopulation should be expressed explicitly in the total need magnitude index.

Based on the preceeding premises, the model's index of total need magnitude is given by

NMTOTAL = (NMINF x WINF) + (NMCHILD x WCHILD) + (NMADOL x WADOL) + (NMWOM x WWOM)

where

NMINF, NMCHILD, NMADOL, and NMWOM are county need magnitudes for each of the MCH subpopulation groups, i.e., infants, children, adolescents, and women of childbearing ages, respectively

and

WINF, WCHILD, WADOL, and WWOM are the corresponding value weights associated with each subpopulation.

In order to make the total need magnitude index operational, it is necessary to further define the subpopulation need magnitudes and value weights. Value weights will be discussed later; first, it is important to define the subpopulation need magnitudes and their components. Subpopulation need magnitudes in the model are functions of subpopulation need intensities and the size of each subpopulation. More formally, the need magnitude for each subpopulation is given by

NMsubpop = NIsubpop x Psubpop

where

NIsubpop is a county's intensity of need for a given subpopulation, and Psubpop is the county population size of the same subpopulation.[3]

Just as subpopulation need magnitudes must be defined in order to make the total need magnitude index operational, so must subpopulation need intensities be defined. Intensity of need for each MCH subpopulation is a function of both the economic status and health status of mothers and children in a given county. Symbolically, the need intensity for each subpopulation, is given by

NIsubpop = EHH x Hsubpop

where

EHH is a county's economic status index and Hsubpop is that county's relative health status index for a given subpopulation.

The economic status index is an adjusted county poverty rate. The base poverty measure is the proportion of families below 125% of the poverty level in 1979.[4] That rate is adjusted for estimated changes in the proportion of poor due to changes in unemployment from 1979 to 1982.[5] In general, the adjustments are relatively small, amounting to only a 3% increase in the base poverty measure if a county experienced an increase in unemployment of 10% over the three year period.[6]

Relative health status for each subpopulation is determined by taking the ratio of a selected health indicator rate for a given county to the same health indicator rate for the state as a whole.[3] In other words, the relative health status indices are constructed as odds ratios, which approximate relative risk.

To illustrate how the relative health status indices are formulated, consider that for the infant subpopulation. The health indicator rate chosen for infants is infant plus fetal deaths per 1,000 live births over the period 1977 - 1981.[7] The inclusion of fetal deaths in the indicator and aggregation over a five-year period are designed to enhance the statistical reliability by increasing the number of events. The relative health status index for infants is computed by dividing a given county's infant plus fetal death rate by the corresponding rate for the state.

The health indicator rates on which the relative health status indices of the other MCH subpopulations are based are:

children -- the 1981 proportion of school age children with selected handicapping conditions;[8]
adolescents -- a combination of the same handicap indicator used for children, plus the 1979 -1981 rate of births to mothers less than 18 years of age;[8,9,10]
women of childbearing ages -- the 1979 - 1981 rate of births to women ages 18 - 44 years.[10]

Turning from definition of subpopulation need magnitudes, subpopulation value weights now may be considered. Defining subpopulation weights is a professional and administrative decision rather than a technical one. However, an attempt has been made to select value weights on as reasonable a basis as possible. The value weights selected are:

infants -- 5.0 (WINF);
children -- 1.5 (WCHILD);
adolescents -- 1.5 (WADOL);
women of childbearing ages -- 1.0 (WWOM).

The chosen weights reflect several considerations. Infants are given the highest weight because on the average, infant health problems -- notably mortality -- are the most serious, and because many MCH services are targeted on infant health. Women of childbearing ages are given the lowest weight because much of MCH treatment given to women is for the sake of their offspring, who are counted in the other subpopulations, and because the MCH services for women are typically (although not always) aimed toward a subset of their health needs. Children and adolescents are given the same weight because both groups have relatively low morbidity and mortality, and many of their health needs have to do with passage through stages of growth and physical development. Adolescents do have additional need caused by their increasing reproductive roles, but it may be argued that such a factor is balanced by the importance of health care in the earlier developmental ages of childhood.

In addition to the total need magnitude index discussed up to this point, the need model includes another central index -- total need intensity. Total need intensity is determined by dividing the total need magnitude for a county by the sum of the county MCH subpopulations. Note that by defining total need intensity in this way, the index is not a direct function of subpopulation need intensities. Instead, just as for total need magnitude, total need intensity is influenced by the value weights assigned to subpopulations and by subpopulation need magnitudes. In contrast to the total need magnitude index, however, total need intensity adjusts for county population size so that larger and smaller counties may be compared with regard to the proportion in need rather than the number in need.

Advantages and Disadvantages of the Model

The advantages of the Ohio MCH needs model can be summarized by observing that it meets all the criteria described earlier. The model assesses need both in terms of numbers of needy persons and in terms proportions of needy persons in the population, with its total need magnitude index and total need intensity index, respectively. Need as measured by the two central indices relates to socioeconomic status as well as to health status, via the component economic status and relative health status indices. The model's value weights assure that need is expressly sensitive to administrative priorities. The analyses of need produced by the model are at the required county level, and the data used are both available and updateable. No particular mix of MCH services is presumed by the model. The model is consistent in that its mathematical structure expresses reasonable concepts concerning MCH need, subpopulation components generally do not overlap, and parallel indicators measure comparable characteristics. The model also is readily interpretable to the lay person because its concepts and arithmetic are straightforward. Not least importantly, the model has produced useful results within the time constraints required for the start-up of the Child and Family Health Service program.

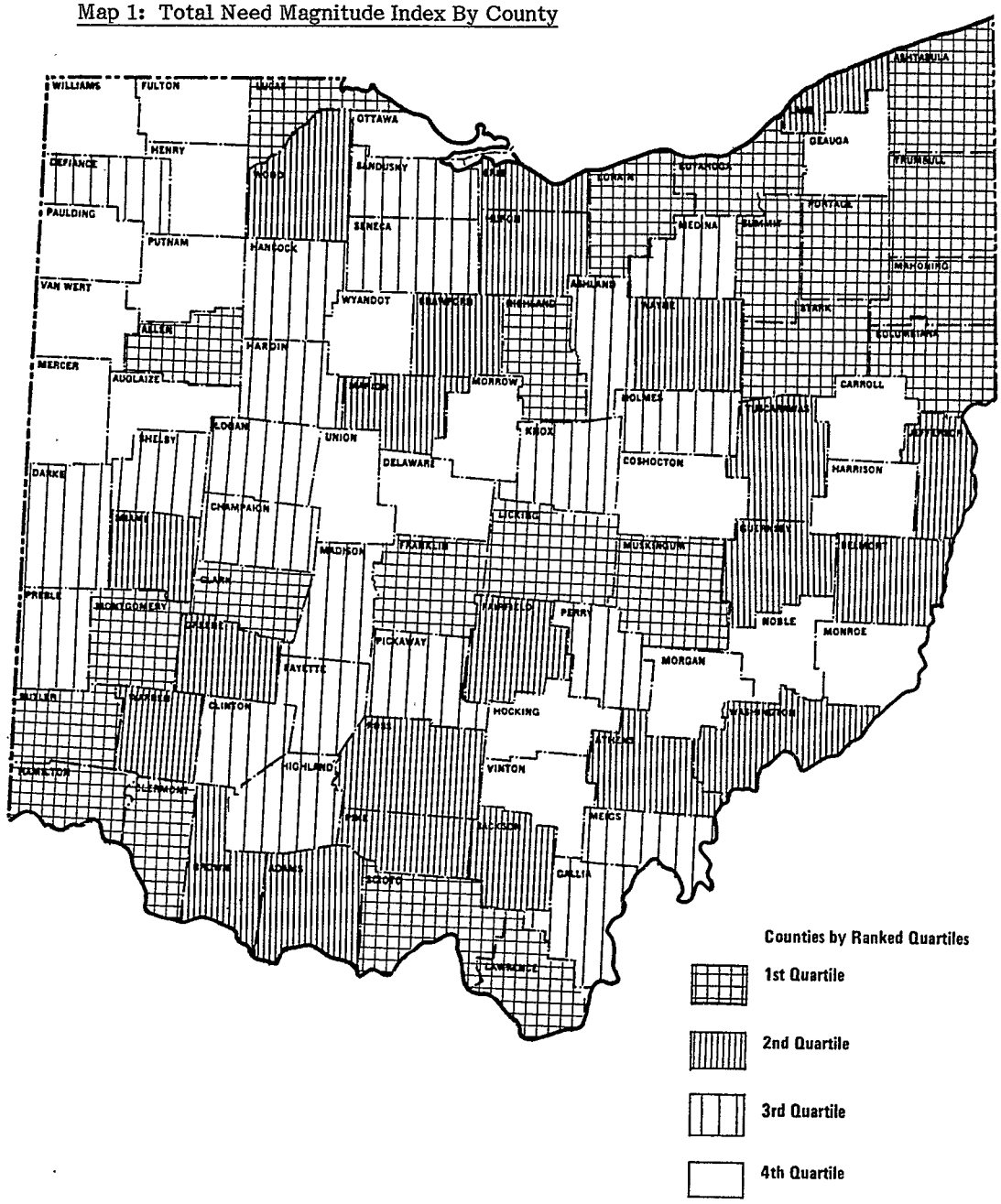
Without itemizing all the potential and actual disadvantages of the needs model, three potential areas of concern may be cited. The first potential area of concern involves the choice of indicators, especially health status indicators, incorporated in the model. One aspect of this concern is validity -- do the indicators truly represent health status problems relevant to the need for comprehensive maternal and child health services? For example, are handicapping conditions in schoolchildren reported consistently? And if they are, are handicaps among children a good predictor of child health services needed in MCH clinics? Another aspect of concern regarding the choice of indicators is reliability -- do variations in the indicator data relate to true rather than random differences in health status? There may be an issue, for example, as to whether a small county has enough infant and fetal deaths to be confident that its five-year infant plus fetal mortality rate gives a good approximation of the "true" underlying rate.

During the development of the needs model, the problems of data validity and reliability were given considerable attention. The chosen indicators are intended to be the most relevant available at a county level in Ohio, although some of the indicators -- such as the handicap data -- are less satisfactory than others. Where appropriate, data have been aggregated over time to improve statistical reliability. The advantages of such aggregation, however, have been balanced against the need for currency of the data: no more than 5 years have been aggregated for any given data element.

A second potential area of concern in relation to the needs model is the choice of value weights. It is not unlikely that another set of actors might place a different emphasis on the need associated with each of the MCH subpopulation groups. Staff of another state's health department, for example, might weight child need at twice infant need rather than at 3.3 times infant need as in the present model. Such value differences are, however, at least as much of an advantage as a disadvantage for the model. No assessment of need can be value-free, and values vary almost by definition with the actors involved. That the value weights used in the Ohio model are unique simply means that the assessment of need is well-tailored to the particular administrative priorities of the Ohio Department of Health.

The third potential area of concern regarding the needs model has to do with its general mathematical and

Map 1: Total Need Magnitude Index By County



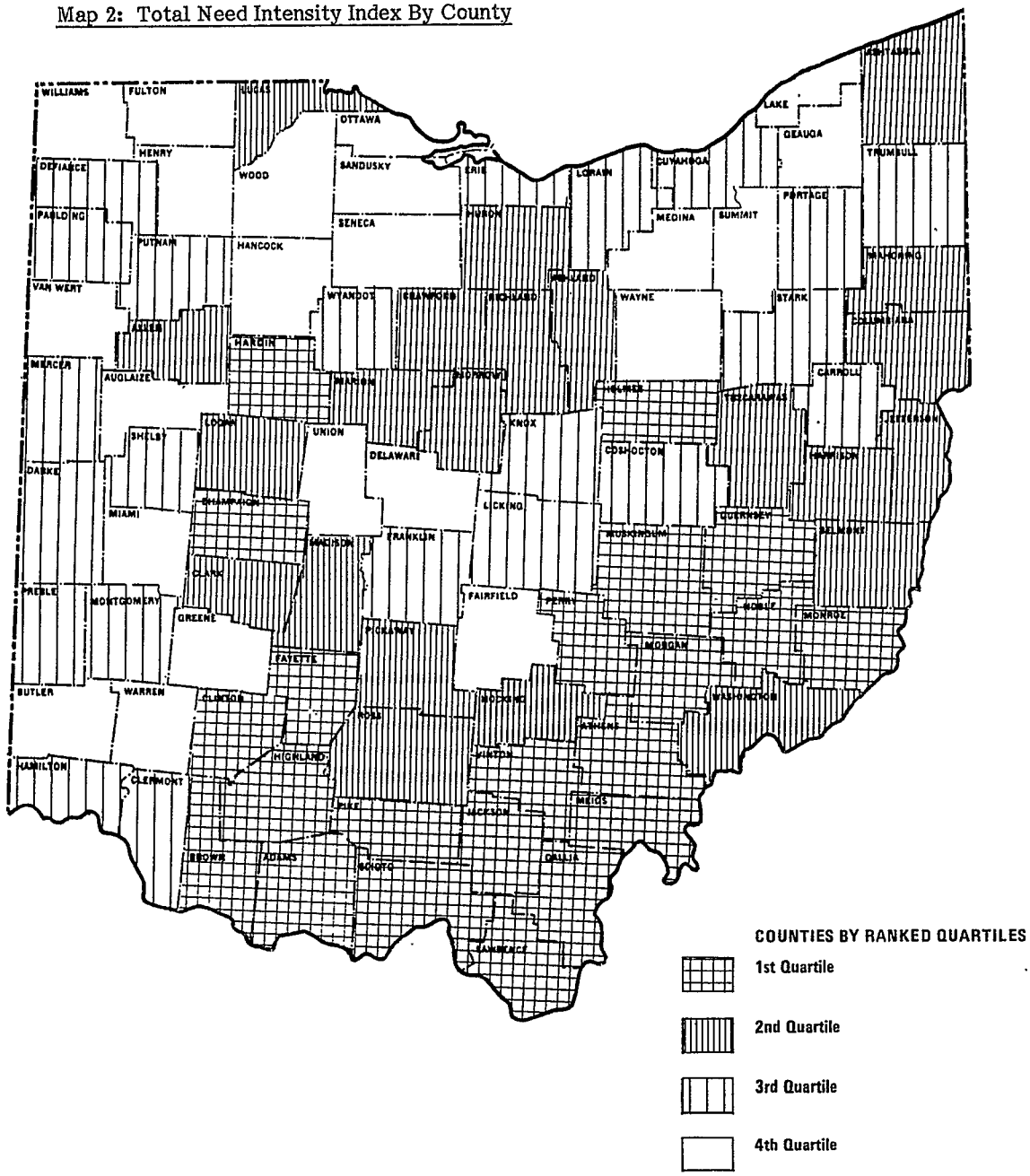
conceptual structure. As mentioned earlier, the model may be characterized as being arbitrary and deductive rather than epidemiological and/or statistical. In other words, the model is not scientific. One consequence is that it is not possible to make direct inferences about the numbers and proportions of needy from the model's results. Nevertheless, the model does produce results that are rational and broadly equitable. Given the criteria used in developing the MCH needs model for Ohio, no more rigorous scientific approach has appeared that would be similarly effective.

Results and Their Use

The results of applying the MCH needs model to Ohio's counties are quite consistent with what is known already about the general patterns of need in the state. Map 1 shows the by-county distribution of the model's total need magnitude index. Although the need model is intended to help avoid a strictly population-based approach to allocating MCH funds, the expectation is that magnitude of need should have a strong relationship to county size. The pattern of need displayed in Map 1 confirms such an expectation, in that 18 of the 22 counties in the first (highest) quartile of need magnitude are also in the top quartile of counties ranked by total population.

Map 2 shows the by-county distribution of the model's total need intensity index. Again, the pattern confirms expectations. On many measures, the Appalachian region of Ohio historically has been considered to be the "neediest" in the state. Of the 22 counties that the map shows as having the highest need intensity, 18 are in the region officially defined

Map 2: Total Need Intensity Index By County



as Appalachia, and several others border the region.[11]

Results of the Ohio MCH needs model are to be used in several different ways to assist in allocating MCH block grant dollars through the Child and Family Health Services program. The first use of the model's results, which has already been implemented, is to estimate maximum county funding levels so that grant applicants may develop appropriate budgets. The total need magnitude index has been employed for this purpose. Total need magnitude is, however, not the only factor used in estimating maximum county funding; also considered are current funding levels, Title X family planning dollar allocations, and estimates of the likely number and distribution of counties to be funded under the CFHS program.

The second use of the model's results will be in the process of objective review to decide whether or not particular applications will be approved. In that process, one-quarter of the review points will be based on the value of the total need intensity index for the county of the applicant. A final use of the model's results will come when actual funding levels for approved applications are determined. The total need magnitude index will be employed to help set funding levels, although its role has not been precisely defined as yet.

Conclusion

Ohio's MCH needs model is proving to be a useful administrative tool in block grant allocation. The model is particularly well suited to determining needs for comprehensive MCH services, such as those to be provided through Ohio's innovative Child and Family

Health Service program. Although the model is not a perfected tool and contains features unique to Ohio's purposes, the concepts it embodies may be useful to other states facing similar block grant allocation problems.

In the era of block grants, the buck literally has been passed to the states. There is an urgent need to develop and improve the knowledge base necessary to spend those dollars in a responsible way. Statewide need models are not the only analytic tools states and localities require to administer health block grant funds effectively, efficiently, and equitably. Other important information gathering and analysis tasks include: 1) detailed assessments of local needs to direct actual service provision; 2) development of local and state program monitoring systems; 3) evaluation studies to chart progress toward program service and outcome objectives; and 4) epidemiological research to help determine the intervention strategies most likely to be effective. The overall goal of such efforts should be to create integrated planning and management information systems, of which statewide need assessments may be one part.

Notes:

1. Fiscal year 1983 health block grant dollars as reported by U.S. Department of Health and Human Services -- preventive block grant information from the Center for Disease Control; information for the other two block grants from the Health Resources and Services Administration. Contacts with USDHSS by staff of the Institute for Health Planning, Madison, Wisconsin.

2. Complete explanation and documentation of the Ohio MCH needs model methodology is available in an Ohio Department of Health draft document entitled "Outline of Model for Assessment of County Maternal and Child Health Needs," Columbus, Ohio, 8/8/83.

3. Population data used in calculating subpopulation need magnitudes and in calculating health status indicator rates taken from: U.S. Bureau of the Census, Census of Population and Housing 1980: Summary Tape File 1A 'Ohio' Washington, D.C.: 1982 (machine readable data file).

4. Data on number of families below 125% of the poverty level in 1980 taken from: U.S. Bureau of the Census, SFT3 Socioeconomic Report, generated from Census of Population and Housing 1980: Summary Tape 3A 'Ohio' (machine readable data file) Washington, D.C., 1982. Prepared by Ohio Data Users Center, Ohio Department of Development, Columbus, Ohio.

5. Data on unemployment rates in 1979 and 1982 taken from: Ohio Bureau of Employment Services, Division of Research and Statistics, 'Estimates of Average Civilian Labor Force, Overall Employment and Unemployment in Ohio by County (for the years 1979 and 1982)', Ohio Labor Market Information, 1980 and 1983.

6. The method of adjusting the poverty data for changes in unemployment is based on national data concerning the relative proportions of employed and unemployed who are poor (see note 2).

7. Data on infant deaths, fetal deaths, and live births 1977-1981 taken from: Ohio Department of Health, Division of Data Services, Statistical Analysis Unit (for Division of Vital Statistics), Vital Statistics Annual Reports for the years 1977 through 1981. Columbus, Ohio. Published 1979 through 1982.

8. Handicapping conditions included in child and adolescent health status indicators are hearing handicaps, visual handicaps, orthopedic or other handicaps, and developmental handicaps. Data have been adjusted to moderate the influence of developmental handicaps on rates (see note 2). Data source: 'Handicapped Child Data' forms for each Ohio school district as of December, 1982. Access to forms provided by the Ohio Department of Education, Division of Special Education.

9. In formulating the relative health status index for adolescents, handicapping conditions data are given a weight of 0.25 and adolescent births data are given a weight of 0.75 (see note 2).

10. Data for 1979-1981 births to females less than 18 years of age and to females ages 18 and over are from: Ohio Department of Health, Division of Data Services, Statistical Analysis Unit (for Division of Vital Statistics), Vital Statistics Annual Reports for the years 1979 through 1981. Columbus, Ohio. Published 1981 through 1983.

11. Information on the officially defined Appalachian counties of Ohio obtained from the Ohio Department of Development, 1983.

EVALUATION OF THE WIC PROGRAM THROUGH LINKAGE OF WIC MANAGEMENT DATA AND THE STATE VITAL STATISTICS REGISTRY

Milton Kotelchuck, Janet B. Schwartz, Marlene T. Anderka, Karl Finison

Massachusetts Department of Public Health

Introduction

The imaginative use of data can open new possibilities and can solve old problems. This report describes how the linkage of two existing data sets allowed for a rigorous evaluation of the Massachusetts WIC Program. By linking two data sets that didn't normally go together, we were able to answer questions that neither data set alone could answer. WIC Management Reports were merged with the State Birth and Death Registry. This information was then used to assess the impact of prenatal participation in WIC on the outcome of pregnancy.

WIC Program Description

The Special Supplemental Food Program for Women, Infants and Children (WIC) was established by Congress in 1972. It is the largest and most specifically targeted public health nutrition program in the country. The goal of WIC is to provide supplemental foods and nutrition education as a part of an individual's health care, during critical times of growth and development. This program was set up to prevent the occurrence of health problems and to improve the health status of those persons it serves.

WIC is administered at the federal level by the Department of Agriculture. Its role is to allocate money to state health agencies and Indian agencies that oversee the operation of the program at the local level. WIC is a voluntary program that ultimately operates through local sites in the community. Funded nationally at $1.06 billion in fiscal year 1983, WIC Programs serve approximately 2.3 million persons per month, including over 400,000 pregnant women.

The WIC Program is targeted to high risk pregnant, postpartum and breastfeeding women, infants and children under five years of age. Two factors determine eligibility for the program: income less than 185% of the OMB poverty guidelines and being at nutritional risk. Pregnant women are considered nutritionally at risk if they are, for example, younger than 19 or older than 35, have a poor obstetrical history, are anemic, have experienced too little or too much weight gain or have an inadequate dietary intake. WIC is the first federally funded program to directly assess nutritional status as a criterion for eligibility. In 1978 and in prior years, geographic eligibility was a third criterion because of the limited WIC funding. WIC is now available statewide.

Once certified as eligible, a pregnant woman receives nutrition counseling and a monthly set of food vouchers which are redeemable at local grocers for specific nutritious foods. These foods are tailored to individual needs. Available foods include milk, cheese, eggs, iron-fortified cereal and 100% fruit juices. In Massachusetts, food vouchers are monitored by a computer system, established to keep track of the issuance and redemption of vouchers for fiscal purposes.

Difficulties of Evaluating WIC

WIC is a complex public health program. In its basic design, which is its strength, WIC is multifaceted. It offers nutrition counseling, nutritious foods and a link to the health care system. Although this is an effective means of delivering services, it makes the program's impact difficult to determine. Prior evaluations have been hindered by methodological issues as well. There have been 3 basic problems encountered. The first problem has been a lack of uniformly collected data. Typical of public health programs, data is collected on all persons participating in the WIC Program. This data is collected by many different people at local sites in the community and there is not a standardized record-keeping system across programs. The second problem has been difficulty in obtaining a proper comparison sample. Who are you going to get to compare these women to? Where do you have a group of women that you have similar information on? The third problem has been small sample sizes. The effects that you are expecting are small and therefore you need to have a large sample size to see them.

Researchers have grappled with these problems since the program's inception. As a result of increased political pressure to determine the usefulness of the program, investigators were encouraged to explore resolutions to these methodological problems.

Design of Massachusetts State-Wide Evaluation Project

At the Massachusetts Department of Public Health, we were trying to solve these methodological problems. We realized that, although WIC records did not have birth outcome data for prenatal participants, the State Birth Registry did contain information on every birth in the State.

This paper is, in part, a revision of a paper that was presented at the Annual Meeting of the APHA, Los Angeles, October, 1981. An expanded version is presently under consideration for publication.

TABLE 1 - Data Available From Massachusetts Birth Registry (1978)

Public Information
1. Infant's Name
2. Infant's Sex
3. Plurality
4. Date of Birth
5. Father's Name, Birthplace and Age
6. Mother's Name, Birthplace and Age
Confidential Information
7. Father's Race
8. Father's Education
9. Mother's Race
10. Mother's Education
11. Number and Date of Previous Live Births
12. Number and Date of Previous Terminations
13. Date of Last Normal Menses
14. Month of Pregnancy Prenatal Care Began
15. Total Number of Prenatal Visits
16. Complications of Pregnancy, Labor and Delivery
17. Congenital Malformations or Anomalies of Infant
18. Birth Weight
19. One and Five Minute Apgar Score

TABLE 1 is a list of data items available on the Massachusetts Birth Registry. If we could figure out how to link the WIC data and Birth Registry Data, we would have all the information needed to perform an effective, methodologically sound evaluation of the WIC program. Specifically, we would have: uniformly collected outcome data, a comparison sample on whom we had similar information, and a large sample size.

Such an evaluation would be able to answer the following questions:

1. Does WIC reach its high risk target population?
2. Is WIC participation associated with improved outcomes of pregnancy?
3. Does duration on WIC relate to relative improvements in birth outcomes?
4. Is improvement in pregnancy outcomes strongest in the highest risk groups?

Methodology

Doing this study involved four steps: determining who was a WIC participant; linking WIC data with birth certificates; finding a matched control group; and comparing birth outcomes of the WIC and controls.

No new data was collected for this study. Data was drawn from two existing, computerized data sources: WIC Management Data, and the State Birth and Death Registry.

The first step was to develop a group of cases (i.e., women who participated in WIC and delivered in 1978, and who could be linked with their babies' birth certificates). 1978 was chosen because when we began the study a few years back, this was the most recent year that we had complete WIC and State vital statistics records available.

Using WIC management records, we made a list of all names of women who participated in WIC and could possibly have given birth in 1978. We recorded information, including: the name, town of residence, months of participation, number of uncashed vouchers, and delivery date for these women. We excluded from the study women who had not cashed any vouchers or who were terminated from the program for cause.

With this information, we went to the Birth Registry and found the birth certificates of the infants born to these mothers. In other words, we linked mother to baby. This linkage was a tedious process that involved actually sitting down with two lists: the list with the names of the WIC women, and a list of birth certificate information sorted by birthdate or last name, and then, by hand, searching for the links. At this point, we had a group of WIC cases on whom we had information on both WIC participation and on pregnancy outcome.

The next step was to find a group of matched controls. There were 68,000 births in Massachusetts in 1978. By removing all birth certificates of babies born to WIC women from the Birth Registry, we were left with 64,000 potential controls. These controls had the same data as the WIC women, but they were not WIC participants.

Using five demographic variables available on the birth certificates, each WIC woman was individually matched to a single control. TABLE 2 lists the matching variables and the ranges within which we matched. These matching variables were: age, race, parity, education and marital status.

TABLE 2 - Matching Criteria

Age:	15 & under, 16-17, 18-19, 20-24, 25-29, 30-34, 35+
Race*:	Black, White, Oriental, Other
Parity:	1,2,3,4,5+
Years of Education:	8 & under, 9-11, 12, 13-16, 17 or more
Marital Status:	Married, Unmarried

*Hispanic ethnicity is coded racially as White on birth certificates following the NCHS convention.

TABLE 3 - Derivation of Study Population

Number of Names of WIC Prenatal Participants		4,898
Number of Excluded Names	525	
Known moved out-of-state	18	
Known abortions and miscarriages	62	
Terminated for cause from program*	353	
Computer errors	82	
Number of Names Eligible for Study		4,373
Number of Omitted and Unlinked Names	252	
Twins omitted	46	
Stillbirths omitted (no birth certificates)	15	
Unlinked (no birth certificate found)**	191	
Number of Unmatched Names	5	
No Control found	5	
Number of WIC Prenatal Participants Linked to Their Infant's Birth Certificate and Matched to a Control		4,126
Percent Study Cases of Eligible Names		95

*Causes for termination include: non-use of issued vouchers, no longer at nutritional risk, violators of regulations, no longer income eligible, possible fraud.

**Reasons include: moved out of state, name changes, possible fraud.

This matching allowed us to control for some of the factors that influence birth outcomes. Other important factors known to influence birth outcomes, such as smoking, maternal weight gain, and material height, were not available to us. However, given the strict matching criteria, there is no reason to believe that these variables would be differentially distributed in the two groups.

Only five women had to be dropped from the study because we could not locate a proper control. At this point we had data on 4,126 women who participated in the Massachusetts WIC Program in 1978, and 4,126 individually matched controls who delivered in 1978 but who did not participate in WIC. The derivation of the study population is summarized in TABLE 3. As you can see, we were able to link and match 95% of the eligible names.

The study design allowed us to evaluate outcome measures that were available on the birth certificates (see TABLE 4). We were able to look at measures of: birth weight, gestational age, morbidity, mortality, and adequacy of prenatal care. Statistical differences in the outcome measures between the two groups (WIC and control) were established using a paired t-test for continuous data items such as

TABLE 4 - Outcome Variables

Birth Weight
Birth Weight
Low Birth Weight $\leq$2500 grams
Birth weight adjusted for gestational age
Small for gestational age
Gestation
Length of gestation
Prematurity
Morbidity
Complications of pregnancy, labor and delivery
Apgar scores (one and five minute)
Malformations
Mortality
Neonatal deaths
Prenatal Care
Number of prenatal visits
Month prenatal care began
Adequacy of prenatal care index
Percent receiving inadequate, intermediate, and adequate prenatal care

TABLE 5 - Selected Maternal Demographic Characteristics: WIC Participants, Catchment Area Residents and All State Residents

Age	WIC Participants	Catchment Area Residents	All State Residents
≤17 Years	12.2%	6.0%	3.8%
≤19 Years	28.6%	16.9%	11.5%
Education			
< 9 Years	14.9%	10.5%	5.1%
<12 Years	49.2%	31.5%	19.0%
Marital Status			
Married	59.3%	76.1%	86.3%
Unmarried	40.7%	23.9%	13.7%
Race			
White	73.6%	81.6%	91.8%
Black	23.8%	16.0%	6.2%
Parity			
1	44.9%	45.9%	44.6%
5+	6.5%	1.1%	1.1%

birth weight and gestational age, and a McNemar Chi-Square for ordinal data items such as low birth weight and neonatal deaths. If data was missing on either the subject or the control, the pair was eliminated from the analysis on that outcome variable. In sum, the basic design was a comparison of the WIC to the control group.

Results

1. Does WIC reach its high risk target population?

The answer is yes, WIC appears to be highly successful in enrolling women whose demographic characteristics indicate that they may be at high risk for nutritional problems. Demographic characteristics of WIC participants were compared with demographic characteristics of all women giving birth in the study year in WIC catchment areas and in the State, using data from the Birth Registry. Higher percentages of WIC women were in the high risk categories. WIC women were more likely to be young, unmarried, Black and to have had less than a high school education.

2. Is WIC participation associated with improved outcomes of pregnancy?

The answer is yes, taken as a whole, WIC appears to be associated with improved outcomes of pregnancy (see TABLE 6). WIC participation was associated with an overall improvement in birth weight, although not quite reaching statistical significance. WIC was significantly associated with: fewer low birth weight infants, an increase in gestational age, a decrease

TABLE 6 - Comparison of WIC and Control Birth Outcomes

	WIC	Control	p-Value
Birth Weight Findings			
Birth Weight (in grams)	3281	3260	.087
Percent Low Birth Weight	6.9	8.7	.005
Percent Small for Gestational Age*	4.6	5.0	n.s
Gestation Adjusted Birth Weight (in grams)**	-52.0	-48.4	n.s.
Gestation Findings			
Gestation Age (in weeks)	40.0	39.7	.001
Percent Premature	5.8	6.8	n.s.
Morbidity			
Percent with Complications of Pregnancy, Delivery and Labor	20.2	21.1	n.s.
Percent with Congenital Malformations	1.7	1.7	n.s.
Percent Low (≤5) Apgar Score (one minute)	8.2	8.2	n.s.
Percent Low (≤5) Apgar Score (five minutes)	.5	1.0	n.s.
Mortality			
Number of Neonatal Deaths	12	35	.005
Prenatal Care			
Number of Prenatal Visits	11.2	10.8	.001
Month Prenatal Care Began	2.7	2.9	.001
Adequacy of Prenatal Care Index***	1.1	1.2	.001
Percent with Inadequate Care	3.8	7.0	.01
Percent with Intermediate Care	26.5	26.6	n.s.
Percent with Adequate Care	69.7	66.4	n.s.

*SGA is defined as weighing below the 10th %ile for gestational age at birth. Figures derived from Battaglia and Lubchecno.
**The gestational correction for birth weight is determined by subtracting the observed birth weight from the mean Massachusetts birth weight for that gestational age.
***Adequacy of care is a 3-point index, combining the number of prenatal visits and month prenatal care began with an adjustment for gestational age.

TABLE 7 - Selected Outcome Measures by Duration WIC Participation

Measure	Length of Participation 1-3 Months	4-6 Months	7-9 Months
Birth Weight Difference (Grams)	-23	10	111.2***
Corrected Birth Weight (Grams)	-28	-3	34
Low Birth Weight Ratio (WIC: Control)	112:106	140:176*	31:78***
Neonatal Deaths (WIC: Control)	7:12	5:15*	0:8*
Gestational Age Difference (Weeks)	.07	.11	.72***
Adequacy Index Difference	-.07***	+.07***	+.23***
Number of Prenatal Visits Difference	-.20	.45***	1.15***

p<.0001 = ***
p< .01 = **
p< .05 = *

in neonatal deaths, and an improvement in the prenatal care index. We found no statistical difference in the morbidity factors. The low birth weight result represents a 21% decrease in low birth weight for the WIC group.

3. Does duration on WIC relate to the relative improvement in birth outcomes?

The answer is yes, there appears to be a direct correlation between increased length of participation and more positive outcomes. This correlation can be seen in FIGURE 1, which shows the difference in the mean birth weight between the WIC subjects and the controls by the length of participation in WIC.

The strongest benefits were associated with participation in WIC for seven to nine months. In the WIC group we see a 111 gram birth weight increase along with a decrease in low birth weight infants and no neonatal mortality (see TABLE 7). There was, in fact, no neonatal mortality in the WIC group for those women participating for six or more months, as can be seen on TABLE 8.

4. Is the improvement in pregnancy outcomes strongest in the highest risk groups?

The answer is yes, the improvements in pregnancy outcomes were strongest in the highest risk groups. An example of this can be seen in FIGURE 2, which shows the mean birth weight by the age of mother for the WIC and control groups. You can see that the younger the women, the more powerful the WIC effect.

Discussion

This evaluation has been very useful to the WIC Program, both nationally and in Massachusetts. WIC administrations in Massachusetts have since become more attuned to using evaluation data for policy and planning. Currently, steps are being taken to make the linkage of WIC data and Birth Registry data a permanent feature of the program in our state. This is part of a larger effort to establish ongoing evaluation.

The ability to link two discrepant data sets (Birth and Death Registry data with WIC data) was the key to evaluating pregnancy outcomes of women on WIC. There are many possibilities that arise from this kind of linking of pre-existing data sets. This study demonstrates that evaluation need not require new data bases; high quality evaluation of public health programs is possible by creatively using existing data sets.

TABLE 8 - Number of Neonatal Deaths by Duration of Participation in WIC

Duration of Participation (In Months)	N	WIC	Control	State
1	216	0	2	
2	508	3	4	
3	642	4	6	
4	654	4	7	
5	584	1	4	
6	611	0	4	
7	535	0	3	
8	317	0	5	
9	58	0	0	
TOTAL	4126	12	35	522

FIGURE 1 - Mean Birth Weight Difference by Length of Participation

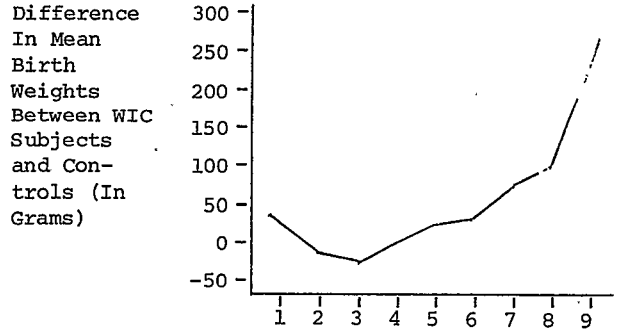
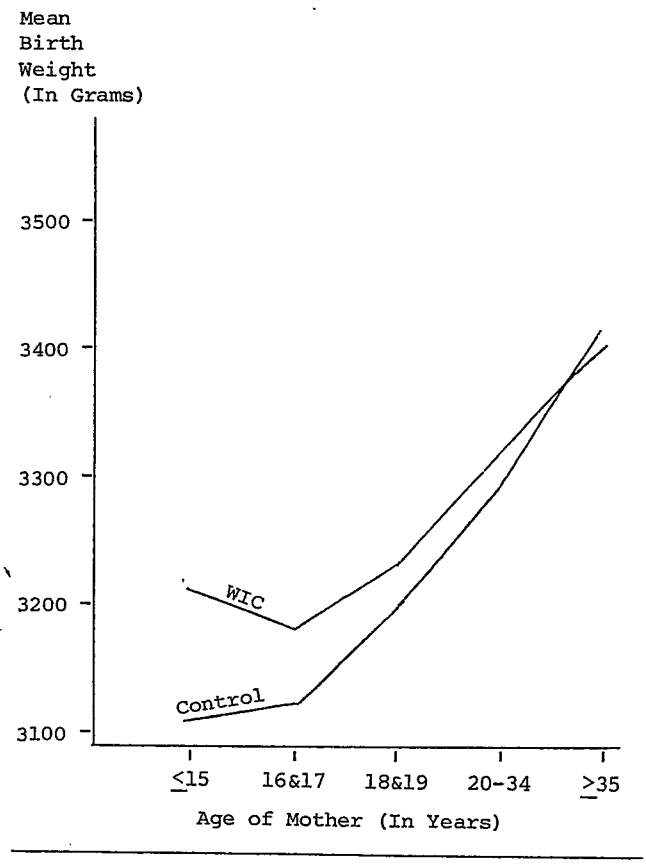


FIGURE 2 - Mean Birth Weight by Age of Mother



USES OF THE COMMUNITY HEALTH INFORMATION POLICY STUDY (CHIPS)

Edward L. Perrine, Florida Gulf Health Systems Agency, Inc.
David L. Bayless, Research Triangle Institute
Richard E. Cairl, University of South Florida

Planning for the delivery of health care services at the local level in the United States is a primary responsibility of each of the Health Systems Agencies (HSAs) as created by the National Health Planning and Resources Development Act of 1974 (P.L. 93-641). Health planning makes use of a wide range and depth of statistical information bearing on the factors affecting the demand for and supply of health care. Relevant to the measurement of health status of the population, service effectiveness, equitability, and cost of this health system are data that are timely, accurate, uniform, economical, and comparable at national, state, and local levels.

Concurrent with these mandates, the National Center for Health Statistics also expressed interest in the evaluation of application of several existing national data collection methodologies for selected small geographic areas. A contract for this purpose was awarded to the Research Triangle Institute in North Carolina; and the Florida Gulf Health Systems Agency in the Tampa-St. Petersburg area of Florida was selected as the test site.

The overall goal of this evaluation was to examine the usefulness at the local HSA level of the content and methodology of four national surveys conducted by the National Center for Health Statistics (NCHS). The four surveys were:

1. Health Interview Survey (HIS);
2. Health and Nutrition Examination Survey (HANES);
3. National Ambulatory Medical Care Survey (NAMCS); and
4. Hospital Discharge Survey (HDS).

Initially, the project was to evaluate and determine a cost-beneficial method of conducting local area surveys of the HIS, HANES, NAMCS, and HDS type that would provide useful data for planning purposes at the HSA level. Usefulness to the national surveys would come primarily from methodological sub-studies.

The study progressed toward these goals through the beginning of September, 1980, at which time a significant change was initiated by NCHS as required by the Office of Management and Budget. The constraint stipulated that only a study which addressed the national design could be implemented. Independent of this constraint, a decision was made to delete the HANES from Phase II based on cost considerations.

Consequently, significant changes were made to the questionnaire design in the final month of the Phase I study to conform to the revised objectives. However, the survey and sampling designs remained substantially the same.

The surveys were conducted, during Phase II of the study, in the four-county area served by the Florida Gulf Health Systems Agency (FGHSA).

Phase III of the study included preparation of data tapes for the National Center for Health Statistics and the Florida Gulf Health Systems Agency (FGHSA), preparation and delivery of routine tabulations for HCHS and FGHSA, and special tabulations and analyses for NCHS. Evaluation of current national survey methodology and performance of required methodological studies were completed during this phase with recommendations to NCHS on improvement of national health surveys.

The surveys, which became known locally as the Community Health Information Policy Study, produced an unprecedented amount of information about the health status and use of health services by the population of the Florida Gulf Health Systems area. They also generated a great deal of interest and concern among staff of the FGHSA, local health care professionals, and consumer constituencies over major issues being faced in their planning and decision-making responsibilities. For this reason, local funding was made available for analysis of several policy issues of greatest concern.

USES OF THE COMMUNITY HEALTH INFORMATION POLICY STUDY (CHIPS)

A major initiative of the CHIPS project was the expressed need on the part of health professionals of the Florida Gulf Health Systems Agency (FGHSA) for health information to make decisions, enhance their understanding of policy issues, and, in general, be better planners of the health service delivery system of the local area.

A health data task force representing key providers and consumemrs volunteered its time to serve in an oversight function to the RTI research team throughout the design, implementation (data collection) and analysis phases of CHIPS.

The six policy issues, which were developed and approved by the Health Data Task Force and the FGHSA Board for the CHIPS policy analysis project were:

1. the effects of seasonal residents on the use of health care services in the region;

2. accessibility of existing health care services to the indigent and medically indigent;

3. use of the primary care system in the region;

4. service needs of the elderly in the region;

5. strategies for health promotion and prevention; and

6. improved planning for acute hospital service needs through identification of bed-service specific patient origin patterns.

This summary highlights the parameters of the six major health policy issues identified above. It represents a brief rationale for focusing upon each policy issue, the uses to which the data/findings of the CHIPS study have been put, and the potential for further uses of the data. Overall, this summary serves to illustrate the extent to which the findings from the CHIPS study have served as a catalyst for dealing with some fundamental health and health related problems and policy issues in the FGHSA community.

1. Policy Issue: Seasonal Residents

Policy Parameters: Two major policy parameters were identified ir relationship to the policy issue of seasonal residency in the FGHSA region:

a. Does the existing health planning process adequately deal with the effects of seasonal residents in the health care system of the region?

b. What are some of the potential alternatives to deal with the affects of seasonal residents of the health care system of the region?

Rationale: The designation of residency status as a major policy issue of the FGHSA region reflects the relatively unique and enviable status of the Tampa Bay area as a popular destination not only for tourists but also for part-time residents. The Tampa Bay health care community expressed reasonable concern that the use of health care services in the region by seasonal residents and tourists may present some special and challenging problems for the health care delivery system of the FGHSA - problems that may not currently be adequately recognized and addressed in the health planning process.

Uses:

(1) The FGHSA has utilized the data to adjust population-based planning estimates and projections for health services. This adjustment has allowed for variations in demand for hospital services during peak months, i.e., periods in which tourists and seasonal residents migrate into the area.

(2) Various media sources have used the data (specifically, the finding that ten percent of the population during the peak season are seasonal residents in the area) to determine the overall impact of seasonal residents on the commerce of the area. These estimates have also served to counter previously held statistically unsubstantiated assumptions regarding a much higher volume of seasonal residents on the health service system.

Potential Uses and Users: From a health planning perspective, the data from the CHIPS study on seasonal residency can be used to adjust within-county estimates of health service demand by seasonal residents.

2. Policy Issue: The Indigent and Medically Indigent

Policy Parameters: Two major policy parameters were identified in relation to the major policy issue of the indigent and medically indigent in the FGHSA region:

a. What are the implications of indigent demand and health service needs (based upon health status) for health planning?

b. What are some of the alternative plans/strategies that may influence increased accessibility of indigents to health care?

For analytical purposes, a distinction was made between indigents and medically indigents. Indigents were defined as persons reporting Medicaid coverage during the time of the CHIPS survey. Medical indigents were those persons ineligible for Medicaid and with no

public or private health insurance coverage and low incomes.

Rationale: In an effort to obtain improved information for planning the allocation of limited resources and accessibility to health services in the FGHSA region, members of the Tampa Bay health care community desired to establish the numbers and proportions of indigent and medically indigent sub-populations in the region.

Uses:

(1) Results from the CHIPS survey were used by the Manatee and Pasco Counties rural health centers to develop grant proposals for federal funds for primary care services for indigent and medically indigent persons. These grant proposals focused specifically on the problem of access to primary care services.

(2) Four workshops were held in each of the four counties comprising the Tampa Bay region on the topic of "Primary Care for the Medically Indigent" wherein the CHIPS data were presented and discussed. As a result of these workshops and the analysis of the CHIPS data, county commissioners from each of the respective counties were persuaded to increase the local funding for health care to the medically indigent.

(3) The Manatee and Pinellas County Commissioners utilized the CHIPS data to justify a change in the standard of income level for welfare eligibility and also to increase reimbursement levels for health care. These changes were prompted by the identification of a large proportion of medically indigent as opposed to indigent and non-indigent sub-populations in these two counties.

(4) In Hillsborough and Manatee Counties, additional census tracts were identified as Medically Underserved areas for further expansion of rural health clinics.

(5) The data assisted in documenting the health needs in an inner city area of Pinellas County which resulted in securing three commitments of financial support from two national church denominations and an approved DHHS grant for an Urban Health Center.

(6) Projections through 1990 of indigent and medically indigent/hospital use were made to assure accessibility of care relating to the sale of Manatee County Hospital.

(7) Data and reports were presented to a Florida Legislative Task Force studying the needs of the indigent and medically indigent.

Potential Uses and Users: From a public awareness/public education perspective, the data may be used to highlight the distinction between indigent (to include medically indigent) and non-indigent sub-populations with regard to health status and health service utilization. These distinctions may be the springboard from which to launch program development activ either in the form of grant development or policy/program changes.

Finally, the methodology used in CHIPS survey to identify these sub-populations may be useful to Florida Department of Health and Rehabilitative Services in performing a state-wide survey of health status and health service utilization patterns.

3. Policy Issue: Primary Care

Policy Parameters: Three major parameters were identified in relation to the major policy issue of primary care in the FGHSA region:

a. What are the implications for health planning of the variations in ambulatory care patterns?

b. What strategies may be introduced to alleviate perceived barriers to care and increase accessibility to primary care?

c. Where should resources be placed in the primary care system (what are the priority areas for program and/or service development)?

Each of these policy parameters focused upon the overall and differentiated utilization of the primary care system in the region.

Rationale: With regard to the primary care issue, the CHIPS data were used to characterize utilization of the primary care system by age-specific cohorts. The rationale for generating these statistics was that they would provide the Tampa Bay health care community with relevant and timely information to permit an assessment of variations in pattern of care and identification of where scarce resources might optimally be

placed. Especially important with regard to this latter consideration was the need for identifying perceived barriers to access to the ambulatory care system based both upon selected patient characteristics and variations in the sources of care.

Uses:

(1) In each of these FGHSA workshops mentioned earlier, the CHIPS data were used as a basis for identifying gaps in the primary care system and areas in which service development and/or program changes should (could) occur.

(2) The district mental health boards of Hillsborough and Manatee Counties utilized the CHIPS data to prioritize their core services. In addition, the district board used the data to develop a needs statement for psychiatric beds for the region.

(3) The Northside Community Mental Health Center utilized the CHIPS data in the development of its annual grant application for core services.

(4) The CHIPS data were analyzed by four firms and resulted, in part, in the establishment of three independent practice associations - a type of Health Maintenance Organization.

(5) Area hospitals are studying the data to explore the development of satellite hospitals and clinics.

(6) The accessibility issue of primary care and the costs of use of hospital emergency departments appear to be a major factor in the rapid expansion of free-standing walk-in clinics - ten in the Tampa Bay area in the past two years.

Potential Uses and Users: From a health planning perspective, the data from the CHIPS surveys on primary care could be used to develop manpower projections for primary care and ambulatory care services both within each of the respective counties and on a regional level. The data also could be used as a base for organizing a plan for primary care services throughout the region. Finally, the CHIPS data could be used in the formal educational process of health and health related professionals.

4. Policy Issue: The Elderly

Policy Parameters: Two policy parameters were identified in relation to this major policy issue:

a. What strategy(s) should be used in the planning for the health and health related service needs of the elderly in the region?

b. What is the variation in the health status and health service utilization of the elderly as compared to the non-elderly in the region and its implications to service/program development?

Rationale: Available demographic information indicates that the growth of the over age 65 population in the FGHSA region far exceeds that of the growth of the aged population in the nation as a whole. With this recognition of demographic growth, health professionals have expressed concern over the development of services to satisfy the health and health service related needs of the aged in the area. One problem which has existed with regard to the identification of the nature and scope of these services has been the absence of primary data which would indicate the relative health status of the aged as compared to the non-aged population and their utilization of health services, particularly acute care hospital services. The CHIPS data were generated in an effort to ameliorate this problem.

Uses:

(1) The CHIPS data on the elderly were used by the Suncoast Gerontology Center of the University of South Florida College of Medicine in developing their annual grant application for long-term care services. These data were used to describe the unique characteristics of the aged in the area as compared to the aged on the state and national levels and to characterize the Tampa Bay region as a laboratory for studying the health and health related service needs of the elderly population.

(2) The CHIPS data were used by the Suncoast Gerontology Center and the Department of Gerontology at the University of South Florida to examine variations in health status and utilization in the older age cohorts. This analysis of the CHIPS data was instrumental in identifying the need for considering the aged population as very heterogeneous, as opposed to homogeneous, sub-population with increasing health service needs with advancing age.

(3) The CHIPS data were used as a base for resource allocation in a joint forum between the FGHSA, the Area Agency on Aging for Hillsborough and Manatee Counties and the District VI Mental Health Board of Hillsborough and Manatee Counties.

(4) The CHIPS data set on the elderly has and will continue to be used by the Florida Health Data Consortium in developing periodic information briefs on characteristics of the aged, specifically health status, functional status and utilization characteristics of the elderly, for distribution to service providers.

(5) The CHIPS data on the elderly have been used effectively to make adjustments in health services utilization projections. The previously assumed "good health status" of the elderly was documented and resulted in lower estimates of projected bed need for hospital and nursing home services.

Potential Uses and Users: The CHIPS data on the elderly also could be used as a base for health and health related manpower projections. Finally, the identification of variations in the prevalence of chronic morbidity among the older age cohorts would be instructive to medical and other health professionals and students.

5. Policy Issue: Health Promotion and Disease Prevention

Policy Parameters: Five major policy parameters were identified in relation to the major policy issue to health promotion and disease prevention:

a. What sub-population should be targeted for health promotion and prevention activities?

b. What chronic conditions are most important and which are amenable to screening services?

c. What strategies could be employed within sub-populations to increase preventive health behaviors?

d. What is the potential role of the educational system?

e. What is the potential role of business and industry?

Overall, these policy parameters represent an effort to address the strategies which the FGHSA community may use to promote the health of its people and prevention of disease.

Rationale: Major strategies available for health promotion and disease prevention include improving the existing health care through the elderly, detection of curable or treatable diseases, improvement of access to health care services, increasing the awareness of healthy lifestyles by education and awareness programs, the control or removal of environmental risk factors, and direct attempts to influence lifestyle behaviors. The CHIPS data set addressed a portion of these strategies through the collection of data on the incidence of chronic disease, levels of immunizations in each of the four counties, and frequency of use of health screening tests. These data were collected as a result of a mounting interest in the FGHSA community on putting more emphasis on health promotion and prevention activities within the health care system.

Uses:

(1) The Tampa area employer-based Health Care Coalition utilized the CHIPS data in preparing a grant to the Robert Wood Johnson Foundation to develop a demonstration project on the value of health promotion activities for employers in reducing the incidence of acute morbidity and absenteeism rates.

(2) The area employer-based health coalition has studied the data to determine the need for and priority that should be given to various health screening tests based on morbidity.

Potential Users and Uses: There are at least two different areas in which the health promotion and prevention data from the CHIPS project may be used. The first of these is for public education and public awareness programming within the community. Based upon the findings in the CHIPS study with regard to use of health screening tests and immunizations, programs could be developed for stimulating greater use of these tests and immunizations. Secondly, the county commissioners could focus on levels of immunization and areas and sub-populations in which increased public health efforts need to be introduced in order to insure needed preventative services among the residents of their respective communities.

6. Policy Issue: Hospital Service Needs

Policy Parameters: Four major policy parameters were identified in relationship to the major policy issue of hospital service needs in the FGHSA region:

a. Does the existing planning method for determining hospital service needs adequately reflect utilization patterns?

b. Based upon patient origin patterns, what are the "rational" health service areas in the region?

c. What strategies can be employed or implemented to reduce duplications or overlap of services within the health service area?

d. What method(s) can (should) be used to project hospital bed service needs?

Rationale: To facilitate improved planning for acute hospital service needs, service specific utilization data are needed and have heretofore been largely unavailable for the FGHSA health planning community. To counter this situation, it was decided that two general sorts of data were necessary to address this policy issue. The first included data which provided a general description of the types of inpatient services used by the FGHSA population. The second included statistics that provided a description of users of inpatient services delivered in the area, including, especially, the geographic origin of patients that use specific facilities and hospital services.

