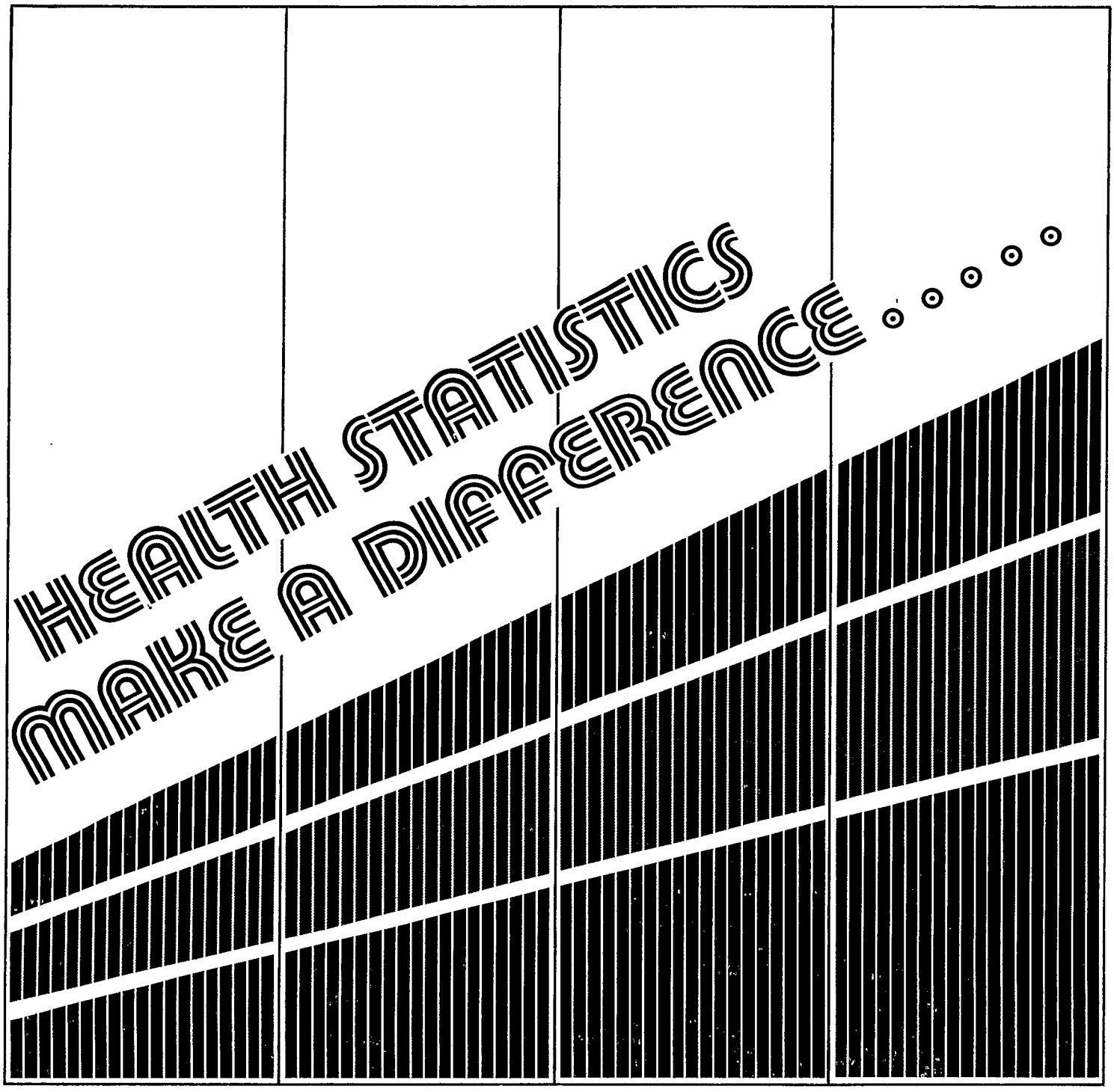
Proceedings of The 1985 Public Health Conference on Records and Statistics

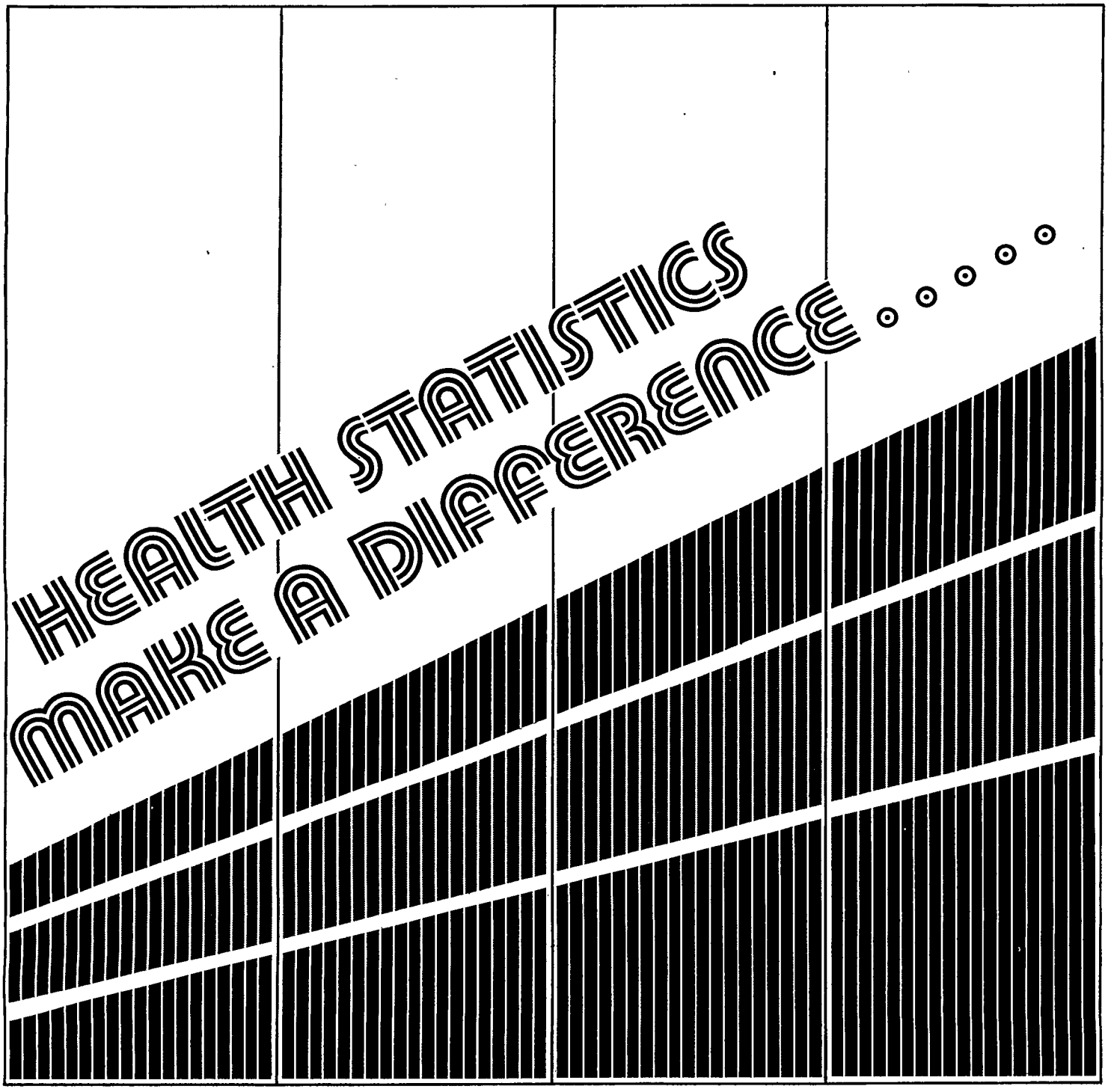


Proceedings of The 1985 Public Health Conference on Records and Statistics

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

Hyattsville, Maryland
December 1985

DHHS Pub. No. (PHS) 86-1214



Each article has been prepared by the author.

TABLE OF CONTENTS

FIRST PLENARY SESSION

CONCURRENT SESSIONS

* Speaker

CONCURRENT SESSIONS

SECOND PLENARY SESSION

CONCURRENT SESSIONS

CONCURRENT SESSIONS

THIRD PLENARY SESSION

CONCURRENT SESSIONS

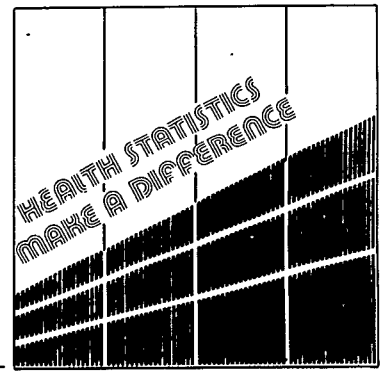
CONCURRENT SESSIONS

First Plenary Session

Good Data and Good Statistics Mean Good Policy and Program (And, Bad Means Bad)

Tracking the Nation's Health: Remarks on the 25th Anniversary of the National Center for Health Statistics

Health Statistics Make a Difference The National Perspective



OPENING REMARKS

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

Good morning, and welcome to the 20th national meeting of the Public Health Conference on Records and Statistics. It is a great pleasure to have you here.