Uses:

(1) The CHIPS data relating to this policy issue have been used for several and somewhat diverse uses. First, the data were used by the South Florida Baptist Hospital to demonstrate the need for obstetrics and psychiatric inpatient services in an effort to submit and gain approval of a certificate of need application for additional beds. Second, the data were used by Tarpon Springs and Mease Hospitals in north Pinellas County to demonstrate that they are providing services to Pasco County and that additional beds were not needed in Pasco County. Finally, in a very general sense, the CHIPS data have been used in virtually all certificate of need applications that relate to expansion of hospital services.

(2) The data have been used to further define hospital service areas, thereby creating sub-county units for projecting the need for additional beds.

(3) A special study was undertaken to develop a methodology for determination of need for rehabilitation services. Hospital ICDA codes relating to rehabilitation services were used to project future need and were presented for state consideration as a reasonable methodology.

Potential Uses and Users: Perhaps the most fundamental use which the CHIPS data provide for hospitals in the Tampa Bay region is patient origin studies which identify primary and secondary service areas of a hospital. Such patient origin studies allow for identification of duplication or overlap of services between hospitals in a given geographic area, as well as influencing the marketing strategy of an individual hospital in relation to competitive, contiguous hospitals in the area.

Conclusions

Planning for the delivery of services will continue to have its primary focus at the state and local levels. This demonstration of the feasibility and use of high quality local health data should encourage the National Center for Health Statistics to further pursue this approach.

Secondly, this approach has proven to be an effective use of health data to study important policy issues at the local level.

Thirdly, community awareness and sensitivity to local health issues can be constructively channelled, by the use of objective data, toward resolution of local health care problems.

Finally, because of the time related relevancy of the data, resurveying at intervals of approximately five years would greatly reduce costs.

Monitoring and Evaluating Health Programs: Statistical Methods

Session U

EFFICIENCY IN PUBLIC HEALTH RESEARCH: THE IMPORTANCE OF A PRIORI SAMPLE SIZE ESTIMATION

Thomas W. Woolley, Quillen-Dishner College of Medicine

Introduction

Perhaps the question most frequently asked of statisticians is, "How many subjects (patients, animals, etc.) do I need to include in my study?". The fact that sample size estimation is afforded minimal (if any) coverage in most introductory-level applied statistics or research design courses or textbooks does not lessen its obvious importance to researchers. The purpose of this paper is to review the concepts underlying a priori sample size estimation and to explore the current state of public health research, from a power-analytic perspective. In addition, the consequences of conducting a study with an inadequate sample are discussed.

A Conceptual Background[1]

Generally speaking, there are two potential types of errors one must take into account when designing a study. The first, a Type I error takes place when the researcher incorrectly concludes (based upon statistical analysis of the data) that a difference exists between study groups (e.g., a difference in group means or proportions). The long-range probability of this error occurring is governed by the significance level, α, which is traditionally set at 0.01, 0.05, or 0.10. The second possible error, Type II, happens when an investigator fails to detect an "important" difference between study groups (again, possibly a difference in group means or proportions). The long-term probability of a Type II error is symbolized by β and, unlike α, has no traditionally established values. It has been suggested (1), however, that a maximum value of $\beta = 0.20$ be adopted in most research contexts. Freiman et al. (2), on the other hand, argue that β should never exceed 0.10 in a controlled clinical trial. Therefore, in addition to specifying a priori the significance level, the researcher has a responsibility to define what would constitute a nontrivial difference between groups, as well as the desired probability of detecting that difference. This latter probability is simply the complement of β, $1 - \beta$, and is referred to as statistical power.

How, then, is a potentially "important" difference between groups specified? Most investigators are well versed in the research literature pertinent to their field of inquiry and have some notion of the magnitude of treatment effects they might reasonably expect in a given study. This knowledge might also be supplemented from years of clinical, laboratory or field experience, previous or ongoing research, or reports of vital statistics. From such information, most researchers can formulate an estimate of what size of group differences they would consider to be of scientific importance. Such an estimate is referred to as the hypothesized effect size (ES). The inability to generate an ES at a given point in time may indicate that a meaningful study is, as-of-yet, not feasible.

Sample size (n) is a function of Type I error rate (α), Type II error rate (β) or statistical power ($1 - \beta$), and the hypothesized effect size (ES). In other words, given that an investigator has appraised the practical consequences of committing each of the potential decision errors (and subsequently assigned values to α and β), the minimal sample size necessary to detect an "important" difference between groups (i.e., a difference in means or proportions at least as large as that specified by the ES) can be calculated.

The interrelationships of α, β, ES, and n can be viewed and understood simply from a conceptual perspective. As fewer errors of the first or second kind are desired (i.e., as α and β are lowered), the minimal sample size needed to conduct a given study will increase (assuming ES is held constant). Likewise, if smaller and smaller group differences are hypothesized to be of importance (i.e., ES is decreased), then the sample size necessary to detect such diminishing effects will escalate. In general, the desire of the researcher to reduce the potential for decision errors or to increase the ability of a statistical test to discriminate between study groups, requires a concurrent boost in sample size.

The most common research scenario is one in which: (a) the research question is formulated, (b) a relatively small sample of convenience is selected (patients, animals, etc.), (c) a treatment is rendered, measurements are made, and the data are tabulated, and (d) appropriate statistical analyses are conducted assuming the traditional (low) significance level of $\alpha = 0.05$. A review of this process reveals that the naive researcher may be unable to statistically acknowledge a scientifically important difference between study groups, even if that difference exists. What happens is simple. Most investigators initially choose a small sample for their study and set $\alpha = 0.05$ at the time of the data analysis. The probability of committing a Type II error and the definition of a scientifically important treatment effect are generally ignored. Assuming most researchers are looking for relatively small group differences (i.e., a small ES), the probability of committing a Type II error must intuitively (and mathematically) be large. Hence, the researcher, by default, becomes trapped in a situation where the likelihood of failing to find an "important" treatment effect is sizable.

The failure to achieve statistical significance in a study, therefore, does not necessarily imply group equivalence. As Fleiss (1) points out, this cannot be emphasized enough

when sample size is small and the resulting statistical power of the test employed in analysis is low (i.e., β is high).

Procedure

A total of 202 papers in volume 72 (1982) of the American Journal of Public Health and volume 97 (1982) of Public Health Reports were reviewed for this survey. One hundred and twenty-eight of these studies were omitted from consideration due to the fact that there were no statistical tests of significance included (or those tests reported were not applicable in this power survey), or the studies were uninterpretable, i.e., necessary information was missing. All statistical tests of significance, with the exception of secondary tests such as reliability estimates, etc., were power-analyzed.

In order to maintain consistency between this and most other reported statistical power surveys, a number of conditions were standardized:

1. Only the most common statistical tests (t, F, X^2, r) were power-analyzed;

2. Alpha was held at a uniform value of 0.05 and a nondirectional alternative was assumed for all studies;

3. Cohen's (2) definitions of small, medium, and large effect sizes were adopted. Thus, three power determinations were made for each test of significance, and an average power for detecting small, medium, and large ESs was calculated for each study. Note that when unequal cell sizes were in evidence, the harmonic mean functions of the cell sizes were used (see Cohen, 2).

Results

Of the 74 papers power-analyzed for this survey, few made explicit mention of statistical power or sample adequacy. Table 1 presents a breakdown by type for the 2635 statistical tests analyzed. Note that better than half of all statistical tests reported as primary tests of hypotheses were made with regard to the magnitude of Pearson Product Moment correlation coefficients. This reflects the all too common practice of testing large matrices of coefficients (and inviting the wrath of capitalization on chance). See Table 1 in the Appendix.

As one would expect in public health research, the distribution of study sample sizes was somewhat skewed. (See Table 2.) The frequent use of large data sets is reflected in the mean sample size, although the median probably represents the typical study more accurately. The median number of statistical tests reported per paper was 12.

Table 3 illustrates that:

1. For small effects, on the average, two-thirds of the studies reviewed had less than a 50-50 chance of detection.

2. For detection of medium effects (roughly twice the magnitude of a small effect), fewer than 15% of the studies had less than a 50-50 chance of detecting such an effect.

3. For large effects (approximately four times the size of a small effect), an average per study power of 0.94 existed, and only 3% had under a 50-50 chance of detection. See Table 3 in the Appendix.

Although strict comparisons are not possible due to an array of confounding variables (number of studies sampled, history, etc.), public health research is substantially more powerful, on average, than research in other disciplines where similar surveys have been conducted (e.g., 3-7).

Discussion

As pointed out in the introduction to this paper, the most appropriate time to consider statistical power is during the research design phase. For example, by the a priori establishment of an alpha level at the traditional 0.05 level, setting power equal to 0.80 (the value that Cohen (2) considers to be minimally acceptable), and deciding on a medium effect size (as defined by Cohen) for a two-independent group t-test, a researcher would find it necessary to acquire 64 subjects per group. Had the investigator desired the potential to more finely discriminate between groups, that is, adopted a small effect size, 393 subjects per group would be required (assuming α and power remain fixed at 0.05 and 0.80, respectively). Obviously, many investigators would find this latter sample size figure difficult to accept, and adjustments to α, power and ES would have to be made (e.g., increase α, decrease power, or increase ES), or a decision could be made to postpone the study until adequate resources could be obtained. Proper planning of a study, including attention to α, β, ES and sample size, allows the investigator to simultaneously vary sample size against the magnitude of effect sizes capable of being detected. This provides the researcher with a means by which to assess the feasibility and/or practicality of a project before it is initiated.

The absence of any consistent mode of presenting statistical results, and a belief by some that researcher and editorial biases toward submitting and publishing only "successful" studies (i.e., rejection of the primary null hypothesis), leads some to question the representativeness of the research literature (4,5,8). Cohen (2) goes so far as to suggest that logic should tell us that the published research, in all likelihood, is more powerful than that which never came to print. Chase and Tucker (4) point to two other potentially

confounding consequences, for the research literature, of conducting investigations with inadequate power. First, a side effect of the potential bias against publishing negative findings is that the actual rate of "false positives" (incorrectly concluding that a difference exists between experimental groups) in the published work may be greater than traditional alpha levels. Second, since the power of tests for interactions is somewhat lower than for main effects, both "(1) erroneous interpretations of significant main effects, and (2) incorrect acceptance of the null hypothesis when an interaction effect is specifically predicted" (p. 38) may result.

Upon the completion of a study, information related to power may be used to supplement the analysis. For example, the obtained effect size (the amount of variability in the dependent variable accounted for by the independent variable) may be mathematically approximated (4), and assuming the research was well designed, this value (confidence-bounded) is a much better estimate of the population effect size than the original hypothesized ES. Only the sample size and the value of the test statistic itself are needed to calculate the obtained ES (9).

To echo Katzer and Sodt (9), however, the most frustrating aspect of this study was the difficulty in determining the statistical methods used in each reported study. Not only are some researchers entering into experiments blindly, often with very little chance of detecting treatment effects even when they exist, but the information reported in the typical results section leaves those readers concerned with interpretability and generalizability without recourse (beyond writing the author(s) for additional information).

Perhaps the best one can hope for as a consumer of the research is that authors supply at least the minimal amount of information demanded for clear and independent evaluation. This would include not only information needed for planning the data collection (α, desired power, hypothesized effect size, and necessary n), but also an adequate accounting of at least the following:

1. actual n used (total, per cell, per factor);
2. alpha;
3. value of the calculated test statistic and its associated value;
4. all cell means and standard deviations; and
5. a complete description of experimental design(s) incorporated in the study.

This additional information serves as the basis for a number of interpretative analyses including the approximation of the obtained effect size and construction of suitable confidence intervals on sample statistics, as well as the potential application of meta-analysis. None of these five pieces of information requires any additional effort on the part of the researcher, as they are included in the standard output of most computer packages. It is the sincere belief of the author that the adoption of such standards of reporting would contribute toward a substantial upgrading of the interpretability of research in public health.

Obviously, this paper serves only as an introduction to the concepts of and issues surrounding sample size estimation. Coverage of many related topics is beyond present limitations, however, it is important to realize that despite methodological changes to suit different research scenarios, the concepts, as a rule, will hold.

Footnote

[1]This presentation assumes the researcher will adopt classical inferential techniques (e.g., hypothesis testing).

References

1. Fleiss, J. L. Statistical Methods for Rates and Proportions, Second Edition. New York: John Wiley & Sons, 1981.
2. Cohen, J. Statistical Power Analysis for the Behavioral Sciences, Second Edition. New York: Academic Press, 1977.
3. Cohen, J. The Statistical Power of Abnormal-Social Psychological Research: A Review. J. of Abnormal and Social Psych. 65:145-153, 1962.
4. Chase, L. J. and Tucker, R. K. A Power-Analytic Examination of Contemporary Communication Research. Speech Monographs. 42:29-41, 1975.
5. Freiman, J. A., Chalmers, T. C., Smith, H., and Kuebler, R. R. The Importance of Beta, the Type II Error and Sample Size in the Design and Interpretation of the Randomized Controlled Trial. New Engl. J. of Med. 299:690-694, 1978.
6. Woolley, T. W. A Comprehensive Power-Analytic Investigation of Research in Medical Education. J. of Med. Educ. 58 (September):1983.
7. Brewer, J. K. On the Power of Statistical Tests in the American Educational Research Journal. American Educational Research Journal. 9:381-401, 1972.
8. Schoolman, H. M., Becktel, J. M., Best, W. R., and Johnson, A. F. Statistics in Medical Research: Principles versus Practices. J. of Laboratory and Clinical Medicine. 71:357-367, 1968.

9. Katzer, J. and Sodt, J. An Analysis of the Use of Statistical Testing in Communication Research. J. of Communication. 23:251-265, 1973.

APPENDIX

TABLE 1

Frequency and Percentage of Statistical Tests Used in Volume 72 of the American Journal of Public Health and Volume 97 of Public Health Reports

Statistical Test	Frequency	Percentage of Total
t	308	12
F	118	4
X^2	834	32
r	1375	52
Total	2635	100

TABLE 2

Distribution of Mean Sample Size of 74 Articles Power Analyzed in Volume 72 of the American Journal of Public Health and Volume 97 of Public Health Reports

Sample Size	Frequency	Cumulative %
1000+	16	100
500-999	9	78
300-499	11	66
250-299	7	51
200-249	3	42
150-199	5	38
100-149	9	31
50-99	6	19
0-49	8	11

Mean = 5294 Median = 295

TABLE 3

Frequency and Cumulative Percentage Distributions of the Mean Power of 74 Articles in Volume 72 of the American Journal of Public Health and Volume 97 of Public Health Reports for Detecting Small, Medium, and Large Effects

	Effect Size					
	Small		Medium		Large	
Power	Freq.	Cum. %	Freq.	Cum. %	Freq.	Cum. %
.99+	7	100	33	100	50	100
.95-.98			7	55	8	32
.90-.94	2	91	8	46	4	22
.80-.89	4	88	9	35	5	16
.70-.79	2	82	3	23	3	9
.60-.69	7	80	2	19		
.50-.59	3	70	2	16	2	5
.40-.49	7	66	4	14		
.30-.39	11	57	2	8	1	3
.20-.29	11	42	2	5	1	1
.10-.19	14	27	2	3		
.00-.09	6	8				
Total	74		74		74	
Mean	.43		.85		.94	
Median	.34		.96		.99+	

ANALYSIS AND SCALING OF SELF-REPORTED HEALTH STATUS MEASURES: APPLICATION OF LEAST-SQUARES METHOD FOR ORDERED DISCRETE RESPONSES

David S. Salkever, Lawrence M. Curcio, Alison S. Jones, Johns Hopkins University, and Robert Seidman, San Diego State University

I. INTRODUCTION

The past decade of research in health services has witnessed dramatic improvement in computer resources and use of sophisticated estimation techniques. Of particular importance is the increased focus on models with qualitative or limited dependent variables. Maximum likelihood procedures (e.g., probit regression) have been used to estimate models with dichotomous dependent variables due to the desired asymptotic properties of the resulting parameter estimates and test statistics. By contrast, ordinary least squares (OLS) estimation of such models is an ad hoc procedure which lacks these properties.[1] This method requires considerably less computing time, though, and may thus be useful in exploratory data analysis if OLS and maximum likelihood estimation yield similar qualitative results.

Recent comparisons suggest that qualitative findings are indeed similar. OLS-probit comparisons of dichotomous dependent variable regressions report similar signs and significance levels for virtually all regressors.[2] Unfortunately, there exists little evidence on whether this similarity holds in the case of models with dependent variables which assume more than two discrete responses. Such a comparison is of interest since ordered polychotomous responses are frequently used in household surveys. In addition, the cost differential between maximum likelihood and OLS techniques increases the greater the number of discrete responses which the dependent variable may assume.

The primary objective of this study is to compare results obtained from OLS and maximum likelihood estimation of models when the dependent variable is characterized by multiple discrete ordered responses. An appropriate maximum likelihood technique and a modified OLS method first proposed by Rubinfeld[3] to estimate this type of model are outlined. While this OLS procedure has been applied previously in health-related contexts,[4] it has never been compared with the relatively expensive but theoretically more appealing maximum likelihood technique. We also describe Rubinfeld's iterative extension of his OLS method which assigns scores to these ordered discrete responses. Self-reported health status is used as the dependent variable in all regression models.

II. ESTIMATING MODELS WITH MULTIPLE DISCRETE ORDERED DEPENDENT VARIABLES

The ordered multinomial probit model may be described as follows. Suppose y denotes a dependent variable, e is a $N(0,\sigma^2)$ random disturbance, and X and ϕ represent vectors of independent variables and coefficients, respectively. We may view y as taking on a particular ordinal value when the quantity $X\phi-e$ falls within a particular range. If $L_1,\ldots,L_r$ are cut-off points defining r+1 ranges of $X\phi-e$ and V_i denotes the ordinal value of y when $X\phi-e$ is in the ith range, then the value of y is determined by

$$y = \begin{cases} V_1 & \text{if } X\phi-e \leq L_1 \\ V_i & \text{if } L_{i-1} < X\phi-e \leq L_i, \text{ for } i=2,\ldots,r \\ V_{r+1} & \text{if } L_r < X\phi-e \end{cases}.$$

The sample of observed data points may be partitioned into r+1 subsets, where subset S_i consists of all observations for which $y = V_i$. The likelihood function may then be written

$$\Lambda = \prod_{S_1} Q\left(\frac{X\phi - L_1}{\sigma}\right) \cdot \prod_{j=2}^{r} \left\{ \prod_{S_j} \left[P\left(\frac{X\phi - L_{j-1}}{\sigma}\right) - P\left(\frac{X\phi - L_j}{\sigma}\right)\right]\right\} \cdot \prod_{S_{r+1}} P\left(\frac{X\phi - L_r}{\sigma}\right),$$

where $P(\ )$ is the standard normal cumulative density function and $Q(\ ) = 1-P$. Following the normalization method adopted by McKelvey and Zavoina, the multinomial probit model is estimated by setting $L_1 = 0$ and $\sigma = 1$, and choosing the values of ϕ and $L_1,\ldots,L_r$ which maximize the likelihood function.[5]

In Rubinfeld's OLS technique, the dependent variable, Y, is assigned a value of one for the highest response category (i.e., V_{r+1}) and zero for all other response categories. A new vector, Z, is defined containing r-1 independent variables, with Z_i equalling one if y falls into the ith intermediate response category. The OLS regression to be estimated is $Y = X\beta - ZD + u$ where u is a random disturbance, X is defined as above, and β and D are the vectors of coefficients.

There are several interesting features of the modified OLS regression procedure outlined above. First, the estimate $\hat{D}_i$ is the mean predicted value of Y (i.e., $\hat{D}_i = \overline{X\hat{\beta}}$) for observations in subsample S_i. Setting the Y value for these observations equal to $\hat{D}_i$ transforms the ordinal dependent variable into an interval scale $(0,\hat{D}_2,\ldots,\hat{D}_r,1)$ and permits quantitative interpretation of intermediate Y values.[6] Second, the appendix to this paper demonstrates that the estimates of D and β would also minimize the sum of squared residuals in an OLS regression of Y on X if $(0,\hat{D}_2\ldots,\hat{D}_r,1)$ are the values assigned to Y. This property offers some justification in using the OLS procedure for predictive purposes when Y is actually continuous but only ordinal data are available.

Once the initial OLS estimation is completed, the results may be used to perform the iterative scoring procedure mentioned above.[7] According to this method, the results of the initial OLS regression are used to rescale the original values of the dependent variable on the 0-1 interval. A second OLS regression is then estimated using this rescaled dependent variable. The rescaling is performed again, and the iteration process continues until convergence of the rescaled

dependent variable scores is achieved. (Rubinfeld points out that these final scores are equivalent to those obtained from the standard canonical analysis of variance method for scaling.)

If $\hat{\beta}^0$ and $\hat{D}^0$ denote the parameter estimates of the vectors β and D resulting from the initial Rubinfeld OLS regression described above, then the iterative procedure may be expressed formally for the jth iteration as follows:

STEP 1 Calculate $\bar{Y}_i^j$ for all observations in S_i, $i = 1,\ldots,r+1$.

$$\text{If } j = 1, \text{ then } \bar{Y}_i^j = \begin{cases} \overline{X\hat{\beta}^0} & \text{, for } S_i = S_1 \text{ or } S_{r+1} \\ \hat{D}_i^0 & \text{, for } S_2 \leq S_i \leq S_r \end{cases}$$

$$\text{If } j \geq 2, \text{ then } \bar{Y}_i^j = \overline{X\hat{\beta}^{j-1}}, \text{ for all } S_i$$

STEP 2 Compute the rescaled dependent variable:

$$\tilde{Y}_i^j = \frac{\bar{Y}_i^j - \bar{Y}_1^j}{\bar{Y}_{r+1}^j - \bar{Y}_1^j} \quad \text{for all } S_i$$

STEP 3 Estimate $Y_i^j = X\hat{\beta}^j + u^j$ by OLS regression.

The calculated value $\bar{Y}_i^j$ in STEP 1 is simply the mean predicted value of Y for category V_i using the previous iteration's regression results. STEP 2 uses these mean predicted values of Y to rescale the dependent variable on the unit interval. Note also that each step of this iterative procedure yields a separate vector of OLS coefficient estimates for the X's which may be compared with the multinomial probit results.

III. DATA DESCRIPTION AND EMPIRICAL RESULTS

The multinomial probit and Rubinfeld OLS procedures were used to estimate the parameters of health status regression models. Two different samples were constructed from the 1978 Health Interview Survey (HIS) and equations were estimated for both samples. A 10% sample was extracted from the HIS of all individuals under 17 years of age with known self-reported health status who were children of the head of household or spouse, who resided in a housing unit, and whose household head had known education and income. A 25% sample was also drawn of females over 64 years of age with known reported health status who lived in a housing unit and who did not suffer from a limitation of activity which has existed since age 25. These two samples contained 2716 and 1404 observations, respectively.

The dependent variable measured whether the individual reported his/her health status as excellent, good, fair, or poor. Since so few children reported poor health, the categories poor and fair health were combined in this sample so only three responses were possible for the dependent variable. The four health status responses were treated separately for the sample of women over 64. Independent variables relate to race, sex, education, age, family size, income, and location of residence. Marital status was also included in the over-64 sample while the respondent to the survey was included as an independent variable in the sample of children.

Detailed variable definitions and estimation results for children and females over 64 are presented in Tables 1 and 2, respectively. These tables report for each sample the results from the probit and initial OLS regressions as well as the first and final OLS iterations. As indicated above, the dependent variable in the initial OLS regression is dichotomous, equalling one if reported health status is excellent. Although Table 2 contains two independent dummy variables for health status, note that there exists only one such variable in Table 1 (i.e., DHLTH2) since the lowest two health response categories were combined in the sample of children.

The results from the OLS and probit regressions on both samples reveal that parameter estimates of the independent variables generally have similar significance levels, and the significant variables have the same qualitative impact on health status in both methods. Lower educational achievement consistently has a negative impact on health status in both samples, while higher family income tends to increase the person's health status. Marital status and family size are also shown to have a significant impact on health status for the sample of older women. A comparison of the DHLTH2 intermediate response dummy variable in the initial OLS regressions on both samples not surprisingly suggests that the relative difference in health status between children who report good and excellent health is typically smaller than is true for older women.

Another interesting result concerns those variables which are statistically significant when estimated by one method but are insignificant when the other procedure is used.[8] Tables 1 and 2 indicate that the discrepancies in significance level between the initial OLS and probit results for these variables are virtually eliminated by the first iteration of the OLS procedure. In fact, all variables that are significant at the five percent level or higher in the probit regression are also significant at the five percent level or higher in the first iteration of the OLS procedure. Further, it is worth noting that there exists an apparent "diminishing return" to iteration in the sense that the differences in significance levels between the initial OLS and probit results are narrowed by a much larger percentage in just the first OLS iteration than from all successive iterations combined.

An ordinal scale for health status responses was derived from the Rubinfeld iterative OLS procedure, and is shown for both samples in Table 3. The iteration process was terminated when none of the scale values changed by more than .001 between two successive iterations. No more than five iterations were required before convergence occurred. The scale values for both children and older women indicate a substantial difference between good and excellent health levels and, for the latter sample, a relatively small difference between fair and good or poor and fair.

IV. CONCLUSION

We have considered maximum likelihood and

ordinary least squares methods for estimating models whose dependent variables are characterized by multiple discrete ordered responses. Our findings suggest that Rubinfeld's modified OLS procedure with a single iteration may be quite useful in preliminary data exploration since it consistently identifies significant explanatory variables despite requiring substantially lower computing time when compared to the maximum likelihood procedure. This technique may also be extended by a series of iterations to estimate an ordinal scale for the dependent variable which will often be useful for descriptive purposes or as an input into further analyses.

APPENDIX

This appendix demonstrates that Rubinfeld's initial OLS procedure is equivalent to finding the values of $\hat{D}$ and $\hat{\beta}$ which minimize the sum of squared residuals when $\hat{D}_i$ is assigned as the value of y for all observations in S_i (i = 2,...,r) and y is regressed on X. We shall consider only the trichotomous case for simplicity, so $\hat{D}$ is a scalar. Extension to the more general case is trivial though tedious.

First, let us partition the y vector into the three subvectors y_1, y_2, and y_3 representing the y values of observations in S_1, S_2, and S_3, respectively, where y_3 is a unit vector, y_1 is a null vector, and y_2 is a vector of $\hat{D}$'s. The sum of squared residuals for any choice of $\hat{D}$ and $\hat{\beta}$ may be written

(A.1)

$$SSR = (y_3' - \hat{\beta}'X_3', \; y_1 - \hat{\beta}'X_1', \; y_2 - \hat{\beta}'X_2') \begin{pmatrix} y_3 - X_3\hat{\beta} \\ y_1 - X_1\hat{\beta} \\ y_2 - X_2\hat{\beta} \end{pmatrix}$$

where X_i (i=1,2,3) denotes the matrix of independent variable values for S_i. Minimization of SSR with respect to $\hat{\beta}$ and $\hat{D}$ requires

(A.2)

$$\frac{\partial SSR}{\partial \hat{\beta}} = -2X'X\hat{\beta} - 2X'y = 0$$

(A.3)

$$\frac{\partial SSR}{\partial \hat{D}} = -2n_m\hat{D} - 2I_M'X_M\hat{\beta} = 0,$$

where D_M is the number of S_2 cases and I_M is a n_Mx1 unit vector. Equation (A.3) implies that $\hat{D}$ equals the mean of $X\hat{\beta}$ for the S_2 cases, while (A.2) yields the standard OLS result that $\hat{\beta} = (X'X)^{-1}X'y$.

In order to describe Rubinfeld's least squares estimation procedure, it will be helpful to partition the n x (k+1) matrix of regressors, Q, into (X Z_2). Let the first n_3 rows of Q correspond to S_3 cases. Noting that the first $(n-n_M)$ elements of Z_2 are zero and the remaining n_M elements are one, we obtain the partitioned cross-product matrix

(A.4)

$$Q'Q = \left(\begin{array}{c|c} X'X & \Omega \\ \hline \Sigma_M & n_M \end{array}\right),$$

where Ω is a k x 1 null vector and Σ_M is a 1 x k vector of summations of X's from the S_2 cases (i.e., $(\Sigma X_{S_2}^1, \ldots, \Sigma X_{S_2}^k)$). The values of $\hat{D}$ and $\hat{\beta}$ are determined from

(A.5)

$$(Q'Q)^{-1}Q'y_3 = \begin{pmatrix} \hat{\beta} \\ \hat{D} \end{pmatrix}.$$

However, a standard result on the inverse of partitioned matrices[9] implies that

(A.6)

$$(Q'Q)^{-1} = \left[\begin{array}{c|c} (X'X)^{-1} & \Omega \\ \hline -\frac{1}{n_M}\Sigma_M(X'X)^{-1} & \frac{1}{n_M} \end{array}\right].$$

Substituting this result into (A.5), partitioning Q' as above, and noting that the first n_3 elements of y_3 are ones while the remaining $n_1 + n_2$ elements are zeros, we find that (A.5) yields the same expressions for $\hat{\beta}$ and $\hat{D}$ as equations (A.2) and (A.3).

NOTES

1. J. Kmenta, Elements of Econometrics, Macmillan, 1971.
2. D. Salkever, et al., "Episodes of Illness and Access to Care in the Inner City: A Comparison of HMO and Non-HMO Populations," Health Services Research 11:252-270, Fall 1976; C. Mallar and C. Thornton, "Transitional Aid for Released Prisoners: Evidence from the LIFE Experiment," Journal of Human Resources 13: 208-236, Spring 1978; and F. Sloan and S. Richupan, "Short-Run Supply Responses of Professional Nurses: A Microanalysis," Journal of Human Resources 10:241-257, Spring 1975.
3. D. Rubinfeld, "Credit Ratings and the Market for General Obligation Municipal Bonds," National Tax Journal 26:17-27, March 1973.
4. B. Friedman, et al., "The Influence of Medicaid and Private Health Insurance on the Early Diagnosis of Breast Cancer," Medical Care 11:485-490, November-December 1973.
5. R. McKelvey and W. Zavoina, "A Statistical Model for the Analysis of Ordinal Level Dependent Variables," Journal of Mathematical Sociology 4:103-120, January 1975.
6. For example, $\hat{D}$=.4 in a simple insurance example where the dependent variable indicates full coverage, some coverage, or no coverage suggests that the typical partial coverage is forty percent of full coverage (i.e., a coinsurance rate of .6).
7. D. Rubinfeld, "Multiple Regression with a Qualitative Dependent Variable," unpublished manuscript, Institute of Public Policy Studies, University of Michigan, March 1977.

8. These variables were BLACK, LIVES IN ONE OF 31 LARGE SELF-REPRESENTING SMSA's, LIVES IN WESTERN USA, and LIVES IN SOUTHERN USA in the sample of children, and AGE and HIGHEST GRADE COMPLETED IS 9-11 in the female sample.
9. A. Goldberger, Econometric Theory, Wiley, 1964, p. 27.

TABLE 1

SELF-REPORTED OVERALL HEALTH STATUS:
RESULTS OF OLS AND N-CHOTOMOUS PROBIT REGRESSIONS FOR CHILDREN
16 YEARS AND YOUNGER FROM THE 1978 HEALTH INTERVIEW SURVEY
(t-statistics in parentheses)

VARIABLE	INITIAL OLS	FIRST ITERATION OLS	FINAL ITERATION OLS	PROBIT
ASIAN	-.00609 (0.179)	-.04749 (1.093)	-.05572 (1.184)	-.28377 (1.295)
BLACK	-.01190 (.957)	-.04941 (3.099)[a]	-.05687[a] (3.294)	-.25428[a] (3.206)
SPANISH	-.00965 (0.637)	-.01913 (.983)	-.02102 (.997)	-.9268 (.939)
OTHER	.00076 (.04472)	-.02490 (1.002)	-.03000 (1.115)	-.15642 (1.252)
AMERICAN INDIAN	.05259 (1.230)	.04355 (.792)	.04175 (.701)	.15096 (.528)
HIGHEST GRADE COMPLETED BY FAMILY HEAD IS 0-8	-.03958 (3.006)[a]	-.06754 (3.996)[a]	-.07310[a] (3.994)	-.32918[a] (3.841)
HIGHEST GRADE COMPLETED BY FAMILY HEAD IS 9-11	-.03044 (2.611)[a]	-.05926 (3.961)[a]	-.06499[a] (4.012)	-.30173[a] (3.945)
HIGHEST GRADE COMPLETED BY FAMILY HEAD IS 12	-.00002 (0.000)	-.00588 (.520)	-.00704 (.574)	-.04303 (.714)
LIVES ON FARM NOT IN SMSA	.01487 (.653)	.01881 (.404)	.01120 (.354)	.03444 (.229)
LIVES IN CENTER CITY OF ONE OF 31 LARGE SELF-REPRESENTING SMSA's	-.04131 (3.191)[a]	-.05312 (3.194)[a]	-.05547[a] (3.080)	-.25490[a] (2.976)
LIVES IN CENTER CITY OF SMALL SMSA	.01541 (1.164)	.00974 (.573)	.00862 (.468)	.02885 (.327)
LIVES IN ONE OF 31 LARGE SELF-REPRESENTING SMSA's	.00826 (.810)	.02618 (2.000)[b]	.02975 (2.098)[b]	.15572 (2.233)[b]
LIVES IN SMALL SMSA	-.00248 (.230)	.00303 (.219)	.00413 (.276)	.02191 (.303)
LIVES IN NORTH CENTRAL USA	-.00528 (.516)	-.01794 (1.352)	-.01773 (1.235)	-.00533 (.077)
LIVES IN SOUTHERN USA	-.01896 (1.837)[c]	-.00313 (.239)	-.00270 (.190)	-.07399 (1.068)
LIVES IN WESTERN USA	-0.2283 (2.009)[b]	-.02243 (1.536)	-.02235 (1.413)	-.09963 (1.301)
AGE (IN YEARS)	-.00126 (.437)	-.00488 (1.312)	-.00560 (1.391)	-.02796 (1.431)
AGE^2	.00007 (.442)	.00025 (1.132)	.00028 (1.191)	.00136 (1.197)
FAMILY INCOME (THOUSANDS)	.00021 (4.537)[a]	.00035 (5.783)[a]	.00038 (5.752)[a]	.01781 (5.659)[a]
FAMILY SIZE	.00106 (.455)	.00256 (.856)	.00285 (.883)	.01409 (.916)
FATHER RESPONDED TO SURVEY	.00438 (.195)	-.00624 (.217)	-.00836 (.268)	-.04867 (.326)
MOTHER RESPONDED TO SURVEY	.00631 (.303)	-.00444 (.167)	-.00658 (.226)	-.04092 (.295)
MALE	.00255 (.358)	.00356 (.390)	.00367 (.379)	.01309 (.275)
DHLTH2 = 1 IF HEALTH STATUS IS REPORTED TO BE GOOD	-.93091 (122.62)[a]	-	-	-
CONSTANT	.91892	.81374	.79282	-.24918
R^2 or x^2	0.857	0.066	0.067	178.97

Omitted categories for independent variables: race=white, education=highest grade completed by family head > 12, geographic area of residence=lives in northeast USA, respondent to survey = other person (not mother or father).

[a] $p \leq .01$
[b] $p \leq .05$
[c] $p \leq .10$

TABLE 2

SELF-REPORTED OVERALL HEALTH STATUS

RESULTS OF OLS AND N-CHOTOMOUS PROBIT REGRESSIONS FOR WOMEN

65 YEARS AND OLDER FROM THE 1978 HEALTH INTERVIEW SURVEY

(t-statistics in parentheses)

VARIABLE	INITIAL OLS	FIRST ITERATION OLS	FINAL ITERATION OLS	PROBIT
ASIAN	.03665 (.401)	-.01905 (.179)	-.04503 (.367)	-.03189 (1.628)
BLACK	-.00522 (.212)	.00962 (.336)	.01720 (.521)	.03783 (.339)
SPANISH	-.03226 (.707)	.00738 (.138)	.02978 (.487)	.10955 (.524)
OTHER	-.00990 (.192)	.05668 (.943)	.09088 (1.311)	.27525 (1.139)
AMERICAN INDIAN	.01273 (.077)	-.01045 (.055)	-.01227 (.055)	.19462 (.257)
HIGHEST GRADE COMPLETED IS 0-8	-.06133 (2.989)[a]	-.11618 (4.896)[a]	-.14338 (5.239)[a]	-.44994 (4.801)[a]
HIGHEST GRADE COMPLETED IS 9-11	-.00032 (0.014)	-.04139 (1.531)	-.06206 (1.990)[b]	-.17838 (1.674)[c]
HIGHEST GRADE COMPLETED IS 12	-.00286 (.138)	-.02166 (.890)	-.03031 (1.079)	-.07720 (.799)
LIVES ON FARM NOT IN SMSA	-.02551 (.612)	-.02904 (.598)	-.02862 (.511)	-.05454 (.290)
LIVES IN CENTER CITY OF ONE OF 31 LARGE SELF-REPRESENTING SMSA's	-.01606 (.748)	-.02622 (1.050)	-.03093 (1.074)	-.09461 (.968)
LIVES IN CENTER CITY OF SMALL SMSA	.00779 (.305)	.02245 (.755)	.03020 (.880)	.09868 (.853)
LIVES IN ONE OF 31 LARGE SELF REPRE-SENTING SMSA's	.01597 (.850)	.01904 (.870)	.01954 (.775)	.04937 (.576)
LIVES IN SMALL SMSA	-.00423 (.195)	-.02206 (.868)	-.03047 (1.039)	-.07649 (.576)
LIVES IN NORTH CENTRAL USA	.01155 (.632)	.00972 (.465)	.01820 (.754)	.03399 (.410)
LIVES IN SOUTHERN USA	-.01101 (.611)	.01218 (.573)	.01206 (.491)	-.00334 (.041)
LIVES IN WESTERN USA	.01307 (.626)	.01548 (.637)	.01576 (.562)	.03667 (.386)
AGE (IN YEARS)	-.00192 (1.823)[c]	-.00024 (.195)	.00063 (.444)	-.00021 (.043)
FAMILY INCOME (THOUSANDS)	.000004 (3.336)[a]	.000005 (4.169)[a]	.000006 (4.228)[a]	.02237 (4.515)[a]
FAMILY SIZE	-.01983 (3.071)[a]	-.02594 (3.452)[a]	-.02913 (3.361)[a]	-.10525 (3.582)[a]
WIDOWED	.04205 (2.569)[b]	.04662 (2.447)[b]	.04784 (2.177)[b]	.15011 (2.028)[b]
NEVER MARRIED	.05288 (1.799)[c]	.07160 (2.093)[b]	.08143 (2.064)[b]	.28971 (2.145)[b]
DIVORCED	.04747 (1.298)	.01769 (.416)	.00482 (.100)	.10939 (.665)
SEPARATED	.06630 (.979)	.11173 (1.420)	.12898 (1.421)	.31021 (.988)
*DHLTH2=1 IF HEALTH STATUS IS REPORTED TO BE GOOD	-.76345 (51.559)[a]	-	-	-
*DHLTH3=1 IF HEALTH STATUS IS REPORTED TO BE FAIR	-.75486 (42.081)[a]	-	-	-
CONSTANT	.89844	.60618	.46424	.58589
R^2 or X^2	.71044	.07690	.07891	109.53

Omitted categories for independent variables: race=white, education=highest grade completed > 12, geographic area of residence=lives in northeast USA, marital status=married.

[a] $p \leq .01$

[b] $p \leq .05$

[c] $p \leq .10$

TABLE 3

HEALTH STATUS SCALE OBTAINED FROM RUBINFELD
OLS ITERATIVE REGRESSION METHOD

	Health Status	Iteration #1	Iteration #2	Iteration #3	Iteration #4	Iteration #5
	Excellent	1.000	1.000	1.000	1.000	
Children 16 and Under	Good	.622	.548	.545	.552	
	Fair-Poor	0.000	0.000	0.000	0.000	
Women 65+ Years	Excellent	1.000	1.000	1.000	1.000	1.000
	Good	.518	.457	.447	.436	.423
	Fair	.375	.243	.197	.179	.167
	Poor	0.000	0.000	0.000	0.000	0.000

THIRD-GENERATION EPIDEMIOLOGY: PROBLEMS AND ISSUES FOR CONDUCTING EPIDEMIOLOGIC ANALYSES OF ARTIFICIALLY-PROLONGED LIFE ON A MASS SCALE -- THE CASE OF END-STAGE RENAL DISEASE

Tai Sugimoto, University of South Carolina

It is generally agreed that the epidemiologic mode of inquiry originated in 17th century Great Britain when John Graunt was the first to analyze mortality records. Since then, the epidemiologic mode of inquiry has surfaced periodically, most notably with John Snow's identification of the relationship between water supply and cholera in 19th century London. The primary factor enabling these pioneers to initiate the epidemiologic mode of inquiry was the large-scale records-keeping system of the British government which kept track of mortality and morbidity statistics. Despite the work of these early pioneers, however, modern epidemiology did not become established until the late 1940's in the United States when it became the methodological vehicle associated with public health concerns such as fluoridation of public water systems and clinical trials for immunization against various communicable diseases.

During the 1960's and 70's, the discipline of epidemiology expanded into investigation of the relationships between chronic disease conditions and environmental and life style related factors. Some of the most prominent topics for investigation have been studies on smoking and health, various studies on correlations between diet and cardiovascular disease, and numerous studies concerning chemical carcinogens and low-level radiation and human cancer.

Organ transplant procedures, renal dialysis, and other medical technologies which prolong and/or sustain lives artificially began in the 60's and proliferated rapidly in the 70's, especially with the 1973 initiation of federal Medicare subsidy of dialysis and kidney transplant therapies for end-stage renal disease (ESRD) patients. Other transplant procedures (e.g., heart, lung, liver, pancreas) and other means of prolonging/sustaining lives which would be otherwise lost in a short period of time have been developed, refined and performed at a rapidly increasing rate despite the lack of government subsidy except for research. A five-year survival rate of 50% or more has already been achieved for kidney, heart and liver transplants (although the actual survival rate varies by organ, related donor versus cadaver donor, histocompatibility factors, etc.) (1, 2,3). With the increased use of immunosuppressive drugs such as Cyclosporin, even higher graft retention rates are projected. Currently, it is not uncommon to see patients who have lived well over ten years with various transplanted organs. It is estimated that, worldwide, there have been more than 100,000 renal transplants alone, of which those in the USA account for approximately one-third. Unfortunately, although there still exists a Renal Transplant Registry, which was transferred from the American College of Surgeons (ACS) to the Health Care Financing Administration (HCFA) in 1977, there are now no registries for heart, lung, liver or pancreas transplants. Therefore, there are no nationwide statistics for transplants of organs other than kidneys. Based on ACS Registry records as of July 1, 1977, there had been 346 heart, 318 liver, 37 lung and 57 pancreas transplants in the world to that point in time. The rate of occurrence of those organ transplants since 1977 is estimated to be substantially higher than the pre-1977 rate. ESRD patients currently maintained on renal dialysis therapy are estimated to be in excess of 60,000 in the United States alone -- again about one-third of renal dialysis patients worldwide (4).

For the first time in history, there appears to be a substantial proportion of the human population whose lives are prolonged/sustained by medical technology for a relatively long period of time (5-10 years on average). These patients have been given literally a second lease on life.

The characteristics of this phenomenon pose a set of problems and challenges to the discipline of epidemiology in terms of both basic approaches and methodologic techniques. These new problems and challenges also transcend the issues unique to this new population group -- in the direction of epidemiologic analyses of the whole spectrum of human disease. However preposterous this suggestion might sound, it may be that the discipline of epidemiology is about to enter a new phase -- that is, third-generation epidemiology.

Some of the specific issues associated with this new dimension in epidemiology, based on studies of ESRD and its therapy (1,4), are given here. First of all, why was the disease category of ESRD, alone among many other competing disease categories, chosen for federal subsidy through the Medicare program? Why not heart disease, liver disease, or cancer? Several factors were involved in this decision: (a) the reliability of renal dialysis and renal transplant therapies beyond the experimental stage; (b) intermittent (dialysis) and one-time (transplant) therapy in terms of frequency; (c) in the case of transplants, the duality of kidneys enhances the availability of live related-donor organs; (d) costs for the treatment are beyond the capabilities of the vast majority of private individuals (especially of ESRD-disabled

At the present time, close to 70,000 ESRD patients are maintained on dialysis therapy in the United States (Figure 1). The incidence of untreated ESRD has shifted in the period 1967-1980, from just under 20,000 to an estimated 1,000 per year. Conversely, the incidence of treated ESRD increased from about 2,000 to close to 19,000 per year in the same period. There has been a corresponding rise in the incidence of loss from the ESRD treatment program; this generally means death. These changes present some problems for epidemiologists. In retrospect, the incidence of ESRD appears to have been relatively constant during the period 1967-1980 (total of incidence of untreated, treated, and loss from treatment program). The most remarkable change is the increase in prevalence of treated ESRD starting in 1973 when the subsidy program went into effect. Before 1973, there were very few dialysis facilities; it is estimated that only approximately 2,000 ESRD patients were maintained on dialysis and transplant therapies in the 60's. The remainder of those who manifested the uremic syndrome were simply untreated and usually died within a few months of onset of the uremic syndrome.

What has happened since is that the duration of the disease has been so prolonged that prevalence of the disease increased radically, without a significant increase in incidence -- that is, $P = I \times D$ (where P = prevalence, I = incidence, and D = duration of disease condition).

If the federal Medicare subsidy for ESRD treatment had not been instituted, the trend of a very gradual increase in dialysis and transplant facilities and personnel would have continued. The incidence of treated ESRD would probably have doubled and the prevalence rate would have shown a sharper rise relative to incidence starting in the mid-70's (Figure 2). However, the increase in both the incidence of treated ESRD and its prevalence would not have been as dramatic as what actually occurred with initiation of the federal subsidy. This particular pattern of occurrence no doubt will be repeated for other procedures of artificially prolonging/sustaining lives, such as coronary by-pass procedures, heart, liver, pancreas, and other organ transplants, as the numbers of patients treated by those procedures increase.

Since Medicare coverage of ESRD, the number of ESRD patients on maintenance dialysis and transplant therapy (or both) has risen rapidly each year, as predicted for the initial patient intake phase. What has been difficult to assess is when the increase would cease and when the annual incidence rate of ESRD would stabilize. Furthermore, it has been more difficult to predict if and when the prevalence rate would stabilize since it depends on the incidence rate as well as on the duration of patient survival. At the same time, it is suspected that more older patients and patients with more severe complications are being put on dialysis therapy -- those not previously considered appropriate candidates for this treatment.

Another way in which ESRD provides a unique opportunity, and problems, for epidemiology is its data base. Because of the requirements of federal subsidy programs, there exists a national registry of this particular disease category, theoretically with 100% coverage. However, in reality there is still a gap between ESRD treatment program enrollment figures and the true incidence of ESRD in the population. This particular dichotomy of a unique opportunity and problems demonstrated in the ESRD treatment program will be replicated as life is prolonged/sustained by artificial means after organ failure. The primary factors causing this gap are: (a) variations in completeness of reporting to the HCFA ESRD management information system; (b) variations in dialysis and transplant facility resource levels; (c) ESRD therapy paid for by other than Medicare, e.g., VA, private insurance, etc.; (d) physician decision to not put ESRD patient on therapy; (e) undetected (undiagnosed/misdiagnosed) disease; (f) variations in demographic characteristics; and (g) variations in environmental factors.

The problem of incomplete reporting to HCFA ESRD MIS by dialysis and transplant centers should be solved when the reporting requirement is tied to reimbursement and strictly enforced.

With regard to geographic variations in availability of dialysis and transplant facilities, the saturation point has been almost reached, as the number of established dialysis units peaked two or three years ago. In 1983, there are few if any providers attempting to establish new dialysis units, according to HCFA figures. Therefore, this factor should no longer be significant in contributing to the gap.

The ESRD MIS does not include patients who are Medicare-ineligible or are treated by other than Medicare services. Most patients in this category, about 4,000, are those treated at VA facilities, a relatively constant figure. According to the HCFA, at any given time approximately 15% of patients in the USA are being supported by funds other than Medicare; however, half of these patients are undergoing a three-month waiting period before they receive Medicare benefits. In order to have a comprehensive picture of ESRD in the United States, VA as well as other ESRD patients who are not supported by Medicare funds need to be accounted for. The HCFA is currently completing a study of patients not entitled to Medicare. Preliminary data show that 92% of all the patients currently being treated are Medicare ESRD program patients; therefore, the ESRD MIS figures represent a relatively slight underestimate of the total of ESRD patients being treated in this country.

There is a segment of the ESRD patient population not dialyzed because their medical and/or psychological problems make them inappropriate patients for this treatment. It is not certain if there is an appropriate method for estimating this figure by using morbidity and mortality figures such as hospitalization and

death certificate information; however, from morbidity studies by Hiatt and Friedman (5) and mortality studies by Chamblee and Evans (6), this method is far from practical. Similar to the above are patients who have never been diagnosed as ESRD, for whatever reason; we have no answer as to how to account for these latter patients.

Although there is a gap between the ESRD MIS figures and the true population incidence of ESRD due to these five factors, the ESRD MIS figures represent the majority, 80-90%, of the true population incidence. This is a higher coverage rate than any other major chronic disease category, including heart disease, malignancies, diabetes, etc.

Finally, assuming we can overcome the problems in accounting for the preceding five categories of patients, or the relative insignificance of the fourth and fifth categories, then the traditional task of epidemiologic study in search of causal factors can be performed.

Furthermore, a problem associated with defining "incidence" and "prevalence" rates should be mentioned. To date, the majority of published reports concerning ESRD rates have tended to utilize rates per million population without clearly defining "who is counted," "for what period," and "how counted." Most typically, the ESRD rates per million population reported so far are called "point prevalence" rates, without precisely defining these rates. The "point prevalence rate" is the number of persons afflicted with a certain (disease) condition in question at one point in time. The longer the time period used, the higher the count. However, in reality the variation among the reporting facilities in the counting method and period makes so-called "point prevalence" rates suspect.

These are inherent methodological problems and issues that the discipline of epidemiology has to deal with when a supposedly complete disease category data base becomes available, especially in the form of a registry compiled as a requirement for federal treatment subsidies of particular illnesses or conditions.

Procedures such as coronary artery by-pass, percutaneous transluminal coronary angioplasty, and various methods of treating malignancy are considered to be beyond the experimental stage; it is standard for the majority of health insurance carriers to cover these services. Since financial coverage can come from multiple sources, including private health insurance, Blue Cross/Blue Shield, and Medicare/Medicaid, there are no centralized, national registries for these diseases or conditions. However, the methodological problems and issues exemplified by the ESRD registry could apply to these other diseases or conditions.

"Third-generation epidemiology" has to broaden its spectrum of concern beyond the traditional concerns of spidemiologic inquiry. Before the phenomenon of artificially-prolonged life on a mass scale, the primary epidemiologic task had been more or less confined to identification of causal relationships between a disease condition and factors in the human environment. Third-generation epidemiology, on the other hand, requires additional tasks, beyond the traditional search for causal relationships -- still the predominant task of epidemiologic inquiry. These additional tasks and responsibilities of third-generation epidemiology include: (a) efficiency and efficacy of treatment/procedure (beyond the normal clinical trial sense) in terms of both length of life prolongation (efficiency) and quality of prolonged life (efficacy); (b) financial requirements in terms of allocation of total economic and medical resources for artificial prolongation of lives on a mass scale; and (c) compilation of records and statistics which could be utilized for examining ethical characteristics of these phenomena.

The efficiency of dialysis and transplant treatment for ESRD is evident from the remarkable length of survival reported by various studies (1-3). Although there is room for technical improvements (such as tissue matching criteria and techniques, immunosuppressive medication, and infection control in dialysis), on balance patient survival is improving. The dominant criterion for renal transplantation is HLA-A and B loci-based tissue matching, and epidemiologic studies have not yet provided more comprehensive criteria for donor-recipient matches (1).

In measuring the efficacy of treatment in terms of quality of life, epidemiology is far from providing basic descriptive statistics; there are very few studies in this area. The importance of epidemiologic inquiry, together with ethical and financial considerations, could perhaps help in determining future policies on mass prolongation of life. Furthermore, there are relatively few studies on the financial aspects of ESRD treatment (as reported at this conference by Dr. Paul Eggers of HCFA). The costs associated with ESRD treatment are changing rapidly as treatment technology changes, as well as from the general inflationary trends within the health care industry. As total health care costs in the United States exceed 300 billion dollars per year and 11% of the GNP, and the ESRD program costs 2 billion dollars per year, the cost factor will no doubt become a highly politicized issue.

In Massachusetts, Blue Cross/Blue Shield recently decided to provide coverage for heart transplants in both group and individual policies. Child liver transplants (recently, highly publicized pleas for donors have been organized) cost even more than heart transplants. An artificial heart was recently implanted and no one yet knows the total costs associated with that procedure. The rule of thumb has generally been that if the procedure is considered beyond the experimental stage, it is now reimbursible as a legitimate medical expense under most insurance policies.

The responsibility of epidemiology is to provide an accurate picture of the efficiency and efficacy of artificial means of prolonging lives and the associated costs.

Lastly, the biomedical ethics associated with this phenomenon are not the exclusive domain of philosophers and ethicists. Before the 1973 subsidy of ESRD teatment, it was standard that a committee, consisting of prominent citizens of a community, debated and made decisions as to who would be treated and who would not -- essentially a death sentence. Along these lines, an important characteristic of the ESRD treatment program is that more and more persons aged 65 and over, and patients with high-risk complications (diabetes and hypertensive nephropathy, for example) are enrolling in the program -- persons who would never have been considered as appropriate candidates for treatment before initiation of the subsidy. Considering the financial burdens and the resource allocations required for artificially prolonging lives on a mass scale, it is conceivable that the community committee system might be re-instituted, but perhaps in a more subtle manner at the federal policy-making level.