The meetings of the Public Health Conference are one of the few occasions when the people who produce and use vital and health statistics gather together from all parts of the United States to discuss and debate their mutual concerns and to share their knowledge and experience. About 1,000 of you are here now, a very good turnout. As of the end of last week, nearly all states had designated a representative to the conference, and I hope that we will fill any remaining gaps before the day is over. Also represented are some of the outlying areas, including Guam and Puerto Rico.

Without exception, the meetings of the Public Health Conference have been fruitful and productive sessions for the participants and for the National Center for Health Statistics as sponsor. Our staff at NCHS have worked hard to ensure a stimulating program for this year. Before I introduce our guests, let me spend a minute broadly outlining the next few days.

The theme of this conference "Health Statistics Make A Difference," was chosen in recognition of the contributions that health statistics have made in the past and will make in the future to the improvement of health.

The program addresses four areas with primary need for health information and statistics. These areas are: Data Access and Availability, Statistical Organization and Planning, Data Analysis, and Methodology and Technology.

In the scheduling, we are following the format that has worked so well at previous conferences. We will begin each morning with a plenary session. On successive days, the speakers at the plenary sessions will address the conference theme, Health Statistics Make A Difference, from their respective perspectives as national, state, or local officials.

After the plenary session, the remainder of each day will be devoted to a mix of concurrent and special sessions. These sessions cover a wide range of areas and interests, and it is not going to be easy to choose the one session you most want to attend at any time.

Let me note for you, also, that the National Center for Health Statistics is observing its 25th anniversary this year, and we have incorporated elements of our celebration into t' is conference, as you will see in your program packet.

We have a full agenda and I am looking forward to every minute of it.

WELCOME TO WASHINGTON, D.C.

Andrew D. McBride, M.D., M.P.H.
Commissioner of Public Health
District of Columbia

It gives me great pleasure to be here this morning and have an opportunity to greet and welcome you. I am here as the Commissioner of Public Health and I bring you greetings from Mayor Marion Barry. The Mayor is very pleased that the Conference chose Washington, DC as the site for its twentieth meeting. Being 25 years old, you are well into adulthood now, into maturity, and we would like to see how this maturity in the statistical area and the record keeping area applies to policy and program development.

I come from a long professional history in public health and often we in public health have done a lot of numbers collecting. But we have a long way to go in getting those numbers together, applying those numbers, and having those numbers affect our policies and programs.

The need for conferences like this is self-explanatory. We have now a growing recognition that goals need to be quantified and that access to services can be measured with outcomes in terms of morbidity and mortality. There are certain statistics that catch our attention all the time. I guess those of us who have been long in the public health discipline think in terms of infant mortality as one of those areas.

We in the District have had the dubious distinction of having a very high infant mortality rate in relationship to other cities. In addition to that, we also recognize that there are trends in infant mortality that concern us all. Those of us in public health recognize that infant mortality still represents a very significant figure.

We also look with alarm at the recent trends in the slowing of the average decline in infant mortality. We would like to understand what all the factors are that led to the improvement of the infant mortality rate and the slowing of the decline in the infant mortality rate.

Obviously, there are some areas that we need to work on, particularly in the area of preventive health. Often the area of preventive health is given short shrift in health and policy planning. Often the untoward results of the lack of preventive programs are overlooked.

We should reflect on how our prevention programs can be developed. The keystone for those programs must be a needs assessment based on statistics and based on quantifiable terms. We know that there are certain quantifiable terms that stand out. Particularly in this country we talk about the differentials between white and non-white infant mortality that are very clear. In fact, the gap between whites and blacks in the infant mortality area seems to be widening. It should be viewed with some alarm by those who are in policy-making positions and those who are in health programs and should stimulate research into why the differentials between white and black persist in this country.

I think all of us have viewed the recent events in South Africa with great alarm. It brings to the fore the whole issue of differences between whites and blacks. We still have in this country some of the same kind of differentials. Of course, in South Africa, the extremes are much greater. It is interesting in terms of dealing with statistics how health statistics can articulate broader problems. The health side of the South African apartheid system articulates very clearly a very severe problem in that country. In this country the differential between white and black infant mortality is some two times, which is in itself a national disgrace. However, South African white infant mortality rates are 13 per thousand, while the infant mortality for black Africans is some 90 per thousand. In certain areas it goes as high as 200 to 300 per thousand.

In other areas, in terms of quantifiable things, you can look at the health manpower development where the physician ratio of white physicians is one in 300 persons and for black Africans it is one to 19,000 persons.

Most of us can understand numbers like that. That is an example of how very basic health statistics can guide our thinking in certain areas. I think that the future of our nation's health will depend on what we can do to prevent diseases and promote health lifestyles.

We all know the serious health problems that AIDS presents. This is clearly an area that represents an epidemiologic challenge to all of us in public health. We are being made aware that the solution to the AIDS situation will be in the prevention of that disease, rather than in the cure after the disease has taken hold.

I think we public health professionals also have to recognize the so-called new morbidities, recognizing that those of us in the health field have to take some responsibility for the other kinds of morbidities and mortalities that affect us. We usually think in terms of accidents, homicides, suicides and other kinds of preventable acts that relate to behavior and lifestyles. We in the District are particularly interested in those so-called new morbidities and the interrelationships between those and others which might be considered non-public, non-traditional public health areas, such as the areas of mental health, and drug and alcohol abuse services.

We in the health field, and particularly in the public health field, have to see the interrelationships; we must be the driving force to pull various disparate groups together and to decategorize certain programs. I think we have been categorized enough. We have suffered in the public health field from being too categorical. We have our sexually transmitted disease clinics, and we have our tuberculosis clinics; then we go on to other kinds of categorical programs, but now we can see more interrelationships.

It is very significant that the increase of tuberculosis is mostly in the homeless population and that accounts for our recent small increase in tuberculosis. We also recognize that well over half of that homeless population are severely, chronically, mentally ill and have other kinds of primary health care needs that should be met. In addition, another 40 percent or more of those have severe alcohol and substance abuse problems. We must have a data system in order to see those interrelationships. We must have a system that recognizes and can articulate needs so that in the implementation phase we can develop programs that work in a coordinated, cooperative and, at times, integrated fashion. I think we have to think in terms of integration of services both within the public sector as well as the private sector.

You have a very rich agenda here. I hope that you all take the opportunity to partake of the Conference, both in an informal and a formal way. The formal way is to attend the sessions; the informal way is relating to one another by sharing information and networking. I hope that you also share the amenities of the city. Washington, DC is a great town, and I hope that you will spend some time enjoying yourselves here.

So, once again I welcome you to the District of Columbia. I hope to participate and meet with some of you on an individual basis and I hope that this Conference proves fruitful for all of you.

Thank you.

TRACKING THE NATION'S HEALTH: REMARKS ON THE 25TH ANNIVERSARY OF THE NATIONAL CENTER FOR HEALTH STATISTICS

Manning Feinleib, M.D., Dr.P.H.

This month of August 1985 marks the 25th anniversary of the National Center for Health Statistics. The Center was established in 1960 by Surgeon General Leroy Burney to "bring together the major components of Public Health Service competence in the measurement of health status ... and the identification of significant associations between characteristics of the population and health-related problems."

We in the Center are marking our silver anniversary in several ways this year, but we wanted particularly to share it with the participants at this Public Health Conference on Records and Statistics. Many of you have had ties with the Center over the entire period of its existence and all of you are users or producers of health data.

On a personal note, I was startled to realize that the 25 years of the Center are half of my own life. When the Center was established in 1960 I was a medical student. Although I had known of the Center's activities and used the data in various roles at NIH and my own research all these years and had even served on some of the advisory groups of the Center, I have been director for only the past 2½ years. Therefore, the achievements of the first quarter century are largely the work of other people -- the Center's former directors, its dedicated staff, and people in and outside the government who offered advice, criticism and suggestions.

For this particular silver anniversary I can speak objectively, because I was not directly involved in most of the activities, and also appreciatively, because like other people in the health field, I was one of the beneficiaries of its work.

Thinking back over what I know of these activities of the Center, it is amazing how NCHS has expanded the national health data base. I believe that this expansion is one of the very considerable achievements of the Center and its staff over the past 25 years.

I would like to use this opportunity to show you how the health of the country has changed during these last 25 years, how the data systems of NCHS have grown to provide the data demonstrating these changes, and a few examples of how, in the nature of the theme of this Conference, health statistics have made a difference.

The charts that I will use this morning are drawn largely from a forthcoming publication called, "Charting the Nation's Health, Trends Since 1960." I will quote liberally from the text of this publication. Your registration kits have an order form; if you complete it and return it, we will be happy to send you a copy which should be off the press in about six weeks. Many people at NCHS were involved in the planning and production of this chartbook. I cannot mention all of them, but I want to commend especially three key people. They are Pat Golden, from our Division of Epidemiology and Health Promotion, who coordinated and wrote most of the report; and our Publications Branch, represented by Rolfe Larson, the editor of this publication, and Pat Vaughn, the designer of the chartbook.

Much of the period 1960-85 was characterized by change and turmoil in this country. The 1960s have been called the decade of rising expectations, and the 1970s the decade of disillusionment. You won't find these terms in the NCHS data base but we can see evidence of the effects of some of the events during the past 25 years. The Medicare and Medicaid programs were implemented to ensure better health care for large components of the population. At the same time we were learning more about the causes of disease and this has led to the recent trend toward healthier lifestyles by many of the population .

We will begin with one of the oldest and most widely used measures of a nation's well-being -- life expectancy at birth (Figure 1). These data are based on the National Vital Statistics System and the registration of deaths and other vital events by the States. The vital statistics system, which dates back to the beginning of the century, was one of the major statistical components of the Public Health Service that became part of the National Center for Health Statistics in 1960. These particular data demonstrate the continuity that is essential in building a data base to mark trends in our health.

Americans are living longer. In 1960, life expectancy at birth was about 70 years. By 1983, life expectancy was almost 75 years. Over the years, white people have had longer life expectancies than black people, as noted by Dr. McBride, but in the period 1960-83 there has been a closing of that disparity.

Life expectancy is an efficient way of summing up what is happening to the mortality rates across the age range. As mortality rates decline, life expectancy increases. Life expectancy in the U.S. did not actually change appreciably between 1950 and about 1965. Like most students of medicine and public health in the late 50's and early 60's, I was taught that we may have already achieved the maximum life expectancy, at least for white persons.

But since the mid-1960s, mortality has turned downward and life expectancy has increased markedly. The decline in mortality has occurred in all ages, from the youngest to the oldest, for both sexes, and for all racial groups, and the decline has been found for most leading causes of death.

The infant mortality rate was reduced by more than 50 percent from 1960 to 1982 (Figure 2). The rate is continuing to decline, let us not lose sight of that, although the rate of decline seems to have slowed. However, despite the decline for both white and for black infants, the race differential in infant mortality has changed very little during the past 25 years.

Mortality trends at older ages are illustrated in Figure 3 for the age group 55-64 years. Heart disease, cancer, and stroke are still the leading causes of death in this group, just as they had been in 1960. But there has been remarkable improvement for the cardiovascular diseases. Stroke deaths have declined by fully 60 percent during these 25 years and heart disease in this age group has dropped by 36 percent. For cancer,

however, the death rate was actually 11 percent higher in 1982 than it had been in 1960.

Respiratory cancer is the major factor in the continued increase in cancer mortality in this age group. If there had been no additional deaths from respiratory cancer, mortality from cancer would have actually shown an 8 percent decrease between 1960 and 1982.

In fact, by 1982 the respiratory cancer death rate for this age group was approaching two fold the rate that it was in 1960 (Figure 4). In the last few years, from 1979 to 1982, the increase was far greater for women -- nearly 20 percent -- much larger than the increase among men of only about 3 percent.

Among teenagers and young adults, there is a different pattern of mortality. Here deaths from violent causes have been a concern throughout the past 25 years. In 1982 as in 1960, accidents were the leading cause of death for persons aged 15-24. Over three-quarters of the deaths in 1982 were due to a violent cause -- homicide, suicide, and accidents.

There has been some improvement in some specific causes of violent deaths in this age group, homicide, for example. Although it is still the leading cause of death for young black males 15-24, it has been declining somewhat since 1971.

Deaths from suicide are a special tragedy among young people. There have been declines for both white and black females, and the rates for white males seem to have stabilized (Figure 5). But among the subgroup of teenagers 15-19 years, only the suicide rates for white males have shown no improvement.

Health status is not as easily quantifiable as is mortality. Various aspects of the effects of illness and injury on the individual -- the limitation of activity that results, the work and school loss time -- have been measured through the last 25 years on a continuous basis by the National Health Interview Survey. This was the first continuing survey started under the National Health Survey Act of 1956, and that National Health Survey program was the other statistical component that existed when NCHS was established in 1960. The Health Interview Survey is conducted annually on a national sample of 40,000 households, covering about 110,000 individuals.

There is no single summary measure of health status, but one of the best that we have is personal perception of health; that is, a person's view of his or her health in comparison with other people of the same age. The significance of this measure lies in the fact that people may act as if they are in good health if they think of themselves as in good health, regardless of any impairments, disabilities, or illnesses that they may actually have.

Over the past decade, 87 percent of the population have stated that their health was good or excellent (Figure 6). The proportions have been quite stable. Among the elderly living in the community in 1981, nearly 70 percent perceived their health as good or excellent compared with their peers and at younger ages we see that there is appreciably better regard of their health status.

This chart on self-perceived health status illustrates another way that NCHS expanded its data base. That is by the addition of new topics and areas of study to ongoing statistical systems. The question on self-perceived health status was added to the National Health Interview Survey in 1972. Since then, research has shown this measure to correlate highly with other measures of health status and with measures of health service utilization.

Some of the risk factors that have become so important in health promotion were studied early on in the surveys, so that trend data are available today. This chart shows trends in serum cholesterol levels, as obtained in what are now the National Health and Nutrition Examination Surveys (Figure 7). These cyclical surveys provide standardized physical examinations to samples of the population by means of mobile examination centers.

Over the time of these surveys, serum cholesterol levels have declined somewhat for every age group of men and women. This trend is encouraging, but many people still have levels above that at which the risk of coronary heart disease begins to rise sharply. Tracking of other risk factors -- hypertension, overweight -- continues in our periodic National Health and Nutrition Examination Surveys.

Over the years 1960-1985 significant changes took place in the way Americans use health care services, as well as in the way these services are organized and provided. Changes in utilization of nursing homes and hospitals were captured in the Center's other surveys and inventories that began in the early 1960s.

The growing size of the elderly population, as well as the aging of the elderly population itself, was a major influence on health care in general and on nursing home care in particular. Although only a small proportion of the elderly live in nursing homes, most of the nursing home population is elderly (Figure 8).

The implementation of Medicare and Medicaid in 1966 and the liberalization of eligibility requirements for these programs in 1972 were the main contributors to the rapid increase in the rate of nursing home use in the mid-1960s and early 1970s. Between 1969 and 1973, the number of beds in nursing homes with 25 beds or more soared 30 percent. The rate rose from 43.4 to 56.8 nursing home beds per 1,000 population. Since then, the number of nursing home beds per capita has decreased somewhat (Figure 9). The National Nursing Home Survey, the source of these data, have been conducted at several intervals since the mid-1960s and we are in the field with the 1985 survey.

Notable change in the availability and utilization of health care resources has not been confined to the nursing home industry, however. Over the years significant changes have also taken place in the way Americans have come to use dentists, physicians, and hospitals, as well as in the way these practitioners make their services available. Although changes in the nature and availability of health care services and resources have affected patterns of mortality and morbidity, changes in mortality and morbidity have, in turn, helped shape patterns of health care utilization.

For example, in contrast to the increase in nursing home beds, the number of beds in specialty hospitals has decreased. In 1982 the number of beds per 1,000 population was one-half of what it had been in 1970. During this same 12-year period the number of beds in general hospitals also decreased slightly.

Discharge rates from short-stay hospitals were actually rather stable over much of this period with only a modest but steady increase in the discharge rates since 1973 (Figure 10).

For the most part, the trend in average length of stay has been in the opposite direction from the discharge rate. The average length of stay rose somewhat between 1965 and 1968, but declined thereafter.

These data are from the National Hospital Discharge Survey, which has been conducted on a continuing basis since 1965. The survey provides diagnostic data on the causes of hospitalizations and enables the monitoring of trends in surgery and in the use of other new technologies.

A major component of hospital use can be ascribed to surgery performed on women and on elderly people. The tracking of surgery by means of the Hospital Discharge Survey showed changes, for example, in the rates that may reflect in part the availability of data on the use of these procedures. For example, hysterectomies are still the most common major surgical procedure. However, the rates have declined since debate about the frequency of such operations began some years ago (Figure 11). Although part of this trend may be due to a decrease in women at risk in the sense that they may have had hysterectomies earlier in life, it is felt that a major contributor to this trend was the documentation that the rates had been increasing at an unsuspected level. This information led to a reevaluation of the indications for the procedure. Undoubtedly these data also led to a greater awareness on the part of women themselves about the frequency of this operation and probably led to dialogues between the patients and the physicians which may have slowed down the rate of increase of this operation.

A similar concern about the increase in the proportion of cesarean deliveries is being expressed now in countless articles and meetings. It will be interesting to see whether knowledge of the sharp increase in rates of cesarean deliveries will influence future trends in the use of this procedure (Figure 12).

The Hospital Discharge Survey also provides essential information to track changes in surgery that represent the application of new medical knowledge and opinion.

For example, Figure 13 shows the recent trends in surgery for breast cancer. Whereas radical mastectomy had been the treatment of choice in this condition 15 years ago, it has rapidly been replaced by the modified radical procedure.

Another dramatic change in the frequencies of certain types of procedures is manifested by the trends in cataract surgery (Figure 14). Lens extraction has always been performed more frequently on the elderly than other age groups but it is not simply the increase in the proportion of elderly people that has led to the marked increase in this operation in the past 20 years. Between 1965 and 1983 the number of lens extraction procedures increased from 142,000 to 630,000 each year. Increasingly, the Hospital Discharge Survey also shows that cataract surgery is accompanied by insertion of new devices such as prosthetic lenses.