As an example of the ethical issues involved, consider that in renal transplants the trend is toward an increased proportion of cadaver donors over living relative donors, and the primary criterion for matching donor with recipient is matching at the HLA-A and B loci. Each locus has two antigens, yielding degrees of matching, i.e., 4, 3, 2, 1, 0. The differences in one year survival between a high match (4 or 3) and a low match (0, 1 or 2) are 4.8% and 11.3% in cadavers and living donors, respectively (1). Does this mean one should wait until a higher match organ is found for a less than 5% increased chance of survival at one year in the case of a cadaver donor? Choosing between a cadaver donor and a living relative donor is not an easy decision. How much is better survival worth in asking such a favor of a relative, especially since the advent of new drugs such as Cyclosporin which has led to substantial improvement in the prevention of organ rejection. Other ethical issues include the fact that, as recently reported (7, 8), 20% of the deaths of transplant patients during a five year period were suicides. Moreover, black and white racial differences in the incidence of ESRD as well as survival characteristics in transplant and dialysis therapies is another controversial item. Comprehensive epidemiological inquiry into these virtually unexplored dimensions of the phenomenon is imperative for those who treat the patients as well as for the policy makers who write the ground rules by means of resource allocation.

These new tasks and responsibilities of third-generation epidemiology illustrate the increased use of the epidemiologic mode of inquiry by a variety of policy makers and health care practitioners in recent years, far beyond the traditional task of epidemiology. Yet the irony is that the ultimate answer probably lies in the purpose of traditional epidemiology -- provision of information for primary prevention.

REFERENCES

1. Rosansky, S. J., and Sugimoto, T. An analysis of the United States renal transplant patient population and organ survival characteristics: 1977-1980. Kidney International 22:685-692, 1982.

2. Krakauer, H., et al. The recent U.S. experience in the treatment of end-stage disease by dialysis and transplantation. New England Journal of Medicine 308:1558-1563, 1983.

3. Vollmer, W. M., et al. Survival with dialysis and transplantation in patients with end-stage renal disease. New England Journal of Medicine 308:1553-1558, 1983.

4. Sugimoto, T., and Rosansky, S. J. The incidence of treated end-stage renal disease in the eastern United States: 1973-1979. American Journal of Public Health (in press).

5. Hiatt, R. A., et al. The frequency of kidney and urinary tract disease in a defined population. Kidney International 22:63-68, 1982.

6. Chamblee, R. F., and Evans, M. C. New dimensions in cause of death statistics. American Journal of Public Health 72:1265-1270, 1982.

7. Kalman, T. P., Wilson, P. G., and Kalman, C. M. Psychiatric morbidity in long-term renal transplant recipients and patients undergoing hemodialysis. JAMA 250:55-58, 1983.

8. Washer, G. R., et al. Causes of death after kidney transplantation. JAMA 250: 49-54, 1983.

FIGURE 1. U.S. ESRD Patient Status Change Through Time
(actual change with Medicare subsidy)

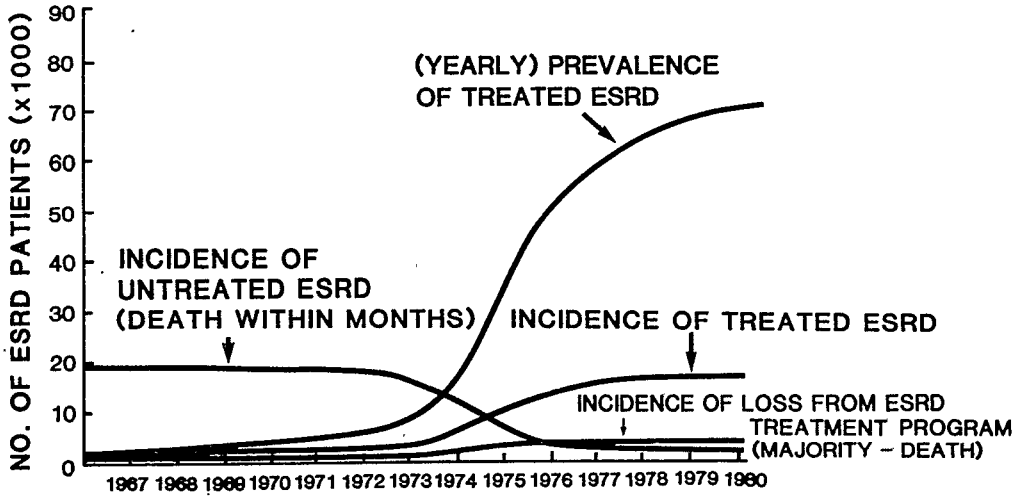
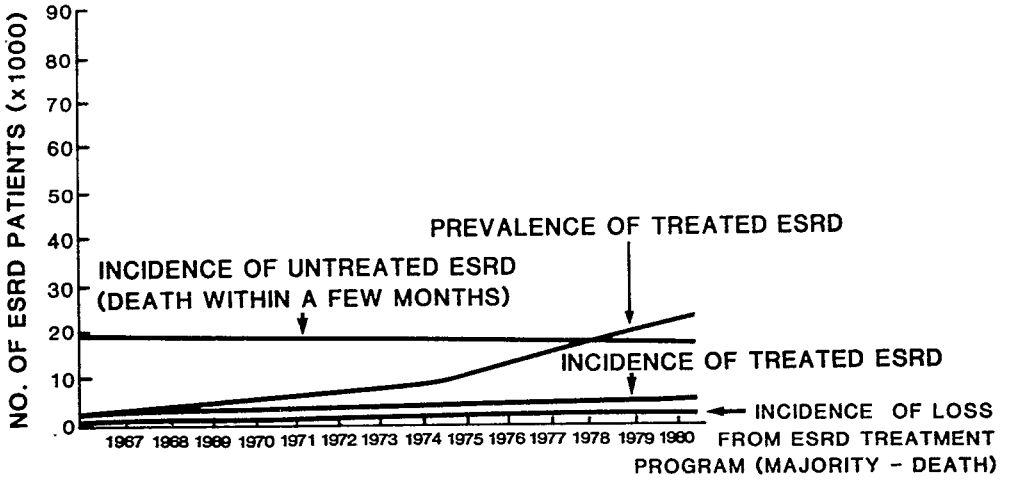


FIGURE 2. U.S. ESRD Patient Status Change Through Time
(hypothetical -- without Medicare subsidy)



Evaluating Long-Term Care: Some Analytical Approaches

Session V

ISSUES IN DEVELOPING A CROSSCUTTING DATA SET FOR EVALUATING COMMUNITY-BASED LONG-TERM CARE DELIVERY SYSTEMS

John Capitman, Berkeley Planning Associates, and Saul M. Spivack, University of Pennsylvania[1]

INTRODUCTION

The 1970's marked the emergence of functional independence and community-oriented care for the aged and disabled as major policy issues in health and human services. Interest in long-term care grew as local, state, and federal officials noted the growth of nursing homes and their rapidly expanding shares of Medicaid budgets. Recognition of the "graying" of America has focused attention upon difficult health care policy choices that require an informed response.

At the center of the current policy debate are widespread criticisms of Medicare and Medicaid coverage for long-term care. Many have asserted that the public insurance programs are biased toward institutional care and away from community tenure. Current benefit structures foster use of medically-intensive institutions for maintenance care, while discouraging use of community-oriented alternatives. Fragmentation and inadequate supply of the community services has also been cited. In most communities, a continuum of long-term care services, including institutional, ambulatory, rehabilitation, and home care options, as well as methods for matching patients with appropriate services is lacking. Inappropriate institutionalization or other patterns of excessively costly health care use by the aged and disabled have been viewed as the outcome of these problems in the delivery system.

The Health Care Financing Administration has fostered development of coordinated community-oriented long-term care systems by granting Medicaid and/or Medicare waivers to selected demonstration projects. The waivers permit testing the relative benefits of changing eligibility requirements for service, as well as the expanding amount, duration, and scope of services available under Medicaid and Medicare

Each of the HCFA projects was designed around a central organization that could coordinate and manage existing and expanded services, while establishing payment mechanisms suited to the particular needs of participants and communities. The projects test whether coordinated delivery of community-oriented health and social services specifically tailored to a client's needs can result in more appropriate and cost-effective use of both institutional and noninstitutional acute and long-term care resources. The programs are intended for individuals who are currently or would soon be medically appropriate for nursing-home-level care.

In order to explore the viability of the case management and service expansion systems, HCFA contracted with Berkeley Planning Associates, and subcontractors the University of Pennsylvania Rehabilitation Research and Training Center in Aging and the Western Center for Health Planning in 1980, to conduct an independent crosscutting evaluation of 15 waiver projects. Included were four projects in California, four in New York, and projects in Connecticut, Georgia, South Carolina, Florida, Oregon, Wisconsin, and Texas.[2]

The overall goal of this perhaps excessively ambitious meta-evaluation was to develop an integrative policy-relevant assessment of the impacts of the Medicaid and Medicare reforms on overall patterns of public expenditures for long-term care patients.[3] At the same time, the project sought to assess the extent to which anticipated beneficial impacts of the programs on health care expenditures could be achieved without sacrificing the quality of life and the quality of medical and social care provided for participants.

THREATS TO VALIDITY

The HCFA coordinated community long-term care projects developed independently and in response to the unique perceived delivery system problems of their communities. Beliefs about the most effective types of interventions varied almost as frequently as ideas about the most appropriate target group for which the new services would be made available. Nevertheless, all projects could be evaluated and compared in terms of their public costs, impacts on health services use, and client outcomes. But the meaning of these effectiveness measures would be misinterpreted without recognizing that the observed differences in the magnitude and direction of impacts would, to a great extent, be determined by differences in what the projects were attempting to acheive and the types of individuals they enrolled.

Accompanying the variation in intervention methods were differences in the research designs, patient assessment procedures, and source utilization data acquisition methods of the demonstrations. Thus, the crosscutting evaluation needed to address issues of validity at two levels: (1) traditional concerns in the assessment of individual project performance, and (2) assessment and comparison of differential performance of projects with varying intervention strategies and client populations. Within any given project, the traditional threats to internal validity included differences between treatment and comparative groups at intake, and differential attrition rates, as well as the customary concerns about the relevance, reliability, and validity of specific measures. Across projects, the central validity issues concerned distinguishing true variation in demonstration performance from the artifacts of research approaches. Thus, performance differences resulting from alternative demonstration methodologies needed to be separated from the impacts of different research designs, variables measured, and the nature of the measurements.

The first step in approaching this meta-evaluation challenge was to focus attention on the nature of measurements and their sources. The variety of measurement domains relevant to research goals

were surveyed for each project. While the projects collected a great deal of data, the crosscutting evaluation faced a number of critical issues in data availability, quality, and most importantly, comparability across projects. These problems were addressed through inclusion in the analytic data set of measurement domains and variables that were available in sufficient numbers of sites and with sufficient compatability to support crosscutting analysis. After describing the basic components of the analytic data set, examples of problem areas and the limitations they imposed are presented.

THE ANALYTIC DATA SET

Keeping in mind that demonstration performance was to be evaluated on differential criteria and at different analytic levels as a reflection of the intervention strategies and target groups, basic issues of cost-effectiveness and service-effectiveness could be examined for all projects. Cost-effectiveness is viewed from the perspective of demonstration impacts on patterns of medical and social service utilization, while service-effectiveness is viewed from the outcome perspective of client morbidity and mortality as well as changes in functioning and informal supports. In all cases, effectiveness was to be gauged by examining differences in experiences of demonstration participants and samples of individuals using the existing long-term care system. Further, in all cases, the possibility of exogenous factors at the level of the long-term care system obscuring the program's impacts would be considered. For some projects, it is exactly such changes at the level of the delivery system, however, that are crucial to testing evaluation hypotheses.

As a reflection of these basic goals, the performance evaluation analytic data set includes the first four files in Figure 1 for all projects. In addition, for all projects where delivery system level characteristics over time were central, all six of the indicated files are included. The first four files are described below.

PARTICIPANT DESCRIPTIVE DATA SET: This includes basic demographic functions status, sensory impairments, cognitive impairments, and a set of risk factors and service need indicators. Risk factors were "significant life events" which have been found to be correlated with institutionalization, as well as indications of recent hospitalization or applications for nursing home placement.

Measures of functional status include mobility and ambulation, activities of daily living (ADL), continence, and the instrumental activities of daily living (IADL). In the psychosocial functioning domain, lack of standardized approaches only permitted a gross measure of cognitive impairment, the Mental Status Questionnaire, which itself is not available in all cases and may not be sufficiently sensitive to temporal changes. Its utility for noninstitutional populations has also been questioned.

Since measures of functional status only reflect raw disability, they are almost meaningless as service need indicators without the context of the individual's informal (or social) support system. Since many potential demonstration participants receive unpaid assistance voluntarily rendered by family and friends, the projects actually serve residual or unmet needs that remain after considering the assistance provided by these "informal caregivers." Measures of unmet need were developed, calculated as the number of personal care and instrumental care activities that the client is unable to perform and for which no informal caregiver provided help at the time of assessment. The measure does not discriminate partial situations, i.e., activities where a caregiver was present but provided insufficient help, but the measures of Unmet ADL and Unmet IADL enable a cleaner picture of formal service needs. Comparisons of unmet with raw needs also serve as indicators of the role and magnitude of informal support systems.

While the internal consistency and interrater reliability of the measures of ADL, IADL, MSQ, Unmet ADL and Unmet IADL were demonstrated in most projects, unfortunate differences in scale construction often preclude meaningful cross-site comparisons at the level of the scale. Individual item comparisons are useful, but they create numerous methodological problems. Other elements of the data set appear particularly variable across sites. For example, risk measures such as recent hospitalization or Skilled Nursing Facility/Intermediate Care Facility (SNF/ICF) applications appear difficult to obtain in many cases, although they may be critical to comparative analyses of client group composition.

Because cost-effective community long-term care demonstrations must be built upon the foundation of existing informal support systems, the crosscutting evaluation sought to supplement existing assessment procedures in this domain. Four projects were able, upon the recommendation of the national evaluation, to collect detailed data about each of 11 personal care and instrumental tasks. For each task in which the client is dependent, the new instrument collected data on the participation of up to two primary caregivers, including the frequency of assistance and the limitations on assistance. The presence of secondary caregivers is also determined. The ability to supplement project protocols in more domains and for more projects would have been desirable.

PARTICIPANT OUTCOME DATA: This data set is intended for examination of service effectiveness or quality of care as viewed from the perspective of client outcomes. The major elements consist of reassessment measures of client functioning and informal supports, and measures of key transitions (such as institutionalization or mortality), as well as the length of project participation prior to their occurrences. The reassessment functioning variables will be used in analysis of covariance and related procedures to establish relative levels of change across experimental groups. Project reassessment procedures varied from initial assessments in a number of ways, often precluding development of fully comparable measures across time frames. The major difficulty in this data set revolves around project differences

in protocols for tracking transitions, and the adequacy of transition tracking data for periods between episodes of participation or after participant-initiated terminations. Resources were not available for studies of the reliability of transition data, and it will be difficult to fully account for artificial censoring of observations.[4]

PARTICIPANT SERVICE USE AND REIMBURSEMENT: This is the basic data on service utilization, including Medicare Part A and Part B use and reimbursements, Medicaid use and reimbursements, waivered service use and associated payments, and use of other public programs. While the Medicare use data has the greatest potential for consistency across projects, only Medicare Part A and Part B home health services were available to the evaluation directly from HCFA for the period 1978 till project termination. In those few cases where all project treatment and comparative samples are tracked by a single Medicare Part B carrier, data is being acquired from the carriers for the period of project participation. In most cases, however, the data were collected by multiple carriers and reliability and acquisition cost concerns have precluded use of this data.

Medicaid data will generally be available through the states. In some instances, this will also provide a method for assessing the reliability of Medicare utilization data. The detail of Medicaid use data, its organization and available time frames vary markedly across states. Data on waivered services are available in good quality from both the projects and for the Medicare demonstrations from HCFA. Data on use of Title XX and Title III are only available in selected sites, and even in those cases its services detail and reliability are limited.

Because of extended intake periods, multiple episodes of client participation, project terminations prior to the completion of client long-term care careers, and differential attrition across experimental groups, censoring of utilization data may seriously bias estimates of length of stay and associated reimbursement estimates. The utilization analytic files include the first episode of client or comparison involvement in the demonstration. Data will be expressed, where possible, as averaged monthly utilization and reimbursement estimates by service and payment category, as well as overall for the first 12 months after project enrollment or until termination by death or lost eligibility for service. Combined with statistical controls for prior use, this approach appears to introduce the lowest level of bias, short of the complex event history analyses applied by Miller and his coworkers.[5]

CASE MANAGEMENT AND SYSTEMS COORDINATION COSTS DATA: The provision of case management services and the performance of other new gatekeeping functions in the delivery system by the community care demonstrations are expensive. The results of some completed demonstrations suggest that the new coordination functions are so expensive that they more than offset incremental savings associated with project impacts on service utilization. Further qualitative analysis of differences between case management approaches in the projects indicated the likelihood of widely varied costs for delivering the administrative service.

This data set consists of averaged monthly case management cost estimates per client. In most cases, data on client utilization of case management was merged with the development of overall costs of the case management system to produce true unit costs of the service.

Accounting review of project internal cost data (at the San Diego, Monroe County, New York City, South Carolina, Project OPEN, On Lok, and Texas projects) suggests that both direct and indirect costs of the demonstrations can be estimated in reasonable ways, given effective separation of replication-relevant costs from artifacts of research and demonstration status, and the definition of comparable timeframes and caseloads estimates. Because of wide variation across sites in the availability of client-level data on case management use, allocation of the case management costs to individual clients was based on either time studies or staff estimates of time spent on various activities.

BARRIERS TO COMPARABILITY

The four basic analytic files just described were developed through attempting to find the most comparable sets of data elements and measurement approaches within the major evaluative domains relevant to policy formation. Nevertheless, a number of major barriers to comparable analysis across projects remain. Some general examples for the four analytic files discussed above are presented.

Participant Descriptive Data

- Information on medical severity or prognosis and physical or social risk factors was often missing.
- There were missing items from the versions of ADL and IADL scales adopted by the demonstrations.
- There was little uniformity in questions addressing the psychosocial functioning domains.

Participant Outcomes Data

- Little uniformity existed in measurement of medical outcomes and basic diagnostic measures were largely unavailable at reassessment.
- The physical, instrumental, and cognitive functioning scales required for comparability were often insufficiently sensitive to temporal change.
- Some important measures of outcomes were largely unavailable, such as measures of change in the living environment and the structure and function of informal support systems.
- Mortality data was often biased by differential attrition for reasons of participant preference

and, in some cases, by terminations because of institutionalization.

Participant Service Use and Reimbursement Data

- Across-project differences in service definitions was a problem.
- Using treatment plans or participant diaries for tracking of service utilization and reimbursements was not successful in the populations served by the demonstrations.
- Medicare Part B data other than home health was either unavailable or from differing sources.
- Differences in service tracking between experimental groups often limited within-project comparisons.
- Utilization data for programs other than Medicaid and Medicare was largely lacking.
- The availability of prior use data from Medicaid varied considerably across projects.
- The effectiveness of regional and temporal controls for price levels was limited.

Case Management and System Coordination Costs Data

- No project collected client-specific case management use data for all clients. In three cases, individual-level data was collected for selected clients and timeframes only.
- Projects varied considerably in methods available for estimating the replication-relevant proportion of staff time.
- The effectiveness of regional and temporal controls variations in price levels was limited.
- Differences in project history often created differences in defining operational timeframes.

IMPLICATIONS: THE NEED FOR STANDARD MEASURES OF PROGRAM IMPACT

The preceding discussion has highlighted reasons why the cross-cutting evaluation of community-oriented long-term care demonstrations must address unique questions related to the validity of project comparisons as well as traditional validity and reliability concerns. Faced with similar barriers to comparability, some prior studies have still attempted to pool data from various demonstrations into single analyses, while others have abandoned all but qualitative comparisons. Based on the approach to cross-cutting data set definition described in this paper, the evaluation of the HCFA-sponsored community care programs has taken a new route.

The crosscutting analysis was developed in two stages. During the first stage, or primary analysis, the best-available approaches to analysis of individual project data were pursued. All appropriate measures within the indicated domains were utilized and qualitative data on program design was used interpreting analyses. In the second stage, or meta-analysis, only those measures that were available across projects were utilized, but the best-available approach to analysis of an individual project was still used. The goal was to find the best estimate of given measures of program impact for individual projects. Variations in the direction and magnitude of effects across projects are then described, using both qualitative data and results from the individual project analyses.[6]

One example of a standard measure of program impact used in the meta-analysis was in the area of combined overall Medicaid and Medicare expenditures. Only program component costs available across projects were included. The impact of given projects on public expenditures was expressed in terms of the number of additional or fewer nursing home days that would have to be consumed by treatment clients to equalize their averaged costs to the averaged costs of serving comparison participants. Where appropriate, this estimate was corrected for intake or prior use differences between groups, although this introduced project-to-project differences in methods for estimating utilization patterns. This measure adjusts for temporal and regional differences in service pricing structures and patterns of inflation. While this and related approaches allow overcoming many program-specific validity problems and allows for cross-project comparisons on standard measures of impact, issues of differential validity of component data elements are not overcome. Use of the primary analyses and qualitative program descriptions are still required to establish judgmental intervals of significance when comparing demonstrations.

Footnotes

[1]This research was funded by the U.S. Department of Health and Human Services under Contract No. 500-80-0073.

[2]The demonstrations included in the crosscutting evaluation are: ACCESS I and ACCESS II of Monroe County Long Term Care, Inc.,; Long Term Care Project of North San Diego County; New York City Home Care Project; On Lok Senior Health Services; South Carolina Community Long Term Care Project; Project OPEN of Mt. Zion Hospital; Florida Pentastar Project; Wisconsin Community Care Organization, Triage, Inc.; Georgia Alternative Health Service Project; Oregon FIG/WAIVER Continuum of Care Project; New York Nursing Home Without Walls Project; Texas Alternative for the Institutionalized Aged Project; and California Multipurpose Senior Services Project.

[3]The crosscutting project required both primary and secondary analysis of data collected by the evaluator and the programs themselves. Both qualitative and quantitative methods were used in the synthesis of findings derived from these many sources. See J. Hunter, et al., Meta-analysis Cumulating Findings Across Studies, Sage Publications, Beverly Hills, CA, 1982.

[4]Censoring occurs when the period of observation

for a case is determined by research methods such as failures in client tracking rather than by the occurrence of the events being investigated. See, for example, B. Brown, et al., "Nonparametric Tests of Independence for Censored Data with Applications to Heart-Transplant Studies," Reliability and Biometry, Philadelphia, 1977, 327-354.

[5]Leonard Miller, et al., The Comparative Evaluation of the Multipurpose Senior Services Project: A Progress Report, U.S. DHHS, HCFA Grant No. 11-P-97553, Sacramento, CA, July 26, 1983.

[6]For a similar approach, see R. Light and D. Pillemer, "Numbers and Narrative: Combining Their Strengths in Research Reviews," Harvard Educational Review, 52(1), 1982, pp. 1-26. It represents a departure from the meta-analytic techniques comparing equivalently calculated effect sizes advocated by G. Glass, et al., Meta-Analysis of Social Research, Sage Publications, Beverly Hills, CA, 1982.

Figure 1

The Performance Analytic Data Set: Evaluation of Coordinated Community-Based Long-Term Care

File 1 Participant Descriptive Data	File 2 Participant Outcome Data	File 3 Participant Service Use and Reimbursement	File 4 Case Management & Systems Coordination Costs Data	File 5* Aggregate Long Term Supply and Changes	File 6* Aggregate Long Term Care Utilization and Changes
Participant I.D. Background Age at admission Location at assessment Source of referral Sex Eligibility for Medicaid/Medicare Living Arrangement Assessment MSQ ADL IADL Mobility Sensory Function Diagnoses Unmet ADL given informal support Unmet IADL given informal support Significant life events Other Service Need Indicators Recent hospitalization ICF/SNF applications Prognosis Client/caretaker preferences Informal Support Caregiving network Configuration (primary caregivers) Characteristics of up to 2 primary caregivers for 11 individual ADL/IADL tasks Caregiving frequency for 11 tasks Presence of additional caregivers for 11 tasks	Participant I.D. Reassessment of Functioning 6, 12, 18 months ADL IADL Mobility Sensory function Unmet ADL given informal support Unmet IADL given informal support Changes in Informal Support Caregiving network Configuration Caregiving frequency Caregiving tasks Dates of Transitions Changed locus of care Changed eligibility for Medicaid/Medicare Changed source of informal care Death	Participant I.D. Medicare Part A Hospital inpatient SNF Skilled home health Home health aide Other (Averaged monthly units and charges) Medicare Part B Skilled home health Home health aide Other (Averaged monthly units and charges) Medicaid Hospital-inpatient SNF ICF Other, depending on project (Averaged monthly units and charges) Waiver Depends on project Other Programs Depends on community	Participant I.D. Total Direct and Indirect Costs for: Intake Assessment Reassessments Care planning Service arrangement Care plan and service monitoring Direct client service	Study Area I.D. Institutional Beds Acute SNF ICF Domiciliary Community Services Available Units Home health-skilled Home health aide Personal care Homemaker/home chore Adult day health Adult day social Medical transportation Nonmedical transportation	Study Area I.D. Aggregate Expenditures by Medicare Acute SNF Home health Other Aggregate Expenditures by Medicaid Acute SNF ICF Other Average LOS by Population Groups Acute SNF ICF Organized home care Admission Sources for: Acute SNF ICF Organized home care Discharge Locations from: Acute SNF ICF Organized home care
				*Particularly for programs with population-level scope of intervention, but may be relevant to all projects' performance.	

Variables listed are only representative. Elements may be missing or additional elements required for individual projects.

ANALYSIS OF LONG TERM CARE PATIENT ORIGIN STUDIES CARRIED OUT IN HEALTH SERVICES AREA # 9 (CALIFORNIA)

Richard W. Ainsley, Central California Health Systems Agency

In an attempt to better understand the local system of long term care, the origin of patients in licensed skilled nursing and intermediate care facilities was surveyed in the summer of 1979 and again in 1981 by the data division of the Central California Health Systems Agency. Analysis of the study data was directed at understanding market-place dynamics rather than re-defining official planning areas.

The most obvious result of the analysis was that nearly all of the patients in the surveyed facilities formerly resided within 20 miles of the facility in which they resided. Tracing patients back to their zipcodes of origin a stricking difference in the utilization of long term care facilities became apparent. Citizens in towns with existing long term care facilities are three to seven times as likely to use long term care services as citizens in towns without long term care facilities. Increases in market penetration were associated with larger numbers of licensed beds (Roomers Law), better public relations, and increasing density of senior citizen population. However, the projected need (additional capacity) for long term care beds does not correlate well.

Changes which took place between the two studies which took place between the two studies were analyzed and yield insights about the transition between rural-style (low penetration) and urban style (high penetration) long term care markets.

I. The Problem

Simple Analysis of Long Term Care markets by Central California Health Systems Agency (CCHSA) reveals great contrasts between the URBAN areas and rural areas and the two major urban areas (Bakersfield and Fresno cities). Previous to the studies reported here, it was commonly assumed that the high percapita use rates in the urban areas were the result of significant inward migration of patients from rural areas. Patient origin studies were carried out in an effort to equitably adjust skilled nursing facility and intermediate care facility bed requirements for this assumed patient flow.

II. Survey Development and Response

Two surveys were conducted by CCHSA, the first in summer of 1979 and the second in the summer of 1981. Information requested relevant to this report was the same in both studies. A variety of other questions were also asked including questions about waiting lists and admission policies. Survey instruments were developed by the author in consultation with several nearby skilled nursing facilities on both occasions. On each occasion, survey instruments, instructions, cover letters and self-addressed, stamped envelopes were posted to each freestanding skilled nursing and intermediate care facility and distinct part skilled nursing facility in HSA-9. After six months, no further attempts were made to collect survey instruments. The following analyses are based on 54 (70%) returns from 76 surveys posted in 1979 and 73 returns (95%) from 77 surveys posted in 1981. Known patient origins within HSA-9 were analyzed for 4,427 patients in 1979 and 5,773 patients in 1981. Data were analyzed on a TRS-80 Model III Microcomputer using custom software made available gratis by Ainsley Interfacing Data Systems of Visalia, CA. The following analyses were carried out:

1. Data were summarized in a large table with columns corresponding to zipcode of patient origin and rows to facilities surveyed.

2. The summary data table above was used to calculate matrices of commitment indices and relevance indices.

3. Net flow of patients was calculated.

4. Percapita use rates in terms of patients per 100 seniors were derived for Tulare and Kings County zipcodes using 1980 census data.

Results and Analyses

Due to the brevity of this report the tables of raw data, commitment indices and revelance will not be reported here, but are available from the author.

The simplist analysis of the data was to determine the geographic size of an average market area. To this end a statute table was developed by the author for all locations with post offices in HSA-9. Assuming that City Halls tend to be located near the greatest concentration of people in each zipcode, on the average about 95% of the patients from HSA-9 originated within 25 miles of a typical reporting facility.

The hypothesis that the high per capita use rates in urban areas were the result of a net inflow of patients was found to be false. There is an inadequate net flow of patients from rural areas to urban areas to account for the large differences between urban and rural areas in terms of patients per 100 seniors (see Net Flow Table I). In an effort to better understand the area wide dynamics of this result, a detailed analysis of the 1981 surveys was developed for Tulare and Kings Counties (see Table II). Detailed analysis was restricted to these two counties because there was 100% return from skilled nursing and intermediate care facilities from both counties and due to the central location of the two counties, all or nearly all patients from Tulare and Kings Counties were reported. Results indicate that there is a very strong positive association between patients per 100 seniors and the total number of available skilled nursing and intermediate care facility beds in a zipcode. (See Table III.) Other

features of the zipcode areas such as population are less tightly associated with changes in market population and HSA shortage - excess calculations are not related to changes in market penetration. This result is similar to the concept developed by Roomer that the availability of beds induces demand. The basis of this result appears to be the result changing cultural attitudes towards skilled nursing and intermediate care facility care as skilled nursing care becomes more visible in small communities.

Analysis of Change to New Beds

Tulare County offered an opportunity to closely examine the result of opening additional skilled nursing facility beds. In the spring of 1980, Linwood Gardens Convalescent Hospital opened 79 licensed beds. Table IV shows the distribution of patients by origin before and after the opening of Linwood Gardens Convalescent Hospital.

Notice that:

(1) Among the three facilities previously established in Visalia (the most distant from Linwood Gardens within Visalia, about three miles away) was negatively affected, Delta Convalescent Hospital was not affected, and Kaweah Manor Convalescent Hospital (the nearest facility to Linwood Gardens Convalescent Hospital) increased its census.

(2) The greatest number of new patients actually originated within Visalia zipcodes as the new facility came on line.

(3) Visalia Convalescent Hospital had a slightly increased proportion of out-of-town patients. The author was fortunate to have monthly occupancy statistics for all the skilled nursing facilities in Visalia City area which date back to four months before the opening of Linwood Gardens Convalescent Hospital. These data substantiate the above general conclusions.

Discussion

Although there have been several patient origin studies carried out by health systems agencies (HSAs) in California (and no doubt others, elsewhere), few were published and none were analyzed in the depth reported here (1). However, an unpublished study carried out by the Middle Tennessee Health Systems Agency (MTHSA) in 1982 does complement the results reported here. In their study, the staff of the MTHSA interviewed relatives of clientele in four skilled nursing/intermediate care facilities in considerable depth and found:

(1) There was a preference for local care when it was available.

(2) The presence of an adjacent general acute care hospital, although important, was not a major factor in the selection of a skilled nursing or intermediate care facility.

3. Awareness of various options to SNF (and likely among SNFs) was directly related to the number of relatives actually in SNF/ICFs.

Thus, both the MTHSA study and that reported here indicate the skilled nursing and intermediate care facilities' market penetration varies greatly from place to place at a local level, and that availability of service is an important predictor of market penetration at the local level.

Bibliography

1. Patient origin studies for long term care patients have been carried out in California by the Golden Empire HSA, Mid Coast HSA, Inland Empire HSA and Central California HSA. Brief summaries of the results of these studies have been published in the proceedings of the Annual Conference on Using California Health Facilities Commission.

2. "Maury County Nursing Home Report," Middle Tennessee Health Systems Agency, December 1982.

TABLE I

1981 PATIENT ORIGIN STUDY

TOTAL NET INFLOW OF PATIENTS INTO FACILITY HFPA FROM OTHER HFPAs

	F A C I L I T Y H F P A													
ITEM*	601	603	605	607	608	609	611	613	615	617	619	621	623	625
HSA FLOW	-6	-33	176	4	-38	8	-32	-44	-27	81	-30	-4	-16	-39
NET FLOW	25	-22	429	24	-34	17	-16	-39	-10	153	-30	-4	-16	-39
% FLOW	6.6	-34.4	21.0	5.4	-26.2	47.2	-2.8	-10.9	-3.1	17.7	-100.0	-100.0	-100.0	-53.4
LOCAL PT	379	64	2046	443	130	36	572	357	320	865	30	4	16	73
PTS ADMT	404	42	2475	467	96	53	556	318	310	1018	0	0	0	34

* HSA Flow is the net flow of patients from known zipcodes within the HSA to the facility HFPA.

LOCAL PT refers to the known total of patients who originated in the facility HFPA.

PTS ADMIT refers to the total patients in participating LTC facilities in the facility HFPA.

% FLOW is 100% times the difference between patients admitted and patients originating in the facility HFPA divided by the patients originating in the HFPA.

TABLE II

MARKET PENETRATION IN KINGS COUNTY
(1981 PATIENT ORIGIN DATA)

LOCATION	POPULATION TOTAL	POPULATION 65+ YEARS	TOTAL LTC PATIENTS	PATIENTS PER 1000 TOTAL POPULATION	PATIENTS PER 100 SENIORS (65+)
Kings County Census Division					
Hanford Division less Armona*	33,683	3,723	233	6.9	6.3
Armona (C.T.5)	3,501	308	10	2.9	3.2
Avenal & Stratford Division	7,814	643	19	2.4	3.0
Corcoran Division	9,182	768	26	2.8	3.4
Lemoore Division	19,558	928	33	1.7	3.6
Kings County	73,738	6,370	321	4.3	5.0
Tulare County Census Division					
Dinuba	19,478	2,274	97	4.98	4.27
Earlimart	4,578	364	13	2.84	3.57
Exeter	17,099	2,104	39	2.28	1.85
Ivanhoe	5,460	551	2	0.37	0.36
Lindsay	15,536	1,757	58	3.73	3.30
Orosi-Cutler	10,415	848	33	3.17	3.89
Pixley	4,471	468	8	1.79	1.71
Porterville	40,888	4,962	242	5.92	4.88
Springville	4,160	633	11	2.64	1.74
Strathmore	5,437	665	12	2.21	1.80
Terra Bella	4,211	362	13	3.09	3.59
Tipton	5,033	360	4	0.79	1.11
Tulare	35,595	3,566	158	4.44	4.43
Visalia	65,047	6,350	356	5.47	5.61
Woodlake	8,730	1,008	13	1.49	1.29
Tulare County	246,138	26,272	1,059	4.29	4.02

* Location with Skilled Nursing and Intermediate Care Facilities.

SOURCES: 1980 Census Population Data (Summary Tape File 2).
1981 Long Term Care Patient Origin Study, Ainsley 1982.

TABLE III

ORIGIN LOCATIONS RANKED BY PATIENTS PER CAPITA (SENIORS ONLY)

CENSUS DIVISION OF ORIGIN	SKILLED NURSING FACILITIES	LIC. BEDS	PATIENTS PER 100 SENIORS	PATIENTS PER 1000 POPULATION
Hanford (*)	3	316	6.26	6.92
Visalia (+)	4	389	5.61	5.47
Porterville (+)	3	273	4.88	5.92
Tulare (+)	2	195	4.43	4.44
Dinuba (+)	1	99	4.27	4.98
Orosi-Cutler (+)	0		3.66	2.98
Terra Bella (+)	0		3.59	3.09
Earlimart (+)	0		3.57	2.84
Lemoore (*)	0		3.56	1.69
Corcoran (*)	0		2.4	2.8
Lindsay (+)	1	51	3.3	3.73
Armona (*)	0		3.2	2.9
Avenal and Stratford (*)	0		3.0	2.4
Exeter (+)	1	19	1.85	2.28
Strathmore (+)	0		1.80	2.21
Springdale (+)	0		1.74	2.64
Pixley (+)	0		1.71	1.79
Woodlake (+)	0		1.29	1.49
Tipton (+)	0		1.11	0.79
Ivanhoe (+)	0		0.36	0.37

* Kings County Census Division

\+ Tulare County Census Division

TABLE IV

THE IMPACT OF 70 NEW SNF BEDS OPENING IN VISALIA
ON ORIGIN OF PATIENTS USING SNF
IN VISALIA CITY

ITEM	ORIGIN LOCATIONS WITH FACILITIES					
	Visalia	Tulare	Dinuba	Porterville	Out-of-County	TOTAL
'79 Study Patients	195	9	1	5	11	286
% Patients	68%	3%	0.3%	2%	4%	
'81 Study Patients	239	5	1	2	29	355
% Patients	67%	1%	0.3%	0.6%	8%	

Statistical Perspectives on the U.S. Mental Health Delivery System

NIMH Special Session

STATISTICAL PERSPECTIVES ON THE U.S. MENTAL HEALTH SERVICE DELIVERY SYSTEM

Ronald W. Manderscheid, Chairperson
Michael J. Witkin
Marilyn J. Rosenstein
Rosalyn D. Bass
Survey and Systems Research Branch
National Institute of Mental Health

Symposium Summary

This symposium presented a description of the current status of the U.S. mental health service delivery system through an examination of trends in data collected by the Survey and Systems Research Branch, National Institute of Mental Health, in collaboration with the States. Data derive from organizational inventories and patient sample surveys conducted by the Branch on a periodic basis as part of the National Reporting Program. Specific topics of discussion were as follows: (1) trends in availability and utilization of services across all specialty mental health service settings, e.g., State and county mental hospitals, private psychiatric hospitals, Veterans Administration psychiatric services, community mental health centers, etc.; (2) trends in the sociodemographic, clinical, and treatment characteristics of patient populations; and (3) trends in staffing, with particular emphasis on the core service disciplines, i.e., psychiatrists, psychologists, social workers, and psychiatric nurses. Patterns in each of these areas are synopsized below.

Service Availability

Presented by Michael J. Witkin

Based on data over the past 25 years and, in particular, the last 10 years, several distinct patterns have been observed.

The composition of the mental health service delivery system has changed throughout the period. In 1955, the year in which the number of residents in State mental hospitals was at a peak, 77 percent of all episodes in mental health facilities were in inpatient settings. During the 1960s and 1970s, the growth of federally-funded community mental health centers, with their emphasis on outpatient and day treatment programs, accelerated, while State mental hospital populations continued to decline. By 1971, only 42 percent of the episodes were inpatient; by 1979, only 24 percent were inpatient, 73 percent outpatient, and 3 percent day treatment.

The number of patient care episodes has increased throughout the years. In 1979, there were 7.4 million patient care episodes in specialty mental health organizations, an increase of 335 percent over the 1.7 million episodes in 1955. In 1955, nearly half the episodes occurred in the inpatient units of State mental hospitals. By 1979, 39 percent occurred in the inpatient and outpatient units of community mental health centers, while inpatient services of State mental hospitals contributed only 8 percent.

The growth in the number of patient care episodes in mental health facilities has exceeded growth of the civilian population. For inpatient and outpatient care episodes combined, the rates per 100,000 population more than tripled from 1,028 in 1955 to 3,251 in 1979. For inpatient service settings, the rates in the same time period only rose from 795 to 801, and actually decreased between 1975 and 1979. As expected, however, the most dramatic growth came in outpatient episodes, which were up over 10-fold from 233/100,000 in 1955 to 2,450/100,000 in 1979.

Expenditures for all mental health organizations combined increased only slightly between 1969 and 1979 in constant dollars (adjusted for inflation), while similar expenditures for State mental hospitals actually decreased. In actual dollars, the expenditures for all mental health facilities combined rose from $3.3 billion in 1969-70 to $8.8 billion in 1979-80 (167 percent). When adjusted for inflation, the increase was only from $3.3 to $4.7 billion (27 percent), or from $16.53 to $19.37 on a per capita basis. For State mental hospitals, the dollar expenditures, when adjusted for inflation actually decreased from $1.81 to $1.78 billion or from $9.11 to $7.98 on a per capita basis, while rising slightly for other types of organizations.

What does the future hold? At this point it is not clear what trends will be experienced in the utilization of mental health facilities over the next few years. Some factors that will influence future utilization include the availability of health insurance for mental disorders, the availability of services, and continued pressures to close State mental hospitals, while at the same time providing adequate services to the chronically mentally ill.

Patients

Presented by Marilyn J. Rosenstein

In the recent past, an increasing emphasis has been placed on alternatives to public mental health services. It has been noted that the public and private sectors often function as separate systems of care, with the public sector providing care to the chronic patient with few resources, and the private sector providing care to the less difficult, more affluent patient. It is of interest to compare the patient groups admitted to publically and privately operated facilities to see if such differences continue to exist and if any major shifts have occurred in the characteristics of patient groups over time.

In 1980 and 1981, the Branch conducted sample surveys of patients admitted to the psychiatric inpatient services of State and county mental hospitals, private psychiatric hospitals, Veterans Administration medical centers, and separate psychiatric units of non-Federal general hospitals. National estimates of sociodemographic, clinical, and treatment characteristics of patients were derived from these surveys.

Based on the results from these four surveys, the distribution of sociodemographic characteristics, such as sex, race, age, marital status and education, and other patient characteristics, such as legal status, prior psychiatric care, diagnostic groupings and payment source, were compared across public and private mental health programs. In general, during 1980, public inpatient programs tended to admit higher proportions of males, minorities, adults, the unmarried, and the less educated, compared with private inpatient programs. Public and private programs also differed considerably with respect to the proportion of people entering treatment on an involuntary basis, with such commitments almost totally centered in the public sector. Although most persons admitted for inpatient psychiatric care had some prior psychiatric treatment, the proportions of admissions with prior inpatient care were generally higher in public than in private services. Diagnostic distributions also differed, with a higher proportion of the public admissions diagnosed with schizophrenia, and a higher proportion of the private admissions diagnosed with affective disorders. Perhaps the most dramatic difference between public and private settings is in the expected primary source of payment. Public programs admitted a relatively large proportion of patients for which the program expects no payment, while private programs tended to admit those with some type of insurance coverage.

Thus, the data from 1980-81 do indicate that there are considerable differences in the types of patients admitted to public and private inpatient programs. Comparison of these data with similar available data from 1975 indicated that these differences tended to remain fairly stable over time.

Future surveys are planned for 1985 to compare patients seen in inpatient, outpatient, and partial care services across a wide range of public and private mental health programs.

Staffing

Presented by Rosalyn D. Bass

Facilities included in the organizational inventories are asked to report the number of staff employed, by discipline and employment status (i.e., full-time, part-time, or trainee status) and the number of hours these staff are scheduled to work during a sample week.

Reported staffing data suggest that the mental health system is expanding in organized settings. During the period between 1972 and 1978, an increase of 13 percent occurred in the number of full-time equivalent (FTE) staff. The resulting change in the staffing mix of the mental health system was toward greater professionalization, as the core mental health professionals--psychiatrists, psychologists, social workers, and registered-nurses--increased their numbers in greater proportions than the aggregate of all other patient care staff or the aggregate of administrative and maintenance personnel.

For the core mental health professions, the increase in number of FTE psychiatrists has not kept pace with increases in FTE psychologists, social workers and registered nurses. Between 1972 and 1978, the number of FTE psychiatrists working in organized settings grew by 12 percent, as compared to 81 percent for psychologists, 71 percent for social workers, and 36 percent for registered nurses. An increasing growth curve for psychologists, social workers, and registered nurses was consistent across all facility types. A more varied pattern was found for psychiatrists, with decreases actually reported in number of FTE psychiatrists in State mental hospitals and the psychiatric services of general hospitals during 1974 to 1978.

These findings suggest the need for research on the impact of these changes in staffing mix on patterns of care in the mental health service system and on personnel substitutability in relation to meeting patient care needs.

Further Information

Publications are available from the Survey and Systems Research Branch on each of the topics covered in the symposium. Of particular interest will be Mental Health, United States: 1983, a compilation of data covering a ten year period. Copies may be obtained by writing Ms Berdie Firestone, Services Research Resources Branch, DBE, NIMH, Room 18C-06, Parklawn Building, 5600 Fishers Lane, Rockville, Maryland 20857.

Monitoring and Evaluating Health Programs: System Design and Program Administration

Session W

USE OF HEALTH STATISTICS IN MONITORING AND EVALUATING HEALTH CARE PROGRAMS FOR MANAGEMENT PURPOSES

Linda S. Chan, Los Angeles County-USC Medical Center

INTRODUCTION

In this era of accountability and financial difficulty, the management of health care delivery systems is continually faced with the need to establish priorities of health care programs for future implementation. Especially when the decision maker is faced with multiple health care programs and multiple geographic areas, and the constant need of selective resource allocation, a systematic approach enabling him to make sound and rational decisions is of great necessity.

In the past two decades, a great deal of effort has been directed toward the development of data collection systems and the derivation of health indicators and indexes for health care planning and evaluation purposes. In many recent planning activities, it was found that tremendous amounts of health data are being collected at local, state and federal levels, but are not being adequately utilized and analysed. A gap between the generation of health data and their actual use is widely felt by many health care professionals.

Past attempts were made to derive health indexes for management purposes. One classic example was the development of the Index of Medical Underservice[1] (IMU) in 1975 which was later adopted by the Congress to designate medically underserviced areas in the nation so that federal health resources could be allocated objectively. The Index had a significant impact in the planning and administration of health care services. However, it somewhat pointed out a serious weakness in the understanding of the concept of the development and use of health indicators and indexes. In his review of the IMU, Wysong[2] pointed out that the major weakness of the Index lay in the lack of definition of the concept of medical underservice which resulted in the unwise mixing of distinct components of health care planning considerations. Among the four parameters used in the Index, one was a measure of the availability of health services to a population, another was a limited measure of the health status of the population and the other two were socioeconomic characteristics of the population that were related to a wide range of system characteristics, such as availability, accessibility, utilization, and health status. The Index, in a sense, was a mixed bag of different dimensions of health planning considerations. Such a vagueness of the conceptual understanding of the use of the Index calls for a need to reiterate the fundamental concepts of health care planning and evaluation.

Introduced here is an analytic approach that takes into consideration the various dimensions of decision making, makes use of readily available health statistics, and produces a decision making matrix that assists management in establishing priorities among health care programs. It demonstrates a rational use of health statistics that builds on the fundamental concepts of health care management.

DESCRIPTION OF THE ANALYTIC APPROACH

The analytic approach is intended to be a tool that would allow the management of multi-service and multi-location health care delivery systems to establish priorities to allocate limited resources in a rational and timely manner. It also helps the management to cope with the dynamic problems of changing budgets and changing need and demand of health care programs in the communities. The organization and management of personal health services systems present problems not common to industrial management. The health service system is engaged in the production of a custom-made product for each consumer. It does not have a homogeneous or standardized output. The paths appropriate for diagnosis and management of care are too numerous to anticipate. Each medical management case is unique or almost unique. It is a complex of personal, social, technical and biological interactions. In the management of personal health services one is dealing with many variables, many of them are related to human behavior patterns which may be contrary to the provider's expectations. This poses a challenge to health care planning and management.

This analytic approach attempts to integrate the different dimensions of health care management in a systematic and meaningful manner leading to rational decisions. It is not an approach to develop appropriate indicators for health care programs. Rather, it demonstrates the use of the already available health data and health indicators for each health care program. It considers each dimension of health planning individually rather than in an aggregate fashion. It is a systematic process that begins with the concept and ends with the ultimate decision.

The approach can be divided into six analytic steps:

Step 1: Identify the purpose and use of the findings.

This is the first and most important step. It represents the purpose of the entire mission. Since the planning and evaluation process is cyclical, it begins with what it ends. If the purpose of the exercise is not clear or if the use of the priorities once established is unknown, one should not even attempt to proceed. It had been too often that an entire evaluation effort is wasted because there is a lack of purpose or use of the product.

Step 2: Identify the management parameters.

The choice of the management parameters to be considered in the evaluation is dependent on the purpose of the evaluation and the nature of the health care programs under consideration. In general, the management of a health care delivery system is concerned with three broad areas: the need of the programs, the performance of the system, and the financial return of the programs.

The need of a program can be determined from different points of view and in many different ways. From the management standpoint, the need

of a program is the measurement of service that is useful and required to serve the target population. This is sometimes termed the true need of the program. The need determined in this way may not be the same as the perceived need of the consumers. The latter is usually expressed as the demand of the program and is measured by the actual utilization rate. It is advisable that both the true need and the perceived need be considered for management decisions.

The performance of a system consists of many parameters. At the minimum, the productivity of the staff, the adequacy of the program in meeting the need, and the quality of the care are the basic ones necessary for health care program management.

The third area of concern is the financial status of the programs. The management of public programs is equally concerned with the financial return of the programs as the private sector, although the ways in which the programs are financed may be quite different.

Very often, the three areas of concern are evaluated independently. It is the intention of this approach to integrate them into the framework from which administrative decisions can be derived.

Step 3: Identify the scope of the evaluation.

The scope of the evaluation in terms of the geographic areas, the type of health programs, and the time period, should be defined at the outset. The choice of geographic areas depends on the level of management, the administrative jurisdiction of the manager of each geographic area, and the manner in which the resources will be allocated. Similar considerations are to be given to selecting or including health programs in the evaluation. They should be in the smallest unit that is administratively meaningful. The time period of evaluation can be retrospective or prospective or both depending on the availability of data.

Step 4: Identify the indicators of each parameter.

Quantitative measurements of the management parameters must be selected for each health program. They should be objective, meaningful, valid, and practical.

The need and demand indicators should be specific to the target population served by the program and to the type of services it provides. The type of indicators can be measures of health status or health system depending on the nature of the program and the type of agency. Utilization rates can be used as demand indicators.

Productivity index should also be specific for each program. The productivity index is generally defined as the ratio of the actual to the expected work load of the program based on a standard staffing pattern. Different expectations for different staffing patterns for each program must be taken into consideration in deriving productivity ratios. Adequacy of a program can be measured as the difference between the demand and supply in terms of visits, or unit of service. Or, it can be measured using proxy indicators such as backlog work load, the number of patients turned away or referred out, the waiting time for the first appointment. Quality of a program is an important parameter but extremely controversial and difficult to measure. The true quality of a program cannot be determined without an extensive evaluative effort which is both expensive and time consuming. While true quality indicators are being developed, broken appointment rates and patient satisfaction can be used as limited measures of quality of care. Since a broken appointment can be related to many other factors, it is advisable to use the broken appointment rate only for appointments given for follow-up care within two weeks from the first appointment. For clinics not requiring follow-up appointments, a simple patient satisfaction questionnaire administered at the end of the clinic visit can provide a limited measure of the quality of the program from the consumer's standpoint.

The financial status of a program is usually not a difficult parameter to measure, although the reimburable mechanism of public programs may be more complicated than the ones in the private sector.

Step 5: Formulate the decision matrix.

The formulation or determination of the criteria for decision making should be done before data are collected or analysed. The formulation can be approached in different ways. The quantitative value of each indicator can be used directly or through standardization or aggregation to arrive at a priority index. Multi-variate statistical analysis can be used to select the most usable parameters similar to the development of the IMU index. However, the results of these methods are frequently difficult to comprehend. The method introduced here employs a systematic and rational approach to derive management decisions for different combinations of levels of the parameter indicators. It avoids insensible and preposterous aggregation of different parameters.

The first step of the formulation is to redefine the indicators into high, moderate or low levels by establishing appropriate cut-points. Although finer breakdown of the levels can be used, two or three levels are usually adequate for management purposes. This step is a way of standardizing the indicators. The cut-points can be based on the actual value of the indicators or the ratio of the indicator to an average of the geographic areas, the county, the state, the nation, or a standard.

Once the levels are established, the different combinations of the levels of the parameters are derived. In order to avoid working with a large number of combinations at a time, it is advisable to divide the derivation of the combination into stages. First, formulate a transitional decision matrix for the level of each of the three general areas of concern: need, performance, and finance. For each area, derive an overall grading of high, moderate or low level based on the possible combinations of its parameters. The next stage involves formulating the final decision matrix based on the grading of the three areas. For each combination of levels of the areas, make a decision on the type of administrative action to be taken, such as discontinuing the program, improving the productivity, improving the quality of the program, etc. Although the number of combinations appears to be large, the actual number of decisive actions is usually manageable. This is because some common decisions

can usually be made irrespective of the levels of one or more parameters or areas. For example, for programs with high need but moderate or low performance, the decision might be to improve the performance regardless of the level of the financial status of the programs. This is equivalent to making a decision for six of the 27 combinations. The process of deciding on the type of action to take for each combination forces the management to think and to rationalize. This is a unique characteristic of this approach enhancing the relationship of management, providers and consumers.

Step 6: Actual derivation of the levels of indicators and the decision matrix.

Once Steps 1 through 5 are accomplished, the collection and analysis of data can proceed as outlined and defined. With the exception of quality indicators, most require data elements that are being collected in existing data systems including population, vital and health statistics, and management information. Caution is necessary, however, to ensure compatibility of the data elements from various sources or systems.