During the past 25 years there have also been marked changes in medical and technical knowledge relating to fertility and childbirth as well as marked changes in our social and legal attitudes about contraception and abortion.

The decline in fertility is a major demographic trend of the past 25 years. As a way of summing up what is happening in age-specific birth rates, the total fertility rate shown in Figure 15 is analogous to the measure of life expectancy in summarizing the mortality experience of a population. In the late 1960s and through the early 1970s until the present American women have clearly been having fewer babies than women in the late 1950s and early 1960s had. In fact, by 1972 fertility levels for the total population and for the white population have fallen below the replacement level of 2,100 births per thousand women.

The decline in fertility has been the major influence on the aging of the population that we have observed since 1960, surpassing, in fact, the effect of the decline in mortality.

Fertility data are routinely available from the vital statistics system. However, data from that system are limited to the information reported on the birth certificates. Therefore, more information is needed on the dynamics of decision-making with respect to childbearing and family size.

With the establishment of the National Survey of Family Growth in 1973, the Center significantly expanded the data base available for the study and projection of fertility and population growth. We have completed three surveys of family growth so far. For the last survey the reference population was enlarged to cover all women of childbearing age, regardless of their marital status. As a result of these surveys, there is a large body of data on how many children a couple wants, when they plan to have them, and how they are controlling their fertility.

Trends show a decline in the proportion of births that were unwanted at conception, as the use of contraception increased between 1965 and 1982 (Figure 16). There were shifts in the methods used. Married couples have shown more than a threefold increase in surgical sterilization in the period 1965 to the present. Over the same period, the popularity of the pill declined sharply, although it is still the method of choice among women who eventually want to have more children. Overall, the surveys have shown, couples remain substantially more effective at controlling the numbers of children that they have than at actually determining when they will have their children.

A data base can also be expanded by analyzing data differently than originally planned. Data on blood lead levels are a good example. These were collected in the National Health and Nutrition Examination Survey to assess the distribution of lead in the population, particularly among young children. These distributions indicated that there were more young children with

excessive exposure to lead than would have been anticipated on the basis of smaller studies, which itself was a very important finding. But when Lee Annest and other researchers at NCHS and CDC examined the data in a different way, they found that blood lead levels had been declining over the four years of the survey from 1976 to 1980, and paralleled remarkedly a change in the amount of lead used in gasoline during that period.

Now when we are having budget sessions and program reviews the staff of the NHANES survey are likely to come in wearing buttons that say "NHANES II got the lead out." When EPA had to decide what to do about lead in gasoline and was prepared to actually drop the standards, they looked at the NHANES data closely and at the evidence of the health risks. They decided to increase the severity of the standards for gasoline lead. Health statistics have made a difference.

Figure 18 sums up health in the past 20-25 years. It shows the years added to life expectancy in two 10-year periods. Gains were greater in the recent 10-year period from 1972-82 than in the preceding 10 years. Over the entire 20 years, females increased their life expectancy much more than men, and minority groups increased their life expectancy much more than did the white population.

The charts that I have just shown represent only a few of the changes and modifications that have been made over the past 25 years in the NCHS data systems. Behind each change and modification lies an extensive process.

It is easy enough to say that we added a supplement or carried a different question on a survey but the process is actually a good deal more complex that that.

Let me illustrate with the work being done at the Center in the area of physical fitness. Relationships between health and physical fitness are not well-documented. National estimates of the physical fitness of American adults based on standardized state-of-the-art assessments of national probability samples simply do not exist. Nor do generally accepted, national representative data on exercise patterns exist.

There is a clear need for data that will help to determine relationships between health and physical fitness such as there are for many other parameters. Data also are needed to measure progress toward the 1990 physical fitness and exercise objectives in health promotion.

Given these needs, the question becomes how to measure physical fitness in the general population. To answer this question, NCHS is involved in an extensive developmental undertaking. As a part of it, we have commissioned the assistance of outside experts who have prepared a series of papers on aspects of physical fitness assessment. These advisors and our own staff are concerned with methodology of measuring physical fitness, with issues about the protection of human subjects in a physical fitness testing situation, with analysis and interpretation, and with the many other considerations that lie behind the simple question of how to measure physical fitness in the general population.

This work is geared particularly to inclusion of a component to measure physical fitness in the next National Health and Nutrition Examination Survey, which is slated to go in the field in 1988. Other aspects of physical fitness may also be studied in the National Health Interview Survey by a series of well-structured questions. This gives you an example of the interlinkage of our various surveys and the entire effort will, I believe, represent a significant contribution to the scientific assessment of health status, just as other developmental projects have been done over the past 25 years.

I am sure that 25 years from now the then-Director of the National Center for Health Statistics will be able to show equally encouraging and diverse trends to the Public Health Conference on Records and Statistics that will be held that year. The data systems will continue to evolve and to enlarge the national health data base.

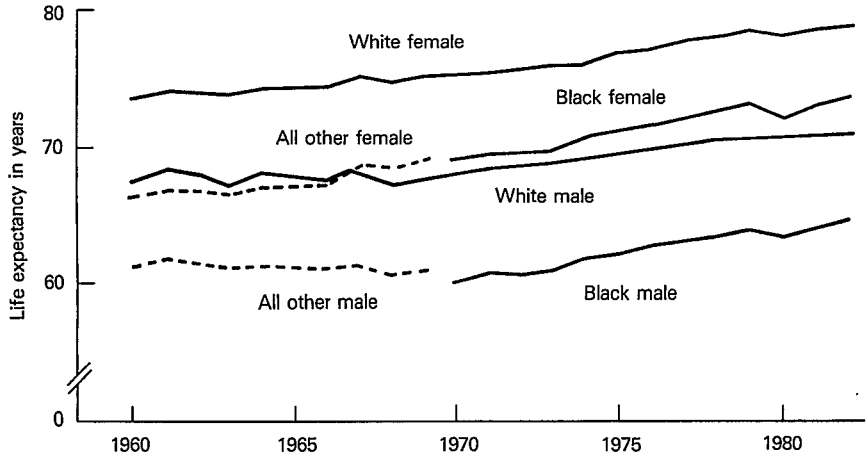
We already have begun some new types of studies. For example, longitudinal surveys that I believe are essential for the future study of health. In the NHANES I Epidemiologic Followup we are interviewing participants in the original survey to find out how their health changes over time and the associations between risk factors and those changes over a 12 year period since we first examined them. We are also improving our data bases on the health of minorities. And we are moving toward a national system of linked birth and infant death records that is an essential component of the data base for tracking this important objective in health promotion.

There are some lines of Walter Lippman's that I would like to close on that many statisticians like to quote in reference to their own work. Lippman wrote, "The printing of comparative statistics of infant mortality is often followed by a reduction of the birth rate of babies. Municipal officials and voters did not have, before such publication, a place in their picture of the environment for those babies. The statistics made them visible, as visible as if the babies had elected an alderman to air their grievances."

By expanding the national data base, by adding to the store of information about health, by widely disseminating its findings, the NCHS has made visible many trends and conditions of health. By describing many aspects of health in the United States, NCHS data have assisted the decisions of policymakers and helped to inform the public. Thus, it is clear that those statistics have made a difference and continue to make a difference. NCHS, I believe, has fully met the objectives envisioned by Surgeon General Burney 25 years ago. We intend to do equally well in the future.

Figure 1.

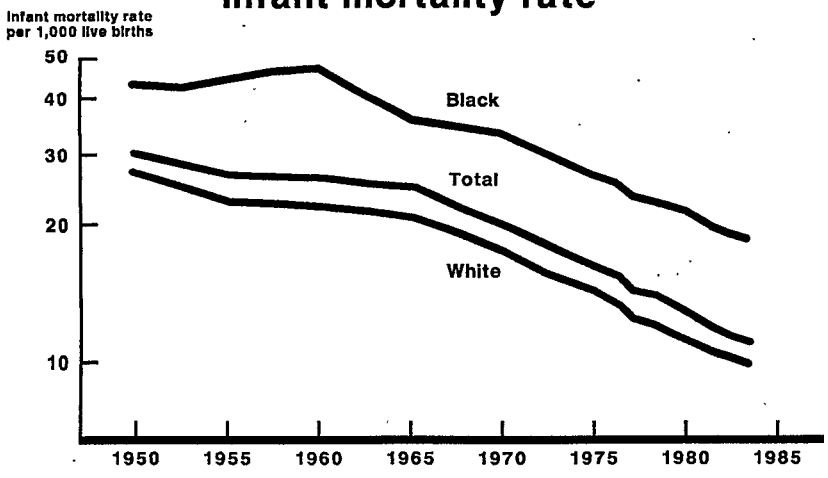
Life expectancy at birth, by race and sex: United States, 1960-82



SOURCE: Division of Vital Statistics, National Vital Statistics System.

Figure 2.

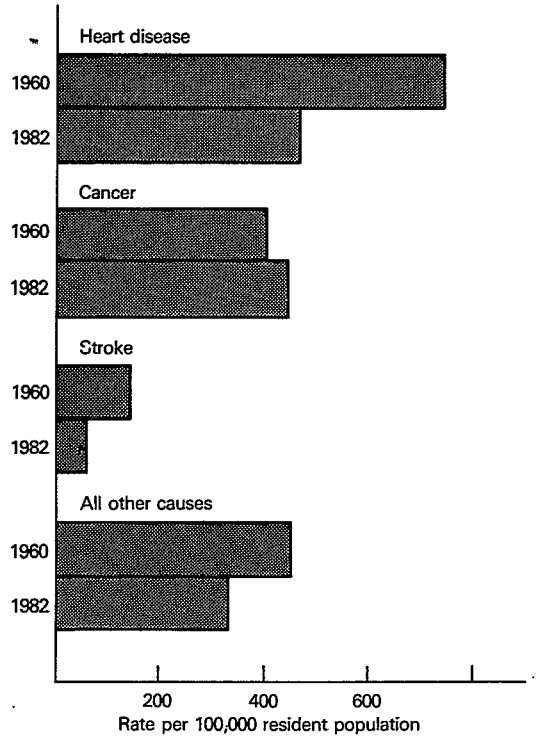
Infant mortality rate



SOURCE: National Center for Health Statistics

Figure 3.

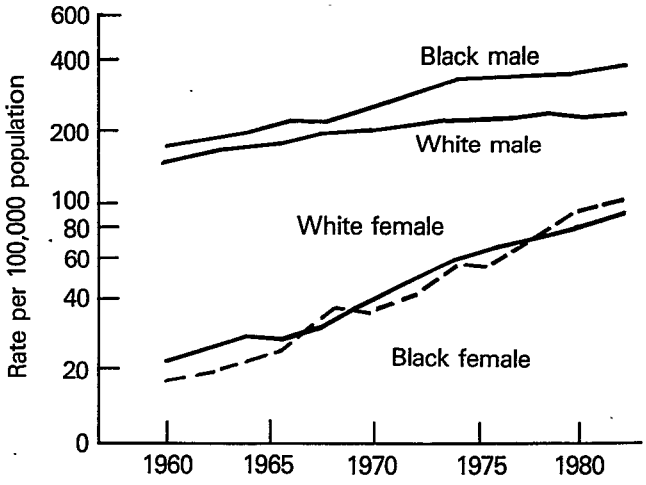
Death rates for heart disease, cancer, and stroke for persons 55-64 years of age: United States, 1960 and 1982



SOURCE: Division of Vital Statistics, National Vital Statistics System.

Figure 4.

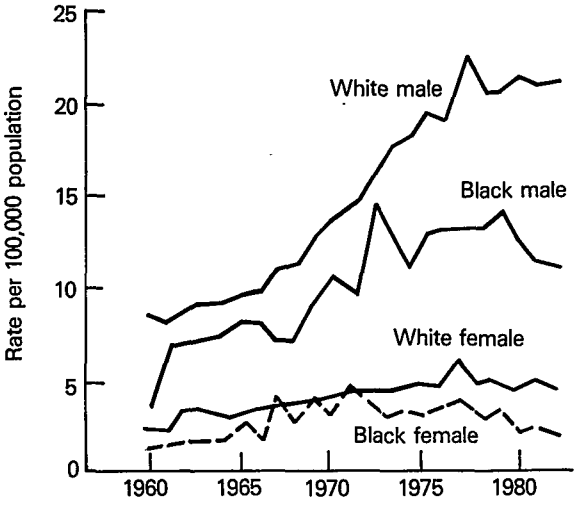
Death rates for respiratory cancer for persons 55-64 years of age, by race and sex: United States, 1960-82



SOURCE: Division of Vital Statistics, National Vital Statistics System.

Figure 5.

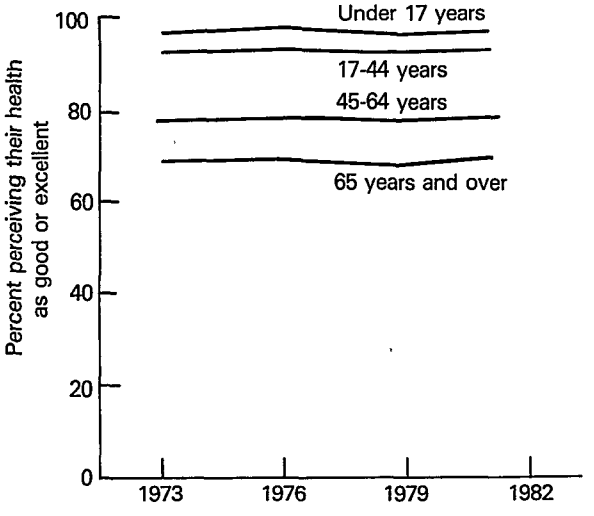
Death rates for suicide for persons 15-24 years of age, by race and sex: United States, 1960-82



SOURCE: Division of Vital Statistics, National Vital Statistics System.

Figure 6.

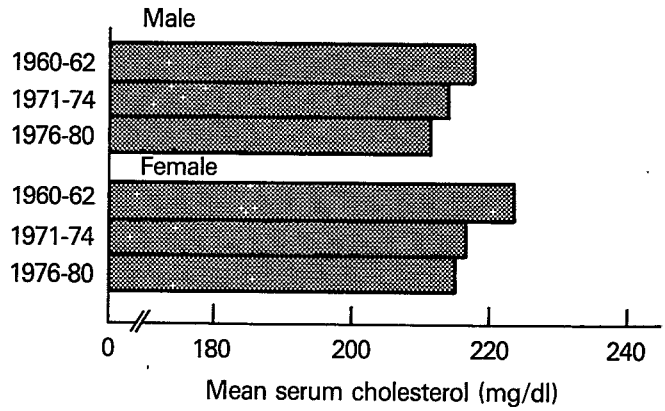
Perceived health status, by age: United States, 1973-81



SOURCE: Division of Health Interview Statistics, National Health Interview Survey.

Figure 7.

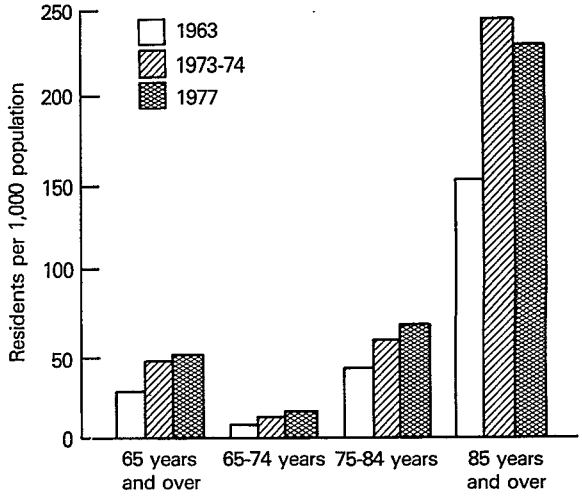
Mean serum cholesterol levels for persons 20-74 years of age, by sex: United States, selected periods 1960-80



SOURCE: Division of Health Examination Statistics, National Health and Nutrition Examination Survey.

Figure 8.

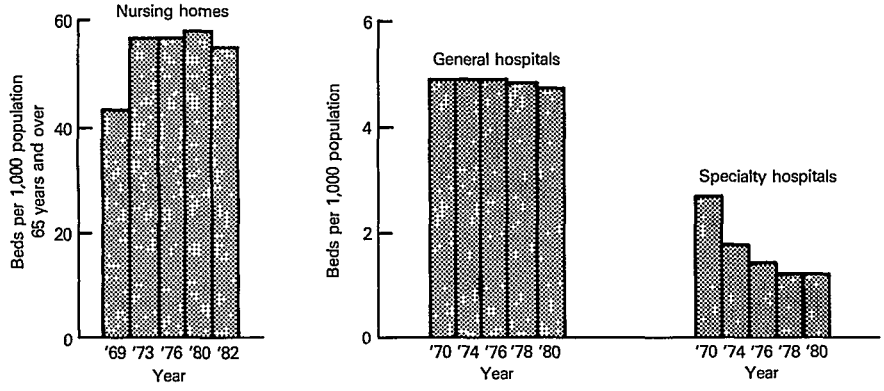
Nursing home use by persons 65 years of age and over: United States, 1963, 1973-74, and 1977



SOURCE: Division of Health Care Statistics, National Nursing Home Survey.

Figure 9.