AN ACTUAL APPLICATION

The demonstration of the approach is for the management of the ambulatory care services of a region in Los Angeles County. The total population of the region was about 1.3 million and the target population, defined as the medically indigent population of the region, was estimated to be 300,000 or 24% of the total population. The region had five health districts, the smallest had a total population of 380,000 and a target population of 76,000. Each of the districts had been providing health care services to its target population. Fourteen different types of clinics were offered in the five districts. The methodology was developed to evaluate the existing need, performance and financial status of the different health programs in each district for the purpose of reallocating resources appropriately. A committee was formed to make decisions on the approach and the criteria to formulate the decision matrix.

The scope of the evaluation consisted of the five health districts, and fourteen health care programs. The management parameters selected included true need and demand for the need area; productivity, adequacy and quality for the performance area, and reimbursable status for the financial area. The indicators chosen are listed in Table 1. With the exception of the indicator for quality, all indicators utilize data that were being collected in the existing systems.

The cut-points used to define the indicators by high, moderate and low levels are presented in Table 2. They represented the decisions of the majority of the committee members. The transitional decision matrices for need and performance are presented in Table 3 and the final decision matrix in terms of actions is presented in Table 4. It should be noted that several common decisions were made in the formulation allowing a reduced number of possible situations requiring administrative actions. After the analysis of the data, the actual decision made for each program in each district is presented in Table 5.

DISCUSSION

It should be emphasized that we are presenting an approach and not a formular for deriving management decisions involving multiple programs and multiple locations. This is because a great deal of flexibility is needed with respect to the number of management parameters, the choice of indicators, the grading of levels, and the formulation of decisions.

As each health care delivery system is unique, it is advisable for management of each system to customize the approach according to its own need and purpose and to test the methodology thoroughly to ensure proper interpretation of the combination of the grading of the levels and the choice of indicators.

The choice of indicators is a very complex matter and has received much attention[3,4]. Each parameter deserves more research efforts in identifying better indicators for management purposes. As each field receives its proper attention and as each program is being evaluated, the findings of the evaluation tool for each program can be incorporated in the overall management decision making process.

In the demonstration, data on quality and adequacy indicators were not systematically collected. Therefore, although data were used, decisions based on the two parameters were not treated definitive. However, because of the need for the management approach, a new data collection system was established.

One should not be discouraged by the number of combinations of the levels of parameters. Most often, a number of situations result in similar decisions and the number of actions to be taken is usually not many. The process of thinking through each combination of levels of the parameters has the advantage of assisting the management to cope with all possible contingencies that may arise, in a systematic and rational manner.

REFERENCES

1. Health Services Research Group, University of Wisconsin: "Development of the Index of Medical Underservice", Health Services Research, Vol. 10, No. 2, Summer 1975, pp. 168-180.
2. Wysong JA: "The Index of Medical Underservice: Problems in Meaning, Measurement and Use", Health Services Research, Vol. 10, No. 2, Summer 1975, pp. 127-135.
3. Freeburg LC, Lave JR, Lave LB, Leinhardt S: "Health Status, Medical Care Utilization, and Outcome; An Annotated Bibliography of Empirical Studies Volume I", NCHSR Research Report Series, DHEW Publication No. (PHS) 80-3263, November 79, National Center for Health Services Research.
4. Hadley SA, Kinsman J, Janovsky AB, Kirlin MH: "Operational Measures for Health Systems Characteristics". Health Planning Methods and Technology Series, No. 17. DHEW Publication No. (HRA) 79-14032, U.S. Bureau of Health Planning, Division of Planned Assistance and Assessment.

TABLE 1: CHOICES OF INDICATORS

Management Parameter	Indicator	
Need	1. Ambulatory Care Clinic:	Percent of medically indigent population
	2. Senior Citizen Clinic:	Percent of population ages 65 and over
	3. Youth Clinic:	Percent of population 11-17 years old
	4. Pediatric Clinic:	Percent of population under 18 years old
	5. Immunization Clinic:	Preventable diseases case rate
	6. Screening Clinic:	Percent of population under 6 years old
	7. Prenatal Clinic:	Birth rate
	8. Family Planning Clinic:	Percent of population 15-44 years old
	9. Pregnancy Testing Clinic:	Percent of female population 15-44 years old
	10. Dental Clinic:	Percent of population under age 18 with decay, missing or filled teeth
	11. Alcohol Clinic:	Alcohol related death rate
	12. Drug Clinic:	Drug related death rate
	13. Tuberculosis Clinic:	Tuberculosis case rate
	14. Venereal Disease Clinic:	Venereal diseases case rate
Demand	Number of clinic visits per 1,000 target population by year. The target population of each clinic is defined by age, sex, and/or health problems as used in the definition of the indicators.	
Productivity	Ratio of actual number of patients seen to expected number based on recommended work load standards during year.	
Adequacy	For clinics requiring appointments, use the number of weeks backlogged for first appointment. For clinics not requiring appointments, use the number of patients turned away or referred out.	
Quality	For clinics requiring follow-up visits, use the broken appointment rate for follow-up visits scheduled within two weeks.	
Financial Status	Percent of visits reimbursable by third party payers.	

TABLE 2: CUT-POINTS USED TO DEFINE INDICATORS BY HIGH, MODERATE AND LOW LEVELS

Parameter	Level	Definition
Need, Demand	High	Ratio to County Average >1.5
	Moderate	Ratio to County Average 1.1-1.5
	Low	Ratio to County Average <1.1
Productivity	High	Productivity Ratio >0.90
	Moderate	Productivity Ratio 0.70-0.90
	Low	Productivity Ratio <0.70
Adequacy	High	Backlog for first appointment under 2 weeks or no patients turned away
	Moderate	Backlog for first appointment 2-4 weeks or patients are turned away in less than 50% clinic sessions
	Low	Backlog for first appointment over 4 weeks or patients are turned away in 50% or more clinic sessions
Quality	High	Under 2 weeks broken appointment rate <20%
	Moderate	Under 2 weeks broken appointment rate 20-40%
	Low	Under 2 weeks broken appointment rate >40%
Financial Status	High	Percent reimbursable visits >75%
	Moderate	Percent reimbursable visits 25-75%
	Low	Percent reimbursable visits <25%

TABLE 3: TRANSITIONAL DECISION MATRICES

(A) Transitional Decision Matrix for Need Levels

Number of Combinations	True Need Level	Demand Level	Decision for Need Level
1.	High	High	High
2.	High	Moderate	Moderate
3.	Moderate	High	Moderate
4.	Moderate	Moderate	Moderate
5.	Low	High/Moderate/Low	Low
6.	High/Moderate/Low	Low	Low

(B) Transitional Decision Matrix for Performance Levels

Number of Combinations	Productivity Level	Adequacy Level	Quality Level	Decision for Performance Level
1.	High	High	High	High
2.	High/Moderate	High/Moderate	Moderate	Moderate
3.	High/Moderate	Moderate	High/Moderate	Moderate
4.	Moderate	High/Moderate	High/Moderate	Moderate
5.	Low	High/Moderate/Low	High/Moderate/Low	Low
6.	High/Moderate/Low	Low	High/Moderate/Low	Low
7.	High/Moderate/Low	High/Moderate/Low	Low	Low

TABLE 4: THE FINAL DECISION MATRIX

Number of Combinations	Need Level	Performance Level	Financial Status Level	Decision on Administrative Action
1.	High/Moderate	High	High/Moderate/Low	Keep
2.	High/Moderate	Moderate/Low	High/Moderate/Low	Improve Performance
3.	Low	High	High/Moderate	Incorporate in Other Clinics
4.	Low	High	Low	Discontinue Program, Relocate Staff
5.	Low	Moderate/Low	High	Incorporate in Other Clinic
6.	Low	Moderate/Low	Low	Discontinue

TABLE 5: PROPOSED MODIFICATIONS OF SERVICES BASED ON NEED, DEMAND, PRODUCTIVITY, ADEQUACY, QUALITY AND FINANCIAL STATUS

CLINICS	DISTRICT A	DISTRICT B	DISTRICT C	DISTRICT D	DISTRICT E
AMBULATORY CARE	KEEP	KEEP	KEEP	KEEP	KEEP
SENIOR CITIZENS	KEEP	KEEP	KEEP	DISCONTINUE	DISCONTINUE
YOUTH HEALTH	DISCONTINUE	KEEP	DISCONTINUE	KEEP	KEEP
PEDIATRICS	IMPROVE ADEQUACY AND PRODUCTIVITY	IMPROVE ADEQUACY AND PRODUCTIVITY	IMPROVE ADEQUACY AND PRODUCTIVITY	IMPROVE ADEQUACY AND PRODUCTIVITY	IMPROVE PRODUCTIVITY
IMMUNIZATION	KEEP	KEEP	KEEP	KEEP	DISCONTINUE
SCREENING	INCORPORATE IN OTHER PROGRAMS	INCORPORATE IN OTHER PROGRAMS	INCORPORATE IN OTHER PROGRAMS	INCORPORATE IN OTHER PROGRAMS	INCORPORATE IN OTHER PROGRAMS
PRENATAL	IMPROVE ADEQUACY AND PRODUCTIVITY	IMPROVE ADEQUACY AND PRODUCTIVITY	IMPROVE ADEQUACY AND PRODUCTIVITY	IMPROVE ADEQUACY AND PRODUCTIVITY	IMPROVE PRODUCTIVITY
FAMILY PLANNING	IMPROVE ADEQUACY AND PRODUCTIVITY	IMPROVE ADEQUACY AND PRODUCTIVITY	IMPROVE ADEQUACY AND PRODUCTIVITY	IMPROVE ADEQUACY AND PRODUCTIVITY	IMPROVE ADEQUACY PRODUCTIVITY
PREGNANCY TESTING	INCORPORATE IN OTHER PROGRAMS	KEEP	INCORPORATE IN OTHER PROGRAMS	KEEP	KEEP
DENTAL	IMPROVE ADEQUACY	IMPROVE ADEQUACY	IMPROVE ADEQUACY	IMPROVE ADEQUACY	IMPROVE ADEQUACY
TUBERCULOSIS	IMPROVE ADEQUACY	KEEP	IMPROVE ADEQUACY AND PRODUCTIVITY	IMPROVE ADEQUACY AND PRODUCTIVITY	IMPROVE ADEQUACY AND PRODUCTIVITY
VENERAL DISEASE	IMPROVE PRODUCTIVITY	DISCONTINUE AND RELOCATE STAFF	KEEP	INCORPORATE IN OTHER PROGRAMS	IMPROVE PRODUCTIVITY
ALCOHOL REHABILITATION	KEEP	KEEP	KEEP	KEEP	KEEP
DRUG	KEEP	KEEP	KEEP	KEEP	KEEP

A COMPREHENSIVE MANAGEMENT INFORMATION SYSTEM FOR A LOCAL HEALTH DEPARTMENT

LINDSON FEUN, OAKLAND COUNTY HEALTH DIVISION

PRESENT DATA PROCESSING SYSTEM

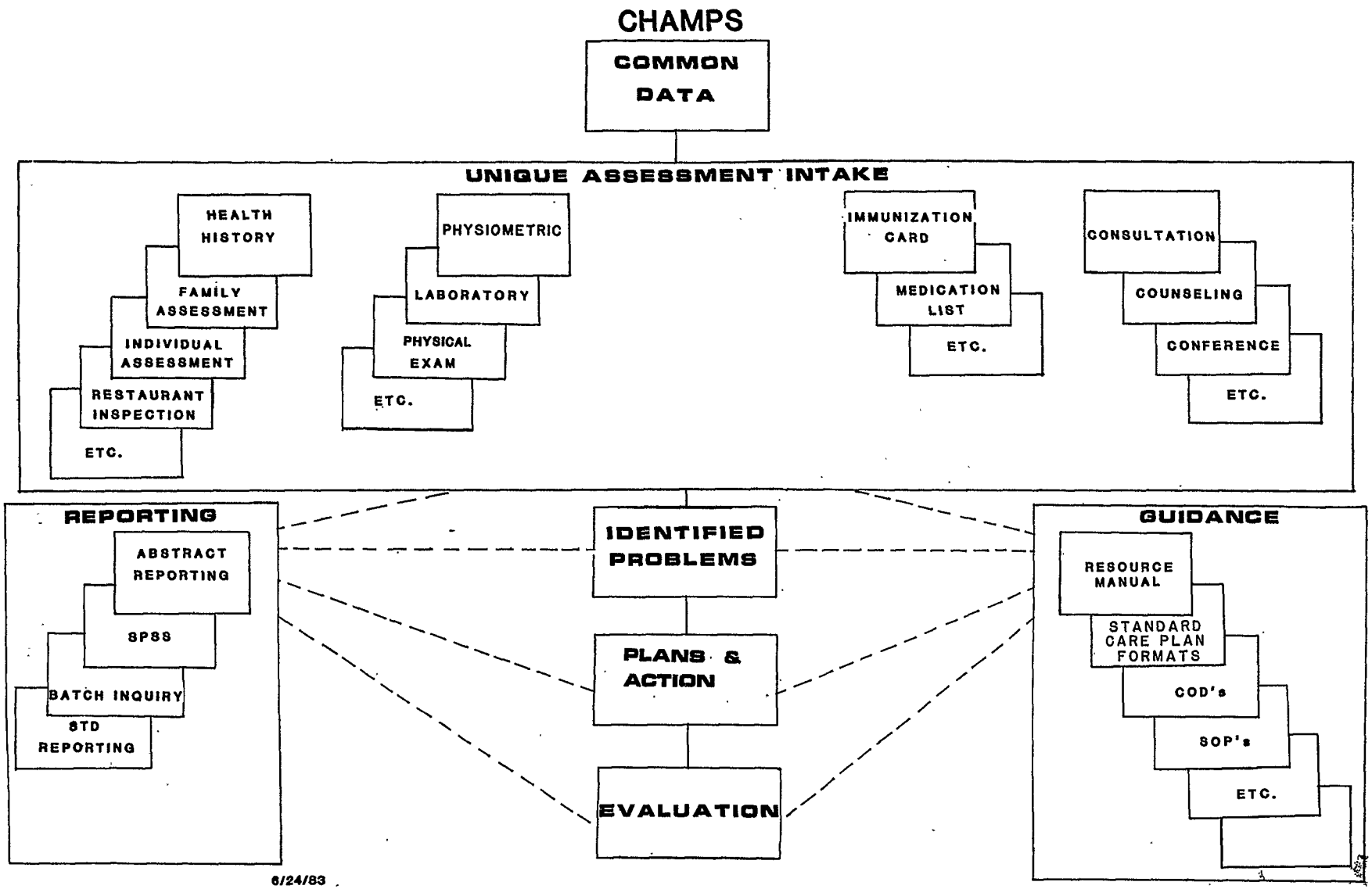
- SEPARATE BATCH MODE SYSTEMS FOR PPHS AND EHS
- COLLECTS PRIMARILY PROFESSIONAL ACTIVITIES

PROBLEMS WITH PRESENT SYSTEM

- INABILITY TO ACCURATELY COUNT SERVICES ACROSS UNITS
- INABILITY TO ASSESS OUTCOMES
- APPROXIMATELY 1,400 FORMS - PAPER BOUND
- DUPLICATION OF EFFORT BY BOTH PROFESSIONAL AND CLERICAL STAFF
- INSUFFICIENT INFORMATION FOR MANAGEMENT PURPOSES
- IN SUMMARY, LACK INFORMATION THAT IS TIMELY, ACCURATE AND COMPREHENSIVE

BASIC CHARACTERISTICS OF CHAMPS

1. CLIENT CENTERED
2. PROBLEM-ORIENTED PROCESS
3. COMPREHENSIVE - CAN BE USED BY ALL UNITS/PROGRAMS
4. MANAGEMENT CAPABILITIES
5. INTEGRATED DATA BASE - CAN BE USED BY ALL DISCIPLINES
6. MODULAR IN NATURE
7. CENTRAL COMPUTER PROTOTYPE - ON-LINE CAPABILITIES


CHAMPS
COMMON DATA
UNIQUE ASSESSMENT INTAKE
HEALTH HISTORY
FAMILY ASSESSMENT
INDIVIDUAL ASSESSMENT
RESTAURANT INSPECTION
ETC.
PHYSIOMETRIC
LABORATORY
PHYSICAL EXAM
ETC.
IMMUNIZATION CARD
MEDICATION LIST
ETC.
CONSULTATION
COUNSELING
CONFERENCE
ETC.
REPORTING
ABSTRACT REPORTING
SPSS
BATCH INQUIRY
STD REPORTING
IDENTIFIED PROBLEMS
PLANS & ACTION
EVALUATION
GUIDANCE
RESOURCE MANUAL
STANDARD CARE PLAN FORMATS
COD's
SOP's
ETC.
6/24/83

```
                OAKLAND COUNTY HEALTH DIVISION    DATE 080983  TIME 1234
              C H A M P S -- SUBSYSTEM SELECTION    HSCC70DX  LUHDNM
          SUBSYSTEM                                 CODE
          COMMON INTAKE MAINTENANCE                  A
          COMMON INTAKE DISPLAY                      B
          COMMON INTAKE MANAGEMENT FUNCTIONS         C


 ENTER SELECTION        --PASSWORD--        CHAMPS ID NBR
```

```
                OAKLAND COUNTY HEALTH DIVISION    DATE 080983  TIME 1235
              C H A M P S  COMMON INTAKE MAINTENANCE   HSCC71DX
          FUNCTION          FUNCTION DESCRIPTION
             A          NAME SEARCH
             B          CROSS REFERENCE SEARCH
             C          MAINTENANCE FAMILY RELATIONSHIPS
             D          MAINTENANCE CLT PRIM DEMO DATA
             E          MAINTENANCE SECONDARY DEMO DATA
             F          MAINTENANCE UNIT(S)
             G          MAINTENANCE EVENT(S) BY UNIT
             H          MAINTENANCE CROSS REFERENCE(S)
             I          MAINTENANCE RELATIONSHIP TYPES
             J          DISPLAY EVENTS CHRONOLOGICAL
             K          MAINTENANCE AKA/ROSTER
FUNCTION   CLIENT #   CLIENT TYPE   UNIT #   UNIT DATE   JURIS ID   X-REF ID
           000000000                          00000     000000
HELLO, L   FEUN, PHD              WELCOME TO THE CHAMPS SYSTEM!
```

UNIQUE INTAKE

CLIENT NAME: MARTHA JOHNSON CLIENT # 12345

LEDGER	FAMILY HISTORY	COMMENTS	DATE 9-10-82 PROF. #122	DATE PROF. #	DATE PROF. #
CHC/FN	FATHER'S FAMILY HISTORY				
	ALLERGY (ASTHMA)		1		
	ANEMIA		2		
	CANCER		1		
	CARDIOVASCULAR DISEASE		1		
	DEAFNESS		1		
	DIABETES		2		
	HYPERTENSION		3		
	KIDNEY DISEASE		2		
	SEIZURES		1		
	TUBERCULOSIS		1		
	OTHER		1		
CHC/FN	MOTHER'S FAMILY HISTORY				
	ALLERGY (ASTHMA)		1		
	ANEMIA		2		
	CANCER		1		
	CARDIOVASCULAR DISEASE		3		
	DEAFNESS		1		
	DIABETES		1		
	HYPERTENSION		2		
	KIDNEY DISEASE		2		
	SEIZURES		1		
	TUBERCULOSIS		2		
	OTHER		1		

1 - NORMAL 2 - FURTHER ASSESSMENT 3 - PROBLEM - ID LIST

8/17/82

OCHD - PROBLEM LIST

CLIENT NAME:_____ CLIENT ID #:_____

DATE	PROBLEM CODE	PROBLEM/NEED/VIOLATION	UNIT	PROGRAM	STAFF	STATUS*	STATUS DATE
2-9-82	59	NUTRITIONAL RISK	FN CHC WIC	0012 0012 0012			
3-10-82	01	SOURCE: SOUND CONDITION, NO SPOILAGE	EH	0070			
3-10-82	02	ORIGINAL CONTAINER, PROPERLY LABELED	EH	0070			
4-11-82	30	FAMILY PLANNING	FN	0029			

PPHS/HENS

ACTIVE PROBLEM LIST

CLIENT NAME: MARTHA SMITH CLIENT ID #: 123456

PROBLEM/ NEED: FAMILY PLANNING (30)

PROF. DIAGNOSIS: LACKS UNDERSTANDING OF BC METHODS (30-1)

#	OBJECTIVE (EXPECTED OUTCOME)	TIME-LINE	RESOLUTION	INTERVENTION	DATE STAFF
30-1-1	VERBALIZES UNDERSTANDING OF BC METHODS	1ST VISIT (DATE)		USING FP DEMO KIT AND APPROPRIATE FORMS, DISCUSS AVAILABLE BC METHODS EXPLAINING SIDE EFFECTS AND THE CONTRA EFFECTS.	

PPHS/HENS

EVALUATION (SUMMARY)

CLIENT NAME:_____ CLIENT ID #:_____

DATE:_____ OCHD STAFF:_____

EXPECTED OUTCOMES	RESOLUTION	COMMENTS

Monitoring and Evaluating Health Programs: Methodology

Session X

SURVEY MEASUREMENTS OF EMOTIONAL WELL-BEING: THE HEALTH OPINION SURVEY AND THE BRADBURN SCALE COMPARED

Ian McDowell, University of Ottawa
Ed Praught, Statistics Canada

INTRODUCTION

Those who conduct community health surveys are confronted with a problem in selecting a method to measure general psychological well-being. Indicators of "general well-being" are normally intended both to screen for low levels of emotional distress (which may or may not benefit from formal psychiatric treatment), and also to reflect more general feelings of happiness or unhappiness. Not only are there relatively few such indicators, but those that are available have been criticised on several grounds.

Most of the measurement scales are old: Bradburn's questions were first published in 1965, Macmillan's in 1957 and Langner's twenty-two item scale was published in 1962; the two latter scales are based on questions developed in the Second World War. A few more recent measurement scales do exist, but these have not replaced the older instruments, which continue to be used.

Macmillan's Health Opinion Survey and Bradburn's Affect Balance Scale illustrate contrasting approaches to assessing emotional well-being in the community. Bradburn's questions ask about subjective feelings, positive and negative, in response to daily life. Indices such as the HOS or Langner's scale record physical and behavioral symptoms of emotional distress.

Strengths and weaknesses have been noted in both approaches, but they have been evaluated in rather different ways, and few studies compare the two. To determine which approach may prove most valuable for a survey instrument, these criticisms will be briefly documented.

Reviews of the subjective scales have mainly concentrated on their psychometric properties. while relatively little attention has been paid to the practical use of the Bradburn scale as a case-detection method or as an indicator of the need for care.

By contrast, and following logically from their original purpose, most discussions of the symptom check-lists concern their validity as screening tests. The HOS appears to perform well in its original purpose as a case-finding method for use in community surveys. The possibility that the inclusion of physical signs of emotional disorder in the HOS will falsely classify physical illness as mental distress had been widely discussed (see Wells & Strickland, 1982). At the same time less attention has been paid to the construct validity of this approach: there is indeed considerable disagreement over precisely what the symptom check-lists measure. Similarly, while Macmillan intended the Health Opinion Survey to screen for "Psychoneurotic and related types of disorder", others have subsequently argued that it is better viewed as an indicator of low level, short term distress than of psychiatric illness. This interpretation of the HOS would appear to bring it conceptually close to the Bradburn scale, supporting the relevance of assessing whether the two methods do, indeed, provide equivalent information, and which approach is preferable for population surveys.

METHODS OF STUDY

The data reported here were drawn from the 1978-79 Canada Health Survey. This was sponsored by National Health and Welfare Canada (1981) and by Statistics Canada and provided nationally representative data from a sample of approximately 12,000 households.

The survey examined health risk factors, current health status and consequences such as disability days and the use of health services. Data were collected through an interview, a questionnaire and a set of physical measurements. The interview had an 86% response rate and collected information on the health problems and use of medical care of each person in the household. Eighty-seven percent of those 15 years and older in the participating households then completed the questionnaire themselves: a total of 20,726 respondents. The Bradburn and the HOS questions were included in the questionnaire component of the survey, but the HOS was added in October 1978, five months after field work began.

The set of HOS questions were answered by 13,111 respondents, and these responses are used in the following analyses of the HOS. However, not all of these respondents answered all of the Bradburn questions, so that analyses comparing the two scales employ the 11,645 respondents who answered all the questions on both sides.

Sixteen HOS questions were used in the survey. Three response choices were used for the Bradburn and HOS questions. A single, overall score was used for the HOS. The Bradburn questions form three scales: indications of negative and of positive affect, and a single score summarizing the answers to all ten questions, calculated as the arithmetic difference between positive and negative scores. This is termed the Affect Balance Scale.

ASSESSING HEALTH STATUS: The interview section of the survey recorded data on health status and the use of medical services during the previous year. The data are based on self report only.

Health problems were coded using the ICD9 system and multiple problems could be coded for each individual. For the analyses which follow, respondents were grouped into three main categories: those reporting no health problems; those reporting one or more physical (but no emotional) problems, and those reporting one or more emotional problems. Because of possible reticence in self-report of emotional problems, those who in the interview stated that they had

no emotional problems, but that in the past two days they had used "tranquilizers or medicine for the nerves or medicine to help you sleep" were added to the group reporting emotional problems.

In all, the emotional problem group comprised 6.5% of the sample, and in certain analyses the group is sub-divided, as follows:

neurotic and depressed	:	1.1% of respondents
nervousness	:	2.1% of respondents
insomnia	:	1.0% of respondents
other mental disorders	:	0.4% of respondents
tranquilizer users	:	2.0% of respondents

RESULTS

DISCRIMINAL ABILITY OF THE SCALES

The original purpose of the HOS was to screen for "psychoneurotic and related types of disorder". To what extent was it able to do this? Did it, indeed, confound physical with emotional symptoms, and how did it compare as a screening test with the Bradburn scale?

Considering the classification of physical illness, the HOS did misclassify slightly more physical problems than did the Bradburn (22.4% versus 21.2%).

PSYCHOMETRIC PROPERTIES OF THE SCALES

Intercorrelations among the HOS items showed positive but often low associations (Kendall coefficients ranged from .06 to .37), although each item contributed modestly to the overall score (item-total tau b coefficients from .32 to .49). Correlations among the Bradburn items were higher (ignoring the sign of the coefficient), running from .17 to .45, with item-total correlations between .49 and .63.

The comparatively low intercorrelations among the HOS items may suggest that they do not measure the same qualities: some may therefore not contribute to the purpose of screening for emotional disorders. Would discarding those items responsible for the physical bias in the HOS improve the scale? Wells and Strickland (1982) and Crandell and Dohrenwend (1967) obtained clinical judgements on how far each item

Table 1

Percentages of Self-Reported Diagnostic Groups Correctly Classified by the Health Opinion Survey (HOS) and Affect Balance Scale (ABS)

Canada Health Survey, 1978-79, N=11,645

	Percentages correctly classified by HOS	ABS	(N)
Percentage of Group Classified as 'Sick'			
Neurotic and depressive disorders	66.7	59.6	118
Insomnia	48.1	35.7	92
Other mental disorders	43.5	29.6	42
Respondents using tranquilizers	59.7	38.3	257
Nervousness	58.3	45.4	293
Percentage of Group Classified as 'Well'			
No health problems	89.8	83.3	4786
Physical health problems	77.6	78.8	5726
Other symptoms, signs and ill-defined conditions (ICD9 Chapter XVI except codes 780.5 and 799.2)	62.1	71.6	331

From cross-tabulation analyses table 1 shows the percentages of respondents in each illness category correctly classified by the two scales. Using a cutting point of 26 or greater - equivalent to that recommended by Tousignant et al. - 66.7% of the neurotic and depressed group were correctly classified by the HOS, as were 89.8% of the well group. The equivalent figures for Bradburn's ABS, at 59.6% and 83.3%, were somewhat lower, using a cutting point for the ABS of zero or below for the sick category. The HOS was more successful in classifying insomnia and nervousness as emotional disorders than was the ABS; neither scale identified the "other mental disorders".

identified physical complaints. In the main, their results corresponded with the results of factor analyses reported by Butler and Jones (1979) and others which were interpreted as distinguishing physical from psychological questions.

The sixteen items in the present study were factor analysed using the PA2 factoring method, working from a Pearson correlation matrix. Communality estimates for the main diagonal of the correlation matrix were begun at unity and iterated to stability. This produced a three factor solution with a strong common factor containing twelve items. Two questions

concerning sleep loaded on the second factor, and two covering general health and work capacity on the third. In an attempt to replicate the conceptual division into psychogenic and physiogenic questions empirically, a two factor solution was enforced. The clearest results were obtained with the oblique rotation shown in Table 2. The results resemble those from other studies and the variables on the second factor correspond to those judged by Wells and Strickland as most likely to be physiogenic.

Table 2

Principal Components Analysis
of Health Opinion Survey Questions in Canada Health Survey
1978-79, N=13,111

(Factor pattern matrix based on an oblique rotation of a two factor solution; delta = 0.0. Correlation between factors - 0.57)

Question	Factor 1	Factor 2
Loss of appetite	0.563	0.068
Lose weight	0.491	0.102
Tired in morning	0.471	-0.039
Nightmares	0.457	-0.103
Trouble getting asleep	0.449	-0.111
Nervous breakdown	0.435	-0.155
Hands or feet sweat	0.352	-0.010
Hands tremble	0.351	-0.165
Upset stomach	0.331	-0.112
Amount of work	0.065	-0.585
Short of breath	0.129	-0.491
Feel healthy enough	-0.119	-0.489
Weak all over	0.271	-0.449
All sorts of ailments	0.213	-0.431
Heart beating hard	0.186	-0.369
Dizziness	0.307	-0.331

However, several considerations counsel against attempting to improve the HOS by eliminating the second factor. First, it seems debatable whether the two factors should be interpreted as psychological and physical: the first includes specific symptoms of distress, while the several variables on the second reflect more generalized responses.

Second, when the responses of sub-groups of respondents reporting physical health problems were compared, acute conditions, such as upper respiratory tract infections, scored low on the "physical" factor and high on the "psychological". Considering the questions included in the first factor, it is reasonable that someone with a cold or influenza will report an upset stomach, loss of appetite and tiredness. Finally, cross tabulations showed that every question in the HOS was successful in discriminating between psychological and physical problems. Thus, although the internal consistency of the HOS is not high, it appears that its performance will not be markedly improved by discarding selected questions.

A COMPARISON OF THE HOS AND THE BRADBURN SCALES

Several commentators have viewed the HOS as an indicator of short-term distress; as such, it may duplicate the information provided by the Bradburn scale making it unnecessary to include both in a survey. If the two scales do measure the same construct, they would be expected to intercorrelate highly and to show similar patterns of correlations with other variables. The results of several analyses failed to support this assumption.

A Kendall correlation of -.14 was obtained between the HOS and the positive affect score, indicating that they share very little common variance. The HOS was more closely, but still not strongly, associated with the negative affect score (tau b = .35). The low overlap between the scales was confirmed when the Bradburn items were pooled with the HOS items in a factor analysis, and questions from the two scales fell on separate factors. Table 3 shows the association of the two scales with selected health indicators. The HOS correlated more strongly than did the Bradburn questions with numbers of health problems and disability days reported.

We conclude that the two scales do not measure the same psychological state; but examining the interrelationship between them further leads to the question of whether there is a logical pattern to the association between the two types of response, the verbal and the somatic.

Table 3

Kendall Tau b Correlations Between Health Opinion Survey Scores, Bradburn Scores and Selected Criterion Variables

Canada Health Survey, 1978-79, N=11,645

	Health Opinion Survey	Bradburn Affect Scores		
Health Effects:		Positive	Negative	ABS
Number of Health Problems	0.27	-0.07	0.09	-0.11
Disability Days (Previous two weeks)	0.16	-0.05	0.07	-0.08

It could be hypothesized that distress will be expressed verbally before being translated into somatic symptoms recorded on the HOS, perhaps where the cause of the distress was not alleviated. Alternatively, certain types of people may express an emotional reaction somatically while others express their reaction verbally. The hypothesis that responses fall into such a sequence would imply that few respondents would be classified as well by the Bradburn questions if they were sick according to the HOS. By contrast, an alternative hypothesis would accept this as a possibility, regarding it as equally likely to occur as the alternative disagreement between the two scales.

Tabulations were drawn between "well" and "sick" respondents as identified by the HOS and by the negative affect questions. Tousignant's (1974) cutting point was used for the HOS, and a cutting point was selected for the Bradburn questions which gave similar marginal distributions of well and sick respondents. Comparing observed cell frequencies with those expected from the marginal distributions showed that fewer than half of the cases expected were observed in the cell defined as sick by the HOS but well by the Bradburn questions. By contrast over 80% of the expected numbers fell into the other disagreement cell. The data lend more support to the first than to the second hypothesis.

DISCUSSION

Debate has long continued over what the scales in fact measure. The present data show tnat several types of mental disorder were identified by both scales: they do not reflect one specific type of disorder. The issue of whether a high score represents "disorder" or milder "distress" is probably best answered by reference to the questions themselves: some do appear to reflect treatable disorders ("depressed or very unhappy", "bothered by shortness of breath"), while others ("bored", or "do you tend to feel tired in the mornings?") reflect lower levels of short-term disturbance. Goldberg's (1972) term "non-psychotic emotional disorders" may describe the former class of question, while Dohrenwend's phrase "nonspecific distress" (1980) aptly describes the latter. Both scales offer broad, screening devices which indicate that something is wrong, but they do not specify what. From the correlations shown in table 3 and from the pattern of associations between the scales, we infer that the HOS questions in general reflect clinical levels of disorder, whereas the Bradburn questions indicate lower levels of distress.

While the Canada Health Survey was not ideal for assessing a possible physical bias in the HOS, the data presented suggest that this was no worse than that shown by the Bradburn scale (which contains no items describing physical symptoms). Attempts to reduce this physical bias by discarding questions will not be successful. These findings concur with those of Wells and Strickland: most such questions also identify emotional illness and every question identified a higher proportion of emotionally than of physically sick.

Several contrasts between the two types of scale are noteworthy. The structure of the Bradburn scale is clear and conceptually appealing, while the HOS showed a lower internal consistency - perhaps a characteristic of symptom check-lists because the various symptoms may represent alternative responses to stress. While this produces an imperfect scale in psychometric terms, the scale is successful as a screening tool. Indeed, in multivariate terms, the structure is appropriate: each item is associated with the dependent variable, and yet associations among the items are low. The drawback is that, lacking in internal consistency, the meaning of a high score is unclear. This is what the critics of the HOS have long argued, and what Dohrenwend et al. expressed so clearly: the symptom check-lists at best indicate non-specific disorder.

Finally the two scales appear to describe overlapping but somewhat different stages in the response to stress. This may reflect a hierarchical concept of emotional illness such as proposed by Foulds and Bedford (1975).

In conclusion, if the purpose is to screen for non-specific illness which may be of clinical importance, the symptom check-list approach appears the more suitable. To reflect positive well-being and milder levels of distress, questions such as Bradburn's are superior.

REFERENCES

Bradburn, N.M. and D. Caplovitz (1965): Reports on happiness: a pilot study of behavior related to mental health. Chicago: Aldine.

Butler, M.C. and A.P. Jones (1979): The Health Opinion Survey reconsidered: dimensionality, reliability and validity. J Clin Psychol 35:554-9.

Crandell, D.L. and B.P. Dohrenwend (1967): Some relations among psychiatric symptoms, organic illness, and social class. Am J Psychiatry 123:1527-37.

Dohrenwend, B.P., P.E. Shrout, G. Egri and F.S. Mendelsohn (1980): Nonspecific psychological distress and other dimensions of psychopathology: measures for use in the general population. Arch Gen Psychiatry 37:1229-36.

Foulds, G.A. and A. Bedford (1975): Hierarchy of classes of personal illness. Psychol Med 5: 181-92.

Goldberg, D.P. (1972): The detection of psychiatric illness by questionnaire. London: Oxford University Press.

Langner, T.S. (1962): A twenty-two item screening score of psychiatric symptoms indicating impairment. J Health Human Behav 3:269-76.

Macmillan, A.M. (1957): The Health Opinion Survey: technique for estimating prevalence of psychoneurotic and related types of disorder in communities. Psychol Rep 3:325-39.

McDowell, I. and E. Praught (1982): On the measurement of happiness: an examination of the Bradburn Scale in the Canada Health Survey. Am J Epidemiol 116:949-58.

National Health and Welfare Canada (1981): The Health of Canadians: Report of the Canada Health Survey. Ottawa, Ontario: Ministry of Supply and Services. (Catalogue No. 82-538E).

Tousignant, M., G. Denis and R. Lachapelle (1974): Some considerations concerning the validity and use of the Health Opinion Survey. J Health Soc Behav 15(3):241-52.

Wells, J.A. and D.E. Strickland (1982): Physiogenic bias as invalidity in psychiatric symptom scales. J Health Soc Behav 23:235-52.

INDICATORS OF NEED FOR CANCER -SPECIFIC SCREENING

Bonnie Morel Fdington, New Jersey Department of Health

Background

This methodology was developed initially for one county in New Jersey. The State has some large, county-wide local health departments (LHDs) but most are small. This county, for example, has eight LHDs seeking to cooperate and coordinate their cancer screening efforts.

The New Jersey Department of Health has a relatively recent regulation requiring that all LHDs screen for colo-rectal, breast and cervical cancer. This county also chose to screen for oral cancer. Since not all the LHDs were enthusiastic about a regulation requiring cancer screening, the State is letting them become involved conjointly in setting goals for subsequent evaluation of their efforts.

The State has published Guidelines giving LHDs some idea of what kinds of people they should be screening:

One risk factor is given for colo-rectal cancer -- age.

Eleven risk factors are given for breast cancer -- age, and 10 factors whose existence in a population would not be known without a special survey, e.g., family history of breast cancer; first pregnancy after age 34; exposure to ionizing radiation; use of oral contraceptives.

Five risk factors are given for cervical cancer -- age; socio-economic status; sexual relations having begun at an early age; pregnancy at an early age; history of venereal disease. Clearly, some of these factors are unlikely to be disclosed in a special survey, and may also be difficult to elicit from people after they have come in for services.

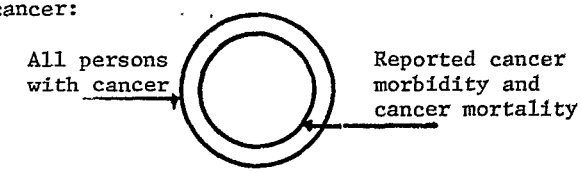
Thus, the LHDs needed to have something beyond the Guidelines to give them an idea of what the relative need for screening was in each area, and to permit their relative success in screening to be evaluated. The purpose of the methodology was to refine and expand the State's Guidelines, using readily available data on indicators, in a manner that small agencies could understand and apply themselves, with pocket calculators and without statistically sophisticated personnel.

The methodology uses: published Census data (although it is made clear that Census tapes can be used for more refinement of these indicators if an agency has the resources; vital statistics; unemployment statistics; and either hospital discharge abstract data or cancer registry data. New Jersey's State Department of Health receives, by regulation, hospital discharge abstract data, including a standardized residence code, on every patient in every New Jersey hospital (more than one million abstracts a year). Also, in New Jersey, cancer is a reportable disease, thus the Cancer Registry is complete and required by State law.

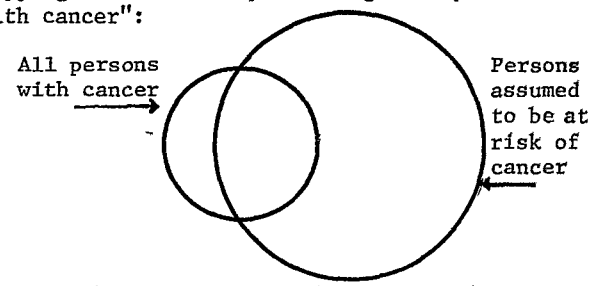
Underlying Principles

Because the methodology was developed for the LHDs, the detailed paper describing it assumed little or no knowledge about indicators and how to use and weight them, and provided a discussion of underlying concepts and principles and some Venn diagrams illustrating them, for example:

Persons with reported cancer morbidity and mortality are a subset of all persons with cancer:



Persons assumed to be at risk of cancer because they have certain risk characteristics are a much larger group, only partially overlapping the circle representing "all persons with cancer":



Even when the research literature gives statistics on risk, it usually reports this in terms of relative risk, rather than prevalence within risk. We may be told that the population group with the characteristic has one or two times the risk of the population without that characteristic, rather than being told how great the risk is within the riskier group -- whether one out of 10, or one out of 1,000 would be expected to have the disease. And we are sometimes told that two or three characteristics are indicators of risk, but do not know what the effect is when they are combined in the same persons -- is the effect less than additive, additive or interactive? Frequently, even when we know the interactive effect, we do not have routine access to data showing the number of individuals with the combined characteristics and may have to use the indicators separately. For example, data may be available on the number of people in an area who are in an ethnic group, in an age group, and in a socio-economic class, but not on the number of people who have all three characteristics. Again, using Census tapes would permit a more refined aggregation of persons with combined risk characteristics, and a refinement of this methodology.

Risk indicators tell us something about gross or unmet need, but these indicators do not decrease when needs are met. That is, two areas could have exactly the same number of persons with exactly the same risk characteristics, such as age and race, even though everbody in one area had received cancer screening services and nobody in the other area had received any. On

the other hand, disease indicators, such as morbidity and mortality, are clearly indicative of some unmet need when they increase, and of needs met when they decrease. However, morbidity increases reported by health care providers -- i.e., hospital discharge abstract and cancer registry data -- which can suggest that some needs are unmet in terms of prevention and early detection, also reflect met need, since all of these cases are already in the health care system. We use this data as indicative of unmet need, understanding that we are presuming that it is indicative of prevalence and that if reported cases are high in number, unreported cases are high in number; rather than presuming that if reported cases are high in number this merely reflects needs being met and unreported cases being low. In general, we make a case for inflating -- or over-weighting -- disease indicators (morbidity and mortality) -- and for deflating or under-weighting indicators of risk.

In the methodology we make a distinction between epidemiological research, and planning for resource allocation or evaluation. Epidemiology looks at rates, and planning for resource allocation requires absolute numbers. That is, areas of greater population would be expected to have more cancers, in absolute numbers, so epidemiologists want to control for or wipe out that difference by looking at rates of disease per 1,000 or 10,000 population. In planning for resource allocation and subsequent evaluation, we do not want indexes that wipe out these distinctions. Two areas could have exactly the same rate of cancer, but if one area is twice as large as the other, it will have to screen twice as many persons, will need twice the resources, and must be evaluated on those terms. So we do not use percentages of persons with a characteristic, we use the number of persons that percentage reflects.

Selection and Weighting of Indicators

An Index of Need for Screening was compiled for each of the four anatomic sites. Each of the 21 towns was given a total score on this index, and the scores were added for the eight multi-town areas and the county as a whole.

Mortality and hospitalization for each anatomic site were used as indicators for each of the four indexes. Hospitalization data was used in lieu of the more preferable Cancer Registry data, since the latter was not available at the time this methodology was developed. The specific cancer as principal diagnosis was used, and the residence of patients regardless of location of hospital.

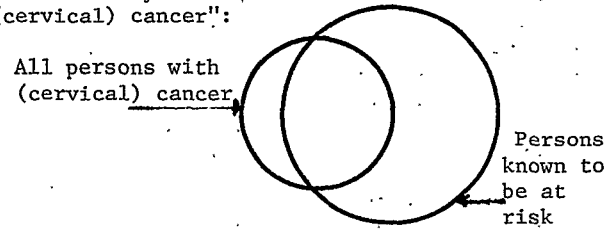
Population Census data was used in each of the four indexes. In the case of colo-rectal cancer, "adults"; in the case of breast cancer, "adult females", and "adult white females", since white women are at higher risk of breast cancer than those of other races; in the case or oral cancer, "adult males", since oral cancer is about three times more common among men than among women; in the case of cervical cancer, "adult females" and "adult black and Spanish females", since black and Spanish females are at higher risk of cervical cancer than other women. In every case, the number was divided by 1,000 to deflate the indicator. In essence, this deflation of the risk indicators weights the mortality and hospitalization figures by 1,000 since the entire number was used for those indicators. As the Venn diagrams and earlier narrative make clear, actual disease is considered to be a far stronger indicator of need in an area than the number of persons potentially at risk.

Socioeconomic Census data was used for breast and cervical cancer, since these cancers are correlated with such factors. In the case of breast cancer, "adult female high school graduates" and "adult females above poverty" were used; in the case of cervical cancer, "adult females less than high school graduates", "adult females in poverty" and "females in poverty heading families" were used. Again, the number was divided by 1,000 to deflate the indicator.

Originally, unemployment data was used in the cervical cancer index as an indicator of socioeconomic status, since people at lower socioeconomic levels are more likely to be unemployed However, subsequent reconsideration has led us to recommend that this indicator be used in all four indexes, not as an indicator of social status, but as an indicator of the relative need for LHD services. In areas where unemployment is high, people are more likely to have lost insurance coverage and be in need of the services of local health departments.

In addition, liver disease mortality was used as an indicator in the oral cancer index. Oral cancer is more common among alcoholics and heavy drinkers than among other persons and about half of liver disease mortality has been attributed to alcohol use, thus, half the liver disease mortality was used in the index and, since these are deaths, this indicator was not deflated.

Also, venereal disease and teenage births were used as indicators in the cervical cancer index. People who have had VD are more likely to have cervical cancer than other persons in the population. Early sexual activity is also correlated with cervical cancer. These figures were divided by 10 for use in the index. That is, they were deflated relative to mortality and hospitalization, but inflated relative to the other risk indicators. These are indicators of the population known to be at risk, rather than assumed to be at risk. Diagrammatically, persons known to be at risk is a smaller circle than persons assumed to be at risk, including fewer low-risk persons, while still overlapping substantially the circle for "all persons with (cervical) cancer":



Standardized and Prevalence-Weighted Scores

Each index permits comparisons of geographic areas for that cancer, but does not permit comparison of that cancer's index with another. To do this, it was necessary to standardize the scores and then prevalence-weight them. The standardizing was achieved by taking each town's total score as a percentage of the total score for the county. The next step required a weighting of the standardized score to reflect, in some measure, the prevalence of the disease. Without this, a standardized score of 5.0 in colo-rectal cancer would be equated with a 5.0 in cervical cancer, the implication being that equal numbers of persons should be screened for each. Since colo-rectal cancer is far more prevalent than cervical cancer, far more people must be screened for colo-rectal than for cervical cancer.

In the absence of Cancer Registry data, mortality and hospitalization were used as measures of prevalence. The county's deaths and hospitalizations were totalled for each of the four types of cancer, all four were totalled, and each cancer's percentage of the total was calculated. Since 46% of the total deaths and hospitalizations for these four cancers were colo-rectal, 40% were breast, 7% were oral, and 6% were cervix, these proportions were used as the weighting factor applied to the standardized score, and produced prevalence-weighted scores.

Prevalence-weighted scores (PWS) apportion the screens and permit comparisons across geographic areas and across anatomic sites. The scores are applied to a bottom-line number chosen arbitrarily and pragmatically by the policymakers and/or those who will be evaluators, exercising either of two options:

Option A -- A feasible, reasonable number may be selected on the basis of the county's population. In the case of this county, the figure chosen was 10% of the county's adult population. This became the base or 100% figure to which the figures in the PWS table were applied, yielding the number of persons to be screened from each town, for each type of cancer. If the selected goal proves, at the end of the year, to be too high or too low, it affects all the LHDs equally and their performance can still be evaluated relative to each other.

Option B -- The dollars available for screening may be expected to pay for a total number of screens and this sets the number that becomes the base or 100% figure to which the PWS are applied. If the dollars increase, the base number increases proportionately; if the legislature cuts funds, the base number is cut proportionately.

With either option, this methodology permits the LHDs to assess need and to know what is expected of them, quantitatively, and permits the State to quantitatively evaluate their performance.

BIASED ESTIMATES OF THE RISKS OF LOW BIRTH WEIGHT

Robert Lewis, Ph.D., Murray B. Hudson, MPH

Introduction

The concept of a critical value of birth weight to identify high-risk infants appears to have had its origin in Europe around the turn of the century. Schwarz and Kohn (1921)[1] refer to Miller (1886)[2] and Ylppo (1919)[3] as having applied 2275 grams (5 pounds) and 2500 grams (5½ pounds) respectively as critical values in creating two groups of infants of local origin for purposes of clinical comparisons.

The idea of using the low birth weight ratio* (LBWR) to compare populations emerged in the 1930's and Dunham (1936)[4,5] led a movement for acceptance of the 2500 gram value as an international standard. This was achieved through adoption by the World Health Organization in 1948[6], thus casting the value in steel for future research and policy applications.

Several investigations have displayed tables of racial/ethnic distributions which include most or all of the possible weights. Verhoestraete and Puffer (1958)[7], Yerushalmy (1967)[8], Chase (1969)[9] and Fryer et al (1977)[10], all reveal racial/ethnic distributional differences. Further, sex differences are shown in Tanner's 1970 commentary on "Standards for birth weight or intrauterine growth".[11] These results suggest that the relative locations of the distributions on the weight scale differ for both sex and race/ethnic groups.

The critical value which appears to have served European and Nordic investigators early in this century is thus brought to question as to its appropriateness in other places. The statistical concept, risk of an individual falling in the low birth weight range, is useful in practice when the critical value reflects the true risk of injury. There is a conceptual continuum of individual responses to exposures to hazards ranging from a healthy and prospering infant through various debilities of increasing severity to death. Some foreshorten the duration of pregnancy and some result in small full-term infants. Still others result in normal weights but otherwise disadvantaged infants (after Lillienfeld and Passamanick).[12] As the number of debilities increases, the population exhibits more low birth weight outcomes, and the skewness of the weight distribution increases.

The fundamental comparison of risks associated with low birth weight between populations is with respect to the skewness of the weight distributions: If the distribution for a given population is essentially symmetrical, one is led to conclude that a situation of "no-excessive-risk" prevails. If the distribution for a given population is skewed, to produce a left or low birth weight tail, one is led to accept that some agents which shorten the gestational period and/or interfer with in utero development are active for some but not all of the mothers.

The LBWR is intended to function as an indicator of the degree of skewness or risk of low birth weight for a given population, but it is a faulty indicator because of the phenomena of differences in normal weights among populations. Differences in relative location of weight distributions between populations (i.e., overall shifts up and down the weight scale) introduce something other than the degree of skewness when comparisons are made between populations.

Rooth (1980)[13], after Fryer's lead, attempted to solve both problems in one stroke by proposing a "biological standard" (the mean minus two standard deviations). Unfortunately the solution has flaws. Assuming that the gaussian distribution holds, the method is applicable to a symmetric normal distribution, but it appears to apply only in the condition of no excessive risk for the population. Given populations with different levels of excess in the left tail, distributions with greater levels of low birth weights have larger standard deviations. This would in turn result in artifactual reductions of their proposed low birth weight ratio.

This study explores an alternative method of converting sex-race specific weight scales to units of deviation from the median values, makes a comparison of subgroups intuitively selected to represent low and higher risk settings, and examines the implications with respect to sensitivity and specificity of a critical value.

Method

The source of information is a public tape of matched birth and death records provided by the South Carolina Department of Health and Environmental Control, Office of Vital Records and Public Health Statistics. For the six-year period, 1975-1980, there were 299,454 live births recorded for mothers resident in South Carolina and reported black or white and known as to sex and weight. Cases over 7500 grams (16.5 lbs.) were excluded.

An anomoly of classification has led to the use of avoirdupois rather than the metric system of weight. The conventions of original measurement are such that the file includes large numbers of births weighed on avoirdupois scales: values that are converted to metric in the process of computerizing the data. This process results in a systematic sequence of possible

*LBWR = $\frac{\text{Number of live births weighing 2500 grams or less}}{\text{Number of live births in time period in area}}$

and impossible values in grams. For any particular size class in grams (say 150 grams) the number of possible ounce-equivalents differs from one class to the next, producing an unnecessary contribution to unexplained error in the weight variable.

The sets are converted to percent distribution in units of deviation from the median in order to examine sex-race specific differences in the low-weight tail independently from disparities in the locations of the distributions. The median weight value is chosen because the several distributions are observed to be skewed to the left, and the median provides a more effective measure of centrality than the mean.

The arithmetic process for producing a median-deviation (m.d.) distribution is as follows:

1. The Median is calculated.
2. The differences (median ounce - observed ounce) are calculated.
3. The resultant values are classified* in five-ounce groups, with the zero class centered on the median value.
4. Example: The white male median is 122 ounces, so that the midpoint of the (-18) category is 122-(5x18) = 32 and the weight category is 30-34 ounces inclusive.

The following analysis is directed to survivors of the first year of life. Removal from the distribution of those failing to survive the first year is based on the view that those individuals are by pathology definable as part of some underlying distributions of risks different from the normal or non-affected population. There are two general types of consequences of prenatal exposure to hazards: foreshortened pregnancy due to conditions influencing pregnant women, and diminished growth and development of organic elements of the fetus leading to reduced functional abilities. Neither of these general types are necessarily fatal, and one is led to suspect that a large proportion of infants exposed and suffering consequences to some degree, are not identified as such at the time of birth, but may be at continuing risk of further injury or continuing diminution of development in the future due to either pathology at the time of birth or continued exposure to risk elements in the environment.

Results

Figure 1 illustrates a clear relationship between mean birth weights and the number of reported prenatal visits for four sex-race groups. Further, a difference in sex-race specific weights is shown. For reported visits between 6 and 15, the differences between groups are consistent and the difference due to sex is about one-half that due to race. The blurring of differences for low numbers of visits is not merely the effect of small numbers: The race-difference of weights for 1-5 visits shrinks, but the sex-difference remains about the same. The overall differences between sex-race groups suggest that the data should be standardized for further comparisons (figure 3 below). The drop at 18 visits is the beginning of a pathology-related downward trend associated with excessive visits appropriate to troubled pregnancies.

Contrary to the common interpretation that the number of prenatal visits exclusively reflects medical care, the interpretation serving this exploration is that for those reporting visits consistent with optimal obstetric practice (10-18 reported visits) the family has demonstrated domestic resources to make repeated visits for medical care. This implies advantages which extend from before the pregnancy, such as economic resources, familial encouragement, fewer competing priorities, a relatively early seeking of care, and potential compliance with the advice received, (14, 15) as well as recognition and concern for potential medical problems. More remotely inferred for the group as a whole are other advantages such as divertable time, control of priorities, effective use of resources in the larger community, good nutritional opportunities, high standards of personal hygiene, beneficial behavioral characteristics, and maintenance of the physical environment. The balance of positive and negative influences on the duration of growth in utero and/or the quality of development of the fetus would be shifted toward the positive side. For such a group the weights at birth should be nearly symmetric in distribution.

The balance of such forces would be on the negative side for those reporting few or no visits, whether the low number was due to early parturition or due to belated seeking of care. If so, the birth weight distribution for this group should be skewed to the left, showing a higher proportion of low birth weights.

Figure 2 shows the birth weight distribution for the four sex-race groups. The distributions confirm the findings in figure 1: the distributions are about equally skewed to the left, each curve showing a small excess in the low birth weight tail compared to the right side. The conventional low birth weight ratios for the four sex-race groups correspond to the areas under the curves to the left of the 5.5 pound value. The obvious differences in low birth weight ratios appear due to locational shifts of the population distributions. In both figure 1 and 2 the race-shift is about twice the sex-shift. But it is left-skewness that should be compared to observe differentials in the risk of low birth weight. Table 1 shows

*SAS Institute, Statistical Analysis System, using a procedure by Cockrell and Cockrell, University of South Carolina, 1980.

basic statistics by sex and race for the birth weight distributions.

Figure 3 shows that the standardized curves are nearly identical: no important sex-race differences in risk (i.e., differences in area under the curve) appear in the low birth weight tails. Figures 2 and 3 provide the basis for concern that the 2500 gram critical value is of doubtful value when applied indiscriminantly. The present analysis is restricted to information from the birth certificate: a review of variables which might be dichotomized to reveal a true risk differential revealed that the reported number of prenatal visits appears most effective. We use this statistic to demonstrate differences among sex-race specific standardized groups.

Figure 4 demonstrates risk differentials. The four sex-race specific curves for 10-18 visit groups are nearly symmetrical, suggesting very little excess risk of low birth weight, as compared to a normal curve. The groups reporting 0 thru 5 prenatal visits do not differ much from the 10-18 groups on the right side after the effects of larger variance are discounted, but the 0-5 groups do have a much higher proportion falling on the low birth weight tails. The effect of removing infant deaths is nil for the 10-18 visit group, insofar as the shapes of the four distributions are concerned. For the 0-5 visit group, the curve for all births is not different on the right side, but extends nearly horizontally at about one percent per class for all four groups from about -50 ounces through -90 ounces, where the curve for survivors (figure 4) falls off to zero.

The area in the low weight tail between the two sets of curves in figure 4 is the standardized weight specific risk differential. The area demonstrates a differential in exposure to hazards for individuals surviving the first year of life. The excess proportions in the tails of the 0-5 groups appear to be drawn from the weight ranges from the median down through about -30 ounces (the crossover point). The implication is that the effects of inadequate resources and attendant hazards preceed the pregnancy, and the additional stress of pregnancy is shared by the fetus. This conclusion leads to the concept of a standardized critical value, based on the observed weight-specific differential. For the purpose of illustration, the standardized value is set at -30 ounces for all four sex-race groups.

Table 2 shows corrected low birth weight ratios together with conventional values. The corrected values are based upon empirical reasoning from inspection of figure 4. The choice of a critical value at the crossover point of -30 ounces considers an obvious practical limit: setting the criterion too high would introduce negative weight-specific differences within a given sex-race group and would be deficient in specificity.

If the criterion is set below the effective upper range of risk differentials in the life settings where fetuses are carried, the LBWR will be lacking in sensitivity, and if the criterion were set appreciably lower the LBWR will fail as to sensitivity. The untenable nature of this dilemma has been expressly realized in clinical practice. In the context of limited resources, a critical value of 1500 grams is sometimes applied in intensive care facilities. In the light of figure 4, the effect of reducing the critical value has been to increase the false negatives unequally because these differ by race and sex, and reduce false positives to near-zero for all groups.

Discussion

A discussion of the two problems associated with the 2500 gram critical value starts from consideration of its qualities of sensitivity and specificity. The first solution, converting to median-deviation population weight distributions, is directed at the comparability of populations by "freezing" false positives and false negatives to the same relative position.

The second solution is to say what the critical value should be, so as to control both types of error, i.e., so as to have an agreeable indexing of effects of exposures to hazardous objects and conditions.[17]

The 2500 gram critical value produces descending LBWR's from black female through black male and white female. The lowest values are for white male. The pattern is the same for the general population and for both subgroups, with the highest LBWR's in the 0-5 visit subgroup.

For the standardized critical value, the corrected LBWR's reveal a sex-related difference in the 10-18 visit subgroups. For the 0-5 visit subgroup the trend observed in conventional LBWR's is reversed so that the LBWR's ascend from black female through black male and white female, with highest value for white male.

Summary and Implications

The sex-race specific weight curves for the general population of births are essentially identical in shape. From this the conclusion is drawn that there are no meaningful sex or race-dependent differences in exposure to agents or conditions that selectively operate to cause shortened gestation, or reduced fetal growth rate among those that survive the first year of life.

The sex-race specific curves for the two subgroups defined on the basis of reported prenatal visits are observed to differ within, but not greatly between, each sex-race group. The distributions of weight-specific differences (the area between the sets of curves) are about the same as to weight-range and magnitude. That

is, the four sex-race groups show similar differentials between subgroups in the risk of a low birth weight infant. The conclusion is drawn that appreciable exposures to conditions giving rise to low birth weight do exist and the exposures indicated by partitioning reported prenatal visits are operating at about the same levels in all four sex-race groups.

More information is required to determine what a generally acceptable standardized critical value should be: the offering of a specific value would be more defendable if it were based on information that identifies causal agents and follow-up to measure the severity of consequences influencing future growth and development of the infant. Such an approach was proposed in 1966 by Abramowitz and Kass[16]. If carried out, the work would lead to "best fit" population-specific predictive models for use in clinical and policy decisions in a given locality.

References

1. Schwarz H, Kohn JL. The infant of low birth weight: its growth and development. Amer J Disease Children, 1921.

2. Miller NT. Jahrb F. Kinderheilk 25:179, 1886.

3. Ylppo A, Ztschr F. Kinderheilk 24:1, 1919.

4. Dunham EC. Mortality among prematurely born infants. J. Pediatr 9:17-22, 1936.

5. Dunham EC. (Participant) Proceedings, 5th Annual Conference on Diseases of Prematurity. J Pediatr 8:104-121.

6. Manual of the international statistical classification of diseases, injuries and causes of death, Sixth Rev. World Health Organization, Geneva, 1947.

7. Verhoestraete LJ, Puffer RR. Challenge of fetal loss, prematurity and infant mortality - a world view. JAMA 167(8):950-959, 1958.

8. Yerushalmy J et al. Birth weight and gestation as indices of "immaturity". Amer J of Disease Children 109:43-57, 1965.

9. Chase HC. Infant mortality and weight at birth: 1960 U.S. birth cohort. Amer JPH 59:9:1618, 1969.

10. Fryer JG et al. Some indicators of maturity monogr. Pediatrics 9:33-85 (Karger, Basel, 1977).

11. Tanner JM. Commentary: standards for birth weight or intrauterine growth. Pediatrics 46(1):1-6, 1970.

12. Lillienfeld A, Passamanick B. The association of maternal and fetal factors with the development of cerebral palsy and epilepsy. Am J Obst & Gynec 70(1):93-101, 1955.

13. Rooth G. Low birth weight revisited. Lancet 1(8169):639-641, 1980.

14. Keet MP et al. Klein-vir-datum babas: etiologiese faktore by die Kaapse Kleurling-bevolking. Sa Mediese Tydskrif 199-203, 1981.

15. Petros-Barvazien A, Behar M. Low birth weight - what should be done to deal with the global problem. WHO Chronicle 32:231-232, 1978.

16. Abramowicz M, Kass E. Pathogenicities and prognosis of prematurity. NEJM 275:875-885, 938-943, 1001-1007, 1053-1059. (October 20, 27, November 3, 10), 1966.

17. Galen RS, Gambino RS. Beyond Normality: The Predictive Value and Efficiency of Medical Diagnosis. NY: Wiley Biomedical, 1975.

Robert Lewis is Professor of Biostatistics and Epidemiology, School of Public Health, University of South Carolina.

Murray Hudson is Director, Office of Vital Records and Public Health Statistics, South Carolina Department of Health and Environmental Control.

Table 1
Descriptive Statistics, Birth Weight in Ounces, South Carolina Resident
Live Births Surviving the First Year of Life, 1975-1980

	Race-Sex Group			
	BF	BM	WF	WM
Total				
General population	55192	56664	80672	85518
10-18 reported visits	25333	25937	57747	60766
0- 5 reported visits	11215	11308	4754	5250
Mean Weight				
General population	108.04	112.30	117.39	121.94
10-18 reported visits	110.87	115.37	119.05	123.88
0- 5 reported visits	102.46	106.55	108.44	112.24
Median Weight				
General population	109	113	118	123
10-18 reported visits	111	116	119	124
0- 5 reported visits	104	109	110	115
Standard deviation				
General population	19.55	20.18	18.82	20.00
10-18 reported visits	18.15	18.65	17.80	18.85
0- 5 reported visits	22.23	22.71	22.24	23.94
Skewness				
General population	-0.43	-0.45	-0.36	-0.41
10-18 reported visits	-0.13	-0.19	-0.18	-0.21
0- 5 reported visits	-0.62	-0.63	-0.55	-0.67

Table 2
Conventional and Empirically Corrected Low Birth Weight Ratios*
for Infants Surviving the First Year of Life

	Race-Sex Group			
	BF	BM	WF	WM
Conventional LBWR				
Critical value = 5.5 lbs.	88 oz.	88 oz.	88 oz.	88 oz.
General population	13.0	10.3	6.0	5.0
10-18 prenatal visits	9.1	6.8	4.3	3.4
0- 5 prenatal visits	21.2	17.6	15.6	14.0
Standardized LBWR				
CV_{ST}=Med-30 oz.	79 oz.	83 oz.	88 oz.	93 oz.
General population	6.7	7.3	6.0	7.3
10-18 prenatal visits	4.1	4.4	4.3	5.3
0- 5 prenatal visits	12.9	13.4	15.6	17.8

*Percent of births $\leq$ critical value

SOURCE: DHEC OVRPHS Public Tapes 1975-1980 Resident Births

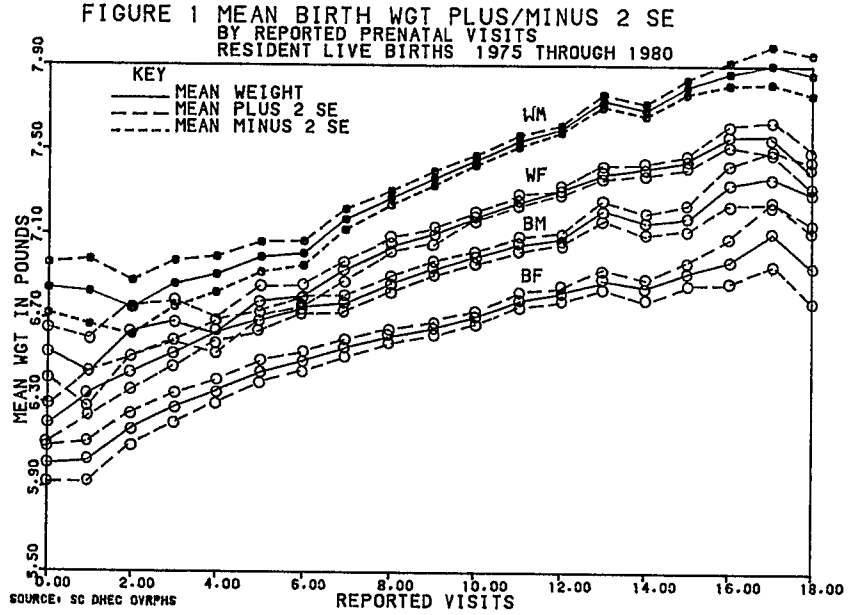


FIGURE 1 MEAN BIRTH WGT PLUS/MINUS 2 SE
BY REPORTED PRENATAL VISITS
RESIDENT LIVE BIRTHS 1975 THROUGH 1980
KEY
MEAN WEIGHT
MEAN PLUS 2 SE
MEAN MINUS 2 SE
WM
WF
BM
BF
MEAN WGT IN POUNDS
5.50
5.90
6.30
6.70
7.10
7.50
7.90
0.00
2.00
4.00
6.00
8.00
10.00
12.00
14.00
16.00
18.00
REPORTED VISITS
SOURCE: SC DHEC OVRPHS

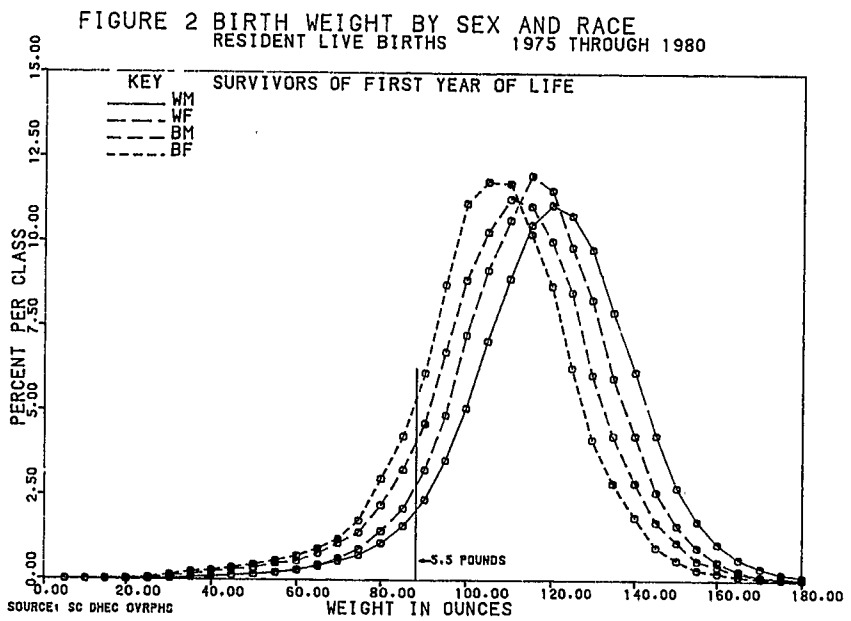


FIGURE 2 BIRTH WEIGHT BY SEX AND RACE
RESIDENT LIVE BIRTHS 1975 THROUGH 1980
KEY SURVIVORS OF FIRST YEAR OF LIFE
WM
WF
BM
BF
PERCENT PER CLASS
0.00
2.50
5.00
7.50
10.00
12.50
15.00
5.5 POUNDS
0.00
20.00
40.00
60.00
80.00
100.00
120.00
140.00
160.00
180.00
WEIGHT IN OUNCES
SOURCE: SC DHEC OVRPHS

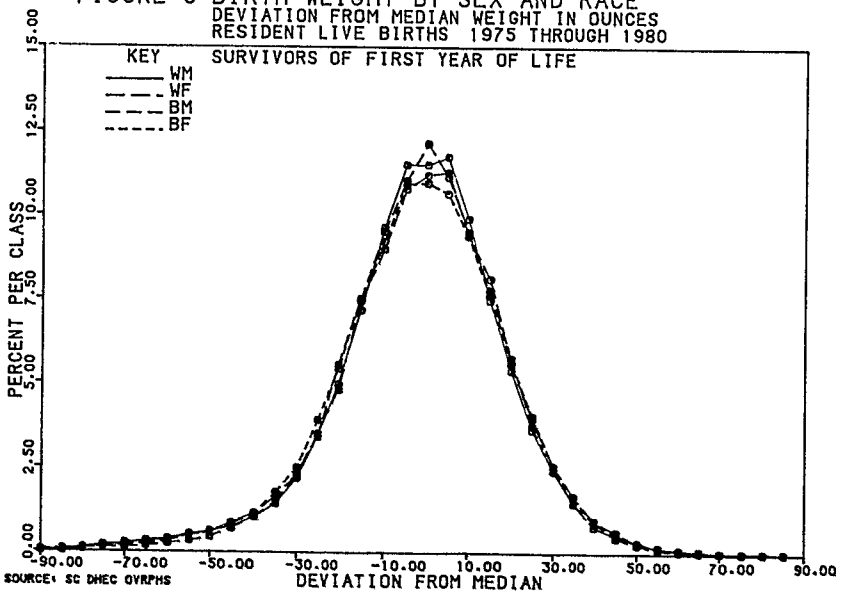


FIGURE 3 BIRTH WEIGHT BY SEX AND RACE
DEVIATION FROM MEDIAN WEIGHT IN OUNCES
RESIDENT LIVE BIRTHS 1975 THROUGH 1980
KEY SURVIVORS OF FIRST YEAR OF LIFE
WM
WF
BM
BF
PERCENT PER CLASS
0.00
2.50
5.00
7.50
10.00
12.50
15.00
-90.00
-70.00
-50.00
-30.00
-10.00
10.00
30.00
50.00
70.00
90.00
DEVIATION FROM MEDIAN
SOURCE: SC DHEC OVRPHS

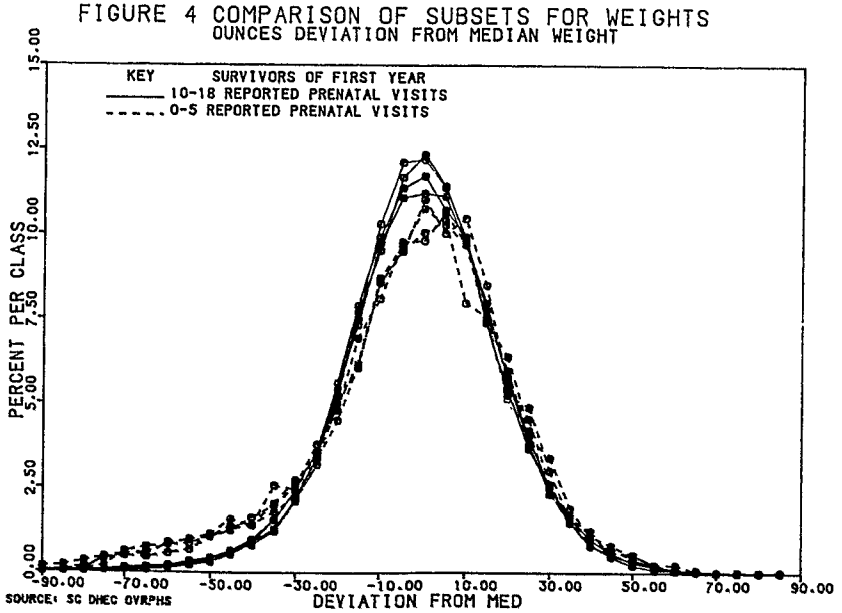


FIGURE 4 COMPARISON OF SUBSETS FOR WEIGHTS
OUNCES DEVIATION FROM MEDIAN WEIGHT
KEY SURVIVORS OF FIRST YEAR
10-18 REPORTED PRENATAL VISITS
0-5 REPORTED PRENATAL VISITS
PERCENT PER CLASS
0.00
2.50
5.00
7.50
10.00
12.50
15.00
-90.00
-70.00
-50.00
-30.00
-10.00
10.00
30.00
50.00
70.00
90.00
DEVIATION FROM MED
SOURCE: SC DHEC OVRPHS

Making Use of Available Resources for Monitoring and Evaluating Health Programs

Session Y

PHYSICIAN UTILIZATION AMONG ELDERLY COHORTS: IMPLICATIONS FROM THE HEALTH INTERVIEW SURVEY*

Fredric D. Wolinsky, Ph.D.
St. Louis University Medical Center

There are two manifest purposes of this article. The first is to present some of the results of applying Andersen's (1968) behavioral model of health services utilization to the elderly population, that is, those aged 60 years or more who were included in the Health Interview Survey. This permits the identification of any barriers that may exist between the elderly and their use of health services. The second purpose is to describe the process by which these same data are being used to search for any cohort and/or aging effects in the access barriers over time. A latent purpose of this article is to point out the difficulties associated with secondary analyses of this sort with the public use tapes available from the Health Interview Surveys.

Identifying Barriers to Access: Predicting the Elderly's Use of Health Services

Although there are a variety of models available to examine the determinants of health services utilization among the elderly, the behavioral model developed by Andersen (1968) and refined by his colleagues (Aday and Andersen, 1974, 1975; Aday et al., 1980; Andersen and Newman, 1973) is perhaps the most widely used and has been recently touted as one of the more policy relevant frameworks with regard to planning services for older people (Ward, 1977), as well as for the general population (Shortell, 1979). Indeed, in addition to various funding sources providing the substantial support necessary to enable Andersen and his colleagues to conduct a series of national surveys demonstrating the utility of his model, policy-makers have placed great currency in the behavioral model's findings on whether or not barriers continue to exist in terms of access to health care.

Briefly, the behavioral model considers the use of health services to be a function of the predisposing, enabling, and need characteristics of the individual. The predisposing characteristics reflect a greater propensity of some individuals to use health services. These propensities can be predicted by three subsets of individual characteristics which exist prior to the incidence of an illness episode: demographic characteristics (such as age, sex, marital status, and family size); social structural characteristics (such as education, occupation, social class, and race); and health beliefs (such as locus of control or medical knowledge). Individuals with different demographic characteristics are thought to have different types and amounts of illness, resulting in different patterns of using health services. Individuals with different social structural characteristics are thought to have different lifestyles, resulting in different patterns of health services utilization. Similarly, these patterns are expected to vary with the salience of health beliefs.

The enabling characteristics in the behavioral model reflect the fact that although an individual may be predisposed to use health services, they are not used unless he or she is able to do so. An individual's ability to use health services depends on both family resources (such as income, health insurance, and having a regular source of care) and community resources (such as variations in the availability of health care providers and their practice patterns, which are reflected in physician and hospital-bed-to-population ratios, place of residence, and geographic location). If there are sufficient family and community resources to enable the individual to use health services, then the individual would be more likely to use those services.

Finally, the behavioral model stipulates that even when one is predisposed and able to use health services, there must be some perceived need for using them. In other words, need is a basic and direct stimulus for the use of health services. Need is usually measured by self reports of symptoms, functional limitations, or perceived health levels.

According to Andersen and his colleagues in an equitable system of health care (that is, one in which there were no socioeconomic, sociodemographic, or sociocultural barriers to the use of health services), only the need characteristics would have a significant impact on the use of health services. If, however, non-need characteristics are found to be significant predictors of the use of health services, then a truly equitable health care system has yet to be achieved. Thus, from a policy and planning perspective, the application of the behavioral model to predict the use of health services among the elderly serves as one method for assessing whether or not Medicare and other programs have brought about a fully equitable health care system.

Using data from the 1977, 1978, and 1979 Health Interview Surveys on all individuals aged 60 years or more, we used the behavioral model to predict physician utilization. Briefly, the predisposing, enabling, and need characteristics were operationalized as follows. (Please note that the following list contains only the indicators common to all of the Health Interview Surveys -- and that

several indicators included solely in the 1978 Health Interview Survey will be identified. This will give some indication of the difficulties in using the Health Interview Survey to address longitudinal or change issues). The family composition dimension of the predisposing characteristics was measured by sex, age, marital status, and whether the respondent lives alone. The sociostructural dimension was measured by race, education, retirement, and labor-force participation. Because previous studies by Andersen (1968) and others (Wolinsky et al., 1983) have failed to find health beliefs to be significantly related to the use of health services and because there were no measures of the respondents health beliefs in these Health Interview Surveys, this dimension of the predisposing characteristics has gone unmeasured. The family resources dimension of the enabling characteristics was measured by having a regular source of medical care, (only in 1978) the presence of a telephone, individual income (only in 1978, family income in all years), having private insurance (only in 1978), and having Medicaid and Medicare (only in 1978) coverage. The community resources dimension of the enabling characteristics was measured by two sets of dummy variables reflecting geographic location and population density, as proxies for practice patterns and provider supply, respectively. The perceived dimension of the need characteristics is measured by the respondent's evaluation of his or her limited activity resulting from health problems, and his or her general health status. The evaluated dimension of the need characteristics is measured by the number of restricted activity, bed disability, and lost work days, as well as by a dummy variable indicating whether or not the respondent is overweight based on national norms for body-mass-ratios (Keys et al., 1972).

Finally, the measure of physician utilization is the natural logarithm of the number of visits to physicians during the previous twelve month period. The natural logarithm of physician utilization was taken in order to normalize its otherwise highly positively skewed distribution. We have addressed this issue in detail elsewhere, and the interested reader is referred to that explanation of this issue (see Wolinsky and Coe, 1984).

To reiterate the comparability problem, several crucial indicators of the family resources dimension of the enabling characteristics are available only in the 1978 Health Interview Survey:

- having a regular source of medical care,
- individual vs. familial income,
- private health insurance coverage,
- and Medicare coverage.

Thus, the assessment of the aging and cohort effects on the parameters of the behavioral model is necessarily limited to the common indicators. However, to assess the equity issue in this article, we use the more fully specified model and apply it solely to the 1978 Health Interview Survey.

Despite the limitations imposed by using only the common incidators, the data do demonstrate the consistency of the common means and standard deviations produced by the 1977, 1978, and 1979 Health Interview Surveys. For example, looking at the means and standard deviations of the common indicators shows, with the exception of reported education and income levels, the stability of these estimates over the three year period. Moreover, the two exceptions are as we would expect, given inflation which has systematically and rapidly increased income levels, and the substitution of "newer-old" cohorts for "older-old" ones which gradually but perceptively increases educational attainment levels.

Table 1 contains the standardized regression and R^2 coefficients of the need, enabling, and predisposing characteristics on the natural logarithms of physician utilization among the elderly in the more comprehensive 1978 Health Interview Survey. Although there are numerous important and interesting relationships identified in Table 1, only the three components of the results most relevant for the discussion of health care policy and the elderly shall be discussed. The first component has to do with the level of explained variance (R^2) obtained by the behavioral model: 22.6 percent of the variance in physician utilization can be explained. This is consistent to moderately larger than those reported in existing literature on the elderly's use of health services. Nonetheless, one is obligated to point out that much of the variance in physician utilization is still unexplained, even after using 26 variables in the predictive equation. There are at least three competing explanations for these low explained variances: (1) the model may not be properly specified; (2) the measurement of the independent and dependent variables may be quite imprecise, attenuating the level of explained variance that may be achieved; and, (3) the use of health services among the elderly may simply be a random phenomena unrelated to the predisposing, enabling, and need characteristics.

Although all three explanations of the nonrobust nature of the behavioral model are plausible, and although proponents of each explanation are easily identifiable, the bottom line is that we cannot accurately predict the elderly's use of health services. The implication of this finding for planning health care policy for the elderly is devastating. That is, we are engaged in the business of planning for the future health care of the elderly in the United States when we are virtually unable to explain why the elderly use health services, even when we apply the most sophisticated conceptual model using the most sophisticated statistical

Table 1. Standardized Regression and R^2 Coefficients of the Need, Enabling, and Predisposing Characteristics on the Natural Logarithms of Physician Utilization Among the Elderly

Independent Variables	Physician Utilization
Need Characteristics	
Limited Activity	.20***
Overall Health	-.19***
Restricted Days	.15***
Bed Disability Days	-.00
Lost Work Days	.01
Over-weight	.02*
Enabling Characteristics	
Regular Source of Care	.17***
Telephone	.02**
Income	.05***
Private Health Insurance	.06***
Medicaid Coverage	.06***
Medicare Coverage	.03**
Northeastern United States	.04***
Western United States	.01
Southern United States	.00
Major-SMSA	.05***
Non-SMSA	-.05
Predisposing Characteristics	
Sex	.07***
Age	.00
Married	.04**
Widowed	.03**
Lives Alone	.05***
Race	.02**
Education	.02*
Retired	-.02
Labor Force	-.03***
R^2	.226***

* = $p \leq .05$

** = $p \leq .01$

*** = $p \leq .001$

methods to the most reliable and valid national data-set available. In short, we are planning for the future even though we can not explain much of the use of health services in the present. One may, indeed, view this as a rather risky business (Blum, 1982).

The second component of the findings shown in Table 1 is more comforting. This component has to do with the effects of the individual variables on physician utilization. For the most part, the individual effects are as we expected. That is, the greater the need for health services the greater the utilization, the greater the ability one has to use health services the greater the utilization, and the more predisposed one is the greater the utilization. The most important individual predictors of physician utilization are, in order of relative importance (i.e., the magnitude of standardized regression coefficients), limited activity, overall health, having a regular source of medical care, and the number of restricted activity days. Although a number of the other independent variables have effects on health services utilization that are significant at the .05 level or beyond, they do not produce standardized regression coefficients greater than $\pm.10$, which has been suggested as a cut-off point for substantive importance in the analysis of large data-sets (Heise, 1969). When taken together, these effects are reassuring, in that they validate the behavioral model of health services utilization, even though the majority of the variance remains unexplained.

The third component of the results shown in Table 1 addresses the issue of equitable access to health care among the elderly. On the one hand, it is possible to interpret these results as providing substantial evidence for the existence of an equitable health care delivery system with regard to the elderly. This support stems from the fact that the need characteristics are the most significant predictors of the elderly's use of health services, and this may be demonstrated in either of two ways. First, an examination of the standardized regression coefficients greater then $\pm.10$ indicates that all but one of the large standardized regression coefficients for physician utilization belong to the need characteristics. Second, calculating the unique contributions of the need characteristics toward explaining the use of health services indicates that they account for more than two-thirds of the variance that can be explained in physician utilization. These data are consistent with those reported in existing studies of both the elderly's use of health services (Branch et al., 1981; Coulton and Frost, 1982; Eve and Friedsam, 1980; Wolinsky et al., 1983) and the use of health services among the general population (Aday et al., 1980; Wolinsky, 1978). According to Andersen and his colleagues, findings such as these are indicative of an equitable health care system.

On the other hand, these same data may be used to support the argument that equity has yet to be achieved, because there are significant effects in predicting the use of health services for having a regular source of care, having access to a telephone, individual income, private health insurance, Medicaid coverage, Medicare coverage, living in the northeastern or western United States, living in a major metropolitan area, being female, being married, being widowed, living alone, being nonwhite, and having more education. In particular, the significant posi-

tive effects of having a regular source of care, individual income, and having private health insurance all suggest the continued existence of access barriers between the elderly and the health services that they need to use. Although the effects of these barriers may be small when compared to the effects of the need characteristics, they do nonetheless exist. This is especially important because Medicare was designed specifically to eliminate socioeconomic barriers, especially the ability to pay for health care. The ability to pay, however, either manifest in terms of individual income or having private health insurance remains an influential factor in the elderly's use of health services. Moreover, having a regular source of care also has a significant impact on service consumption; thus, not having a regular source of care represents a sizeable access barrier. Accordingly, although Medicare may well have diminished the impact of the ability to pay on the elderly's use of health services, economic and noneconomic barriers to accessing health care have not been entirely eliminated for the elderly.

Assessing Cohort and/or Aging Effects in the Access Barriers Over Time

This article turns now to a brief overview of how the behavioral model and the 1977, 1978, and 1979 Health Interview Surveys will be used to look for aging and/or cohort changes in the effects of the predisposing, enabling, and need characteristics on the use of health services among the elderly. Basically, the parameters of the behavioral model will be estimated (i.e., the unstandardized regression coefficients) separately within each one-year age cohort within each Health Interview Survey. Table 2 helps to graphically portray the approach. In Table 2 the entries in the cells are the number of individuals of that age in that Health Interview Survey. For example, there are 1,047 individuals aged 60 in the 1977 Health Interview Survey, 997 of that age in 1978, and 1,070 of that age in 1979; and so on.

The general approach that will be used is a cross-sequential design, in which artifical cohorts are constructed using a series of cross-sectional data-sets, that is, the Health Interview Surveys. Given the known sampling veracity of the Health Interview Surveys, each of the one-year cohorts identified in Table 2 is representative of the non-institutionalized population from which they were drawn. Given this, the cross-sequential design assumes that the 1,047 respondents aged 60 in the 1977 Health Interview Survey, the 912 respondents aged 61 in the 1978 Health Interview Survey, and the 985 respondents aged 62 in the 1979 Health Interview Survey are all from the same cohort born in 1917. Similarly, representative samples of the 1916 cohort are the 1,025 individuals aged 61 in 1977, the 972 individuals aged 62 in 1978, and the 964 individuals aged 63 in 1979. Following Erdman Palmore (1978, 1981) this allows one to identify and isolate the aging and cohort effects as indicated in Figure 1. Here the horizontal string comparisons will reflect cohort effects, the diagonal string comparisons will reflect aging effects, and the vertical string comparisons will reflect cross-sectional or aging plus cohort effects. We can do this in these data, however, only because between 1977 and 1979 there were no changes in the period or environmental effects on the use of health services among the elderly. We could not assume this now because of the period or environmental effects of Diagnostic Related Groups.

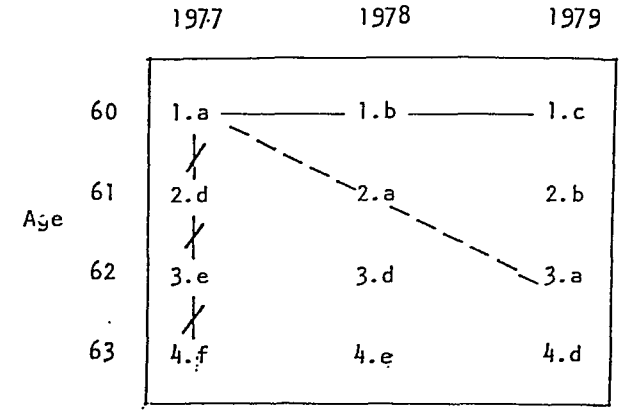
Table 2. Number of Respondents in the Age Cohorts to be Analyzed in the 1977, 1978, and 1979 Health Interview Surveys (Abbreviated for Illustrative Purposes)

	Health Interview Survey		
Age	1977	1978	1979
60	1,047	997	1,070
61	1,025	912	957
62	1,040	972	985
63	987	906	964

What actually will be compared in Figure 1 will be the partial, unstandardized regression coefficients obtained in estimating the model for each cell of the figure. This will yield as many tables that look like Figure 1 as there are independent variables in the behavioral model. For example, one table will contain the unstandardized regression coefficients obtained for income for each cell. What we will see then is whether the effect of income on physician utilization is subject to aging or cohort changes. That is, is income more important in using health services for older cohorts than for younger ones, and does income become more important in using health services as a single cohort ages? For planning purposes this is a rather important issue, because although the elderly population pyramid is continually growing, there are rapid shifts occuring in the age and cohort structure of that population pyramid. Thus, knowing what effects, if any, that shift will have on our projections for the elderly's demand for health services will facilitate accuracy in the planning process.

The statistic to be used in making the comparisons of the regression coefficients along the horizontal, vertical and diagonal strings is called the Relative Instability Ratio, or the RIR (see Wolinsky, 1980). Its numerator is calculated by assuming that the regression coefficients in the string to be compared are estimates of the same underlying population parameter. That is, following the

Figure 1. Matrix Representation of the Partial Unstandardized Regression Coefficient Comparisons (Abbreviated for Illustrative Purposes)



Key: —— —— horizontal string comparisons; cohort effects

------- diagonal string comparisons; aging effects

—/— —/— vertical string comparisons; cross-sectional effects (cohort plus aging)

null hypothesis, we assume that $b_i = b_j = b_k$, with any differences due to sampling error rather than actual changes. The RIR then uses this mean to calculate the standard deviation around it, which is an estimate of the average deviation from the estimated population parameter. The denominator of the RIR is a pooled estimate of the standard error of the mean regression coefficient. Because the sample sizes are generally large, and the means and standard deviations are nearly identical, the RIR simply takes the mean of the standard errors of the regression coefficients.

Essentially, the RIR takes the standard deviation of the observed coefficients and norms it to the pooled estimate of their standard error, an analog to the common F-ratio of the variance between groups to the variance within groups. The value of the RIR is a metric indicating how many times the standard deviation is greater than the standard error. As such, $RIR^2 \times (N - 1)$ is distributed as Chi-squared at N - 1 degrees of freedom, where N is the number of observed regression coefficients in the string to be compared. Thus, for the aging and cohort strings, an RIR of 2.0 or greater would indicate that the differences between the observed regression coefficients are statistically significant at the .02 level or beyond. This would indicate that even within the age-specific elderly group, projections for the demand of health services should be made separately for each one-year age cohort, because the effects of the predictor indicators are not the same for all ages and/or cohorts of the elderly.

References

Aday, L.A. & Andersen, R. A framework for the study of access to medical care. Health Services Research, 1974, 9, 208-220.

Aday, L.A. & Andersen, R. Access to medical care. Health Administration Press, Ann Arbor, Michigan, 1975.

Aday, L.A., Andersen, R., & Fleming, G. Health care in the U.S.: Equitable for whom? Sage, Beverly Hills, 1980.

Andersen, R. A behavioral model of families' use of health services. Center for Health Administration Studies, Chicago, 1978.

Andersen, R. & Newman, J. Societal and individual determinants of medical care utilization in the United States. Milbank Memorial Fund Quarterly, 1973, 52, 95-124.

Blum, H. Planning for Health, 2d ed. Human Sciences Press, New York, 1982.

Branch, L., Jette, A., Evashwick, C., Polansky, M., Rowe, G., & Diehr, P. Toward understanding elder's health services utilization. Journal of Community Health, 1981, 7, 80-92.

Eve, S.B. & Friedsam, H.J. Multivariate analysis of health services utilization among older Texans. Journal of Health and Human Resources In Aging, 1980, 169-191.

Haug, M.R. Age and medical care utilization patterns. Journal of Gerontology, 1981, 36, 103-111.

Heise, D. Problems in path analysis and causal inference. In E. Borgatta, ed., Sociological Methodology, 1969. Jossey Bass, San Francisco, 1969.

Palmore, E. When can age, period, and cohort be separated. Social Forces, 1978, 57, 282-295.

Palmore, E. Social patterns in normal aging. Duke University Press, Durham, 1981.

Shortell, S. Theories and trends in health services utilization. In S. Williams & P. Torrens, eds., Introduction to Health Services. Wiley, New York, 1979.

Ward, R.A. Services for older people: An integrated framework for research. Journal of Health and Social Behavior, 1977, 18, 61-70.

Wolinsky, F.D. Assessing the effects of predisposing, enabling, and illness-morbidity characteristics on health services utilization. Journal of Health and Social Behavior, 1978, 19, 374-396.

Wolinsky, F.D. The problems for academic and entrepreneurial research in the use of health services: the case of unstable structural relationships. The Sociological Quarterly, 1981, 22, 207-223.

Wolinsky, F.D., Coe, R.M., Miller, D.K., Prendergast, J.M., Creel, M.J., & Chavez, N.M. Health services utilization among the non-institutionalized elderly. Journal of Health and Social Behavior, 1983, 24, in press.

Wolinsky, F.D. and Coe, R.M. Physician and hospital utilization among the non-institutionalized elderly: an analysis of the Health Interview Survey. Journal of Gerontology, 1984, 39, in press.

---

*This research was supported in part by a grant from the AARP Andrus Foundation.

USING PUBLISHED VITAL STATISTICS IN COMPARING THE QUALITY OF HEALTH CARE AMONG POPULATIONS: NEONATAL MORTALITY IN FLORIDA, 1977-1981

Isaac W. Eberstein, Florida State University
Jan R. Parker, Florida State University

Public health statistics are important tools for monitoring patterns of neonatal mortality within and among populations. However, for precise comparisons among geographic areas or population groups and, equally important, for appropriate interpretations of observed differentials, crude rates based on aggregate vital data are not without severe limitations. Nevertheless, since individual-level data are not generally available on a timely basis for small geographic areas or different population groups within such areas, aggregate data must often be used by default. This is especially the case for the analysis of neonatal mortality, where birth and death certificates must be linked to develop a complete person-record.

Fundamentally, interpreting aggregate mortality differentials based on crude neonatal rates necessitates assumptions concerning the degree of similarity of both mortality risk factors and health care among the groups being compared. Nevertheless, analysts are often unable to empirically distinguish factors associated with varying risk from those indicating dissimilar care. Consequently, interpretations of observed mortality differentials seem all too frequently to be overly ambiguous and of little utility for research or application.

To illustrate, white-nonwhite inequalities in rates of infant death have long been observed in the U.S. Indeed, "for as long as such statistics have been kept, a nonwhite newborn has been almost twice as likely to die before his first birthday as a white" (Bouvier and van der Tak, 1976: 19). While research has suggested that a wide range of social, economic, demographic, and health care factors are to varying degrees implicated, the racial differential persists, although declining in absolute magnitude.

The general opinion regarding the existence and persistence of the white-nonwhite infant mortality differential seems to be that it results mainly from a greater prevalence of high risk factors among nonwhites (see, e.g., Eberstein and Parker, 1983), although in many cases the equivalence of health care can not be assumed. Consequently, assessing mortality differentials necessitates considering both risk and care factors in the etiology of infant death, and, additionally, necessitates doing so comparatively among both whites and nonwhites. However, the relative inaccessibility of individual-level data for small geographic areas and the limited range of information on standard vital records often results in there being little objective basis for such a comprehensive assessment.

In this context, methodological procedures have recently been discussed (National Center for Health Statistics, 1982) which may enable analysts to decompose crude neonatal mortality rates (NMRs) into portions roughly along the lines of risk (risk component) and quality of care (care component). Succinctly, this is based on two assumptions: (1) that birthweight is a valid summary indicator of mortality risk, and (2) that birthweight-specific mortality rates are the same among the populations being examined as in a "standard" population for which linked birth and death records are available. Using these assumptions, a schedule of birthweight-specific neonatal mortality rates can be combined through indirect standardization with the distribution of births by weight within a local area and/or for specific population groups (such as racial groups) to generate an "expected" level of neonatal mortality given the risk characteristics of the particular population. Comparisons of expected neonatal mortality rates (ENMRs) would then serve to indicate variations in risk, and comparisons of actual (crude) and expected neonatal mortality rates (as a "standardized" neonatal mortality ratio, or SNMR) would then be indicative of variations in quality of care. Such procedures, if valid, would reduce many of the difficulties noted above and thereby expand the utility of published aggregate vital data for research as well as application.

However, as with any other standardization-based methodology, the validity of such procedures depends wholly on the two assumptions noted above. Although clearly reasonable in the main, these assumptions may be problematic in some instances. First, while birthweight may be the best single summary of the probable mortality risks to which infants are exposed, it may not be a complete description of those risks for all populations. To illustrate, there are other characteristics associated with risk (e.g., gestational age, parity, maternal age, race, and numerous others [see the discussion in Williams, et al., 1980: 560-1]) which, although also associated with birthweight, are to varying degrees independent. To the extent that these factors affect mortality independently of birthweight and, further, to the extent that the form and magnitude of these independent relationships vary among geographic areas and/or population subgroups, then the assumption that the birthweight distribution accurately identifies the risk component of neonatal mortality within a population would be weakened.

Alternatively, it is not clear that the other risk factors affect mortality independently of birthweight and, therefore, that the use of birthweight to indicate the risk component is fallible. Lee, et al., have argued that "once one controls for birthweight, socio-demographic factors are of little importance" for neonatal mortality, because birthweight is a mediating variable in each of these relationships (1980: 19).

The second assumption on which this methodology is grounded is that birthweight-specific mortality is the same among all the populations being examined as in the standard. Two potential sources of error may be identified here: a) a time lag since the period to which the standard rates refer, and b) heterogeneity among populations being compared. Declines in neonatal mortality in recent years have been substantial overall, with the most pronounced declines evident among those of low birthweight (Lee, et al., 1980: 18). The longer the time lag since their observation, the less applicable are the standard rates as indicators of current mortality schedules. This may be complicated by heterogeneity in birthweight-specific mortality among the populations being examined. When a lengthy time lag is combined with relatively great heterogeneity, the validity of using

ENMRs to indicate the risk component of neonatal mortality may be reduced to varying degrees for particular local areas or population subgroups.

In view of these considerations, research is necessary in regard to the assumptions on which this methodological procedure is based. In particular, one important question concerns the extent to which the birthweight distribution is inclusive of factors accounting for differences in mortality risks and, further, the extent to which the comparison of actual and expected mortality reflects differences in quality of health care. Using Florida counties as units of analysis, we test these assumptions through examination of neonatal mortality during 1977-1981, separately for whites and nonwhites.

First, we examine the extent of intercounty variation in NMRs, ENMRs, and SMRs. This is done separately for whites and nonwhites. Next, we attempt to explain variations in county-level neonatal mortality by race in terms consistent with the assumed (hypothesized) properties of each measure. That is, to the extent that county-level variations in ENMR reflect differences in risk, then factors considered important for such risks should be associated with this measure. Further, to the extent that neonatal risk factors are the same for whites and nonwhites throughout the state, then the relationships between these variables and ENMR should be the same for both racial groups. Similar expectations exist for variation in SNMR, to the extent that this measure reflects intercounty differences in quality of care, and for NMR, to the extent that this measure reflects both risk and quality of care. Finally, we assess the utility of this strategy for using published aggregate data for small geographic areas and for population groups in the study of differential mortality.

DATA

Live births by weight and the number of neonatal deaths during 1977-1981 were obtained from Florida vital statistics publications for each of the sixty-seven counties in the state (Department of Health and Rehabilitative Services, 1977-1981). Data were recorded separately for whites and nonwhites. All data refer to place of usual residence, not place of occurrence.

State-level neonatal mortality rates specific to birthweight and race refer to the 1975 cohort of live births in Florida. These rates were computed using the same birthweight categories as are provided in the published reports (0-999, 1000-2499, 2500-3999, 4000+ grams). A small number of cases did not have complete data; these were omitted from the analysis.

The three measures of neonatal mortality of interest, NMRs, ENMRs, and SNMRs, were computed for each county in the state. These are listed at the bottom of Table 1.

Several variables were included to explain intercounty variations in neonatal mortality. Three variables indicate the availability of care: a) the natural logarithm of total hospital beds per 100,000 (from unpublished data provided by the State of Florida, Department of Health and Rehabilitative Services), b) the natural logarithm of physicians per 100,000 (U.S. Bureau of the Census, 1978), and c) the percentage of births occurring in hospitals and attended by a physician (Department of Health and Rehabilitative Services, 1977-1981). The former two variables apply to the entire county population, while the latter variable is race-specific. Although other indicators of availability and quality of care might have been developed (e.g., the extent of prenatal care programs or the level of prenatal care provided in proximate hospital facilities), those employed here are easily assembled in standard and reliable form from available data sources.

As general socio-demographic indicators of risk, three variables reflect socioeconomic status and lifestyle. These are a) the percentage of persons with incomes below poverty (U.S. Bureau of the Census, 1980), b) the percentage of births to mothers younger than age 19, and c) the percentage of births to mothers who are unmarried (Department of Health and Rehabilitative Services, 1977-1981). Each of these variables is specific to race.

Additionally, the general industrial structure of the county is indicated by the total percentage of the labor force employed in agriculture, forestry or fishing (U.S. Bureau of the Census, 1980). This taps a primary dimension of heterogeneity among Florida counties which may be related both to neonatal mortality and the other variables mentioned above.

FINDINGS

Basic descriptive statistics for all variables in the analysis are provided in Table 1. There are no surprises in these data. NMRs are higher for nonwhites than for whites, 13.5 deaths per 1000 live births in comparison with 8.0. Similarly, ENMRs are also higher for nonwhites than whites, 14.8 in comparison with 8.6. However, when actual and expected neonatal mortality are compared in the form of SNMRs, the mean ratio is slightly higher for whites (92.3) than nonwhites (91.0). Nonetheless, before over-interpreting this slight difference, it is important to emphasize that there may be methodological difficulties inherent in such precise comparisons.

Specifically, there is a time lag between 1975, when the birthweight-specific neonatal mortality rates were observed, and the 1977-1981 "cohort" of births by weight to which these rates have been applied in generating the ENMRs referred to in Table 1. Though brief, during this period it is likely that birthweight-specific mortality rates declined, so that the ENMRs as computed may actually overestimate the risk component of neonatal mortality. Further, if these declines most affect nonwhites, this would account for the slightly lower average SNMRs among this group than among whites observed in Table 1.

Finally, none of the other variables listed in Table 1 exhibit unexpected distributions by race. In each case nonwhites exhibit values associated with higher neonatal mortality—a lower percent of hospital births, as well as greater percents of persons below poverty, births to young mothers, and births to unwed mothers. Physician density, hospital bed availability, and percent agriculture are not race-specific, referring instead to the county population as a whole.

Three patterns stand out from inspection of the detailed county mortality data (not shown) which more fully describe the nature of intercounty variation. First, the range for each of the measures of neonatal mortality is quite wide for both whites and nonwhites, and thus there is a great deal of overlap in mortality levels among these two groups. Second, counties with extremely high or low NMRs, ENMRs, or SNMRs are, with few exceptions, both nonmetropolitan and located in the North Florida panhandle or in the central

portion of South Florida. These are the areas where urban expansion has been most limited. Third, neonatal mortality rates for whites vary in an inverse relationship to those of nonwhites. That is, there is a general overall tendency for counties with high neonatal mortality among whites to exhibit low rates among nonwhites, and vice versa.

To illustrate, Figure 1 identifies three groups of counties, those with SNMRs higher than anticipated, those with SNMRs lower than anticipated, and those where the SNMRs are approximately equal to the values which would be expected if the NMRs and ENMRs are equal (100). For graphic purposes, confidence intervals were computed around the SNMRs to take small numbers of vital events and, thus, instability of the computed rates, into account. Thus, if the 95% confidence interval for the SNMR for a particular county contained the value 100, it is assumed in Figure 1 that the SNMR for this county is as expected (100). If the interval was entirely above or below 100, then the SNMR was considered higher or lower than expected, respectively.

On the whole, the SNMR summary categories pictured on the maps in Figure 1 suggest conclusions consistent with the patterns already identified. Namely, counties with high SNMRs are nonmetropolitan (there is one exception for whites, Marion County) and are concentrated in the northern panhandle or the south-central portions of the state. Also, SNMRs for whites and nonwhites tend to take on extreme values in different counties, although there are exceptions to this pattern.

To this point we have accomplished one of the two analytical objectives, description of intercounty variations in NMR, ENMR, and SNMR by race. There is substantial and interesting intercounty differentiation in neonatal mortality by race in Florida. The next, and most important, analytical question concerns accounting for these variations. Hypotheses in this regard are that risk factors should explain variation in ENMR, care factors should explain variation in SNMR, and NMR should vary with both care and risk factors among both whites and nonwhites. These hypotheses underlie the methodological strategy employed here.

Considering whites first, Table 2 presents OLS regression coefficients relating each independent variable with the three mortality indicators. In brief, the table suggests the conclusion that intercounty differences in neonatal mortality are essentially unrelated with the independent variables included here. Only one of the indicators of care, the natural logarithm of hospital beds per 100,000, exhibits a non-zero relationship with white neonatal mortality. As expected theoretically, bed availability is associated with both NMR and SNMR, and the direction of the relationship is such as to imply that counties with higher neonatal mortality also have a greater number of available hospital beds. (This is not surprising given that the fundamental purpose of the Certificate of Need process in the state is to insure that care is most available in the areas where medical need is the greatest.)

However, the finding that bed availability is the only one of the seven variables in Table 2 which evidences a non-zero relationship with NMR and SNMR, and, further, that none of these variables, even those reflecting risk, are related with ENMR, is inconsistent with expectations. This is discussed in more detail following a review of the estimated relationships for nonwhites (Table 3).

Among nonwhites, county differences in neonatal mortality are associated with the independent variables being considered, although the pattern of relationships is not fully consistent with expectations. First, percent agriculture is positively related with both NMR and SNMR, but not ENMR. However, this variable is a control for heterogeneity among counties and is not of primary theoretical concern.

Second, each indicator of the availability of care is related with ENMR among nonwhites, even though ENMR is assumed to reflect risk, not care, and the relationships between these variables and actual NMR are marginally significant. However the indicators of availability of care are unrelated with SNMR, even though this measure is assumed to be sensitive to variations in quality of care.

Further, the direction of the linkages between the care variables and the mortality indexes with which they are associated vary among the three measures of care. While physician density and percent hospital births vary as might be expected, the associations between the availability of hospital beds and NMR and ENMR are opposite these expectations. Additionally, the relationship between bed availability and NMR, the single variable found to be significant among whites in Table 2, is of an opposite direction among nonwhites. This suggests that current rules for assessing need may result in distributing beds away from areas of greatest need among nonwhites. Given that the Florida bed-need formula does not include population data for subgroups other than by age and sex, this possibility is consistent with current procedures. Further research needs to reexamine this finding carefully, and, if borne out, the inequitable effects of existing certificate of need procedures should be brought to the attention of policy makers for redress.