Bed rates for nursing homes, general hospitals, and specialty hospitals: United States, selected years 1969-82



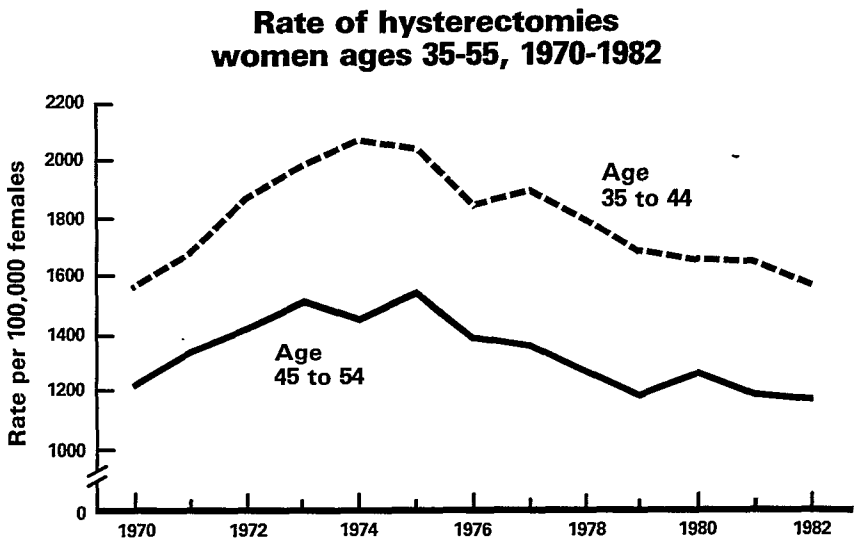
SOURCE: Division of Health Care Statistics, National Master Facility Inventory.

Figure 10.

Discharge rates and average length of stay in non-Federal short-stay hospitals: United States, 1966-83

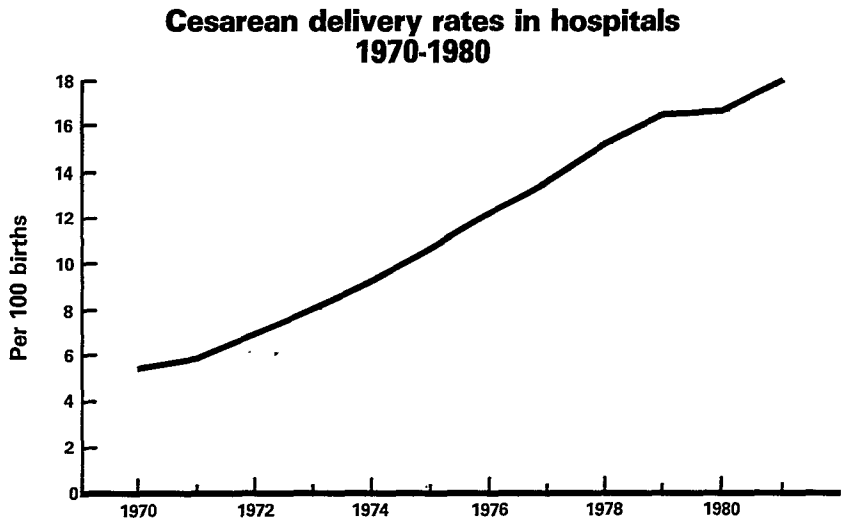
SOURCE: Division of Health Care Statistics, National Hospital Discharge Survey.

Figure 11.



SOURCE: National Center for Health Statistics, National Hospital Discharge Survey, Division of Health Care Statistics

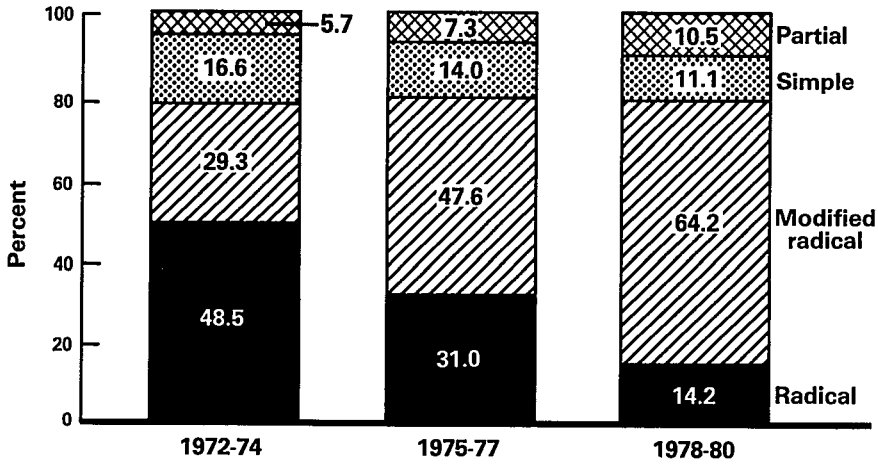
Figure 12.



SOURCE: Taffel and Placek: 1983.

Figure 13.

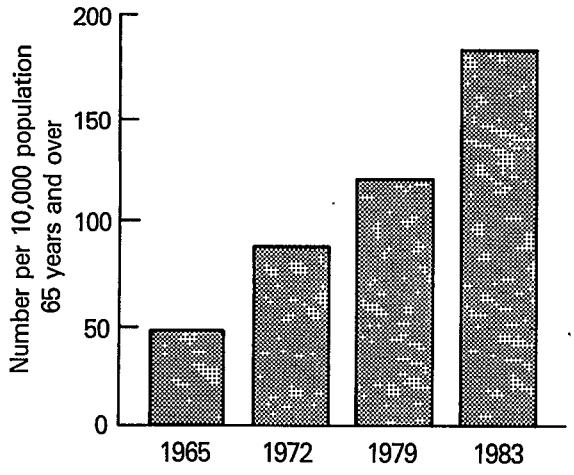
Surgical procedures for breast cancer patients in hospitals 1972-1980



SOURCE: Kleinman, et al.: 1983

Figure 14.

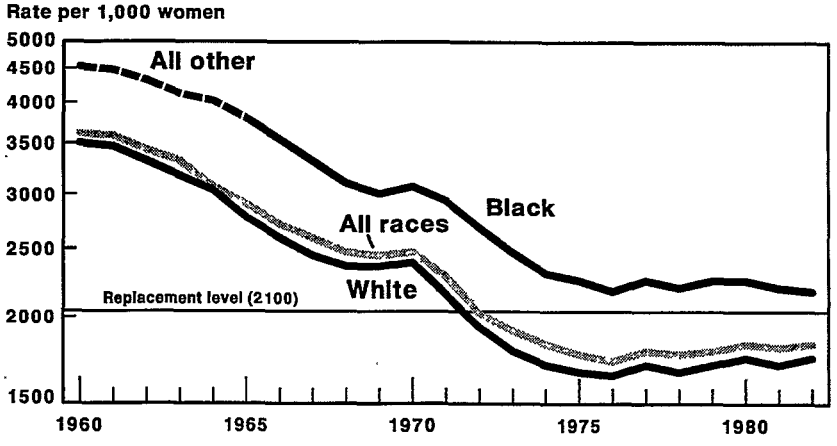
Rates of lens extractions for persons, 65 years of age and over: United States, selected year 1965-83



SOURCE: Division of Health Care Statistics, National Hospital Discharge Survey.

Figure 15.

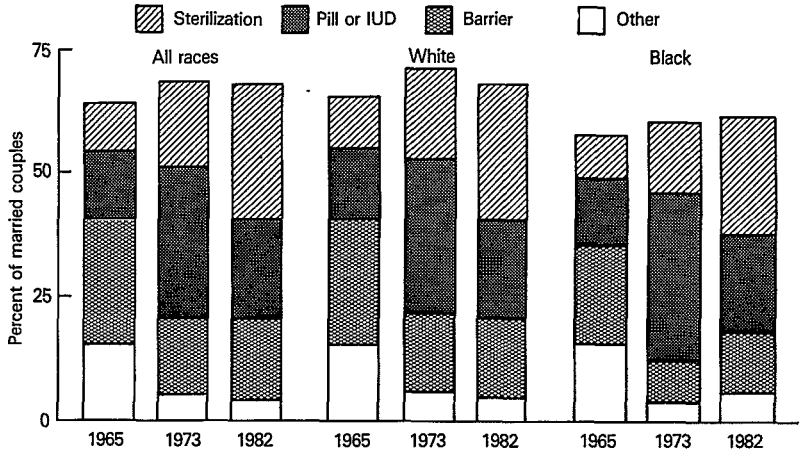
TOTAL FERTILITY RATES BY RACE: UNITED STATES, 1960-82



SOURCE: National Center for Health Statistics

Figure 16.

Contraceptive use, according to method and race: United States, 1965, 1973, and 1982



SOURCE: Division of Vital Statistics, National Survey of Family Growth.

Figure 17.