The second set of independent variables in Table 3 are those reflecting socioeconomic and lifestyle factors. These variables were expected to be closely related with NMR and ENMR but to be essentially unrelated with SNMR. As is clear from the table, percent below poverty is unrelated with both ENMR and NMR and is related with SNMR, contrary to expectations. Also, the percentage of births to young mothers is unrelated with ENMR and NMR, again inconsistent with expectations. However, the percentage of births occurring to unwed mothers is positively related to ENMR, consistent with expectations, and is marginally associated with NMR.

To summarize the regression results in Tables 2 and 3, then, intercounty variation in neonatal mortality among nonwhites is moderately associated with the variables indicating relative risk and quality of care, while neonatal mortality among whites is essentially independent of all the variables except one, an indicator of care availability. Clearly, inconsistencies were observed between the estimated relationships and theoretical expectations. For neither group were the findings wholly inconsistent with expectations, but, similarly, in neither case were expectations fully realized.

DISCUSSION

Turning to a more general discussion of these analytical findings, three distinct interpretations seem plausible. First, it may be that the variables used to indicate the extent of risk or quality of care are invalid measures of these concepts. This seems more

likely for the white population, with very low levels of NMR and ENMR, than among nonwhites. Among nonwhites, higher neonatal mortality implies that traditional indicators of risk/care (such as are employed here) will likely continue to be applicable. While the measures of care are indeed merely gross indicators of relative availability and use of medical facilities and personnel, the indicators of risk are essentially standard. Among whites, though, when mortality levels are quite low, perhaps these gross and traditional measures are no longer important bases for differences among counties.

A second plausible interpretation of these findings is that the methodological assumptions employed in computing ENMR and SNMR may not be sound. That is, perhaps county differences in ENMR do not accurately reflect risk, and variations in SNMR do not accurately reflect care. To the extent that other variables affect risk independently of birthweight, the validity of using the measure of ENMR in this regard is attenuated. Also, to the extent that birthweight-specific mortality actually varies among the county populations being examined, rather than holding constant as is assumed in computing ENMR, then the lower is the validity of the SNMR as an indicator of variations in quality of care. In other words, the mortality measures may only roughly reflect their conceptual properties, thus accounting for negative findings.

A third basis for interpreting these findings recognizes the small size of the analytical units. That is, due to the relatively small number of vital events occurring in some counties during the 1977-1981 period, the reliability of the computed rates may be adversely affected. If instability is problematic, then the estimates of the relationships in Tables 2 and 3 may be attenuated. This might account for the findings for whites.

At this point it is perhaps useful to restate the substantive objectives of the study. Essentially, this research focused on employing aggregate vital statistics data and indirect standardization to enable observation and analysis of variability in quality of health care among counties and its correlates. However, the findings are not clear either substantively or methodologically. Likely each of the three interpretations noted above is at least partly applicable—measurement error could arise from each of these sources.

Consequently, the primary conclusion of this research is that the method employed seems of relatively limited analytical utility in comparing the quality of health care among populations in such small geographic areas as counties. This is due in part to problems of underlying assumptions, but it additionally is a result of the high sensitivity of the neonatal mortality rates to relatively small numbers of vital events. Applications of the technique to larger populations will likely facilitate more positive empirical results. Ironically, this indirect technique will be less necessary for larger areas, due to the likely greater availability of linked birth and death records within such areas which may be directly observed to answer questions concerning relative risk and quality of care.

Figure 1a. Standardized Neonatal Mortality Ratios: Nonwhites, 1977-1981.

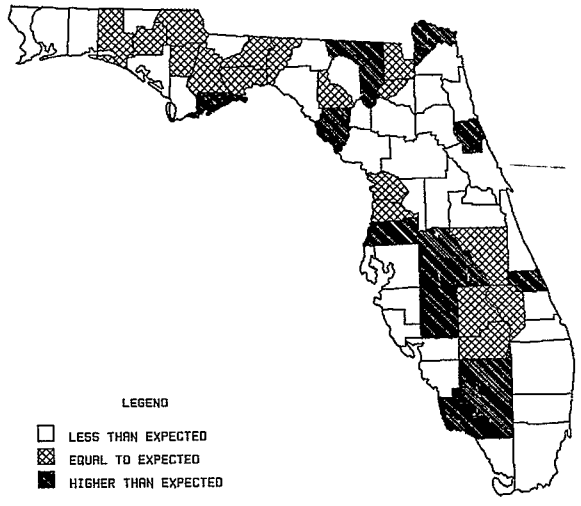
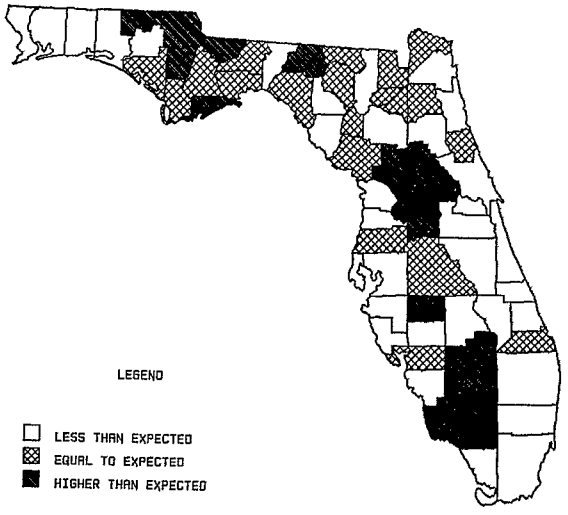


Figure 1b. Standardized Neonatal Mortality Ratios: Whites, 1977-1981.



REFERENCES

Bouvier, L. F. and J. Van der Jak. 1976. "Infant Mortality—Progress and Problems." Population Bulletin 31, No. 1.

Department of Health and Rehabilitative Services. 1977-1981. Florida Vital Statistics, Jacksonville: Office of Vital Statistics.

Eberstein, I. W. and J. R. Parker. 1983. "Racial Differences in Infant Mortality by Cause of Death." Paper presented to the Southern Sociological Society, April 6-9.

Lee, K., N. Paneth, L. Gartner, M. Pearlman, and L. Gruss. 1980. "Neonatal Mortality: An Analysis of the Recent Improvement in the United States." American Journal of Public Health 70(1):15-21.

National Center for Health Statistics. 1982. "Adjusting Neonatal Mortality Rates for Birth Weight." Vital and Health Statistics. Series 2-No. 94. PHS. Washington: GPO.

U.S. Bureau of the Census. 1978. County and City Data Book, 1977. Washington: GPO.

U.S. Bureau of the Census. 1980. Census of Population, (Summary Tapes). Washington: Bureau of the Census.

Williams,R., G.Cunningham, F.Norris, and M.Tashiro. 1980. "Monitoring Perinatal Mortality Rates: California, 1970 to 1976." American Journal of Obstetrics and Gynecology 136(5):559-568.

Table 1. Descriptive Statistics and Definitions of Variables

Variable		White	Nonwhite	Total
NMR	$\bar{X}$	8.0	13.5	
	(S)	(3.2)	(7.9)	
ENMR	$\bar{X}$	8.6	14.8	
	(S)	(1.7)	(5.2)	
SNMR	$\bar{X}$	92.3	91.2	
	(S)	(33.3)	(46.6)	
% Agr.	$\bar{X}$			8.6
	(S)			(6.7)
LN Phys. Den.	$\bar{X}$			4.2
	(S)			(1.6)
LN Hosp. Beds	$\bar{X}$			5.3
	(S)			(2.0)
% Hosp. Births	$\bar{X}$	96.3	95.7	
	(S)	(2.9)	(4.6)	
% Below Pov.	$\bar{X}$	13.2	39.9	
	(S)	(4.7)	(8.4)	
% Young Mothers	$\bar{X}$	11.6	24.4	
	(S)	(3.2)	(4.8)	
% Unwed Mothers	$\bar{X}$	9.3	56.3	
	(S)	(2.3)	(8.1)	

NMR = Neonatal deaths per 1000 live births, 1977-1981.
ENMR = "Expected" NMR 1977-1981 (given 1975 birthweight-race-specific mortality rates for the Florida birth cohort and the distribution of births by weight within each county, 1977-1981).
SNMR = Standardized NMR ((NMR/ENMR)*100).
% Agr. = Percent employed in agriculture, 1980.
LN Phys. Den. = Physicians per 100,000 1980 (LN).
LN Hosp. Beds = Hospital beds per 100,000, 1980 (LN)
% Hosp. Births = Percent of births occurring in a hospital and attended by a physician, 1977-1981.
% Below Pov. = Percent of persons below the poverty level, 1980.
% Young Mothers = Percent of births to mothers younger than 19, 1977-1981.
% Unwed Mothers = Percent of births to unwed mothers, 1977-1981.

Table 2. OLS Regression Coefficients: Whites (N = 67)[1]

Variable	NMR b	NMR β	ENMR b	ENMR β	SNMR b	SNMR β
% Agr.[2]	.10 (1.32)	.22	.03 (.73)	.12	.67 (.83)	.14
LN Phys. Den[2]	-.14 (-.36)	-.07	.30 (1.43)	.27	-3.50 (-.90)	-.16
LN Hosp. Beds[2]	.64 (2.40)	.40	-.05 (-.31)	-.05	7.61 (2.80)	.46
% Hosp. Births	.03 (.23)	.03	-.09 (-1.21)	-.15	1.12 (.80)	.10
% Below Pov.	.11 (.87)	.16	.02 (.29)	.05	1.32 (1.05)	.18
% Young Mothers	-.09 (-.45)	-.09	-.01 (-.12)	-.03	-.39 (-.19)	-.04
% Unwed Mothers	-.17. (-.79)	-.12	-.15 (-1.27)	-.19	-.24 (-.11)	-.02
Intercept	2.5 (.18)		17.4 (2.32)		-57.6 (-.41)	
F	1.3		1.06		1.77	
$\bar{R}$	.03		.01		.08	

[1] t-values are in parentheses.
[2] These variables refer to the entire county population. The other variables are race-specific.

Table 3. OLS Regression Coefficients: Nonwhites (N = 67)[1]

Variable	NMR b	NMR β	ENMR b	ENMR β	SNMR b	SNMR β
% Agr.[2]	.33 (2.27)	.28	.11 (1.17)	.14	2.45 (2.74)	.35
LN Phys. Den[2]	1.26 (1.63)	.25	1.17 (2.40)	.35	5.44 (1.14)	.18
LN Hosp. Beds[2]	-.92 (-1.64)	-.23	-1.40 (-3.99)	-.54	1.25 (.36)	.05
% Hosp. Births	-.59 (-3.12)	-.34	-.34 (-2.89)	-.30	-1.70 (-1.46)	-.17
% Below Pov.	.08 (.63)	.08	-.11 (-1.45)	-.18	1.36 (1.81)	.24
% Young Mothers	.10 (.41)	.06	-.05 (-.33)	-.05	1.19 (.78)	.12
% Unwed Mothers	.22 (1.66)	.23	.20 (2.41)	.31	1.36 (.17)	.02
Intercept	49.0 (2.51)		43.9 (3.57)		112.5 (.94)	
F	3.7		4.8		2.8	
$\bar{R}$	.22		.29		.16	

[1] t-values are in parentheses.
[2] These variables refer to the entire county population. The other variables are race-specific.

ON THE USES AND LIMITATIONS OF ZIP CODED HEALTH DATA, WITH EXAMPLES FROM WISCONSIN

Russell S. Kirby, Wisconsin Bureau of Health Statistics

In recent years, the United States Postal Service (USPS) ZIP Code has come into increasing use in the collection, reporting, geocoding and analysis of vital records and health data. While the use of ZIP Codes in health data analysis opens the door for several applications which cannot be performed with other types of areal units, their use also entails a number of limitations and constraints which are not readily discernible by the casual user. This paper focuses on five issues and methodological problems which are likely to be confronted in using ZIP Codes for geocoding health data in a statewide or multi-county data system. Examples drawn from the Wisconsin experience in mapping and using ZIP Codes in several research contexts serve to illustrate these problems and some potential solutions where applicable.

ZIP Code Areas as USPS Delivery Units

Most vital registration and health data systems collect data over areas with finite, exhaustive and mutually exclusive boundaries. In most cases these areal units have a legal basis as political subdivisions, voting precincts, or school districts. All places in Wisconsin are located in one and only one minor civil division (MCD). These MCDs exhaust the territory of the state, and except for a small number of incorporated places which contain portions of two or more counties, are wholly contained within one of the 72 counties in the state. Data collected by census tract, while collected over areal units defined solely for the purpose of collecting and reporting census data, are also coded to finite areal units.

ZIP Codes do not fall into either of the above categories. ZIP Codes were developed by USPS solely for the purpose of improving and speeding the delivery of the mails. The concept of the ZIP Code had its origins in the free delivery of mail in major cities beginning in the mid-nineteenth century and the full implementation of rural free delivery after 1900. Prior to these developments, postal service was a central function for customers, who had to physically travel to a post office both to send correspondence and to take delivery of the mails received at that post office. As the idea of free delivery evolved, it became necessary to think in terms of delivery areas at least for the purpose of sorting the mail for distribution by the carriers. Over the years, the volume of mail to large cities grew to the point where subdivision into zones became necessary. Eventually, a nationwide system of Zone Information Processing (ZIP) Codes was developed for all post offices, and all mailing addresses were required to include the relevant five-digit ZIP Code for prompt and efficient delivery of the mail.

When analyzing data collected by ZIP Code, it is necessary to think in terms of ZIP Code areas rather than ZIP Codes per se. A ZIP Code can exist and be perfectly useful for all USPS purposes without having any readily definable spatial extent. As some ZIP Codes have been assigned solely to post office boxes, and others serve small rural post offices or communities without providing delivery from that post office, no definitive areal delimitation is possible for these ZIP Codes. In Wisconsin, these nested ZIP Codes, which must be combined with surrounding ZIP Code areas in any data analysis, constitute about twenty percent of the 875 ZIP Codes used for residential mailings. Figure 1 shows the definable ZIP Code areas and the location of the five nested ZIP Codes in Walworth County, Wisconsin. ZIP Code area boundaries in this county correspond with less than ten percent of all MCD boundaries. As can be seen from Figure 1, most ZIP Code areas make more sense in terms of service areas than the arbitrary grid of MCD political units. There are also a number of special use ZIP Codes which serve certain institutions, government agencies (including USPS), and corporate mailers, but only those with residential populations appear with regularity on health records.

The boundaries of ZIP Code areas must be mapped with data obtained from local postmasters concerning postal routes handled by each ZIP Code.[1] Although boundaries may be subject to change, most future changes will fall into one of two categories. First, some ZIP Codes in rural areas will be discontinued, with their former delivery areas assigned to one or more adjacent ZIP Code areas. Some of these ZIP Codes will be downgraded from delivery to nested status. A second category of changes will occur in larger urban areas as the more populous ZIP Code areas are subdivided into two or more delivery areas, sometimes in connection with the creation of new branch post offices. While ZIP Code areas have evolved over the years, pending the introduction of nine-digit ZIP Codes the outer boundaries of five-digit ZIP Code areas should stabilize and the rate of change in ZIP Codes and their delivery areas should slow considerably. However, any ongoing ZIP Coded data system should include procedures for incorporating these changes on a regular basis.

ZIP Code delivery areas have been shaped by decades of local politics, and rarely correspond to counties or even to MCD boundaries within counties. Examples of postal service overlap abound. Residents of the unincorporated community of Pella in rural Shawano county, Wisconsin have no post office of their own. However, as of June, 1981 the rural carriers from no less than three surrounding ZIP Code areas (54166, 54929, and 54950) delivered mails on demand to any and all residents of the community. In rural Oconto county, Wisconsin, the postal delivery routes of ZIP Codes 54139 and 54154 are so intertwined that it proved almost impossible to delineate even an approximate boundary between these two delivery areas in Spruce township. Residents of newly created ZIP Codes often receive mail under the old ZIP Code for years, or even indefinitely. In Wisconsin several ZIP Codes serving Illinois

post offices actually deliver as the sole source of supply to residents across the state line. These problems are by no means unique to Wisconsin, but result from the essential nature of ZIP Codes as postal delivery units designed solely for processing the mails. These issues do not arise, by and large, when sub-county units are used for reporting vital and health records.

ZIP Code areas vary considerably both in areal extent and population size. The smallest ZIP Code areas contain only a few square miles and less than 100 persons, while the largest in area cover well over a hundred square miles, and several ZIP Code areas have populations in excess of 60,000. Thus, for some types of statistical analysis observations must be summed for adjacent ZIP Codes into units with enough observations to calculate stable estimates and rates.

Problems in the Attribution of Health Records to Places or Areas

A determining factor in choosing among types of reporting units for vital and health records is often the mechanism for reporting the information. Vital record certificates are filed as legal documents, and these data are almost always collected both by place of residence and place of occurrence by MCD, county and state. On vital records, ZIP Code data are rarely used. In many states including Wisconsin, birth certificates require a mother's mailing address for use in notification that the record has been filed. As mailing addresses are the principal source of ZIP Code information the use of these data from birth certificate files should be explored. Death certificates have the problem of accurate reporting, especially in the case of the elderly or veteran decedent. As the place of residence from the death certificate may bear no relation to the decedent's last normal habitation, the validity of any ZIP Code attribution from these records is questionable.

Many health data systems do not collect data solely over political units. Indeed, when observations are created through the act of hospital discharge, from a cancer diagnosis, or through completion of a health manpower survey, mailing addresses or ZIP Codes alone may form the basis for attribution to place of residence, occurrence or normal practice. A number of problems can arise in attempting to convert these ZIP Coded records back to the more traditional political units.

Three health data systems in Wisconsin collect data primarily by county and mailing address or ZIP Code. The Cancer Reporting System (CRS) collects data on all new cases of cancer diagnosed in Wisconsin hospitals, with attribution to place by state of residence, county and mailing address (which includes ZIP Code). When analyzing these data at the county level, no serious difficulties arise. However, counties are a gross scale of aggregation, particularly in the emerging fields of environmental and geographical epidemiology. On the CRS, the only units for finer resolution are ZIP Codes. As the ZIP Code of residence is derived from a mailing address, it is subject to the following limitation. ZIP Codes serve areas which are usually composed of several MCDs, often in more than one county. On a mailing address, however, the city-state-ZIP of common parlance actually designates only the post office name. Residents of MCDs other than that from which the post office takes its name generally do not have their MCD designated on the mailing address. Thus, one cannot differentiate residents of a given city from the balance of a ZIP Code area solely on the basis of a mailing address.

A recent request for data on cancer incidence in the city of Eau Claire, Wisconsin serves to illustrate this problem. An investigator wished to know whether an abnormally high incidence of cancer was occurring in that city. Data from the CRS, except in exceptional circumstances, can be reported only for aggregate analysis by county and ZIP Code. The city of Eau Claire includes parts of two counties, while the entire city, much of Eau Claire county and portions of several adjacent counties were served by the ZIP Code area 54701 (effective in July of 1983 that ZIP Code area was divided by the creation of the new ZIP Code 54703--this is the only major change in a Wisconsin ZIP Code in the past two years). It is impossible for the CRS to provide data on the area in question because the mailing address fails to differentiate the city of Eau Claire proper from the balance of the 54701 ZIP Code area.

Similar problems arise on the Wisconsin Health Manpower Surveys. While the reporting instruments vary by health profession, the general problem on these surveys is the incomplete specification of county, ZIP Code, MCD name and type of MCD (city, village or township). When the ZIP Code is missing from a record, if the MCD is given and it has a unique ZIP Code that code can be assigned. In cases where more than one ZIP Code area serves the MCD, the rule that 85 percent of the MCD population is served by a single ZIP Code is used to assign records with missing codes to that ZIP Code. When the rule is not met, the record is given an artificial ZIP Code with zero in the fourth and fifth digits. ZIP Codes ending in 00 are never used as valid codes by USPS, and in most cases all ZIP Codes serving a given MCD will begin with the same first three digits. Reporting of MCD data on these surveys is also often incomplete, but most analyses are done by county or ZIP Code of practice.

A third health data system on which geocoding problems are frequently encountered is the Wisconsin Hospital Discharge Survey. On this survey, conducted over two months of a year for almost all Wisconsin hospitals, the reporting instrument contains only ZIP Code of residence without any other areal identifiers obtained at the time of abstracting. Even analyses by county require some form of recoding of ZIP Codes into counties. As many ZIP Codes cross county boundaries, some assumptions regarding the distributions of populations by ZIP Code by counties are necessary.

The solution to these problems is conceptually simple, but politically inexpedient. On the CRS, data on MCD of residence could be collected, but if this information is not normally included on a hospital's abstracts a considerable expense might be incurred. As the CRS is presently operating as a voluntary reporting system, the geocoding issues are unlikely to be resolved in the near future. The Health Manpower Surveys have a more difficult problem, as on most of the professional surveys the full range of geocodes are collected. A statewide automated address matching system could assign MCD, ZIP Code and county from unique street addresses, but the cost of developing such a system would be prohibitive and if several fields of data are missing, the initial geocoding problem still remains. On the Hospital Discharge Survey, the simple solution is to add county of residence for use in analyzing data by county, but once again the addition of this field is problematic if the variable is not collected on the hospital abstracts from which the discharge data are transcribed.

Clearly, a variety of problems complicate the use of ZIP Codes in geocoding of both vital records and health data. Despite these problems, the use of ZIP Codes represents a viable alternative to MCDs in a number of situations. For most persons, encounters with MCD of residence if different from post office name are tenuous at best. Encounters with a ZIP Code, however, occur as a daily phenomenon. ZIP Codes also provide the analyst with a valuable mid-range scale of spatial aggregation. In terms of the range in order of magnitude, while Wisconsin counties have a population range of about 950,000 to 3,000 and MCDs range from 650,000 to under 100, ZIP Codes range only from about 70,000 to under 100 in population. For special analyses, ZIP Codes can be used for intra-county differentiation, especially for large cities, without recourse to expensive techniques for converting data to census tract units. Finally, with the recent release of the 1980 Census Summary Tape File (STF) 3b files, estimates of many attributes of the population and housing stock of five-digit ZIP Codes are available for use in conjunction with ZIP Coded health data.

Urban-Rural Differentiation by ZIP Code

Many researchers have studied urban-rural differences in health manpower, incidence of disease, health service provisionment and demographic trends. The terms "urban" and "rural" are subjective, but two objective measures have been devised to separate areas into urban and rural categories. The first is the traditional Bureau of the Census criterion of places of 2,500 or more persons. In recent years the concept of the "urbanized area", which includes densely populated areas adjacent to a large city has expanded the original definition. A second measure is the MSA (Metropolitan Statistical Area, formerly SMSA), a county-based measure which includes a central city or cluster with a population of 50,000 or more, the county in which it is located, and adjacent counties with strong employment or economic linkages to the central city. Both of these measures classify political units primarily on the basis of population thresholds into nominal urban and rural categories.

Using ZIP Codes, urban-rural differentiation is far more difficult. ZIP Code areas exist only as the spatial manifestations of postal delivery codes designed to speed the processing and handling of the mail. While post offices can be classified as urban or rural, based on the physical location of the post office building, the ZIP Code areas served by these post offices cover the full spectrum from urban to rural, with many shades in between.

Within large cities there are often several ZIP Code areas located entirely within the major urban center and its contiguous suburbs. These ZIP Codes are "urban" according to the 2,500 person threshold, and should be so classified even if their population falls below that level. At the opposite end of the spectrum are those ZIP Codes associated with post offices in small cities, villages or unincorporated places with populations of less than 2,500. While these ZIP Code areas should be classified as rural, it is possible that the total population of the area will exceed 2,500 persons. In Wisconsin, 50 ZIP Code areas, containing about 26% of the state population, can be classed as urban, while 507 ZIP Code areas, containing about 22% of the state population, would be considered rural by this definition. The balance, more than half of the population of the state, falls into two intermediate categories which are neither urban nor rural. Thirty urban-rural fringe areas, which contain parts of large urbanized areas and adjacent rural districts, account for 15% of Wisconsin's population, while 121 ZIP Code areas with 37% of the state population serve incorporated places with populations of 2,500 or more and surrounding rural areas. These latter areas are particularly difficult to classify, as the rural area often exceeds the urban center in both areal extent and population. The many nested ZIP Codes further complicate the differentiation of ZIP Codes into urban and rural categories based on the Census threshold of 2,500 or more persons.

Attempts to build ZIP Code areas into MSA and non-MSA units are somewhat more successful. As ZIP Codes rarely respect county boundaries, the major stumbling block is the classification of ZIP Codes by county. When ZIP Codes are assigned to one and only one county on the basis of the proportion of each ZIP Code area's population in each county, in most instances each ZIP Code can be accorded either MSA or non-MSA status. The major exceptions will be those ZIP Code areas with nearly equivalent populations in two counties, but this is a very rare occurrence, and only bears on the situation in which the two counties involved have differing MSA status. More problems arise when aggregation is to county rather than MSA/non-MSA, but with a smaller threshold proportion, most problems can be resolved without major loss of information.

Linking ZIP Coded Health Data Sets with Census and Other Statistical Sources

Health data by place or area often requires additional data from other statistical sources for comprehensive analysis. For example, even the calculation of crude demographic rates by county or ZIP Code area requires a population denominator normally available only from Census or other independent sources. Age-specific or age-sex-standardized rates require data on age and sex by ZIP Code area. In general, there are two approaches to obtaining these data, a direct estimation method utilizing data from the Census STF 3b tapes, or an indirect method involving some type of conversion of county or MCD data to ZIP Code areas.

While the direct method is preferable, it involves considerable data manipulation nonetheless. First, the data must be acquired from a consortium of data providers who paid for the processing of the Census data.[2] Then, as the Census provides data for all ZIP Codes, the information must be converted to the ZIP Code area units in the same way that health data are handled. This is very important, as the Census data for nested ZIP Codes cannot be taken as reliable indications of the population characteristics of these ZIP Codes. An additional complication is the suppression of data for some or all cells when the total population of a ZIP Code is not large enough to support a detailed breakdown by a demographic characteristic. While some guesswork is necessary, this obstacle can often be surmounted without a major sacrifice of reliability of the data.

Due to the slow dissemination of the 1980 Census estimates by ZIP Code, the indirect method for linking health data to Census and other statistical sources by ZIP Code is more widely used. An example will serve to illustrate both the methods and some of its weaknesses. In converting data by age from MCD to ZIP Code area, one makes the assumption that population is distributed uniformly by age across each MCD. Thus, regardless of the proportion of an MCD handled by a given ZIP Code area, the method assumes that that area has the same age distribution as the entire MCD. Then, the percent of the MCD population in each age category is multiplied by the proportion of the MCD population served by each ZIP Code area. This product is then multiplied by the MCD population, and the results are summed across MCDs by ZIP Code to yield the age distribution by ZIP Code area. If the initial assumption is valid, this method allows for a straight-forward conversion from MCD to ZIP Code area.

In Figure 2, the ratio of the population under five years of age to that aged 65 and older has been calculated for all MCDs in Walworth county, Wisconsin. This map shows that population distributions by age vary considerably, even among adjacent political subdivisions. Most extreme is the ratio of 0.580 for the town of Lyons, contrasted with 1.319 for adjacent Spring Prairie township. If this is the pattern observed within a single county, it seems highly improbable that at the intra-MCD scale populations are distributed uniformly by demographic characteristic.

The problem of data linkage clearly requires careful consideration at the outset of any research project involving ZIP Coded health data. Depending on the purpose, the indirect method may be preferable to direct manipulation of the Census estimates. However, both approaches have methodological limitations which must be weighed in formulating a research design.

Potential Applications of Nine-Digit ZIP Codes

It has now been several years since the subdivision of five-digit ZIP Codes into nine-digit units was first proposed. While it is highly likely that these geocoding units will eventually come into use, their applicability to the collection and analysis of health data is difficult to gauge. Nine-digit ZIP Codes will consist of the current five-digit code plus four digits which divide the current units into sectors (digits six and seven) and individual addresses or blockfaces within sectors (the eighth and ninth digits). This will permit the identification of individual or groups of mailing addresses directly from ZIP Codes.

Potential benefits of nine-digit ZIP Codes for health researchers include the solidification of the outer boundaries of five-digit ZIP Code areas. As the new codes will subdivide the current areas, the expense of adjusting the outer boundaries of present five-digit ZIP Code areas may outweigh any savings from rationalizing five-digit ZIP Code area boundaries. Nine-digit ZIP Code areas also have the potential to resolve problems of county or MCD to ZIP Code conversion, for the most part. ZIP Code sectors are designed to respect county boundaries, and will only rarely cross MCD boundaries. Thus, records coded by nine-digit ZIP Code will be convertible to county or MCD with little loss of accuracy. An added benefit will be easy urban-rural differentiation by nine-digit ZIP Code.

On the other side of coin is the question of whether nine-digit ZIP Codes will come into universal use by residential mailing addresses. The new codes were designed primarily to handle the large volume of business-generated mail. Just as residential addresses eventually receive most mail even using the wrong ZIP Code, the same is likely with the longer codes. Also, it may be many years before most residential users form the habit, by choice or necessity, of using the nine-digit codes. Thus, health researchers have a long wait before realizing the potentials of the new and soon-to-be-implemented nine-digit ZIP Codes.

Conclusion

In this paper some of the uses, problems and limitations of ZIP Coded health data have been briefly explored. The nature of ZIP Codes as USPS delivery units affects every stage of any analysis based on these areal units. For a number of purposes, ZIP Codes are a natural choice for data collection. While urban-rural differentiation by ZIP Code is possible, no fully satisfactory method can be devised with currently available data. Data from independent statis-

tical sources are often necessary for health analysis, and researchers should carefully consider the choice of direct ZIP Code estimates from Census and related sources or the conversion of MCD or county data to ZIP Code areas. Finally, the nine-digit ZIP Code, which has the capacity to solve most of the problems discussed above, will not be in wide use for at least several years. Thus, the limitations imposed on researchers by ZIP Coded health data must be borne in mind as health planners and analysts use these data in epidemiological and demographic research, and in studies of health facilities marketing and utilization.

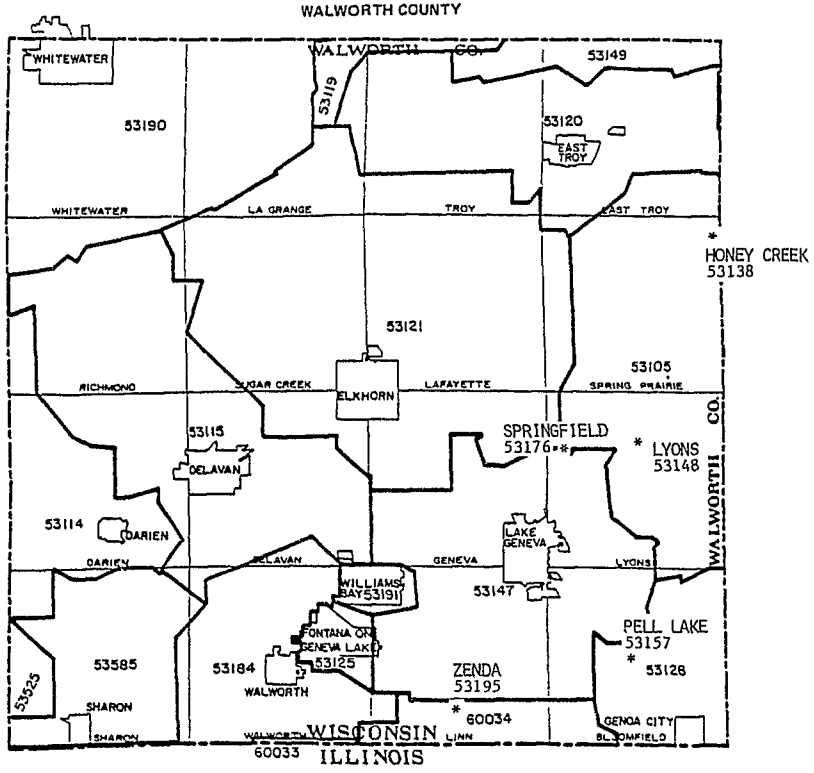
Notes

1. For a description of the method used for Wisconsin, see Russell S. Kirby, "The Wisconsin ZIP Code Mapping Project," Special Libraries Association, Geography and Map Division Bulletin No. 132 (June, 1983), 27-35.

2. For information on the prices, formats, and structures of these data files, contact the National Planning Data Corporation, P. O. Box 610, Ithaca, NY 14850.

Figure 1

U. S. POSTAL SERVICE ZIP CODE AREAS
WALWORTH COUNTY



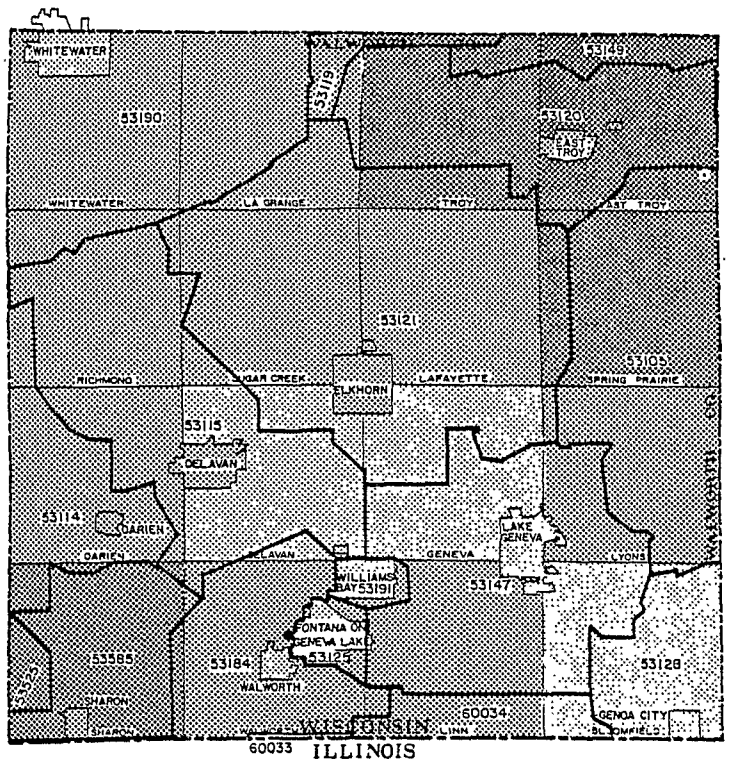
NOTE: THE FOLLOWING ZIP CODES ARE USED FOR CENTRALIZED POST OFFICE BOX DELIVERY ONLY. AS SUCH THEY POSSESS NO MAPABLE SERVICE AREA.

ZIP CODE	POST OFFICE
53138	HONEY CREEK
53148	LYONS
53157	PELL LAKE
53176	SPRINGFIELD
53195	ZENDA

Source: Compiled by SEWRPC during spring of 1978 from published U. S. Postal Service documents and information supplied by local postmasters.

Figure 2

WALWORTH COUNTY
RATIO OF MCD POPULATION UNDER 5 TO 65+, 1980



NOTE: THE FOLLOWING ZIP CODES ARE USED FOR CENTRALIZED POST OFFICE BOX DELIVERY ONLY. AS SUCH THEY POSSESS NO MAPABLE SERVICE AREA.

ZIP CODE	POST OFFICE
53138	HONEY CREEK
53148	LYONS
53157	PELL LAKE
53176	SPRINGFIELD
53195	ZENDA

RATIO
- 1.50+
- 1.00 - 1.49
- 0.50 - 0.99
- LT 0.50

Source: Compiled by SEWRPC during spring of 1978 from published U. S. Postal Service documents and information supplied by local postmasters.

Sources and Uses of Data for Evaluating Medicare and Medicaid Reimbursement, Eligibility, and Coverage Demonstrations and Experiments

HCFA Special Session

"DATABASE FOR THE NATIONAL HOSPITAL RATE-SETTING STUDY: AN EVALUATION OF FIFTEEN STATE HOSPITAL PROSPECTIVE PAYMENT PROGRAMS"

Richard Yaffe, Health Care Financing Administration
Frederic Pratter, Gary Gaumer, Craig Coelen, Abt Associates

Introduction

In 1978, the Health Care Financing Administration awarded a contract for the National Hospital Rate-Setting Study (NHRS) to Abt Associates of Cambridge, Massachusetts to evaluate the impact of 15 hospital prospective payment programs. The study was undertaken in order to determine more definitively the effects of hospital prospective reimbursement programs in eight major areas:

1. hospital revenue and expenditure;
2. volume of services produced;
3. hospital staffing and payroll;
4. quality of hospital care;
5. hospital capital and investment;
6. organization and management of hospitals;
7. accessibility of services; and
8. health systems utilization and costs.

The basic study design for most of the analyses employs a multivariate regression model with the outcome variable of interest, e.g., hospital expenditure per day, as the dependent variable and a wide range of hospital and area specific variables as the independent, or right-hand-side (RHS) variables. The basic unit of analysis is the hospital-year, that is, each case in the regression analysis reflects the variables for a given hospital in a given year from 1970-1979. The intervention, that is a prospective payment program being in effect, enters the model a 0-1 dummy variable, taking values of 1 for those hospitals that are under a prospective reimbursement program in a given year. Separate PR variables are specified for each of the 15 prospective rate-setting States and for significant variations of programs within States. All non-Federal, acute care hospitals in the 15 States are included in the model as well as a 25 percent random sample of similar hospitals in the remainder of the 48 contiguous States and the District of Columbia, which act as the "control group" in the model.

This description of the evaluation design is somewhat oversimplified, but it is a four-way design that includes both cross-sectional experimental/control group comparisons, as well as before and after comparisons.

Some of the major outcome variables include:

- measures of hospital revenue and expenditure, and financial status, (e.g., operating margin and debt/asset ratio);
- measures of hospital output such as outpatient visits, admissions, patient days, average length of stay and occupancy rate;
- hospital staffing levels, payroll expenses, staff mix;
- measures of quality, such as case fatality rates and JCAH accreditation status;
- hospital spending for plant and equipment.

The major independent variables in the model include:

- hospital characteristics such as:
  - ownership;
  - teaching status;
- area variables, such as:
  - demographic characteristics,
  - labor force variables;
  - health system supply characteristics;
  - third party health care coverage;
- information on regulatory programs, such as:
  - PSRO activity;
  - Certificate of Need programs;
- and finally, the variable reflecting the intervention--the presence of a prospective reimbursement program.

The study has relied very heavily on secondary data, although the task of putting together over ten years of data from various sources in a logically consistent fashion has been formidable. The remainder of the presentation will describe the content of the major analytical data files used in the study, the sources of data used to construct them, and the procedures used to process and edit these files to assure that they are logically consistent and comparable from year to year.

Data Sources for the Study

Two major analytical files were constructed:

- the hospital master file; and
- the county master file.

The hospital master file contains data for 2,673 hospitals (see Table 1). In order to be able to carry out certain analyses on a per capita basis--e.g., determining the impact of prospective reimbursement not only on hospital costs per day but also on

hospital costs per capita for a defined geographic area, the county master file was created. This file contains many of the same data elements as the hospital master file, but here the data are aggregated to the county level. The file contains data for each of the 1,317 counties, in which the 2,673 hospitals are located.

Many different types of data elements from various sources were used to create the Master File (see Table 2). The two most important sources were the American Hospital Association's (AHA) Annual Survey and the hospitals' individual Medicare cost reports (MCRs).

The AHA annually surveys all of the hospitals in the U.S. The survey includes a wide range of questions on hospital costs, financial status, volume of services provided, types of services available, and basic characteristics of the hospital such as ownership, bed capacity and teaching status.

The Medicare cost reports are annual, uniform reports that are required from every Medicare certified hospital and are submitted to the hospital's Medicare intermediary. Under the existing Medicare cost reimbursement method, these reports are used to determine the portion of each hospital's allowable costs that are payable by Medicare. The Medicare cost reports contain information on the number of services provided to Medicare and non-Medicare patients, the number and type of employees, and the hospital's expenditures and revenues.

In addition to the data obtained from the AHA Annual Survey and the Medicare cost reports, data for the hospital and county master files were obtained from a wide variety of other sources, as shown in Table 3. Particular data items were obtained from these sources and inserted in the hospital-year or county-year records as appropriate.

Data Editing and Imputation of Missing Values

The major efforts expended in creating the analytical files involved the abstraction and computer processing of the required data elements from the Medicare cost reports and the subsequent editing and imputation of data from both the AHA surveys and the Medicare cost reports to assure the logical consistency of data over time. This is particularly important and difficult to do when one takes a ten year series of annual reports, where specific data elements may have changed from time to time and attempts to put them together in a longitudinal data base where calculated annual change variables are of primary analytical importance.

As indicated earlier, the detailed financial data contained in the NHRS master files were collected from the information reported by the sample hospitals on the Medicare cost reports (MCR) for the period 1970 to 1979. During this period, the format of the standard cost reporting forms has changed and the MCR has grown in length as more information was requested. In addition, hospitals had the option of submitting data on nonstandard forms (such as computer output from hospitals MISs) where these utilized a format comparable to the Medicare supplied forms.

Abt Associates, Inc. (AAI) developed coding conventions to deal with each of these formats. Working from microfilm, staff transferred data from the MCRs to standardized coding forms. The transfer sheets were then keyed to tape, resulting in approximately 500,000 card images, representing a total of 25,471 hospital-years of data. AAI microfilmed and abstracted data from eight worksheet types. Together, these eight forms contain information on hospital characteristics, financial condition, and on the allocation of costs to reimbursable and nonreimbursable cost centers.

Once data were computerized, the next steps were editing and imputation of missing values. Computer routines were developed to detect potentially erroneous values and to impute missing data elements. The primary technique used to identify potentially erroneous data items was "temporal" editing, the comparison of each year's data with the values of that element immediately before and after. In general, a given data value was considered suspect if it was more than twice or less than half the average of its neighbors. If the item seemed to be in error based on this technique, the value was set to missing.

Following temporal editing, a second data cleaning step involved the comparison of different data elements within the same case for reasonableness and consistency. This "contrast" editing technique required the computation of ratios such as length of stay or occupancy rate, and the comparison of overlapping ratios, in order to determine which of a series of data elements was likely to be in error. The erroneous value was then set to missing. A third technique, "cross-validation" editing or the comparison of one hospital's values with those for other similar hospitals, was not used. We felt that it was more accurate to use the available time series for each hospital than to use values from one hospital to impute values for another hospital. In general, we relied more on automatic error detection rules than on subjective judgments based on visual inspection of printouts. This approach runs the risk of introducing errors (that is, replacing real values with imputed ones), but has the virture of being systematic and documentable.

Following the editing step, missing values were estimated where possible. If no more than two adjacent values in a hospital's time series were missing, linear interpolation was used to impute the values of the one or two missing entries. If missing entries occurred at the start or end of a hospital's time series, the two later (or earlier) values were linearly extrapolated to fill in one and only one missing entry.

During data editing, the accounting or reporting interrelationships of the data items were tested. The presence of imputed values in the data base has implications for the reliability of some of these algorithms. Since each variable was estimated based on the available data for that item, it was very possible that the derived values would not agree across each record. Consequently, following imputation, the contrast edits were rerun in order to detect anomalies introduced. Because some of the data items in the hospital master file were available from more than one source, (e.g., analogous items were contained in both the AHA Annual Survey and the

MCRs) alternative versions of some data elements were present that could be used for editing the data or for analyses of the content and completeness of the sources. These cross-source checks were used to choose the most consistent and reasonable set of items, but data from different sources were not mixed in one time series of observations.

The resulting analytical files have been described in detail in documentation prepared by the NHRS staff. This documentation is divided into three sections: a User's Guide to the data contains item specific definitions of all the elements in the files; technical documentation includes file layouts and data tape specifications; and an Appendix details the activities required to convert ten years of Medicare cost reports to machine readable form.

Other Data Used in the Study

The hospital and county master files were used to carry out the principal analyses in four of the impact areas:

- o revenue and expenditure;
- o volume of services;
- o staffing and payroll;
- o capital and investment.

Several other data bases were constructed for use in other analyses. A description of the four other data bases and an indication of how they were used in the analyses follows:

1. In order to study the impact of prospective reimbursement on access to health services, a longitudinal data base using the National Center for Health Statistics Annual Health Interview Survey (HIS) data was constructed. This data file contains aggregate health services utilization data for each of the 376 primary sampling units included in the survey for each year between 1973 and 1979. This represents the first attempt of which we are aware for putting together a longitudinal file from the HIS survey.

2. To study the impact of prospective reimbursement on quality of care, a longitudinal file for the years 1974-1979 was constructed to reflect in-hospital case fatality and 180-day post discharge case fatality for a sample of Medicare hospital cases. The cases included were for a set of "care sensitive" diagnoses that were judged by a panel of experts to be sensitive to changes in resource levels expended by the hospitals.

3. To study the impact of prospective reimbursement on health care utilization and expenditures for noninpatient types of health care services, a longitudinal county-based file for years 1974-1979 was constructed. This file contains per capita utilization, charges and reimbursements by type of service for a sample of Medicare beneficiaries.

4. To study the impact of prospective reimbursement on the organization and management of hospitals, a survey of NHRS sample hospitals was carried out by the AHA in 1981. This file contains the hospitals' responses about a range of organization and management issues.

Summary and Conclusion

We have attempted to provide some insight as to how a large amount of secondary data from many different sources was merged to create a longitudinal ten-year data base of hospital experience in order to evaluate a broad range of possible impacts of 15 different hospital prospective payment programs.

For those readers interested in the results of the NHRS, several reports and articles have already been published. These include a report based on case studies done in 1978, which compares the organization and operation of nine of the fifteen prospective reimbursement programs. 1/ Four articles have also appeared in the HCFA Review. These present some preliminary findings on the impacts of prospective reimbursement on hospital expenditure, volumes of services, payroll costs and employment, and the adoption and sharing of various hospital services. 2--5/These analyses were based primarily on data from the AHA Annual Survey for the period 1970-1978.

Reports of the results in each of the eight impact areas referred to earlier are currently being prepared. An executive summary of the major findings in all eight areas will be available by December 1983. In addition, Abt Associates will add two years of additional data to the analytical files in order to permit them to update, through 1981, evaluation results in four areas:

1. revenue and expenditure;
2. volume of services;
3. staffing and wages;
4. quality of care.

Data reflecting hospitals' experience through 1981 will be added from the AHA Annual Survey and Medicare cost reports as well as many of the other secondary sources. The Medicare cost report data to be added will be limited to those items that are available from HCFA in computerized form for the years 1980 - 1981.

In addition to updating the analyses, Abt Associates has been asked to develop a Hospital Data Book, utilizing the unique longitudinal Medicare cost report data to write a summary descriptive report on trends in hospital operating and financial characteristics and hospital utilization and reimbursement for the Medicare program. The types of measures that will be examined are shown in Table 4. This data book should be of great utility not only to the Medicare program, but to the hospital and health care industry in general.

Finally it is important to indicate that HCFA plans to make the NHRS data base available to the public so that it may be utilized for further research. The fully documented data base with the additional data for 1980-81 included will be delivered to HCFA in the summer of 1984.

Plans will be made to make the data tapes and documentation available to the public--probably through the National Technical Information Service. It is important to mention that the AHA annual survey data is owned by the AHA and that they make it available for use through contractual licensing arrangements. It is hoped that an agreement can be negotiated with the AHA that will allow the AHA data in NHRS files to be included in the public use tapes.

References

1. Hamilton, D. et al., Comparative Review of Nine Prospective Rate-Setting Programs. First Annual Report of the National Hospital Rate-Setting Study prepared for DHHS under Contract No. HCFA-500-78-0036, January 1980.

2. Coelen, C. and D. Sullivan. "An Analysis of the Effects of Prospective Reimbursement Programs on Hospital Cost." Health Care Financing Review:2:3, Winter, 1981.

3. Kidder, D. and D. Sullivan. "Hospital Payroll Costs, Productivity, and Employment Under Prospective Reimbursement." Health Care Financing Review:4:2, December 1982.

4. Cromwell, J. and J. Kanak. "The Effects of Prospective Reimbursement Programs on Hospital Adoption and Service Sharing." Health Care Financing Review:4:2, December 1982.

5. Worthington, N. and P. Piro. "The Effects of Hospital Rate-Setting Programs on Volumes of Hospital Services." Health Care Financing Review:4:2, December 1982.

Table 1: National Hospital Rate-Setting Study Sample, 1969-1979

Prospective Reimbursement Programs:	Number of Hospitals Total	Sample*
Arizona	69	62
Connecticut	36	35
Maryland	50	48
Massachusetts	122	121
Minnesota	164	152
New York	324	303
New Jersey	113	105
Western Pennsylvania	92	89
Washington	112	106
Colorado	83	77
Indiana	120	117
Kentucky	127	116
Nebraska	108	97
Rhode Island	15	13
Wisconsin	147	137
PR group total	1,680	1,576
Control Group**	4,730	1,095
Total--PR and control	6,412	2,673

*The NHRS sample was drawn by selecting a 25 percent simple random sample of the community hospitals in the 48 contiguous States that were operating for at least one year between 1969 and 1977 and then adding all remaining community hospitals in the study area in any year from 1969 to 1977. Due to problems of missing data, opening of a few new hospitals in 1978 and 1979, and miscellaneous other factors, the final sample is not a pure 25 percent sample in control States nor a complete census in study areas.

**NHRS sample hospitals in the remainder of the 48 contiguous States.

Table 2: Data Items Included in the NHRS Hospital Master File

1. Medicare Cost Reports
   a. statistics page
   b. balance sheet
   c. income statement
   d. patient revenue
   e. trial balance
   f. cost after stepdown
   g. inpatient routine cost
   h. Medicare reimbursement
2. American Hospital Association Annual Survey
   a. expense
   b. admissions
   c. beds
   d. inpatient days/outpatient visits
   e. facilities and services
   f. personnel
   g. ownership and control
3. AAI Blue Cross Survey
   a. number of Blue Cross discharges
   b. number of Blue Cross covered inpatient days
   c. Blue Cross hospital payments
4. PSRO Hospital Level Variables
   a. hospital covered by review in year
   b. date of binding review
   c. delegation status
5. Area Level Variables (SMSA/County/State)
   a. population characteristics
   b. physician distribution and health care system characteristics
   c. labor market
6. Regulatory Variables
   a. certificate of need dates
   b. professional standards review organization dates
   c. prospective reimbursement program dates

Table 3: Supplemental Data Sources Used In Creating the Hospital and County Master Files

Abt Associates, Inc. Case Studies
American Medical Association
Area Resources File
Blue Cross Survey of Plans
Blue Cross/Blue Shield Fact Book
Bureau of Health Insurance
Bureau of Labor Statistics
Department of Health Planning
Health Care Financing Administration
Health Insurance Institute
Joint Commission on Accreditation of Hospitals
National Cancer Institute
Office of Health Maintenance Organizations
Sales and Marketing Management Magazines

Table 4: Issues to be Presented in the Hospital Data Book

Expenses

- operating and nonoperating expense
- expense by types (e.g., inpatient routine, ancillary, outpatient)
- expense by function (e.g., salaries, interest, depreciation)

Revenues

- charges (total and composition by type of services, e.g., routine versus ancillary)
- reimbursement (total and composition by payer)
- nonoperating revenue (e.g., gifts and contributions)
- net income (from patient services and total)
- cost/charge ratios (routine, ancillary)

Financial Status

- return on equity
- debt and equity ratios (e.g., long-term debt/fund balances)
- liquidity ratios (e.g., current assets/current liabilities)
- total operating margins (e.g., net income/fund balances)
- receivables ratios (e.g., receivables/current assets)
- bad debt ratios/(uncollectibles/receivables)

Capital Formation

- gross and net fixed assets (undeflated and deflated) per hospital, per bed
- rates of capital formation: total, buildings and fixed equipment, movable equipment

DATA SOURCES FOR EVALUATING MEDICARE COMPETITION DEMONSTRATIONS: CURRENT KNOWLEDGE AND FUTURE NEEDS

Alan S. Friedlob, Health Care Financing Administration

On October 1,1983, the Health Care Financing Administration (HCFA) begins a 50 month evaluation of the efforts of approximately 50 prepaid health plans in 21 states to enroll Medicare beneficiaries and provide them with health care services on a capitated, at-risk basis. Throughout this paper the demonstration sites are referred to as "alternative health plans" or AHPs to broadly designate the many forms of organized health care delivery systems included in this study. This paper discusses the data requirements and research issues concerning the evaluation of this demonstration program.

The effectiveness of a competitive strategy to contain Medicare costs based on the growth of AHPs hinges on a series of sequential events:

- o AHPs perceive a financial incentive to enroll Medicare beneficiaries and to choose to participate in the Medicare program on a risk basis.
- o Medicare beneficiaries are willing to select an alternative health plan instead of standard, fee-for-service Medicare coverage.
- o Medicare beneficiaries choose to enroll in sufficient numbers to induce price competition among traditional fee-for-service providers and insurors and AHPs.
- o Increased competition to enroll Medicare beneficiaries among AHPs and traditional providers/insurors and cost-efficient changes in medical practice induced by this competitive activity, results in reduced Medicare program costs for HCFA and reduced out-of-pocket medical expenses for Medicare beneficiaries.

Until recently, the Medicare program was unable to contract with AHPs on a completely prospective payment basis. HMO reimbursement provisions contained in Section 1876 of the Social Security Act have often been criticized by prepaid providers for contradicting HMO management principles believed to be associated with HMO operational efficiencies.

Under Section 1876, HMOs receive interim monthly capitation payments based on either cost or risk contracts. The primary problem with these contractual arrangements is that the reimbursement and/or cost finding procedures used by Medicare differ substantially from an HMO's usual accounting procedures (i.e., rely on retrospective adjustment of costs). These contracts also may fail to provide the HMO with sufficient financial incentives necessary to generate savings or profit. For example, under the Section 1876 risk contracting option, Medicare reimbursement is based on a comparison of the HMO's actual costs with its Adjusted Average Per Capita Cost (AAPCC) determined on a retrospective basis. The AAPCC is HCFA's method of estimating what HMO Medicare enrollees would have cost under fee-for-service. If the risk-based HMO's costs are less than the AAPCC, it must share these "savings" with the Medicare program. The alternative health plan may receive savings of only up to 10 percent of the AAPCC. These HMOs are not required to provide additional services withg their savings.

To overcome these barriers to Medicare AHP contracting, Congress passed provisions in the Tax Equity and Fiscal Responsibility Act (TEFRA) in September 1982 authorizing prospective reimbursement under risk-sharing contracts with HMOs and other eligible alternative health plans at a rate equal to 95 percent of the AAPCC. Under this new contracting arrangement, which has yet to go into effect, if the Medicare payment (i.e., 95 percent of AAPCC) exceeds the AHP's "adjusted community rate" (ACR), the AHP must use this savings to provide its members with additional benefits or reduced cost-sharing. The ACR is prospectively determined by the AHP and should return to the AHP the same margin of profit or loss for its Medicare enrollees as for its under-65 commercial enrollees. The ACR represents the AHP's cost of doing business with respect to its Medicare enrollees. By contrast, under the Medicare competition demonstration HCFA does not restrict the AHP's use of the difference between 95% of AAPCC and the plan's actual costs.

The Medicare program based this "95 percent of AAPCC" prospective reimbursement approach for AHPS contained in TEFRA on eight Medicare prospective risk capitation demonstration projects that began enrolling Medicare beneficiaries in 1980. The eight demonstrations are located in Worcester, Massachusetts (Fallon Community Health Plan); Lansing, Michigan (Health Central); Minneapolis-St. Paul (HMO Minnesota, MedCenter, Share Health Plan, and Nicollet-Eitel); Portland, Oregon (Kaiser/Portland); and Marshfield Wisconsin (Greater Marshfield Community Health Plan). With the exception of Marshfield which ceased being a demonstration September 1982, the remaining seven sites continue to serve Medicare beneficiaries on a capitated at-risk basis and have a combined enrollment of approximately 40,000.

HCFA awarded a contract to Jurgovan and Blair, Inc. to evaluate these demonstrations in March 1981. The evaluation will end August 1984. The major evaluation objectives of the Jurgovan and Blair study are (Kahn et. al., 1983):

- o To measure HMO versus fee-for-service differences in utilization and cost patterns for Medicare beneficiaries, standardizing for population differences.
- o To assess the accuracy of HCFA's method of estimating what HMO enrollees would have cost under fee-for-service (i.e., for the AAPCC)
- o To measure the extent to which either favorable or adverse selection has occurred, and the cost impact of selection bias in enrollment.
- o To assess the cost-effectiveness of different marketing methods to induce Medicare beneficiaries to enroll in an HMO.

- o To assess the fiscal impact of the demonstrations for HCFA, for the HMO, and for beneficiaries.
- o To examine the organizational changes in HMO administrative and delivery systems conditioned by the addition of Medicare coverage.

Evaluation questions of interest to HCFA in the evaluation of the Medicare competition demonstrations build on the Jurgovan and Blair evaluation and can be divided into four areas:

- o Medicare competition-- Imapct an health services use,cost, and quality.
- o Medicare competition-- Beneficiary choice and AHP marketing
- o Medicare competition--Impact on the fee-for-service sector
- o Medicare competition between AHPs

The remainder of this paper discusses the evaluation research issues and data requirements of each study area.

Medicare Competition--Impact on Health Services Use, Quality and Cost

o Controlling Use Under Medicare Competition

The demonstration assumes AHPs will provide standard Medicare benefits, and possibly additional non-covered benefits, at a cost to Medicare that is less than what HCFA would experience under fee-for-service. AHP utilization is not only a function of efficiencies resulting from the AHP's response to a financial incentive, (i.e., the 95 percent of AAPCC capitation payment) but is also related to enrollee selection. Enrollee selection refers to characteristics of beneficiaries' health care needs (i.e.,health status) and demand (i.e., how an enrollee chooses to use the system) that are beyond an AHP's ability to control.

Variations in utilization rates between AHP enrollees and beneficiaries who remain in fee-for-service can only be partially accounted for by the underwriting factors used by Medicare in calculating the AAPCC (i.e., age, sex, welfare status, and institutional status). Eggers and Prihoda (1982) found that after adjusting the utilization experience of a non-enrollee comparison group for these underwriting factors, Medicare beneficiaries enrolled in the Kaiser-Portland and Fallon Community Health Plan HMO demonstrations had a 21 percent lower reimbursement rate for Medicare Part A and Part B services over a 4-year pre-enrollment period. Eggers' analysis for Marshfield showed no statistically significant differences in reimbursements.

In examining the ability of alternative health plans to control health care use and thus costs, the evaluator must first attempt to control for any enrollee selection effects. To this end, the evaluator will collect data on enrollees' and fee-for-service beneficiaries' health status to test whether the predictive validity of the AAPCC calculation is improved by introducing health status measures. Through a telephone interview, data on self-reported health status, physical functioning, medical conditions, and social support networks will be obtained from enrollees and comparison group non-enrollees. For enrollees, this interview will occur near the point of enrollment. The evaluator will collect health status data data over a period of months so as to be able to compare the health status of early joiners with beneficiaries who choose to enroll later.

This data will provide HCFA information on the feasibility and desirability of adding a health status adjustment to the AAPCC. Constructing such an adjustment presents two major challenges to health services researchers-- what medical conditions or functional limitation measures would improve the average performance of the AAPCC in predicting health care costs and how can these measurements be made with a minimum of time and cost ? The importance of a health status adjustment to the AAPCC is magnified if alternative health plans consider assuming risk for both acute and long-term care, as envisioned in the social/health maintenance organization model of prepaid health care for the elderly (Diamond, et.al., 1983).

To support utilization analyses, HCFA will supply Medicare Part A and Part B beneficiary-specific claims data reflecting enrollees' pre-enrollment utilization experience and health care use in fee-for-service control groups of Medicare beneficiaries. AHPs will supply post-enrollment use data for enrollees. Merging this use data with health status data, the evaluator will test hypotheses concerning the ability of AHPs to control use by Medicare beneficiaries whose health care needs are ostensibly greater and qualitatively different than prepaid health plans are accustomed to treating. For example, the lower health care costs of AHPs for enrollees under-65 are generally attributed to lower hospitalization. AHPs appear to achieve cost savings by controlling health care use through outpatient surgery, pre-admission testing, pre-screening of hospital admissions, and controlling elective surgery rates. After controlling for health status upon entry into the plan, do AHPs control health care use among persons over 65 through applying similar utilization control mechanisms ?

Of particular interest is how AHPs manage health care use among chronically and terminally ill beneficiaries. Fee-for-service Medicare expenditures are highly concentrated among these beneficiaries, with less than 5 percent of Medicare beneficiaries accounting for over 50 percent of yearly Medicare costs (Lubitz, et.al., 1981). As part of the overall utilization analyses, the utilization and cost-experience of AHP enrollees with chronic and/or terminal illness will be compared with comparable fee-for-service beneficiaries.

o Maintaining Quality of Care Under Medicare Competition

The assumption that AHPs achieve lower costs by underserving or skimping on quality is not generally supported by research. The quality of care in AHPs measured along structural, process, and outcome dimensions appears comparable to, if not slightly better than, the community average (Luft, 1981). Quality of care is not uniform across HMOs, it varies just as in the

conventional, fee-for-service setting. However, given the fiscal incentive for cost and utilization control in the Medicare competition demonstrations, HCFA is extremely interested in evaluating the relationship between increased price competition and the quality of care.

Through case studies, the evaluator will collect and analyze information about the quality assurance mechanisms AHPs use to monitor the care they deliver, documenting how AHPs' quality assurance mechanisms may change over time in response to caring for an aging and generally sicker population.

The evaluator will also directly determine the quality of care delivered to Medicare beneficiaries by AHPs. Applying currently acceptable methods for measuring quality of care (e.g., sentinel event, tracer disease, staging methodologies),, the evaluator will test the hypothesis that there is no relationship between Medicare competition and under-service resulting in deleterious health outcomes.

The direct determination of quality of AHP medical care will focus on two broad disease categories: (1) medical conditions in which hospitalization is generally required to correct the problem but which remain elective. These conditions, while interfering with quality of life, are not directly life-threatening (e.g., cataract, hernia, joint deterioration); (2) those conditions where lack of appropriate treatment may result in untimely death (e.g., cancer, diabetes, hypertensive disease, pneumonia). To perform these analyses, the evaluator will have to rely on inpatient and ambulatory medical record abstracting at a sample of AHPs while comparing these results with normative criteria and data contained in secondary data bases in the fee-for-service sector. Study resources make it impossible to collect medical records data on control group beneficiaries.

The quality of care in AHPs will also be evaluated from the perspective of beneficiaries' satisfaction. The TEFRA legislation requires HCFA to conduct a study evaluating the extent of, and reasons for, the termination by Medicare beneficiaries of their AHP memberships. In this study, Congress mandated HCFA to examine the quantity and quality of care provided in AHPs in comparison with the quality of such care when provided on a fee-for-service basis.

In the context of a Medicare competitive strategy, low disenrollment rates can be taken as an indicator of aggregate enrollee satisfaction. Under competition, the decision to remain in the plan can be interpreted to mean that the beneficiary perceives the AHP as relatively superior to available alternatives. Theoretically, plan disatisfaction could ultimately lead to disenrollment, a return to fee-for-service Medicare, or selection of another AHP, if available. If the number of disenrollees grew to a sufficient number, the competitive model would hypothesize that a new AHP would emerge with a potential market for these disaffected beneficiaries.

The evaluator will conduct a limited telephone survey of beneficiaries who disenroll from AHP demonstrations to validate information collected by the AHPs on enrollee grievances and reasons for disenrollment. HCFA's experience with its risk-based HMO demonstrations indicates Medicare disenrollment rates of less than 5 percent, excluding deaths.

o Containing Medicare Costs Under Competition

AHPs are reimbursed by HCFA at 95 percent of the "adjusted average per capita cost" (AAPCC). Continued reliance on 95 percent of AAPCC reimbursement ignores the actual cost and utilization experience of the AHP, which is a function of both the efficiency of the plan and enrollee selection effects. Under a Medicare competitive strategy, there is no reason for HCFA to adhere to the 95 percent of AAPCC reimbursement approach in future years. For example, based on experience gained from these demonstrations, the average cost of standard Medicare coverage may be determined to be 85 percent of AAPCC. If this were the case, HCFA might lower the amount of the capitation. Alternatively, HCFA might allow the market to set the rates through a competitive bidding process. Price and non-price (i.e., quality) competition would drive AHPs to alter benefit packages, adjust premiums, or impose cost-sharing for selected services. If price/quality changes are unacceptable to the beneficiary, disenrollment would follow, resulting in loss of plan revenue.

Evaluating the cost savings potential of a Medicare competition strategy, (i.e., determining the actual costs experienced by the AHPs in providing services) may prove the most difficult task in the study. The objective of this analysis is to assist the Medicare program in determining what a fair market price is for standard Medicare coverage when purchased from an AHP. HCFA wants to identify the range of actual costs of standard Part A and Part B benefits provided by AHPs as a percentage of an AHP's AAPCC revenue.

Each plan varies in the way it accounts for costs. Except in plans serving Medicare beneficiaries exclusively, AHPs allocate costs between Medicare enrollees and those members who are under-65. Allocation methods will differ among plans. Risk-sharing arrangements of plan components which have direct implications for measuring plan costs will also vary by plan. For example, physician risk-sharing arrangements in group model HMOs may make it very difficult for the evaluator to identify the actual costs of physician services since these amounts are found in the accounts of the capitated medical group and may not be reflected in the capitation paid by the HMO to the group. Without access to the medical group's actual costs, excessive use or high unit costs may surface as increased premiums ultimately passed on to the beneficiary. Determining the relationship between these price increases and actual costs can prove extremely difficult.

HCFA has not required the demonstration AHPs to submit uniform cost reports. The evaluator faces the task of developing broad cost categories into which cost data can be aggregated. Alternatively, the evaluator will compare expected to actual AHP use, relying on price/cost weights such as relative value units uniformly applied across sites to "back into" measuring costs.

Medicare Competition -- Consumer Choice and AHP Marketing

There is considerable interest in the question of who chooses to join an AHP and why they select one plan over another. This issue is important because of the possibility that subtle differences between people enrolled in AHPs and those who remain in fee-for-service may explain some of the differences in AHP cost and utilization. As part of the Jurgovan and Blair study, Research Triangle Institute conducted a household interview survey of 3000 Medicare beneficiaries (i.e., HMO enrollees compared with non-enrollees) which covers six of the eight demonstration sites; Fallon, Marshfield, and four HMOs in the Twin Cities. Specific factors underlying beneficiary choice which were examined in this survey include:

- The role of health insurance purchased by Medicare beneficiaries to supplement Medicare's coverage (i.e., "Medi-Gap" policies) in the enrollment decision. Do HMO enrollees retain Medi-gap coverage after they join the demonstration? How similar is the Medi-gap coverage among non-enrollees to the pre-enrollment supplemental coverage of enrollees? For those HMO enrollees who had supplemental coverage and gave it up upon enrollment, what is the net savings (cost) to the beneficiary of switching to the HMO?

- Understanding HMO enrollment decisions among Medicare beneficiaries. To what extent do Medicare beneficiaries compare the HMO choice with other available health insurance options, including standard Medicare coverage? What are the main reasons Medicare beneficiaries enroll in these HMOs? How aware are non-enrollees of the demonstration program? If these non-enrollees considered joining the demonstration, what are their reasons for choosing not to enroll?

- The role of usual source of care in the enrollment decision. Is there a significant difference between the proportion of enrollees and non-enrollees with a usual source of care? To what extent were enrollees' usual source of care physicians' associated with the HMO they subsequently joined? What proportion of non-enrollees would have to give up their usual source of care in order to join an HMO?

In its collection of health status data and emphasis on studying beneficiary choice in market areas where two or more AHPs will compete to enroll Medicare beneficiaries, the evaluation of the Medicare competition demonstrations extends this initial research of beneficiary choice. A beneficiary choice survey will be administered to those AHP enrollees and comparison group non-enrollee providing health status data. This telephone survey will occur in approximately half of the market areas in the demonstration, involving upwards of twenty AHPs. The availability of AHP enrollee use data with which the survey data will be linked and the opportunity to study plan choice where beneficiaries have two or more options will influence final survey site selection. The sample size of beneficiaries included in this survey is sufficient for the evaluator to develop a general model of AHP beneficiary choice that controls for the idiosyncratic characteristics of AHP market areas.

In addition to better understanding the plan selection process from the beneficiary's perspective, the evaluator will describe the marketing strategies used by AHPs and examine the relationship between these Medicare marketing approaches and beneficiaries' demand for the AHP option. Using case study and quantitative marketing research methods such as content analysis of advertising literature and focus groups of beneficiaries to investigate the relative importance of plan attributes in the enrollment decision, the evaluator will address the following questions:

- What marketing approaches are the most cost-effective (i.e., greatest number of enrollees less drop-outs for the fewest dollars expended)?

- Assuming market segmentation exists among Medicare beneficiaries, do AHPs pursue different segments of the market?

- In market areas where two or more AHPs compete, how do advertising expenditures vary among these plans? What market conditions are associated with aggressive advertising campaigns? With situations where marketing is not aggressive?

- What does the information contained in Medicare AHP advertising convey to the prospective enrollee ? Does the AHP's message neutrally inform beneficiaries about the AHP option or pointedly influence plan selection ?

- What percentage of Medicare enrollees represent market saturation ?

Medicare Competition-- Impact on the Fee-for service Sector and Competition Between AHPs

There has only been limited research on the effects of competition in the health care sector. The evidence on competitive effects of HMOs is inconclusive (Luft, 1981). Some data suggests HMOs lead conventional fee-for-service providers to reduce their hospital use, while other studies offer alternative explanations for this finding. HMOs have not shown a much lower inflation rate than conventional insurors in the same market areas and expenditures for medical care are not markedly lower in areas with substantial HMO enrollment (Langwell and Pauly, 1982).

One reason HCFA has sponsored the Medicare competition demonstrations is to test whether AHPs achieve a large enough market share to create a competitive effect on the traditional fee-for-service system. The number of beneficiaries who must choose AHPs in order to stimulate price competition sufficient to prompt traditional insurors and providers to become more cost-efficient or to develop their own AHPs is not known. Evidence from three of the four Twin Cities' HMO capitation demonstrations indicates that plans

may expand their Medicare membership slowly fearing adverse selection. Many HMOs have not been able to achieve large market shares quickly among persons under 65 who belong to group contracts.

HCFA expects that the effects of AHP Medicare enrollment on the fee-for-service sector can only be quantified where plan market penetration is high. It is probably unlikely, with the possible exception of three or four market areas, that market penetration of AHPs among Medicare beneficiaries will exceed 10 percent of eligible beneficiaries.

The data needs for studying issues of Medicare competition's impact on the fee-for-service sector and on competition between plans are difficult to specify because standard micro-economic models may not adequately reflect this market behavior. The evaluator's first task will be to construct such a model based on available theory and apply this model to formulating hypotheses concerning the following questions:

- What impact does Medicare competition have on areawide Medicare costs?
- What impact does Medicare competition have on the economic behavior of hospitals ? on the pricing of physician services ? on the market for supplemental insurance coverage?
- Will competitive pressures from AHP enrollment of Medicare beneficiaries result in fee-for-service insurors (i.e., Blue Cross/Blue Shield and commerical insurors) or providers (i.e., hospitals) forming AHPs (e.g., growth in preferred provider organizations)?

In contrast with examining AHP use, cost, quality, and beneficiary choice, the unit of analysis in this study area shifts from individual beneficiaries to market areas and measures of hospital, physician, and insuror performance examined over time. The evaluation will attempt to measure the impact of AHP competition on economic indicators of the medical marketplace such as hospital occupancy rates and patient days, changes in physician pricing patterns and Medicare assignment rates, and the number of Medicare supplemental insurance policies in force.

Measuring change in these indicators and attributing variation to Medicare competition may be complicated by the confounding effects of at least two major alterations in the Medicare program occurring during the course of the evaluation. These changes are the national implementation of DRG-based prospective hospital reimbursement and the TEFRA provisions which will allow 95% of AAPCC risk-contracting with AHPs. Many market areas that presently do not have risk-based AHPs marketing to Medicare beneficiaries may attract such plans once TEFRA is implemented. The entry of TEFRA AHPs into the Medicare market will limit the ability of the evaluator to use the concept of comparison market areas (i.e., no risk-based AHPs) against which the effects of AHP competition on the fee-for-service sector can be assessed. While DRG-based hospital reimbursement will be implemented nationally, the effects of DRGs on individual hospitals and on particular regions or market areas is unknown. It may prove particularly difficult to disentangle the effects of DRGs on hospital performance from the effects of Medicare AHP market penetration, particularly where beneficiary enrollments are particularly high.

In addition to examining the impact of Medicare competition on the fee-for-service sector, it is important to understand how AHPs compete with each other. Using a case study approach, the evaluator will study how AHPs recognize the impact of their actions on their competitors. Over the course of the 50 month evaluation, the evaluator will track whether certain types of AHPs consistently initiate changes in premium, benefit package, or quality of product while other AHPs characteristically react to change. Once the TEFRA legislation is implemented, the evaluator will examine the demonstration AHPs response to the pricing policies and benefit packages of these new entrants into the Medicare market. Through interviewing key AHP personnel and leaders in the market area, the evaluator will also try to understand how AHPs maintain their competitive position. Physician and institutional reimbursement practices, organizational sponsorship, AHP size, and utilization control mechanisms will be investigated as to their effect on an AHP's ability to remain competitive.

Summary

The Medicare competition demonstrations have been designed as a national test of the desirability and feasibility of a major new policy direction for the Medicare program; a shift from fee-for-service and cost-reimbursement to pre-paid capitation arrangements with organized health care delivery sysytems. The evaluation of the Medicare competition demonstrations creates a number of technical and administrative challenges. Primary data from a beneficiary choice and health status survey will be linked with secondary data sources obtained from the Medicare program and a diverse group of alternative health plans. The alternative health plans will not maintain their use and cost data in a uniform format and difficulties can be anticipated in mapping common data elements across sites.

In addition to the quantitative analysis of the data bases constructed for examining AHP use,cost,quality, and beneficiary choice, many of the insights to understanding Medicare competition will come from carefully constructed qualitative case studies which critically document the dynamics of AHP organizational and market behavior, using these more qualitative impressions as a context for interpreting quantitative findings.

REFERENCES

Diamond,L., L. Gruenberg, and R. Morris. "Eldercare for the 1980s: Health and Social Services in One Prepaid Health Maintenance System." Gerontologist 23(2): 148, April 1983.

Eggers,P. and R. Prihoda. "Pre-enrollment Reimbursement Patterns of Medicare Beneficiaries Enrolled in 'At-Risk' HMOs." Health Care Financing Review 4(1): 55-73, September 1982.

Kahn,H.A., R. Leighton, and A. Friedlob. Summary of Observations , Medicare HMO Demonstration Project. Health Care Financing Administration, Office of Research and Demonstrations Working Paper, March 1983.

Langwell,K.M. and M. Pauly. "Research on Competition in the Financing and Delivery of Health Services: Problems and Prospects." Invited Paper, 1982 Meeting of the American Public Health Association, Montreal, Canada, November 1982.

Lubitz,J., M. Gornick, and R. Prihoda. "Use and Costs of Medicare Services in the Last year of Life." Health Care Financing Administration, Office of Research and Demonstrations Working Paper, August 1981.

Luft,H. The Operations and Performance of Health Maintenance Organizations: A Synthesis of Findings from Health Services Research. National Center for Health Services Research Contract Report Series, October 1981.

THE INTEGRATION OF SOCIAL SURVEY AND PROGRAM DATA TO EVALUATE THE MUNICIPAL HEALTH SERVICES PROGRAM

Tony Hausner, Health Care Financing Administration

Ron Andersen and Gretchen Fleming, University of Chicago

INTRODUCTION

In 1978, the Health Care Financing Administration (HCFA) awarded a contract to the University of Chicago to evaluate the Municipal Health Services Program (MHSP). This paper will describe the different data systems that are being used to evaluate this program. The purpose of this discussion is to indicate that an extensive array of data is needed to comprehensively evaluate MHSP and to describe the procedures used to integrate the data, particularly social survey data and program data. The first section will describe the MHSP.

Municipal Health Services Program

In 1978, the Robert Wood Johnson Foundation (RWJF) and the Health Care Financing Administration (HCFA) agreed to collaborate on the Municipal Health Services Program (MHSP). The chief aim of MHSP is to assist cities in providing health care to the medically underserved by expanding their public health department programs. The emphasis is on delivering primary care and preventive services in ambulatory clinic settings. The unique feature of this program is the involvement of city governments. A major goal of the program is to reduce the total cost and utilization of inpatient and emergency room services.

RWJF Activities

The Robert Wood Johnson Foundation awarded $15 million to five cities to conduct these programs. The cities are Baltimore, Cincinnati, St. Louis, Milwaukee, and San Jose. The foundation grants of $3 million to each city for up to seven years partially covers administrative expenses and bad debts. The funds are not to be used for capital expenditures, renovation or program operations. Each city operates three to four new or expanded clinics which use primary care physicians, physician assistants, and nurse practititioners to provide comprehensive primary and preventive services. The RWJF is administering this program through a contract with Johns Hopkins University and is supporting Columbia University in an evaluation of MHSP's effect on the organization and financing of health services.

HCFA Activities

HCFA is responsible for administering Medicare and Medicaid waivers to the cities and for administering the University of Chicago's Center for Health Administration Studies (CHAS) evaluation of the impact of MHSP on access, use, and costs of health services. The waivers represent exceptions to the provisions of the Medicare and Medicaid programs.

The Medicare waivers which HCFA authorized from 1979 to 1984 include: elimination of coinsurance and deductibles; cost-based rather than fee-for-service reimbursement; direct reimbursement of nurse practitioners and physician assistants, rather than requiring the services to be performed incident to a physician's services; and reimbursement for 16 currently non-covered primary and preventive services.

HCFA approved Medicaid waivers for 1981 to 1984 for four of the five States: Maryland, Wisconsin, Missouri, and California. The waivers for three States consist of changes in reimbursement procedures from their traditional fee-for-service procedures to reimbursement on a cost basis. The waivers for two States consist of some limited additional services; e.g., outreach. Ohio did not apply for a Medicaid waiver, since it is already reimbursing Cincinnati clinics on a cost basis under State law.

Evaluation Hypotheses

To provide further background to the data sources, this section will describe the key evaluation hypotheses (see Table 1) and the research design.

Table 1

Key Evaluation Hypotheses

1. Cost - MHSP will achieve decreased per capita (i.e., per person) cost for each payor, i.e., Medicare, Medicaid, private insurance, clients' out-of-pocket expenses and municipal governments.

2. Primary Care Patterns - MHSP will promote use patterns which include more preventive services, greater use of primary care practitioners (as compared to specialists), and greater use of physician assistants and nurse practitioners (as compared to physicians) relative to other sources of care.

3. Shifts in Use Patterns

   a. MHSP will promote an increase in the use of municipal clinics and a decrease in the use of hospital inpatient services for all client groups.

   b. MHSP will promote a shift in the use of services from hospital emergency rooms and outpatient departments to the municpal ambulatory clinics.

4. Quality - MHSP will facilitate improved access to, continuity of, and quality of ambulatory care services as measured by variables such as travel time, identification of a regular source of care, and client satisfaction, respectively.

Research Design

To address these research hypotheses, this study has implemented the following research design features:

1) a before and after design, also referred to as a baseline follow-up design; and 2) a comparison group design consisting of user and nonuser groups.

Baseline - Follow-up Periods

CHAS has divided this study into three time periods: T1 from 1979 to 1980; T2 from 1980 to 1981; and T3 from 1981 to 1982. T1 is referred to as the baseline period when the program was in the early stages of operation. T3 is the follow-up period when the program became fully operational at select clinics. The major comparisons in this study will be between T1 and T3, particularly with the social survey. T2 is an interim period for which program data is available that allows for trend effects to be observed from T1 to T2 to T3.

Comparison Groups

The comparison groups consist of persons who are regular users of MHSP (program users); persons who use other sources of health care, such as hospitals, private physicians, etc. (non-MHSP users); and persons who do not use any health services during the study period (nonusers).

DATA SOURCES

To address the research hypotheses previously described requires data on individual patients. This study, therefore, focuses on patient level data. However, as discussed later, this study also involves analyses of data aggregated only at the clinic level.

The secondary sources of data such as Medicare and Medicaid claims files and hospital and clinic records which are available in this study only partly address the above hypotheses. Therefore, CHAS has focused this evaluation on a baseline-follow-up social survey design, which more fully addresses the above hypotheses. Nevertheless, this study uses the secondary data either by integrating them with the survey data or by analyses that complement the survey. CHAS has already collected most of the survey and secondary program data that will now be described.

Social Survey

The survey consisted of questions on demographic characteristics, health status measures, use of different health services, the cost of these services, the sources of care, access measures, satisfaction measures, and sources of payment. There were also questions on the more unique parts of this program, such as the use of preventive services, health education services, and the use of physician assistants/nurse practitioners.

Chilton Research Services, the subcontractor to this study, administered the survey to a sample of residents living in the service area surrounding selected MHSP clinics, one in each city. Chilton administered the survey through telephone interviews, using random digit dialing within the telephone exchange of the service area. A screening questionnaire permitted oversampling of MHSP users. Chilton asked each respondent to answer questions on all family members. For phase one, Chilton interviewed about 6,000 respondents who provided data on a total of about 18,000 family members; and for phase two, they interviewed 5,000 respondents who provided data on about 15,000 family members.

Integration Procedures

The procedures that integrate the social survey and program data consist of imputation techniques and verification.

Imputation

Andersen, Kasper, and Frankel (1976) have shown that many social survey respondents are unable to accurately estimate health care expenditures. Medicaid beneficiaries, in particular, do not generally see copies of the submitted bills. CHAS, therefore, has developed imputation techniques for estimating expenditures from the respondents' utilization estimates. The utilization estimates are combined with estimates of the unit costs for these services from a number of other sources such as the American Hospital Association, Maryland Health Services Cost Review Commission, Medicare fee screen, RWJF quarterly reports, the National Medical Care Expenditures Survey, Current Price Indices and clinic fee schedules. The respondents' expenditure estimates are then compared with the imputed estimates and using a set of decision rules a "best estimate" of patient expenditures is determined.

Verification

Verification is another procedure that integrates the social survey data with program data. CHAS conducted tests on T1 data to verify the utilization and expenditure estimates obtained from the social survey. To verify, CHAS compared these estimates with utilization and expenditure data obtained from Medicaid claims files and hospital and clinic records.

This section will now report the results of the verification from two State Medicaid agencies. The States located data on about 60 percent of the social survey respondents. CHAS (1983) found that the total per capita costs were generally consistent between the social survey estimates, the bill files and national per capita figures. The only discrepancy involved hospital inpatient costs, for which the social survey indicated much higher figures than Medicaid bills. CHAS felt that this difference was due primarily to more stays being reported in the social survey than in the bills. The per diem inpatient costs were similar for the two data sources. The social survey probably reported more stays than the Medicaid bill file because of name or Medicaid ID number changes, and stays that were not billed to Medicaid.

These findings suggested that the "best estimates" of Medicaid expenditures obtained from the social survey were reasonable and did not need to be adjusted. However, because of the limited proportion of cases for which Medicaid claims were found, CHAS plans to analyze Medicaid data on the universe of beneficiaries

in two of the service areas for the T3 period as discussed below.

CHAS was only able to perform the verification of hospital and clinic utilization and expenditures on a limited number of hospital and clinic records. Of the available records, CHAS found that the social survey and record data were comparable on utilization figures, but that the social survey estimates were higher and appeared to be more accurate than the record figures for average charges per persons and unit charges for specific services.

These findings suggested that the social survey estimates of clinic and hospital utilization and expenditures were valid and did not need to be adjusted. Since the clinic and hospital data did not add information to the survey, we decided not to collect record data in T3.

Complementary Analyses

In addition to integrating the data, this study analyzes secondary data from claims files and clinic level reports to complement the social survey.

Claims Files

As a complement to the analyses of social survey data, CHAS plans to analyze Medicare and Medicaid claims data for a sample of beneficiaries. These analyses will provide more detailed information on the use and costs of health services for the Medicare and Medicaid programs. The samples will be selected from the five survey service areas. However, in contrast to the survey, data will be collected on the universe of beneficiaries in the service area.

CHAS plans to compare the total use and costs of Medicare and Medicaid services for beneficiaries who used the waiver program with beneficiaries who used only the regular Medicare (and Medicaid) program and, thus, received care from non-MHSP providers. Since the beneficiaries who used the waiver program may also have received care from non-MHSP providers, the waiver bills will be combined with the regular bills for these persons to determine their total care.

Clinic Level Data

This study will also analyze data that is only available aggregated at the clinic level in contrast to the patient level data so far described. Such data is available from Medicare cost reports, program statistics contained in the State Medicaid Agencies quarterly reports and the RWJF program management reports. These sources provide information on program issues such as: the number of clients served; the extent to which different services are used; and the actual costs of providing the services. This data will be used to study trends over time, and differences among clinics, etc. These data, therefore, complements the assessment of MHSP provided by the patient level data.

Data Array

Table 2 provides a synopsis of the different data sources according to the different time periods of the study. This table demonstrates that data has been collected from many different sources for this evaluation.

Table 2
Data Sources by Time Period

	Time Periods		
Data Sources	Baseline T1 1979-1980	Interim T2 1980-1981	Follow-Up T3 1981-1982
Community Social Survey (and National Imputation Sources)	x		x
Medicare			
Cost Reports	+	+	+
Waiver Bills	o	o	o
Regular Bills		x	x
Medicaid			
Program Statistics			+
Waiver Bills			o
Regular Bills	x		o
Clinic and Hospital Records	x		
RWJF Program Management Reports	+	+	+

x - sample
o - universe
+ clinic level

SUMMARY

In summary, this presentation has indicated that an extensive array of data has been used to evaluate MHSP. This extensive array is needed to address: 1) the many research hypotheses developed for this study; 2) the many components involved in this program, such as Medicare, Medicaid, the cities, the clinics, the foundation, the clients, and their respective contributions; and 3) the complex aspects of the research design which are needed to deal with the research hypthoses.

This paper has indicated that the evaluation focuses on the use of a before and after social survey data collection strategy. The survey is integrated and complemented by individual patient level data from claims files and hospital and clinic records. It is further complemented by clinic level data to provide a comprehensive assessment of the effects of MHSP on access, use, and costs of health services.

As indicated earlier, CHAS has already collected most of the data for this study; the analyses of T1 and T2 data are complete and the analyses of T3 data and comparisons between T1, T2, and T3 data are in process. All of these analyses will provide tests of the different hypotheses formulated to evaluate MHSP. The final report for this study is due in early 1984.

NOTES

1. The work described in this paper was performed by the University of Chicago, Center for Health Administration Study, under HCFA contract No. 500-78-0097 and RWJF grant No. 6798. Dr. Hausner is the government project officer for this contract.

REFERENCES

1. Andersen, R.A., J. Kasper, M.R., Frankel and Associates: Total Survey Error. San Francisco: Josey-Bass, 1979.

2. Center for Health Administration Studies, University of Chicago: Evaluation of Municipal Health Services Program Phase II Report. Submitted to HCFA and RWJF under HCFA contract No. 500-78-0097 and RWJF Grant No. 6798, 1983.

Standby Presentations

INFORMING HEALTH PROMOTION POLICY MAKING

David Ross Netherton, Consultant in Health Policy Research

The evolution of health promotion policy has come as a response to developments in medical and health sciences, an increasing level of education and awareness among the public, and the acceptance by government of the widespread feeling that one is entitled to good health as a basic right. Health promotion is now a fundamental element of public health concerned with fostering positive changes in individual habits and the environment which affect health, including such issues as pregnancy and infant care, family planning, immunizations, surveillance and control of infectious diseases, occupational safety and health, accident prevention, smoking cessation, fluoridation and dental health, high blood pressure control, control of stress, improved nutrition, and physical fitness and exercise[1]. Those policymakers concerned with areas of behavioral change that reduce risk to health are seeking more complete information on many aspects of individual activity.

Yet the availability of an increasing variety of data has not by itself met the information needs of policymakers. The more familiar and accessible sets of data -- for example, required data, such as vital events, census information or state health planning data -- can be over-utilized or poorly applied to questions of public policy. Specific data on individual attitudes and practices are what is commonly needed by programs in health education. Information at the community level also becomes more crucial under the "new federalism" system of awarding block grants to states. Innovative uses of customary data together with the development of new data resources can enhance the policy process and strengthen support for prevention programs.

Policy decisions rest on available information. To reduce the likelihood of a wrong policy decision, the decisionmaker wants to consider that information which best answers specific questions. There are several things that can limit the usefulness of data: including how, under what conditions and for what purpose the information was collected, and whether the information can be corroborated by other existing data or data collected in the future. Essential questions about data quality, ease of access and utility of application need to be asked early in the policy process.

Issues in Policy Development

The concept of health information will differ according to the viewpoint of the user. If we consider the perspective of the data producer as largely operational, and that of the manager as tactical, then the policymaker's perspective is strategic[2]. Reaching agreement on and gaining political support for a particular policy choice usually involves defending that choice over alternatives on the basis of some objective criteria, and being able to predict (or at least strongly suggest) an outcome. The policymaker begins with a clear and detailed definition of a problem in its context.

There are three general issues to keep in mind in developing health promotion policy initiatives. The first is to have a clear understanding of the nature and magnitude of the problem. Second, to be reasonably certain that the policy initiative will have the desired impact. And third, to be assured that a mechanism exists to evaluate that impact. A brief discussion of each issue follows.

It is impossible to enact a successful and efficient policy without a clear understanding of the nature and magnitude of the problem one is trying to solve. To understand the problem, the decisionmaker most often turns to existing information, looking for data on its natural history: how the problem occurs, when and how often it occurs, the numbers and kinds of persons affected, and some insight into trends occuring over time. These dynamics are often an important clue to the most effective intervention strategy, or even whether it is more prudent to allow the situation to develop and change according to its own character.

Before undertaking a policy the decisionmaker should be reasonably confident that a new policy or program will succeed. The sorts of questions to be addressed here include: Is the relationship between symptoms and causes of the problem well understood? Has this initiative been tried before by others, and, if so, was it successful? Is there enough time available for the program to work? Can this initiative be targeted to those most affected by the problem? What are the costs and benefits of doing nothing? Do I have the necessary political support to implement this policy?

A problem may be well-defined, a population at particularly high risk revealed, and a new policy initiative supported, and still the result may be as dismal as if no planning at all were done. A well-executed program having no chance of success -- either because it occurs at the wrong time, does not address the underlying causes of the problem, or cannot be sustained long enough to have an impact -- is often no more effective than having no program at all.

At some point in the experience of policymaking someone usually asks the question: How do you know that your policy is working? (Such questions often seem to arise in direct proportion to the amount of resources committed to the policy.) Decisionmakers need to have some information about a policy that has consumed

resources, and the ability to measure the outcome of a program ought to be a consideration in all but the most unusual circumstances.

The policymaker should know how he can demonstrate that an initiative has worked, or whether he can be blamed if it has not. It will also be important to know how long he will have to wait before seeing the impact of his policy. Carefully weighing the applicability of a policy or program evaluation may also help avoid falling into the trap of promising too much too soon.

To summarize, health promotion policy initiatives are best undertaken when three broad considerations have been resolved:

- A clear understanding of the severity and magnitude of the problem, including a(n)
  - knowledge of absolute numbers involved,
  - knowledge of intensity and rates of the problem,
  - ability to identify the population at highest risk,
  - understanding of the secular trends and economic impact,
  - ability to separate the perceived and real problem;
- The power to alter an outcome, including a(n)
  - understanding of the sources and causes of the problem,
  - ability to target resources to those at high risk,
  - understanding of the consequences of doing nothing,
  - knowledge that time needed for the intervention to work is available; and
- The applicability of an impact evaluation, including a(n)
  - knowledge of how long one will have to wait to feel the impact of a policy,
  - knowledge of who is most interested in an evaluation,
  - understanding whether the results can be seen over a long period of time,
  - ability to demonstrate whether a policy has worked.

Developing Resources for Health Promotion Data

The pace of data collection is accellerating. The base of industrial and commercial data becoming available is increasing, with particular attention to environmental data, occupational health and safety studies, employee health and fitness programs and specific hazard monitoring studies. All this is added to the burgeoning national and regional information widely distributed. When policymakers consider their resources, emphasis should be placed on types of information that describe the situation and the problem environment as well as information that can suggest some response.

Federal and state government agencies, municipalities, commercial businesses and non-profit organizations all collect data for two basic reasons: first of all to comply with the law in reporting mandated information, and second, for specific organizational uses.

Certain data are required of government agencies to produce national statistical measures which in turn are standardized and made widely available. These data include vital events (births, deaths, marriages), reportable injuries and conditions, communicable diseases and census data. Other specific information is also required of state governments, such as hospital utilization and licensure reports, rate setting data, records of internal commissions and state and regional health planning data. Protective services is a newer category of reportable data available at the state level and of special interest in health promotion policy.

Mandated information constitutes what might be called the traditional or customary data resources for health promotion policymakers. Each set of data is accessible and frequently used but cannot in all cases resolve the issues previously outlined as policy development. Because much of this information appears in a standardized form, it is often aggregated according to less useful categories and is not very flexible. Accordingly, one of the best uses of these data when applied to health behavior or self-reported data is to use its high reliability to verify other, more specific but less firm information. Under the right conditions, these mandated data may be used either as a primary source of information, or as a corroborating source for new information collected by organizations with special interests.

Data collected for uses other than those required by law have many purposes and are frequently accessible by policymakers. National institutes and service organizations, for instance, are valuable sources that routinely produce data on specific health issues, regions and populations. Registries of cancer and systemic diseases generate large databases and may contain regional or community evaluations and summaries. Hospital associations are now emerging as sources of monitoring data including discharge information, patient origin, and case-mix and utilization studies.

In an effort to facilitate its use, certain agencies have been created at the state level in order to compile, organize, store and distribute data. The Center for Massachusetts Data, for example, is an executive office of the governor which catalogs and makes available data contributed from various public and private sources.

A major reason for the rapid increase in the sheer volume of data available to policymakers is simply that they have found they can best define a problem and focus on probable responses by generating new data according to specific interests. Some of these data are produced

routinely as part of a health promotion program. Program data are usually concerned with measures of process, impact and outcome, and are normally directed toward a single activity. However, this information may be used to verify other data or predicted outcomes of related programs. For example, information on an environmental quality monitoring program may effectively be applied to questions of occupational safety and health, or surveillance of infectious diseases. Evaluation data are a standard component of most programs and will tend to have a low marginal cost.

Although there are limitations in terms of control, selection and procedures, the informal observational studies and interviewing methods can be used to reflect points of interest in the larger community. An increasingly popular methodology related to these two is survey research. This design has the favorable features of lower cost, short time requirements and the ability of agencies to standardize research procedures. There is usually a generalizability of results and a great deal of information collected in a short time, allowing for several phases of study over a longer period. Survey methodology can be correlated with national measures using familiar sources, and data from small groups can be applied to much larger communities and even states[3]. Survey data at the federal level have also increased during the last ten years, in both volume and type. More than a dozen datasets are now available as part of the National Health Survey[4].

The data needs of health promotion policymakers can be better met by expanding information resources from those more familiar datasets to include other, special interest information and newly-generated data. Knowing how to evaluate available data can help in making policy choices based on either new or standard information.

Criteria for Evaluating Data

None of these points is really startling but should remind policymakers of their dependence on data to develop and evaluate policy and program decisions. The wealth of new information available to health policymakers -- and equally important, the availability of high speed technology including computerized databases -- greatly enhances the flexibility of the decisionmaker. In the face of this powerful new technology and greatly expanded resource, policymakers should always evaluate the strength and reliability of the data they uae to support or defend policy initiatives.

There are three general criteria that should be invoked to assess the usefulness of a data source. First, one needs to assess its quality. Researchers use the term reliability as a major aspect of data quality, and this has important implications for decisionmakers. Inherent in the notion of reliability are questions such as: For what purpose were the data originally collected? How were the data collected? Were operational definitions used consistently throughout the data collection? Did those who collected it gather only data of importance to them in achieving their purpose? Is there some independent source which can be used to verify the information? For example, scepticism as to the reliability of self-reported data on physical exercise is compounded in the absence of consistent definitions of activities, frequency of participation, duration and levels of exertion.

A second consideration in assessing information is the practical problem of its accessibility. What is the cost of gaining access to the data? Will extracting these data require a large investment of staff and time? How long will the data be available? Are there plans to continue to collect this information, so that program outcome can be evaluated in the future? Are there costs associated with storage of the dataset? Are there alternatives which are less costly?

Finally, it is important to look at the utility of the data. This is measured in abstract terms with respect to its usefulness to the decisionmaker. It is quite possible to locate information of very high quality only to discover that the definitions used in collecting the data do not correspond to the problem at hand. If, for instance, you wish to assess the utilization of emergency room services by Medicaid-eligible children under age five, data that reports emergency room services only for children under fourteen is of limited utility. Similarly, cross-sectional data collected at only one point and not repeated over a period of time is of questionable value in helping to establish trends.

Table 1 summarizes these criteria as applied to examples of data sources that could be used in the initial stage of developing a health promotion policy. The example policy initiative is to design a policy to reduce the birthrate among teenagers under age 17 years.

TABLE 1. A GUIDE FOR EVALUATING INFORMATION FOR HEALTH PROMOTION POLICY DEVELOPMENT

ISSUE IN POLICY DEVELOPMENT	EXAMPLE OF POLICY INITIATIVE	DATA CATEGORY	FACTORS LIMITING USEFULNESS OF DATA		
			Quality	Accessiblity	Utility
Understanding the severity and magnitude of the problem	Design a policy to reduce the birthrate among teenagers under 17 years of age	a Birth Certificate	Good	Excellent	Excellent
		b School Census	Poor	Poor	Poor
		c WIC Data	Good	Poor	Fair
		d Protective Services Data	Poor	Poor	Fair
		e Hospital Utilization Data	Fair	Poor	Fair
		f Survey Research	Poor	Good	Fair

Summary

The availability of an increasing variety of data has not alone met the information needs of health policymakers. A richness of data can be confusing and such abundance can lead to poor utilization and misinterpretation. In a policy development sequence of first understanding the problem, then assessing the likelihood of a successful intervention and evaluating its impact, standardized and familiar data can be used to verify and strengthen program data. It may also be more desireable and even more efficient to generate a new set of specific information. In any case, essential questions about data quality, ease of access and utility of application need to be asked early in the policy process.

NOTES

1. These topics are among the major areas of program focus published by the Department of Health and Human Services, in Prevention '82, USDHHS (PHS), Office of Disease Prevention and Health Promotion, DHHS (PHS) Pub. No. 82-50157.

2. Concept of health information systems as organized at the Fourth World Congress on Medical Informatics, Amsterdam, August, 1983.

3. See results of the Massachusetts health interview survey and its correlation with census data, in Lambert, et.al., Risk factors and life style: a statewide health-interview survey, New England Journal of Medicine, v. 306, 1048-51.

4. Public Health Reports, v. 96, nr. 3, 1981, 200-1. This article describes 14 separate datasets available through the National Center for Health Statistics.

"A STATE HEALTH AGENCY'S COLLABORATIVE EFFORTS TO EVALUATE AND MONITOR MATERNAL AND CHILD HEALTH PROGRAM EFFORTS --UTILIZING VITAL STATISTICS"

G.L. Sandifer, S.C. Department of Health and Environmental Control
Murray B. Hudson, S.C. Department of Health and Environmental Control

Introduction:

As Federal and State funds for public health services diminish, the prioritizing and allocation of available funds becomes more important. In a State which ranks near the bottom on practically every health, economic, and social indicator, the use of appropriate and timely health status indicators provides assistance to the State decision makers in determining where scarce resources should be directed.

The intent, description, and outcome of four collaborative efforts between Vital Statistics and Maternal and Child Health (MCH) Programs are discussed. A composite computerized cohort file, linked with the MCH maternity client record system, is described and its utilization by these evaluative efforts is demonstrated.

The results confirm that while the approach of each program evaluation activity may not be sensitive enough to completely diagnose and explain a problem; collectively together with the use of vital statistics, they provide invaluable assistance in targeting scarce resources in an effective and efficient manner. What follows is a brief description of these efforts.

I. High-Risk Perinatal Program

In 1976, the State Health Department implemented a High Risk Perinatal Program to reduce perinatal mortality. As the years progressed, the program grew but no evaluation measures were developed to attest to its success. We intuitively knew that high risk groups have worse outcomes than non-high risk, but we did not know what the difference was between treated and non treated high risk patients. With justification for future funding becoming more acute, the need for a good program evaluation was imminent. Therefore, in 1980, the Department contracted with the University of South Carolina to evaluate the program.

In South Carolina, Vital Records files have been cohort matched since 1975. This file provided the critical link necessary to conduct the program evaluation. Fortunately, the Department had created a special linked file combining the cohort and Maternity program files. All that was left was to add the patient information regarding their high risk scoring status which consisted of approximately 25 variables. Using this data file, 564 control and study patients were matched using 2 1/2 years of data on 8 characteristics known to be associated with pregnancy outcome: (1) mother's race, (2) mother's age, (3) pariety, (4) previous fetal death, (5) previous neonatal death, (6) previous low birth weight baby, (7) previous spontaneous abortion, and (8) previous caesarean section.
(see Table I)

This study found that women who participated in the Program had approximately one-half as many infant deaths as the women who did not participate. (see Table II)

The study also found there was no significant difference between the birth weight distribution of the two population groups. The average birth weight difference was only 32 grams (2966 vs. 2998). The study therefore attributed the main difference in mortality outcome to the fact that high risk care was provided program patients in a hospital setting appropriate to their need.

HOSPITAL LEVEL OF CARE USAGE
By No. and % of Patients

	Program		Non-Program	
	No.	%	No.	%
Level I -	15	3	120	21
Level II or - Level III	549	97	444	79

II. Priority Infant Tracking

In 1980, the MCH Division recognized the imminent decrease in both Federal and State funding and decided that a priority system was needed to determine those patients most in need of services. The Division developed a system to prioritize patients while maintaining the objectives of (1) decreasing the infant mortality rate (primarily the postneonatal rate) and (2) reducing childhood morbidity.

The priority process started with selection of 5 prenatal and birth risk factors having a high association with infant mortality. These risk factors are:

1. Low Birth Weight ($\leq$ 2500 grams)
2. $<$ 18 years of age
3. $<$ high school graduate
4. Prenatal care inadequate ($<$5 visits or started care after 6th month)
5. Previous live born now dead

A review of 1976-78 cohort data showed the following relationship between these factors and infant mortality.

No. of Risk Factors Present	Infant Mortality Rate (Per 1,000 Live Births)
0	5.3
1	20.7
2	40.8
3	75.1
4	121.2
5	101.3
One or More	33.7

It is also important to note that the decision to prioritize infants was consistent with the Divisions objective of having 75% of the children served in child health clinic be below age 2.

For the system to be effective, one of the critical components was a means for early public health intervention. With backing from the Department's Commissioner as well as early coordination with the Office of Vital Records and Public Health Statistics, we were able to get the local county vital records staff to screen birth certificates for presence of these risk factors. Those births having one or more risk factors had a tracking card initiated. Births born outside their normal county of residence were identified by a follow-up computer report to their local county health department. Those families known to DHEC through some previous encounter were immediately contacted and encouraged to enter a health care system. Those not previously known to DHEC were not contacted until one month after their date of birth in order to minimize the occurrence of contacting parents who may have experienced a neonatal death.

Since its inception, significant changes have taken place in the child health (CHC) population. The percent of CHC infants enrolled in WIC increased from 56% in 1980 to 82% as of June 30, 1982. The percent of patients less than two years of age in CHC increased from 48% to 78%. The average annual number of visits for CHC infants rose from 2.3 to 4.1. Other measures examined provide similar findings.

The use of information, staff, and expertise available through the Office of Vital Records and Public Health Statistics has provided critical components to Maternal and Child Health's successful efforts in prioritizing, identifying, and serving these high risk infants.

III. Regionalization of Hospital Care

When examining the prospects and need for regionalization of hospital care, South Carolina is both fortunate and unfortunate. Fortunate in that there are only approximately 61 hospitals with delivery capability, and unfortunate in that the State is primarily rural (requiring long distances to transport patients and family) and many physicians are reluctant to transfer care of their patients. The Department's position is to encourage that delivery and neonatal care be conducted at the site most appropriate for the individual patient.

To promote this concept, a special perinatal report was developed and is produced annually showing birth and mortality outcomes by hospital, hospital delivery size, location, etc. (see attachments A and B). This data file and corresponding reports were developed using cohort data and attributing the death back to the hospital of birth - regardless or referral status (a current revision to the file will allow for determining the number of deaths referred to another hospital). Information from this report is used to provide consultation to hospitals and physicians and as a basis for program planning.

IV. Improved Child Health Project

In FY'81 the Improved Child Health Project (a 4-year Federally funded project in a 3 county area) undertook a prenatal survey to:

1. assess the unmet need for prenatal care in the 3 counties comprising the health district;
2. validate information regarding adequacy of care recorded on the birth certificate; and
3. define appropriate strategies for health care and health education activities.

The survey need became evident after looking at the poor mortality statistics for the three county area over time and the percent of births receiving 0-5 prenatal visits in these counties for the years 1976-1979. (see Table III)

The survey methodology included a stratified random sample of the 2,207 resident births in calendar year 1979. Interviews were conducted with each mother in the survey as well as contacting their care provider to obtain information from their medical records and then validating this information with birth certificate data. Thirty-six individual physicians and/or group practices were contacted and participated in the survey. Only one practitioner elected not to cooperate.

An eleven item questionnaire was developed to identify barriers to receiving prenatal care. Two hundred fifty-three women were contacted and interviewed; birth certificates were matched with survey records for 249 (98.4%). There were 22 women for whom it was not possible to verify prenatal care information. For the purposes of this study, inadequate prenatal care was defined as less than five prenatal visits or prenatal care that started after the sixth month of pregnancy. The results of the survey provided the following significant insights into the actual health problems within the district which were not apparent by simply examining birth certificate data.

As regards the unmet need for prenatal care; the majority of women surveyed (63%) waited more than 4 weeks before seeking care after they thought they were pregnant. Less than 15% of the women sought care within two weeks of the time they thought they were pregnant. Also, a majority of the women (88%) reported they sought care as soon as they thought they needed it. Only 28 (11.2%) of the respondents did not seek care when they thought they needed it. Of the 28 not seeking care; inability to pay was the primary reason given for not seeking care.

It was generally assumed that in these particular counties there was a reluctance of providers to serve indigent and medically poor families. Results from this survey indicated that 95% of the women interviewed reported they were able to receive prenatal care when they first sought it. This finding obviously nullified the earlier assumption.

The second purpose of the survey was to

validate birth certificate information. Cross tabulations of the data from the birth certificate and data documented from the care providers were analyzed to determine the accuracy of information on the number of prenatal visits and the month in which prenatal care was started. The survey showed that most of the information regarding the number of prenatal care visits and month care began was incorrect. In one county data regarding the number of prenatal visits was 45% correct; in another the data was 14% correct and in the third county 25% was correct. The birth certificate data on these both over and understated information retrieved from files of the medical provider. Similar conclusions were found in the analysis of "month prenatal care began".