Lead used in gasoline production and average NHANES II blood lead levels (February 1976-February 1980)

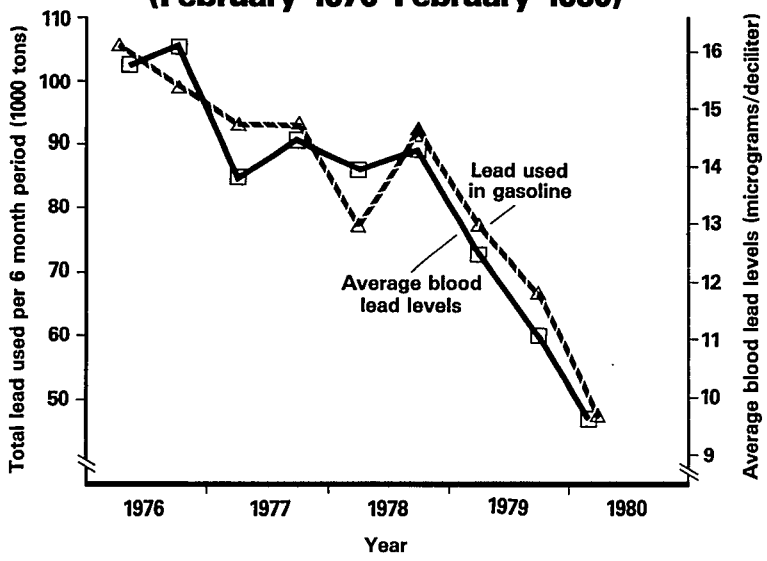
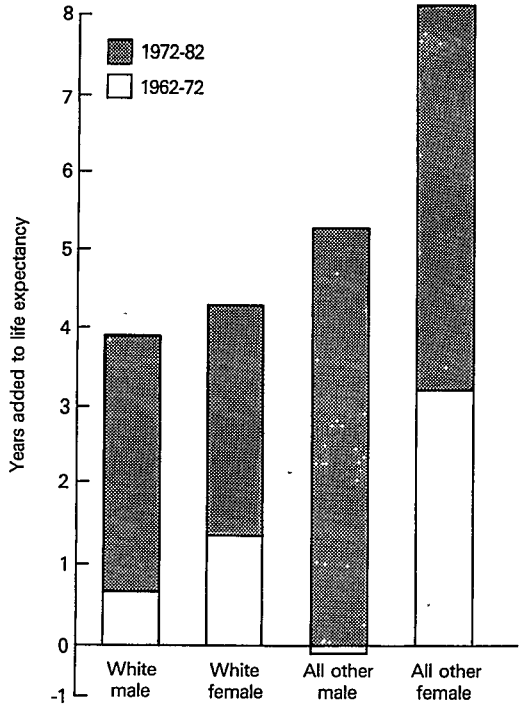


Figure 18.

Years of life added to life expectancy at birth, by race and sex: United States, 1962-72 and 1972-82



SOURCE: Division of Vital Statistics, National Vital Statistics System.

GOOD DATA AND GOOD STATISTICS MEAN GOOD POLICY AND PROGRAM (AND, BAD MEANS BAD)

Charles D. Baker
Under Secretary for Health and Human Services

We are in the anniversary business right about now. This is the silver anniversary of the National Center for Health Statistics and we are engaged today in our 20th biennial session. In fact, along about 25 years ago, the Surgeon General, one LeRoy Burney, put together what is now the National Center for Health Statistics. Now don't ask me to explain the numerics of a 20th biennial and a silver anniversary. Take those numbers on faith.

As a matter of fact, earlier this morning I had the pleasure of joining Martha McSteen at the National Archives where we opened some collections dealing with the 50th anniversary of Social Security and we talked about some very interesting people and things like Francis Perkins, and Styles Bridges, the Townsend Plan and a lot of things that only Dr. Feinleib and I are old enough to remember.

So anniversary celebrations are high on today's agenda and this one is particularly nice because it gives me the opportunity to underscore the significance of what you are collectively up to. One of the really amazing things that is going on in the entire public sector today and, indeed, in the health industry in general, is the fact that we are now really getting on with the business of deciding what is going on, in short, we are facing facts.

Now sometimes it is said that statistics are simply a group of numbers looking for an argument or, if you prefer, as Mark Twain once observed, "The first thing we do is get the facts. We can distort them later." Now, the truth of the matter is that such commentary is just so much baloney and is the kind of rhetoric put forth by people who really don't understand what numbers, what data, what statistics, in short, what information generally can do. I think it is unarguable that the availability of the kind of data, the kind of information that we increasingly have the facility to collect, is focusing us evermore in the directions we ought to go.

Simply put, there are simply too many instances on record that you all know about of the contributions that data collections and analysis have made. In this particular instance, they have helped focus development and expansion of medical science in order to address the problems that we have.

Every year the National Center for Health Statistics demonstrates, unarguably, what, in fact, we can do with this kind of information. This year, for example, we note again that life expectancy is at an all-time high and infant mortality continues its decline, which, needless to say, encourages us to believe that we are doing some things right.

On the other hand, it points out to us where things are not so right. It tells us that there are many challenges remaining. The data that we are collecting on the incidence of lung cancer among women makes the picture agonizingly clear, but some of the correlations are leading us to indications of causes. Our data on theincidence of Alzheimer's indicates what this portends for society at large as well as dimensioning the medical challenge before us.

The statistics on AIDS are painfully well-known to everybody in the room and make clear to us the assignment that medical science has before it as well as the problems facing public health policy.

The data we collect causes us to see things that we would like sometimes to sweep under the rug or perhaps ignore or pretend we don't understand. We see disparities in health conditions -- ethnic disparities, racial disparities, economic disparities. Our attention is called to the priorities we must have, to the challenges that we must face, in short, to the things that we really must get on with.

I would argue that in telling us all these things, NCHS data and other resources of the same kind, instruct us where to focus our efforts and our resources. At the same time, data allows us to measure the progress that we are making in meeting our challenges.

It is not stretching the point to argue that everything we do to preserve and improve health follows from a numerical description of a particular problem, the identification of a trend or, if you will, the detection of a threat.

NCHS told us this year that in California, Connecticut and Wisconsin there are consistently low heart disease death rates for all segments of the population. On the other hand, we discovered that Georgia, Illinois, Kentucky and Louisiana have consistently high rates. Why? What, indeed, should public policy do about it? What should be the health industry's response?

By using these types of statistics, not only for the nation as a whole, but for particular areas or particular regions, we can inform ourselves of trends, needs, priorities, and stimulate the focus and address where it belongs.

Of course, the next year we will gather the data and we will see how we have done. Have we as Government, individuals, the private sector, had any real impact? What more should we do? Where should we shift?

Creating, disseminating, making available this kind of information, among other things, causes governments at all levels, (Federal and local) to do that which governments least like to do: face up to the facts, determine what is really going on, and what the needs really are.

It is essential that this progression from numbers to action continue and, I would urge, be expanded if we in government at any level are to be effective guardians of the nation's well-being. It also goes without saying that it is essential that the numbers be as accurate as possible, their collection and analysis be sound, and as quickly available as we can all make it.

I remarked at the outset, the pundits notwithstanding, there is no question that health statistics indeed make a difference. Patently and

clearly they do. The only question for us here today is how to amass and better utilize them so that their obvious worth can make even more of a difference.

I picked up the program for the Conference which includes topics like accessing data bases, microcomputer applications, linked records, U.S. standard certificates and reports, evaluating the cost and use of health statistics methodology, and health statistics on health behavior.

It then goes further and says why we are meeting, why we are exchanging thoughts, ideas and information. It is because we are concerned about data access and availability. We are concerned about organizing and planning the development and use of statistics. We are concerned about analysis. Simply collecting data, while sometimes an enjoyable thing to do, is of not much use if, in fact, it is not analyzed and applied. We are concerned about methodology and technology.

In sum, what we are doing at this biennial conference is getting together, taking a bit of account of what we are up to, why we are up to it and what it all really means, and most importantly, how we can do still more in the future.

It is commonplace to say that we are now living in the information society. Whether or not personal computers are actually going to wind up in every home is a matter that Wall Street is wrestling with at the moment and has concluded maybe not. But the fact remains we have as a society an ability to amass, utilize and access information that is literally different, not in degree, but in kind from what existed only a decade ago. The question, of course, is how are we going to exploit this marvelous technological event? Arguably the greatest revolution now going on in health and health care is that we are identifying ever more surely what is taking place. We are identifying ever more accurately what, in fact, the responses are that science, that society-at-large is making, and, perhaps most important of all, we are identifying what all this is doing. What is advancing the good. What is not. We have a capability to do this now unlike anything we have had in time past.

So, the challenge before us is to get the information, array it, structure it, and access it for the government and the health industry, so that "we" can use it to shape where we go. We are not a nation of unlimited resources, even if at times it would appear -- from the way we spend -- that we think we are. The kind of information development that is the subject of this Conference allows us to focus resources where they are most needed, where they can do the most good and -- on an ongoing basis -- to evaluate what is occurring. This opportunity for information management to contribute to the overall success of society in meeting its challenges is unparalleled in our history.

As we look back with pride on what our contributions have been over the last decade or two, we can take great satisfaction. On the other hand, the challenges on the table today are not only large, in most instances they are growing larger. Thus, the question before us is whether or not we will rise to that challenge, develop the data, develop the information, structure it in ways and approaches so that "it" will make the major contribution that it can (and must) to public policy, to medical science and health care.

I am obviously confident that that is exactly what we will do. I am confident that in the course of this gathering, this biennial gathering, we will advance the state of play so that our contributions will indeed be major. It is a very great privilege for me, on behalf of Ronald Reagan, Margaret Heckler, 127,000 people in the Department of Health and Human Services, Gooloo Wunderlich, and Manny Feinleib, to stand up here and welcome you to this Conference, . . . and to the challenges before you!

Thank you very much.