The third purpose of this survey was to define strategies for health care and health education activities. The problem identified by this survey was lack of knowledge regarding pregnancy and the need for early prenatal care. Although 88% of the women reported seeking care as soon as they thought they needed it, 63% of these women waited more than four weeks to seek care after they thought they were pregnant.

Of the reasons given for delay in seeking care, one-third of the responses reflected a lack of knowledge regarding signs of pregnancy and normal/abnormal reproductive functions. In practically all the responses, a lack of knowledge regarding the need for early prenatal care was evidenced.

From the survey, it was concluded that there is not a problem in this three county area with unmet need for prenatal care because the majority of women reported that care was available and accessible. The problem is one of lack of education by women who do not know when they should seek care.

SUMMARY

This brief review illustrates four ways in which collaborative efforts have been made to evaluate and monitor MCH program efforts. These efforts have resulted in using vital records information to:

1. assess the health impact of certain programs (High Risk).
2. determine high risk infant variables and use these variables to identify and serve high risk infants.
3. conduct studies which help determine program direction.
4. determine the reasons for poor maternal health practices in a locale.

As the need for benefit/cost analysis and priority setting continues, continuation of these and other collaborative efforts will be necessary.

TABLE I

HIGH RISK PERINATAL EVALUATION FILE

IMPROVED PREGNANCY OUTCOME FILE

COHORT FILE

BIRTH	DEATH	MATERNITY PROGRAM	FETALS	HIGH RISK SCORE INFO.

TABLE II

MATCHED PAIR ANALYSIS	PROGRAM PATIENTS		NON-PROGRAM PATIENTS	
	No.	Rate	No.	Rate
FETAL MORTALITY	8/573	14.0	15/573	26.2
NEONATAL MORTALITY	10/564	17.7	17/564	30.1
POSTNEONATAL MORTALITY	6/564	10.6	12/564	21.3
INFANT MORTALITY	16/564	28.4	29/564	51.4

TABLE III

PERCENT OF BIRTHS
0-5 PRENATAL CARE VISITS
FOR SELECTED COUNTIES & S.C.
1976-1979

YEAR	CHESTERFIELD	DARLINGTON	MARLBORO	STATE
1976	18.4	25.1	19.9	13.1
1977	13.7	21.6	23.3	11.2
1978	17.0	18.1	40.0	11.5
1979	12.7	17.9	28.7	11.4

ATTACHMENT A

SOUTH CAROLINA DEPARTMENT OF HEALTH AND ENVIRONMENTAL CONTROL

DIVISION OF BIOSTATISTICS

DELIVERIES (*), PREMATURE BIRTHS, FETAL, NEONATAL, PERINATAL DEATHS
BY HOSPITAL, BY RACE, BY PERINATAL MORTALITY RANK (**)
1982 DATA
(HOSPITALS > 1999 DELIVERIES)

RANK/HOSPITAL	RACE	NO. DEL.	PREMATURES NO.	PREMATURES RATE	FETALS NO.	FETALS RATE	NEONATALS NO.	NEONATALS RATE	PERINATALS NO.	PERINATALS RATE
1 S.C. BAPTIST	TOTAL	2412	154	64.2	12	5.0	13	5.4	25	10.4
	WHITE	1758	94	53.7	7	4.0	9	5.1	16	9.1
	NONWHITE	654	60	92.4	5	7.6	4	6.2	9	13.8
2 ANDERSON	TOTAL	2114	145	68.9	10	4.7	27	12.8	37	17.5
MEMORIAL	WHITE	1617	81	50.4	9	5.6	15	9.3	24	14.8
	NONWHITE	497	64	129.0	1	2.0	12	24.2	13	26.2
3 GREENVILLE	TOTAL	4630	402	87.6	43	9.3	43	9.4	86	18.6
MEMORIAL	WHITE	3578	252	70.9	25	7.0	31	8.7	56	15.7
	NONWHITE	1052	150	145.1	18	17.1	12	11.6	30	28.5
4 MCLEOD	TOTAL	2370	244	104.1	26	11.0	26	11.1	52	21.9
REGIONAL	WHITE	1301	91	70.2	5	3.8	7	5.4	12	9.2
	NONWHITE	1069	153	146.0	21	19.6	19	18.1	40	37.4
5 SPARTANBURG	TOTAL	2174	238	110.5	20	9.2	40	18.6	60	27.6
GENERAL	WHITE	1392	134	96.9	9	6.5	25	18.1	34	24.4
	NONWHITE	782	104	134.9	11	14.1	15	19.5	26	33.2
6 MEDICAL	TOTAL	3196	520	164.6	37	11.6	69	21.8	106	33.2
UNIVERSITY	WHITE	1021	148	145.7	5	4.9	19	18.7	24	23.5
	NONWHITE	2175	372	173.6	32	14.7	50	23.3	82	37.7
7 RICHLAND	TOTAL	2495	517	209.7	30	12.0	54	21.9	84	33.7
MEMORIAL	WHITE	715	140	198.6	10	14.0	20	28.4	30	42.0
	NONWHITE	1780	377	214.2	20	11.2	34	19.3	54	30.3

(*) DELIVERIES = LIVE BIRTHS (ALL) + FETAL DEATHS (OVER 500 GRAMS)

(**) RANK = 1 FOR LOWEST TOTAL PERINATAL MORTALITY RATE

ATTACHMENT B

LIVE BIRTHS AND NEONATAL DEATHS
BIRTHWEIGHT BY HOSPITAL SIZE AND DELIVERY HOSPITAL
SOUTH CAROLINA, 1980-1982
RATES CALCULATED PER 1000
THIS REPORT UTILIZES PROVISIONAL DATA
(MCH060183A-IPOHOD82)

----------HSIZE=MORE THAN 1999 DELIVERIES HBIRTH=ANDERSON MEMORIAL HOSPITAL----------

	White Live Births	Nonwhite Live Births	Total Live Births	White Neonatal Deaths	White Death Rate	Nonwhite Neonatal Deaths	Nonwhite Death Rate	Total Neonatal Deaths	Total Death Rate	Level 3 Death
0- 500 Grams	4	10	14	4	1000.00	10	1000.00	14	1000.00	*
501- 749 Grams	10	10	20	9	900.00	9	900.00	18	900.00	2
750- 999 Grams	12	9	21	8	666.67	3	333.33	11	523.81	3
1000-1499 Grams	21	13	34	4	190.48	*	*	4	117.65	4
1500-1999 Grams	48	35	83	7	145.83	*	*	7	84.34	4
2000-2499 Grams	186	114	300	1	5.38	2	17.54	3	10.00	1
2500-3999 Grams	4260	1284	5544	9	2.11	5	3.89	14	2.53	5
4000 gm & above	660	58	718	1	1.52	*	*	1	1.39	1
501-1499 Grams	43	32	75	21	488.37	12	375.00	33	440.00	9
501-2499 Grams Total	277	181	458	29	104.69	14	77.35	43	93.89	14
Hospital Births	5201	1533	6734	43	8.27	29	18.92	72	10.69	20

----------HSIZE=MORE THAN 1999 DELIVERIES HBIRTH=MEDICAL UNIVERSITY HOSPITAL----------

	White Live Births	Nonwhite Live Births	Total Live Births	White Neonatal Deaths	White Death Rate	Nonwhite Neonatal Deaths	Nonwhite Death Rate	Total Neonatal Deaths	Total Death Rate	Level 3 Death
0- 500 Grams	11	45	56	11	1000.00	45	1000.00	56	1000.00	*
501- 749 Grams	16	52	68	11	687.50	47	903.85	58	852.94	*
750- 999 Grams	23	45	68	11	478.26	13	288.89	24	352.94	*
1000-1499 Grams	80	129	209	12	150.00	11	85.27	23	110.05	*
1500-1999 Grams	133	236	369	9	67.67	5	21.19	14	37.94	*
2000-2499 Grams	178	584	762	2	11.24	8	13.70	10	13.12	*
2500-3999 Grams	2297	5350	7647	8	3.48	19	3.55	27	3.53	*
4000 gm & above	241	219	460	1	4.15	3	13.70	4	8.70	*
501-1499 Grams	119	226	345	34	285.71	71	314.16	105	304.35	*
501-2499 Grams Total	430	1046	1476	45	104.65	84	80.31	129	87.40	*
Hospital Births	2979	6660	9639	65	21.82	151	22.67	216	22.41	*

----------HSIZE=MORE THAN 1999 DELIVERIES HBIRTH=MCLEOD REGIONAL MEDICAL CENTER----------

	White Live Births	Nonwhite Live Births	Total Live Births	White Neonatal Deaths	White Death Rate	Nonwhite Neonatal Deaths	Nonwhite Death Rate	Total Neonatal Deaths	Total Death Rate	Level 3 Death
0- 500 Grams	5	11	16	5	1000.00	11	1000.00	16	1000.00	*
501- 749 Grams	6	15	21	6	1000.00	15	1000.00	21	1000.00	*
750- 999 Grams	10	20	30	4	400.00	11	550.00	15	500.00	*
1000-1499 Grams	33	65	98	7	212.12	7	107.69	14	142.86	*
1500-1999 Grams	54	86	140	3	55.56	2	23.26	5	35.71	*
2000-2499 Grams	156	236	392	2	12.82	3	12.71	5	12.76	*
2500-3999 Grams	3093	2386	5479	6	1.94	5	2.10	11	2.01	*
4000 gm & above	464	122	586	3	6.466	2	16.393	5	8.532	1
501-1499 Grams	49	100	149	17	346.939	33	330.000	50	335.570	*
501-2499 Grams Total	259	422	681	22	84.942	38	90.047	60	88.106	*
Hospital Births	3821	2941	6762	35	9.422	56	19.041	92	13.605	4

STUDIES OF HOSPITALIZED TRAUMA USING THE NATIONAL HOSPITAL DISCHARGE SURVEY

Jack Goldberg, University of Illinois
Paul S. Levy, University of Illinois
Mary Moien, NCHS
Judith L. Goldberg, University of Illinois

1. Background

The measurement of injury severity is recognized as a crucial component of any study which attempts to assess the effectiveness of medical care among victims of traumatic injuries. The need for a severity index is highlighted when one attempts to examine outcomes among institutions. Hospital performance in treating trauma is frequently judged by the comparison of mortality rates; yet, such a comparison will be biased if the severity of injured patients is not taken into account. For instance, Trauma Centers commonly report mortality rates which are much higher than that seen at community hospitals .

The study described here was designed to develop a new index of injury severity. The overall objectives involve: 1) the development and testing of an index of severity based on the anatomic description of trauma contained in the International Classification of Diseases Adapted for Use in the United States (8th revision) and; 2) the exploration of the impact demographic, hospital and medical factors have on patient mortality after adjustment for the severity of injury. Each of these specific activities are described in succeeding sections.

2. Development of the Revised Estimated Survivial Probability Index

The conceptual model for the Revised Estimated Survival Probability index (RESP) is that the severity of multiple injuries with respect to "threat to life" can be expressed as a function of the severity of the individual injuries in terms of the threat to life existent if each of the injuries were present as single conditions.

This conceptual model implies a product function where each individual injury sustained by a patient would be represented by a single condition survival probablity. A single condition survival probability would be the probability of a patient surviving if he/she had only one particular injury (say fracture of the vault of skull) and no others. The basis of such a model lies in the multiplicative law of probability; that is, if each trauma condition were acting independently of the other trauma conditions present with respect to threat to life, then the probablity of a patient with multiple trauma surviving hospitalization would be equal to the product of the individual single condition survival probabilities.

We recognize that the product function is likely an over-simplification of the effects of multiple trauma, especially if several organ systems are involved in the injury. However, the rationale for this approach is that the product of the single condition survival rates might be good enough approximation for purposes of stratification of patients into groups.

With this conceptual model, our strategy was to obtain estimates of these single condition survival rates from the National Center of Health Statistics (NCHS) National Hospital Discharge Survey (NHDS). In particular, we used standardized microdata tapes from the NHDS covering the years 1971-1975. From this data set, a working tape was compiled consisting of all patients hospitalized with one or more trauma conditions (defined here as ICDA-8 codes 800.0-939, 950.0-959.9 and 991.0-996.9). We did not include burns (940-949), adverse effects of chemical substances (960-989), effects of radiation (990) or complications of surgery or medical care (997-999) in our set of trauma conditions. If a patient had trauma diagnoses in both the inclusion range (800.0-939.0, 950.0-959.9 and 990.0-996.9) and the exclusion range (940-949, 960-989, 990, 997-999) that cases was excluded from the working tape.

From this data source single condition survival probabilities were calculated for patients having only one trauma condition. For each single condition we estimated three survival probabilities based on patient age (under 45 years, 45-64 years, 65 and over). Because of the large number of trauma codes, some codes representing similar conditions or contiguous anatomic sites were grouped together and the single condition survival rates were computed for the grouped codes.

A total of 531 single condition survival rates were computed (177 codes by 3 age groups), and a sample of these are shown in Table 1. The complete set is available from the authors. Again, these single condition survival rates are the building blocks on which the RESP index is constructed.

TABLE 1 SELECTED AGE SPECIFIC SINGLE CONDITION SURVIVAL RATES

ICDA Codes	ESTIMATED SINGLE CONDITION SURVIVAL RATES Under 45 Yrs	45-64 Yrs	65 Yrs and Over
801.0	.9898	.9466	.9225
808.0	1.0000	.9568	.9320
806.0) 806.2) 806.4) 806.6)	.9587	.8304	.7590
811.0	1.0000	1.0000	1.0000
851.0	.9350	.8852	.8575
865.0	1.0000	.8359	.7423
868.1	.9839	.6182	.4145

The conceptual model which the RESP is based upon is tested by examining the linear relationship between RESP and mortality. A data set of 14,824 records with multiple trauma from the 1971-75 NHDS was identified to validate the RESP index.

One method of examining the relationship between a dichotomous dependent variable (such as mortality) and a continuous independent variable (such as the RESP index) is logistic regression analysis. In logistic regression analysis the dependent variable is specified as the natural logarithm of the odds of death from multiple trauma. The validation of the RESP index focuses on the slope parameter estimated from the logistic regression procedure; a slope parameter which is negative and statistically significant would indicate a linear relationship between the RESP index and mortality, (i.e., as RESP decreases mortality increases).

The logistic regression procedure was performed separately for selected subsets of multiple trauma patients; models were fit for patients with two, three, four and five trauma conditions, patients with intracranial injuries (ICDA 850.0-854.9) and patients with cerebral laceration and contusions, without mention of open intracranial wound (ICDA 851.0). The objective of these separate analyses was to determine the ability of RESP to discriminate among groups of patients even when matched on the number or type of traumatic conditions.

Table 2 presents the results of seven separate logistic regressions between RESP and mortality. For each of the separate analyses the coefficient of slope is large and negative. Likewise each of the slope parameters is found

Table 2 LOGISTIC REGRESSION ANALYSIS OF RESP INDEX AND MORTALITY IN MULTIPLE TRAUMA PATIENTS

Data Set Analyzed	Number Discharged Alive	Dead	Slope (beta)
All Multiple Trauma	14,455	369	-10.8097
Two Trauma Conditions	10,399	192	-11.5004
Three Trauma Conditions	2,921	90	-9.2010
Four Trauma Conditions	854	55	-9.4708
Five Trauma Conditions	281	32	-7.5716
Intracranial Injury	4,602	199	-8.9037
Cerebral Lacerations & Contusions	443	85	-7.7249

*All betas are significant at less than the .001 level.

The set of properties which a good severity index should possess has been recently debated by participants at a conference on trauma severity indices sponsored by the National Center for Health Services Research and the American Trauma Society. The consensus among the participants of this conference is that outcome prediction (i.e., correlation with indicators of outcome) is perhaps the most important property of a severity index. With this in mind, the main thrust of the validation studies described above has been to examine how well the RESP index correlates with a specific outcome measure, namely mortality. More specifically, since the intended use of the index is an adjustment for patient mix in restrospective studies based on hospital medical records, it is felt that demonstration of a strong monotonic relationship between RESP scores and mortality provides a good test of criterion validity with respect to this intended use of the index. Clearly, the strong linear relationships observed between the index and mortality among all patients with multiple trauma, among those with multiple trauma involving intracranial injuries and among those having cerebral lacerations and contusions along with other trauma conditions provides evidence that the index can be used to adjust for patient mix in studies using mortality as the outcome variable.

3. Mortality From Hospitalized Traumatic Injuries

The data source for the mortality study is the Hospital Discharge Survey conducted by the National Center for Health Statistics. Computer tape transcripts of the NHDS were purchased from the NCHS for the years 1977 and 1978. These transcripts contain individual cases of patients discharged from over 400 hospitals nationwide. A total of approximately 462,000 unweighted records are contained on the NHDS for 1977 and 1978.

From the data available on the NHDS tape transcripts four types of factors were identified for investigation. The factors analyzed include:

1) Demographic factors
   a) Sex
   b) Race
   c) Medicaid admission
   d) Weekday-weekend admission
2) Hospital Factors
   a) Number of hospital beds
   b) Ownership of hospital
   c) Region of U.S.
   d) Mean RESP score for hospital
3) Patient volume factors
   a) Number of hospital discharges
   b) Number of trauma discharges
   c) Number of sever trauma discharges (RESP < .90)
4) Medical Factors
   a) Number of non-trauma conditions listed
   b) Type of non-trauma conditions
   c) Number of operative and non-surgical procedure listed
   d) Type of surgery performed
   e) Number of trauma conditions listed

Of particular interest are the patient volume factors. Recently several papers have demonstrated a negative relationship between patient outcomes from specific types of surgery and the number of cases treated. These papers have attempted to show that a hospital's proficiency in performing certain types of surgical procedures, as demonstrated by a low mortality rate, improves as the number of patients treated increases.

All trauma patients who were discharged as dead were selected as cases. These deaths were then placed into age, RESP and most severe ICDA categories. The variable age was grouped into three categories: <44, 45-64, 65+; RESP was grouped in eight categories: <.60, .61-70, .71-80, .81-.85, .86-.90, .91-.95, .90-.99, 1.0. The most severe ICDA code is selected by examining each of the single condition survival probabilities and determining the lowest. The ICDA code associated with this survival probability is the most severe injury. The combination of age group, RESP category and most severe ICDA code provide a set of index strata from which the selection of controls can be accomplished.

Controls are selected from patients who did not die and were routine discharges to home. In each index stratum controls were randomly selected to equal the number of cases in the index stratum. This yielded a control group one to one frequency matched to the case group on age category, overall, RESP category and ICDA code of most sever injury. From the above selection procedure a total of 566 cases and 566 frequency matched controls were identified.

The analytical approach is directed to determining the estimated relative risk of not surviving (i.e. being a case) if an independent variable is at a particular level. An estimate of the relative risk is obtained by the odds ratio. The strength of the association between each independent variable, adjusted for possible confounding by other variables, is given by the odds ratio calculated from the unstandardized regression parameters.

The analysis of the case-control data set is performed in two stages: the first stage uses the total data set (n=1132) while the second stage selects only those patients who have very severe injuries. Patients with very severe injuries are defined as those with RESP scores of less than .90 (n=336). The primary reason for this separate analysis is that the volume-outcome relationship, if present, should be apparent in this subset of patients.

Table 3 presents the results of a multiple logistic regression analysis performed on the full data set of 1132 cases and controls. The variables selected for inclusion to the logistic procedure were, in general, those which exhibited statistically significant associations in the unadjusted analysis. Departures from this general rule fell into two classes: first, though the variables, number of hospital beds, total number of discharges, and number of trauma discharges were significantly associated with mortality it is likely that they were taping into the same underlying construct of hospital size. Therefore only one of these three variables, the number of trauma discharges, was entered into the model. Second, several factors which a priori were expected to have an impact on the risk of mortality but failed to demonstrate a significant association in the unadjusted analysis were maintained in the regression analysis. These factors include diabetes, neurosurgery and orthopedic surgery.

TABLE 3 MULTIPLE LOGISTIC REGRESSION ANALYSIS OF SELECTED DEMOGRAPHIC, HOSPITAL, PATIENT VOLUME AND MEDICAL FACTORS FOR FULL DATA SET

Demographic, Hospital, Patient Volume and Medical Factors	Factor Adjusted Logistic Odds Ratio	Unadjusted Odds Ratio
Sex (Males vs. Females)**	1.51	1.50
Average RESP for Hospital***		
RESP > .985	1.00	1.00
RESP .985 - .975	1.32	1.14
RESP < .975	1.54	1.59
Number of Trauma Patients***		
< 750	1.00	1.00
750 - 1,499	1.17	1.32
1,500 +	1.60	1.54
Number of Trauma Patients with RESP less than 90		
< 20	1.00	1.00
20 - 99	1.01	1.24
100 +	1.02	1.61
Number of Non-Trauma Conditions		
One	1.00	1.00
Two	1.04	1.58
Three	1.08	1.86
Four	1.12	2.27
Number of Operative or Non-Surgical Procedures	1.16	4.86
None	1.00	1.14
One	1.04	1.00
Two	1.09	1.69
Three	1.13	2.39
Diabetes	1.12	1.18
Ischemic Heart Disease & Other Forms of Heart Disease***	3.95	3.22
Malignant Neoplasms	3.70	3.06
Influenza and Pneumonia	4.63	4.87
Cerebrovascular Disease***	4.13	3.43
Pulmonary Embolism***	6.29	4.46
Neurosurgery	1.54	1.38
Thoracic Surgery*	2.19	2.35
Orthopedic Surgery	1.46	1.40
Emergency Trachcotomy or Tracheostomy**	4.16	4.49
Exploratory Laparotomy or Celiotomy***	3.95	2.89

The regression analysis yields results which are quite similar to the unadjusted analysis. Comparing the adjusted and unadjusted estimates of the relative risk, little change is noted for the factors sex, average RESP for hospitals, number of trauma patients, ischemic heart disease and other forms of heart disease, malignant neoplasms, influenza and pneumoia, cerebrovascular disease, pulmonary embolism, thoracic surgery, emergency tracheotomy or tracheostomy, and exploratory lapartomy or celiotomy. It is clear from the constancy of these results that confounding is not occuring in the factors examined above. In contrast to this stability are the results for the factors: number of non-trauma conditions, number of operative and non-surgical procedures, and volume of trauma patients with RESP <.90. These factors are not statistically significant in the logistic regression analysis.

Multiple logistic regression analysis is also used to adjust for potential confounding in the case-control data set of patients with RESP scores of less than .90. A more conservative approach to the selection of what factors to include is used than in the previous logistic model involving the full data set. All factors chosen for entry into the model were those that were statistically significant in the unadjusted analysis. This more conservative approach is used because the very severe data set is approximately one third the size of the full data set. It was felt that the inclusion of many non-significant factors in the model would unduly influence the stability of the estimated beta coefficients.

The results from the logistic model are exhibited in Table 4. Overall, only one factor remains statistically significant at the .05 level - ischemic heart disease and other forms of heart disease (w=2.34). Factors such as sex (w=1.58) malignant neoplasms (w=2.76) and influenza and pneumonia (w=2.88) are marginally significant with p values between .05 and .07. Factors which in the unadjusted analysis were strongly associated with mortality were substantially weakened after adjustment. Especially interesting were the changing relative risks for the hospital and volume factors. Whereas before a fourfold increase in risk was noted for high average severity hospitals now this risk is slightly less than twofold and not significant. Of even greater importance is the substantially reduced gradient in relative risk observed for the volume of trauma patients with RESP <.90; after adjustment the relative risk of high versus low volume hospitals was 1.48 and not significant. As was observed for the full data set no significant residual confounding for age and RESP index was found in the very severe data set.

Table 4 MULTIPLE LOGISTIC REGRESSION ANALYSIS OF SELECTED DEMOGRAPHIC HOSPITAL, PATIENT VOLUME AND MEDICAL FACTORS FOR VERY SEVERE TRAUMA

Demographic, Hospital, Patient Volume and Medical Factors	Factor Adjusted Logistic Odds Ratio	Unadjusted Odds Ratio
Sex (Males vs. Females)	1.58	1.63
Average RESP for Hospital		
RESP > .985	1.00	1.00
RESP .985 - .975	1.22	3.81
RESP < .975	1.71	4.17
Number of Trauma Patients with RESP less than .9000		
< 20	1.00	1.00
20 - 99	1.10	1.69
100 +	1.48	2.42
Ischemic Heart Disease & Other Forms of Heart Disease***	2.34	2.35
Malignant Neoplasms	2.76	2.96
Influenza & Pneumonia	2.88	2.92
Emergency Trachostomy or Tracheostomy	3.12	3.13

Appendix

PROGRAM SESSION ORGANIZERS

Myrna J. Aavedal, Ph.D. (G)
Environmental Studies Branch
Division of Epidemiology and Health Promotion
Analysis and Epidemiology Program
National Center for Health Statistics*

Mr. E. Earl Bryant (2nd Plenary)
Associate Director for Interview and Examination Statistics
National Center for Health Statistics*

Thomas Drury, Ph.D. (B)
Chief, Health Status Measurement Branch
Division of Epidemiology and Health Promotion
Analysis and Epidemiology Program
National Center for Health Statistics*

Ms. Pennifer Erickson (D)
Chief, Clearinghouse on Health Indexes
Health Status Measurement Branch
Division of Epidemiology and Health Promotion
Analysis and Epidemiology Program
National Center for Health Statistics*

Mr. Marshall C. Evans (K)
Chief, Technical Services Branch
Division of Vital Statistics
Vital and Health Care Statistics Program
National Center for Health Statistics
P.O. Box 12214
Research Triangle Park, North Carolina 27709

Ms. Janeanne Ferlazzo (E)
Planning, Evaluation, and Legislation Staff
Office of Program Planning, Evaluation and Coordination
National Center for Health Statistics*

Mr. Robert R. Fuchsberg (J,N,R)
Director
Division of Health Interview Statistics
Interview and Examination Statistics Program
National Center for Health Statistics*

Mr. George A. Gay (O)
Chief, Registration Methods Branch
Division of Vital Statistics
Vital and Health Care Statistics Program
National Center for Health Statistics*

Ms. Patricia Golden (F)
Special Assistant to the Director
Division of Epidemiology and Health Promotion
Analysis and Epidemiology Program
National Center for Health Statistics*

Ms. Alice Hetzel (S)
Deputy Director
Division of Vital Statistics
Vital and Health Care Statistics Program
National Center for Health Statistics*

() Indicates Session

Mr. Robert Heuser (T)
Chief, Natality Statistics Branch
Division of Vital Statistics
Vital and Health Care Statistics Program
National Center for Health Statistics*

Mr. Peter L. Hurley (3rd Plenary)
Associate Director for Vital and Health Care Statistics
National Center for Health Statistics*

Mary Grace Kovar, Dr.P.H. (H,L,P)
Special Assistant for Data Policy and Analysis
Interview and Examination Statistics Program
National Center for Health Statistics*

Ronald W. Manderscheid, Ph.D. (NIMH, W)
Acting Chief
Survey and Systems Research Branch
National Institute of Mental Health
5600 Fishers Lane, Room 18C-05
Rockville, Maryland 20857

Mr. Sam Marcus (C)
Assistant to the Director
Office of the Center Director
National Center for Health Statistics*

Ms. Kate McGuire (A)
Chief, Environmental Studies Branch
Division of Epidemiology and Health Promotion
Analysis and Epidemiology Program
National Center for Health Statistics*

Mr. Henry Mount (Y)
Program Development Staff
Office of Program Planning, Evaluation, and Coordination
National Center for Health Statistics*

Mr. James A. Smith (NCHS)
Program Development Staff
Office of Program Planning, Evaluation, and Coordination
National Center for Health Statistics*

Mr. Sheldon Starr (3rd Plenary)
Special Assistant for Technical Policy and State Relations
Vital and Health Care Statistics Program
National Center for Health Statistics*

Owen Thornberry, Jr., Ph.D. (X)
Deputy Director
Division of Health Interview Statistics
Interview and Examination Statistics Program
National Center for Health Statistics*

Ms. Joan Van Nostrand (V)
Deputy Director
Division of Health Care Statistics
Vital and Health Care Statistics Program
National Center for Health Statistics*

Andrew A. White, Ph.D. (U)
Survey Design Staff
Office of Research and Methodology
National Center for Health Statistics*

Mr. Paul D. Williams (I,M,Q)
Chief, Data Applications and Research Staff
Interview and Examination Statistics Program
National Center for Health Statistics*

Mr. Ronald Wilson (1st Plenary)
Director, Division of Epidemiology and Health Promotion
Analysis and Epidemiology Program
National Center for Health Statistics

Mr. Richard Yaffe (HCFA)
Acting Director, Evaluation Studies Staff
Office of Demonstrations and Evaluations
Health Care Financing Administration
Oak Meadows Building
6325 Security Boulevard
Baltimore, Maryland 21207

*3700 East-West Highway
Hyattsville, Maryland 20782

PROGRAM SESSION CHAIRMEN

Mr. Walter P. Bailey (Y)
Chief, Cooperative Health Statistics
State Budget and Control Board
1000 Assembly Street
Columbia, South Carolina 29201

Mr. Bruce M. Brock, M.P.H. (B)
Research Investigator and Survey Director
School of Public Health
University of Michigan
109 S. Observatory Street
Ann Arbor, Michigan 48109

Mr. E. Earl Bryant (2)
Associate Director for Interview and Examination Statistics
National Center for Health Statistics
3700 East West Highway
Hyattsville, Maryland 20782

James P. Cooney, Jr., Ph.D. (M)
Professor and Chairman, Department of Health Administration
Duke University Medical Center, Box 3018
Durham, North Carolina 27710

Gordon DeFriese, Ph.D. (3)
Director, Health Services Research Center
University of North Carolina
Chapel Hill, North Carolina 27514

Mr. Allen Dobson (R)
Division of Beneficiary Studies
Health Care Financing Administration
6325 Security Blvd., 2C-15 Oak Meadows
Baltimore, Maryland 20217

Manning Feinleib, M.D., Dr.P.H. (1)
Director
National Center for Health Statistics
3700 East West Highway
Hyattsville, Maryland 20782

Jacob J. Feldman, Ph.D. (1)
Associate Director for Analysis and Epidemiology
National Center for Health Statistics
3700 East West Highway
Hyattsville, Maryland 20782

Mrs. Mary Anne Freedman (S)
Director, Public Health Statistics
Vermont Department of Health, Box 70
115 Colchester Avenue
Burlington, Vermont 05402

Mr. Robert R. Fuchsberg (J)
Director
Division of Health Interview Statistics
National Center for Health Statistics
3700 East West Highway
Hyattsville, Maryland 20782

() Indicates Session

Bernard G. Greenberg, Ph.D. (H)
Kenan Professor
School of Public Health
University of North Carolina
Rosenau Hall, 201-H
Chapel Hill, North Carolina 27514

Trevor R. Hadley, Ph.D. (W)
Director, Office of Community Programs
Pennsylvania Office of Mental Health
Room 303, Health and Welfare Building
Harrisburg, Pennsylvania 17120

Mr. Daniel G. Horvitz (N)
Vice President, Statistical Sciences Group
Research Triangle Institute
Box 12194
Research Triangle Park, North Carolina 27709

Mr. Robert A. Israel (NCHS)
Deputy Director
National Center for Health Statistics
3700 East West Highway
Hyattsville, Maryland 20782

Mr. Frederick L. King (O)
State Registrar
Minnesota Department of Health
717 Delaware Street, S.E., Box 9441
Minneapolis, Minnesota 55440

Mrs. Ruth I. Knee (V)
Long Term/Mental Health Care Consultant
8809 Arlington Boulevard
Fairfax, Virginia 22031

Marshall Krueter, Ph.D. (E)
Director, Division of Health Education
Center for Health Promotion and Education
Centers for Disease Control
Building 3, Room 121
Atlanta, Georgia 30333

J. Richard Landis, Ph.D. (U)
Professor, Department of Biostatistics
School of Public Health
University of Michigan
109 S. Observatory Street
Ann Arbor, Michigan 48109

Judith R. Lave, Ph.D. (L)
Professor of Health Economics
Graduate School of Public Health
University of Pittsburgh
DeSoto Street
Pittsburgh, Pennsylvania 15261

Bryan R. Luce, Ph.D. (HCFA)
Director, Office of Research and Demonstrations
Health Care Financing Administration
330 Independence Avenue, S.W., Room 4222
Washington, D. C. 20201

Peter McMenamin, Ph.D. (I)
Senior Economist
Mathematica Policy Research, Inc.
600 Maryland Avenue, S.W.
Washington, D.C. 20024

Ronald W. Manderscheid, Ph.D. (NIMH)
Acting Chief, Survey and Systems Research Branch
National Institute of Mental Health
5600 Fishers Lane, Room 18C-05
Rockville, Maryland 20857

Sanford A. Miller, Ph.D. (C)
Director, Bureau of Foods
Food and Drug Administration
Room 6815, Federal Building 8
200 C Street, S.W.
Washington, D. C. 20204

Edward B. Perrin, Ph.D. (F)
Director, Health and Population Study Center
Battelle Human Affairs Research Centers
4000 N.E. 41st Street
Seattle, Washington 98105

Lori Ramonas, Ph.D. (A)
Office of Disease Prevention and Health
Promotion
Office of the Assistant Secretary for Health
300 7th Street, S.W., Room 613
Washington, D. C. 20201

Philip N. Reeves, D.B.A. (T)
Professor, Department of Health
Care Administration
George Washington University
2018 Eye Street, N.W.
Washington, D. C. 20052

Mark Rosenberg, M.D. (G)
Chief, Violence Epidemiology Branch
Center for Health Promotion and Education
Centers for Disease Control
1600 Clifton Road
Atlanta, Georgia 30333

Mr. Charles J. Rothwell (K)
Director, State Center for Health Statistics
Department of Human Resources
P. O. Box 2091
Raleigh, North Carolina 27602

Mrs. Anne A. Scitovsky (P)
Chief, Health Economics Division
Palo Alto Medical Research Foundation
860 Bryant Street
Palo Alto, California 94301

H. Denman Scott, M.D. (X)
Deputy Director of Health
Rhode Island Department of Health
75 Davis Street
Providence, Rhode Island 02908

John Ware, Ph.D. (D)
Senior Research Psychologist
The Rand Corporation
1700 Main Street
Santa Monica, California 90406

Gail R. Wilensky, Ph.D. (Q)
Project Hope
Health Sciences Education Center
Milwood, Virginia 22646

Ms. Michelle C. Adler, M.P.H. (W)
Office of the Assistant Secretary for Planning and Evaluation
Department of Health and Human Services
200 Independence Avenue, SW, Room 436G
Washington, D. C. 20201

Richard W. Ainsley, Ph.D. (V)
Director of Data Management
Central California Health Systems Agency
208 W. Main Street, Suite 9
Visalia, California 93291

Ms. Marlene Anderka (T)
Special Projects Manager, Statistics and Evaluation Unit
Division of Family Health Services
Massachusetts Department of Public Health
39 Boylston Street
Boston, Massachusetts 02116

David R. Anderson, Ph.D. (Q)
Benefits Evaluation Consultant
Control Data Corporation
Box 0, HQS02J
Minneapolis, Minnesota 55440

Jennifer J. Anderson, Ph.D. (I)
Assistant Professor
Boston University Medical Center
720 Harrison Avenue, Suite 1102
Boston, Massachusetts 02118

Joe Anderson, Ph.D. (R)
Project Manager
ICF, Inc.
1850 K Street, NW, Suite 950
Washington, D. C. 20006

Ms. Rosalyn D. Bass, M.A., M.P.H. (NIMH)
Senior Statistician
Survey and Systems Research Branch
National Institute of Mental Health
5600 Fishers Lane, Room 18C-05
Rockville, Maryland 20857

Charles E. Bennett, Ph.D. (Q)
Chief, Division of Health Information and Evaluation
Office of Health Planning
Department of Public Health
535 West Jefferson Street
Springfield, Illinois 62761

Ms. Bee Biggs, R.N. (K)
State Registrar and Chief, Vital Statistics
Department of Health and Welfare
450 W. State Street
Boise, Idaho 83720

Gordon S. Bonham, Ph.D. (N)
Director, Survey Research Unit
Urban Studies Center, University of Louisville
Gardencourt Campus
1020 Alta Vista Road
Louisville, Kentucky 40292

Ms. LaVonne A. Booton (D)
Department of Economics, 12D
Western Illinois University
Macomb, Illinois 61455

Ms. Linda Bowman (O)
Descriptive Statistician
State Health Data Center
Department of Health, P.O. Box 90
Harrisburg, Pennsylvania 17108

Edward N. Brandt, Jr., M.D. (1)
Assistant Secretary for Health
Department of Health and Human Services
200 Independence Avenue, S.W.
Washington, D. C. 20201

Lester Breslow, M.D. (1)
School of Public Health
University of California at Los Angeles
Los Angeles, California 90024

Charles H. Brooks, Ph.D. (M)
Department of Epidemiology and Community Health, School of Medicine
Case Western Reserve University
2119 Abington Road
Cleveland, Ohio 44106

William M. Burke, M.D. (P)
Associate Professor
Family and Community Medicine
University of Massachusetts Medical Center
55 Lake Avenue North
Worcester, Massachusetts 01605

Robert Casady, Ph.D. (J)
Office of Research and Methodology
National Center for Health Statistics
3700 East West Highway
Hyattsville, Maryland 20782

Linda S. Chan, Ph.D. (W)
Director, Research, Planning and Evaluation,
Los Angeles County-USC Medical Center, Room 12-900
1200 N. State Street, General Hospital
Los Angeles, California 90033

Larry S. Corder, Ph.D. (R)
Statistician
Health Care Financing Administration
6340 Security Blvd., Oak Meadows Bldg.
Baltimore, Maryland 21235

Brenda Cox, Ph.D. (J)
Senior Research Statistician
Research Triangle Institute
P.O. Box 12194
Research Triangle Park, North Carolina 27709

Mr. James L. Dallas (P)
Chief, State Health Planning and Development
Department of Social and Health Services
Mail Stop ET-27
Olympia, Washington 98504

() Indicates Session

Mr. Allen Dobson (L)
Division of Beneficiary Studies
Health Care Financing Administration
6340 Security Boulevard, 2C-15 Oak Meadows
Baltimore, Maryland 20217

Isaac W. Eberstein, Ph.D. (Y)
Research Associate/Assistant Professor
Center for the Study of Population
Florida State University
Tallahassee, Florida 32306

Ms. Bonnie Morel Edington, M.A. (X)
Health Planning Specialist
Health Planning Services
New Jersey Department of Health
John Fitch Plaza, Room 403
Trenton, New Jersey 08625

Paul W. Eggers, Ph.D. (L)
Social Science Research Analyst
Office of Research and Demonstrations
Health Care Financing Administration
6340 Security Boulevard
Baltimore, Maryland 21207

Halley S. Faust, M.D., MPH (A)
Associate Medical Director
HealthCare of the Bluegrass
212 N. Upper Street
Lexington, Kentucky 40507

Manning Feinleib, M.D., Dr.P.H. (1,NCHS)
Director
National Center for Health Statistics
3700 East-West Highway
Hyattsville, Maryland 20782

Lindson Feun, Ph.D. (W)
Assistant to the Director for
Planning and Evaluation
Oakland County Health Division
1200 N. Telegraph Road
Executive Office Building
Pontiac, Michigan 48053

Marvin Feuerberg, Ph.D. (R)
Research Sociologist
Division of Biometry and Epidemiology
National Institute of Mental Health
5600 Fishers Lane
Rockville, Maryland 20857

Mr. Michael H. Fleming (T)
Head, Analysis Section
Office of Health Information and Research
Ohio Department of Health
246 N. High Street, Box 118
Columbus, Ohio 43215

Mr. Alan Friedlob (HCFA)
Evaluative Studies Staff
Office of Demonstrations and Evaluations
Health Care Financing Administration
6325 Security Blvd., 2-F-3 Oak Meadows
Baltimore, Maryland 21207

Mr. Richard H. Friedman (W)
Office of the Deputy Assistant Secretary
for Health Planning and Evaluation
Room 740-G, Humphrey Building
Washington, D. C. 20201

Jack Goldberg, Ph.D. (*)
Instructor in Epidemiology
Epidemiology-Biometry Program
School of Public Health
University of Illinois at Chicago Circle
P. O. Box 6998
Chicago, Illinois 60680

David S. Greer, M.D. (H)
Dean, School of Medicine
Brown University
Box G
Providence, Rhode Island 02912

Ms. Margery Gudes, M.P.H. (F)
Biostatistician
Detroit Health Department
1151 Taylor Street
Detroit, Michigan 48202

David M. Gute, MPH, Ph.D. (A)
Director of Research and Epidemiology
Department of Public Health
Commonwealth of Massachusetts
80 Boylston Street, Room 335
Boston, Massachusetts 02116

Tony Hausner, Ph.D. (HCFA)
Office of Demonstrations and Evaluations
Health Care Financing Administration
6340 Security Blvd, Rm 2F3 Oak Meadows
Baltimore, Maryland 21207

Jack Hegenauer, Ph.D. (C)
Associate Research Biologist
Department of Biology, B-022
University of California
La Jolla, California 92093

Paul A. Hensleigh, M.D., Ph.D. (E)
Associate Professor, Gynecology and Obstetrics
Stanford University Medical Center
School of Medicine
Stanford, California 94305

Ms. Embry M. Howell (N)
Senior Research Analyst
SysteMetrics, Inc.
4520 East West Highway
Bethesda, Maryland 20008

Mr. Murray Hudson (K)
Director, Office of Vital Records and
Public Health Statistics
South Carolina Department of Health and
Environmental Control
2600 Bull Street
Columbia, South Carolina 29201

Ms. Nancy J. Hurwitz, MPH (P)
Independent Consultant
5308 Ventnor Road
Bethesda, Maryland 20816

Mr. Kirby L. Jackson (E)
Director, Health Sciences Computing Lab.
Department of Epidemiology and Biostatistics
University of South Carolina
Columbia, South Carolina 29208

Ms. Janine Jagger, M.P.H. (G)
Instructor, Neurosurgery
University of Virginia Medical
Center, Box 473
Charlottesville, Virginia 22908

Russell S. Kirby, Ph.D. (Y)
Research Analyst, Wisconsin Department
of Health and Social Services
P.O. Box 309
1 West Wilson Street, Room 480
Madison, Wisconsin 53701

Mary Grace Kovar, Dr.P.H. (C)
Special Assistant for Data Policy and Analysis
National Center for Health Statistics
3700 East West Highway
Hyattsville, Maryland 20782

Mr. Karl Kronebusch (G)
Health Program
U.S. Congress
Office of Technology Assessment
Washington, D. C. 20510

Mr. Garland Land (O)
Director, Missouri Center for Health Statistics
Missouri Department of Social Services
Broadway State Office Bldg., Box 570
Jefferson City, Missouri 65102

Mr. K. Per Larson (D)
Chief Consultant, Health Economics
American Health Foundation
320 East 43rd Street
New York, New York 10017

Mr. Michael R. Lavoie (O)
Director, Vital Records Service
Georgia Department of Human Services
47 Trinity Avenue. S.W.
Atlanta, Georgia 30334

Monroe Lerner, Ph.D. (I)
Professor, Department of Health Policy
and Management
School of Hygiene and Public Health
Johns Hopkins University
615 North Wolfe Street
Baltimore, Maryland 21205

Robert Lewis, Ph.D. (X)
Professor, Department of
Epidemiology and Biostatistics
School of Public Health
University of South Carolina
Columbia, South Carolina 29208

Ms. Susan A. LeBailly (F)
Director, Division of Health Services
Development
American Academy of Pediatrics
1801 Hinman Avenue, Box 1034
Evanston, Illinois 60204

Mr. James Lubitz, M.P.H. (H)
Statistician
Office of Research and Demonstrations
Health Care Financing Administration
6340 Security Boulevard
Baltimore, Maryland 21207

Mr. Henry Malin (S)
Laboratory of Epidemiology and Population
Studies
National Institute on Alchol Abuse and
Alcoholism
5600 Fishers Lane, Room 14C-27
Rockville, Maryland 20857

Mr. James S. Marks (E)
Chief, Epidemiology Branch, Nutrition
Center for Health Promotion and Education
Centers for Disease Control
Building 3, Room SB44
Atlanta, Georgia 30333

Ms. Gillian Marsden, M.S.P.H. (1)
Seattle Division Director
Seattle-King County Department
of Public Health
610 3rd Avenue, 1500 Public Safety Building
Seattle, Washington 98104

Mr. Edward J. Martin (K)
Assistant Registrar
Division of Vital Statistics
Rhode Island Department of Health
75 Davis Street
Providence, Rhode Island 02908

James O. Mason, M.D, Dr. P.H. (1)
Executive Director
Utah Department of Health
P. O. Box 2500
Salt Lake City, Utah 84110

Thomas J. Mason, Ph.D. (S)
Environmental Epidemiology Branch
National Cancer Institute
Room 3C-29, Landow Building
Bethesda, Maryland 20014

Ian McDowell, Ph.D. (X)
Director, Health Care Research Unit
Department of Epidemiology
University of Ottawa
451 Smyth Road
Ottawa, Ontario, CANADA K1H 8M5

Paper presented by:
Mr. Ed Praught, B.S.E.
Health Division
Statistics Canada
17-K, R.H. Coats Building
Ottawa, Ontario, Canada K1A 0T6

J. Michael McGinnis, M.D. (1)
Deputy Assistant Secretary for Health
Department of Health and Human Services
200 Independence Avenue, S.W.
Washington, D. C. 20201

Ms. Barbara E. Merrill (D)
Evaluation Manager
Control Data Corporation
8100 34th Avenue South (HQC02S)
Minneapolis, Minnesota 55440

Paper presented by:
David R. Anderson, Ph.D.
Benefits Evaluation Consultant
Control Data Corporation
Box 0, HQS02J
Minneapolis, Minnesota 55440

Andrew C. Montgomery, Ph.D. (H)
Survey Research Laboratory
University of Illinois, Rm. 2300
Formfit Building, Box 6905
Chicago, Illinois 60680

Ms. Barbara Moser (N)
Assistant Manager
Computer Applications Center
Research Triangle Institute
P.O. Box 12194
Research Triangle Park, N.C. 27709

Ms. Ellen Naor (O)
Director, Division of Data and Research
Maine Bureau of Health Planning and Development
151 Capitol Street
Augusta, Maine 04333

Mr. David R. Netherton, M.S. (*)
Consultant
1366 Broadway
Sommerville, Massachusetts 02144

Donald L. Patrick, Ph.D., M.P.H. (3)
Associate Professor of Social Medicine
Box 3 Wing D 208H Medical School
University of North Carolina
Chapel Hill, North Carolina 27514

Harrold P. Patterson, Ph.D. (B)
Supervisor, Data Management Section
Bureau of State Health Planning and
Research Development
Texas Department of Health
1100 West 49th Street
Austin, Texas 78756

Mr. Edward L. Perrine, M.P.H. (T)
Executive Director
Health Councils, Inc.
10051 Fifth Street No., Suite 253
St. Petersburg, Florida 33702

Mr. Kirk T. Phillips, M.S.W. (Q)
Policy Analyst, Health Policy
Corporation of Iowa
Carriers Bldg., 601 Locust, Suite 330
Des Moines, Iowa 50309-4090

Mrs. Dorothy P. Rice (2)
Regent's Lecturer
Aging Health Policy Center
University of California, N 631 Q
San Francisco, California 94143

Basil Rifkind, M.D., F.R.C.P. (C)
Chief, Lipid Metabolism-Atherogenesis Branch
National Heart, Lung, and Blood Institute
National Institutes of Health
Federal Building, Room 401
Bethesda, Maryland 20205

Ms. Marilyn J. Rosenstein, B.A. (NIMH)
Supervisory Statistician
Survey and Systems Research Branch
National Institute of Mental Health
5600 Fishers Lane, Room 18C-05
Rockville, Maryland 20857

William A. Rush, Ph.D. (A)
Health Analyst
Meidinger-Health Risk Assesment
3500 W. 80th Street
Minneapolis, Minnesota 55431

M. Nagi Salem, Ph.D. (M)
Research Analysis Specialist
Minnesota Department of Health
717 Delaware Street, S.E.
Minneapolis, Minnesota 55440

David S. Salkever, Ph.D. (U)
Professor, School of Hygiene and Public Health
The Johns Hopkins University
615 North Wolfe Street
Baltimore, Maryland 21205

Mr. G. L. Sandifer, M.P.H. (*)
Administrator, Bureau of Maternal and
Child Health
South Carolina Department of Health and
Environmental Control
2600 Bull Street
Columbia, South Carolina 29201

Ms. Pam Schnagl, M.S. (M)
Deputy Section Chief, Resource Data Section
Wisconsin Division of Health
Department of Health and Social Services
1 West Wilson, P.O. Box 309
Madison, Wisconsin 53701

Mrs. Anne A. Scitovsky (H)
Chief, Health Economics Division
Palo Alto Medical Research Foundation
860 Bryant Street
Palo Alto, California 94301

Mr. Jack C. Smith, M.S. (G)
Chief, Research and Statistics Branch
Division of Reproductive Health
Centers for Disease Control
1600 Clifton Road, N.E., Building 1-4409
Atlanta, Georgia 30333

David Spivack, MRP (I)
Senior Planner, Mount Sinai
Medical Center of Greater Miami
4300 Alton Road
Miami Beach, Florida 33140

Mr. Saul M. Spivack (V)
Rehabilitation Research and Training Center in Aging
Hospital of the University of Pennsylvania
Box 590, 3400 Spruce Street
Philadelphia, Pennsylvania 19104

Thomas Stephens, Ph.D. (B)
Executive Director
Canada Fitness Survey
506 - 294 Albert
Ottawa, Ontario, CANADA K1P 6E6

Tai Sugimoto, Ph.D. (U)
Assistant Professor, Department of Preventive Medicine and Community Health
University of South Carolina
Columbia, South Carolina 29208

Mr. Gene D. Therriault, M.S.P.H. (S)
Research Scientist IV, Biostatistics
New York Department of Health
Empire State Plaza, Tower Building
Albany, New York 12237

Hugh H. Tilson, M.D., Dr.P.H. (1)
Department Head, Product Surveillance and Epidemiology
Burroughs Wellcome Company
3030 Cornwallis Road
Research Triangle Park, North Carolina 27709

Mr. George Van Amburg (S)
State Registrar and Chief, Office of Vital Statistics
Michigan Department of Public Health
3500 N. Logan Street, Box 30035
Lansing, Michigan 48909

Lois M. Verbrugge, Ph.D. (V)
2659 Heather Way
Ann Arbor, Michigan 48104

Mr. Daniel Waldo
Bureau of Data Management and Strategy
Health Care Financing Administration
6340 Security Boulevard
Baltimore, Maryland 20217

James Wiley, Ph.D. (B)
Assistant Director, Survey Research Center
University of California at Berkeley
2538 Channing Way
Berkeley, California 94720

Ronald L. Williams, Ph.D. (K)
Director, Health Data Research Facility
Community and Organization Research Institute
University of California
Santa Barbara, California 93106

Gwynne R. Winsberg, Ph.D. (3)
Adjunct Associate Professor
GRW Associates
5533 N. Glenwood Building
Chicago, Illinois 60640

Mr. Michael Witkin (NIMH)
Supervisory Statistician
Survey and Systems Research Branch
National Institute of Mental Health
5600 Fishers Lane, Rm. 18C-05
Rockville, Maryland 20857

Fredric D. Wolinsky, Ph.D (Y)
Assistant Professor and Director of Health Services Research
St. Louis University Medical Center
3525 Caroline Street
St. Louis, Missouri 63104

Thomas W. Woolley, Ph.D. (U)
Assistant Professor
Section of Medical Education
Quillen-Dishner College of Medicine
East Tennessee State Univ., Box 19 720A
Johnson City, Tennessee 37614-0002

Mr. Theodore D. Woolsey (F)
Health Statistics Consultant
8121 Rayburn Road
Bethesda, Maryland 20817

Mr. Robert A. Wright (J)
Division of Health Interview Statistics
National Center for Health Statistics
3700 East West Highway
Hyattsville, Maryland 20782

Mr. Richard Yaffe (HCFA)
Acting Director, Evaluation Studies Staff
Office of Demonstrations and Evaluations
Health Care Financing Administration
6340 Security Blvd., Oak Meadows Bldg.
Baltimore, Maryland 21207

*Backup Speaker

1983 PHCRS PLANNING COMMITTEE

National Center for Health Statistics:

Mr. Robert Israel, Chairperson
Dr. Edward Bacon
Mr. Phillip Beattie
Mr. E. Earl Bryant
Dr. Jacob Feldman
Mr. Robert Fuchsberg
Mrs. Mary Gabriel
Mr. Kenneth Harris
Mr. Peter Hurley
Dr. Joel Kleinman
Mr. Garrie Losee
Mr. Robert Murphy
Mr. John Patterson
Mrs. Nancy Pearce
Mr. James Smith
Mr. William Stewart
Ms. Barbara Weisel
Mr. Ronald Wilson
Dr. Alvan Zarate

STAFF OF THE PHCRS SECRETARIAT

Program Development Staff
Office of Program Planning, Evaluation
and Coordination
National Center for Health Statistics

Mr. William F. Stewart, Acting Chief

Mr. James A. Smith, Executive Secretary

Mrs. Catherine Barrett
Mrs. Mary Gabriel
Ms. Barbara Weisel

EXHIBITS AND PUBLICATIONS SUPPORT STAFF

Jarmila G. Frazier
Annette F. Gaidurgis
Patricia A. Vaughan
Stephen L. Sloan
Geraldine K. Cherry
Naomi M. Forester

*U.S. GOVERNMENT PRINTING OFFICE : 1984 0-429-012/364

FILE