HEALTH STATISTICS MAKE A DIFFERENCE -- THE NATIONAL PERSPECTIVE

Karl D. Yordy, Institute of Medicine

I am delighted to participate in this 20th Public Health Conference on records and statistics. I am especially pleased and honored to be here on the occasion of the 25th Anniversary of the founding of the National Center for Health Statistics. I remember well the birth of NCHS; my boss at the time at NIH, Joseph Murtaugh, served as a member of the Surgeon General's Study Group on Mission and Organization of the Public Health Service, which recommended the creation of the National Center through the merger of the National Office of Vital Statistics and the National Health Survey. As an individual who deeply believed that health statistics should make a difference, Joe Murtaugh was very proud of that recommendation. In the succeeding years I have had the pleasure of knowing and working with a number of the directors at NCHS and other NCHS staff.

During these 25 years, and continuing to the present, NCHS has solidified its position as a significant source of health data and analysis, widely respected and admired not only in this country but throughout the world for its technical competence and objectivity. Many people in this room have played important roles in the emergence of NCHS and have every reason to be proud.

Yet, many persons active in the health field would question the validity of the title of this presentation. They would argue that national policy decisions are guided by political and value judgments rather than sound health statistics. In our system of government these political values should guide decisions. However, as a long time student and observer of national policy formulation, and an occasional participant, I believe data often plays an important role in informing and structuring the political debate. I will describe some examples where health statistics have made a difference and will discuss why I think that difference will be even greater in the future. I will also discuss some of the limitations and barriers of the effectiveness of our current array of health statistics as a determinant of public and private policies in this era of profound change.

Since my assignment is to look at this topic from the national perspective, I should begin by noting the vast changes in federal responsibilities for health that have taken place since 1960. At a time when much attention is focused on leveling off or shrinking of some federal functions in health, it is easy to forget that a major role in many aspects of health activities is a very recent phenomenon. In 1960 the federal role was confined primarily to control of infectious disease, the support of biomedical research, the regulation of food and drugs, and modest roles in environmental health, the financing of health care for the poor, and community health facilities construction. Direct medical care was provided by the Federal Govermnent to certain beneficiary groups such as the veterans and the Indians, and, of course, a health statistics function was already in place with the National Health Surveys having recently joined the Vital Statistics program as important sources of health data. This set of activities, including health statistics, reflected a view of the federal role that had persisted from the founding of the Republic--that the federal role in domestic activities should be limited to those activities where there is a clear national public purpose.

Still to come were Medicare and Medicaid, grant programs for health services innovations, major expansion of environmental health programs, a powerful federal role in health manpower, the support of health services research, and technology assessment. These new roles emerged in the subsequent years through broad changes in the national political perspectives that extended the federal presence into health activities that had been forbidden grounds throughout most of our national history.

While changes in the political climate were the dominant factor in bringing about these important policy changes, I believe that health statistics played an important instrumental role. Political events during such a period of rapid policy change may seem to have little foundation in quantitative description and analysis, but I believe that closer examination reveals a more complex situation in which the use of health statistics are important means to policy ends. Let me cite some examples from this era of rapid policy change in the 60's and 70's.

First, statistics frequently played a powerful role in revealing the existence and nature of health problems. This problem analysis is often the basis for a justification of a proposed policy or program. For example, analyses of maternal and child health problems, drawing heavily on data about differentials in infant mortality rates, were an important stimulus for the establishment of project grant programs under Title V of maternal and child health section of Social Security Act. These programs had as their intent the focusing of special efforts on target populations at high risk.

Similarly, data showing differentials and access to health services by income were an important part of the justification of the Medicare and Medicaid programs, as well as the health programs launched as part of the war on poverty.

Differentials in health status by race were important factors in justifying programs targeted toward concentrations of poor blacks and hispanics.

The growth of food programs during the 70's drew on analyses of differences in nutritional status among the population. These differences were identified in special surveys and then through the inclusion of nutritional status as a component of the Health Examination Survey.

The interest in prevention programs, such as

those aimed at heart disease and cancer risk factors, were stimulated by data showing the rise of chronic diseases as the primary sources of death and disability. Health statistics often served as a broad focusing mechanism for the design and implementation of more detailed epidemiological studies.

While the growth of environmental awareness in the society had a number of dimensions other than health, health concerns remained the most important driving force in the growing determination of this society to limit pollution of the water and the air through strong public policy actions. Water pollution standards and motor vehicle emission standards acquired and retained firm public support in spite of substantial costs and the opposition by some on economic and ideological grounds. These arguments about environmental issues often revolved around interpretations of health statistics.

Policy concerns about the availability of health resources, especially health manpower, led to numerous federal actions based on analyses of the distribution and availability of those health resources.

While federal policy in the health field was expansionist during this period, concern about health care costs came to the fore as the rapid rise in the proportion of the GNP devoted to health was documented by federal statistics. Furthermore, analyses of health statistics raised questions about the effectiveness of the expansion of medical care in altering health status. Such an analysis was performed by our neighbor to the North and resulted in the LaLonde Report, which called for new emphases in improving the health status of Canadians.

During this period health statistics helped shape the policy agenda by raising political awareness of issues. Health statistics were also important tools in the specifics of program design, justification, and evaluation. Data is often the language of arguments among competing policy analyses in the Executive agencies and the Office of Management and Budget. The use of data to justify a policy position became more important as the size of the federal commitment to health grew and began to compete with other uses of scarce federal resources. By the early 1970's evaluation of program effectiveness and alarm about projections of costs began to change the policy debates. These debates were further intensified by splits between the Executive and Legislative branches that led to a desire by the Legislative branch to create its own independent agencies for program analysis and decision. Congressional staff expanded dramatically. When NCHS was created, the principal House committee handling health program authorization had one staff person devoted to health programs. At last count, the equivalant health subcommittee has 16 staff, plus interns. Many of these staff have backgrounds in health related fields. An expansion of committee staff is only part of the story. During the 1970's Congress also created the Office of Technology Assessment and the Congressional Budget Office, as well as greatly strengthening the Congressional Research Service and the program evaluation activities of the Government Accounting Office.

In this highly politicized Congressional environment, usually associated with trade-offs among competing political interests, an eager group of new consumers of health data have become well entrenched.

In this new era of more careful program evaluation, conflicts among public policy objectives were revealed through the use of health statistics--for example, a conflict between further improvements in access to health services and the rising costs of these services. Health statistics generated by the Health Interview Survey, the National Medical Care Expenditure Survey, and other survey instruments demonstrated that access had, indeed, been improved through the expansion of federal and state programs, but the data on health expenditures showed that this success had been achieved at a high price. Likewise, the federal health manpower programs, along with counterpart activities at the state level, succeeded in stimulating a large expansion in the health manpower supply. That data also revealed continued inequalities of distribution and questions began to be raised about overall manpower surpluses and their possible effects in stimulating further rises in health expenditures. Success in reducing the death rates from chronic diseases, specifically cardiovascular disease and cancer, showed that something was working, although arguments continued about the relative role of prevention and services in achieving these declines. To the extent that risk factor modification proved efficacious in improving health status, the policymakers were increasingly plunged into the controversial arenas of personal behavior modification where trade-offs between traditional American values of freedom of choice conflicted with the continued acceptance of known factors that increased health risks. Progress was also achieved in reducing environmental pollutants, but questions were raised about the ratio of costs and benefits, and the issues of cost-benefit analysis remain full of both technical and political disagreements.

In the present we have a period of intense policy ferment, during which the policy directions of the last fifty years are being reconsidered. This reconsideration began before the current administration, reflecting more intense competition for public resources, a slow down in the rate of real growth in the economy, the impact of an unfavorable trade balance, and frustration with the complexities and lack of progress from government interventions. During this time of ferment and rapid change in health policy, health statistics take an increased significance as guides to decisions by government, health organizations and individuals. The determination of national policymakers to constrain the growth of health expenditures and to decentralize decision-making leads to debates about access, quality, and cost of health care services, and the effectiveness of interventions for prevention, be they changes in the environment or in personal behaviors. Perceptions of the issues in these policy arenas, and many decisions, are now more data-driven than ever before. The participants in these data based policy discussions now include many private health

care organizations, third party payers, and self insured companies. These private organizations influence the outcome of national and state public policy decisions, as well as being generators and consumers of health statistics for their own purposes. In both the public and private sectors there are more individuals involved in these debates with at least some formal training in statistics and quantitative analysis. Kerr White used to refer to most physicians as "a-numerate" and I believe he would have said the same for most policymakers in the earlier eras. No more. Effective lobbyists and policy advocates have become skilled at arraying data to support their causes. I believe that the policymaking world is more statistically literate than was true at the time of the NCHS founding. The wide availability of computers will certainly increase further the use of detailed analysis in policy decisions.

Let me cite several examples of current policy issues where data is playing an important role in the formulation of public policy and achievement of political consensus.

The strong reappearance of concern about health care for the uninsured has been fed by data from the National Medical Care Expenditure Survey and its successor surveys, as well as by the surveys supported by The Robert Wood Johnson Foundation. Many states have been studying this problem and some have taken action. This issue will surely reappear as a national issue.

Long term care is rising as a public policy issue after years of neglect. Demographic projections are the usual starting point for these discussions. Policymakers are very aware of the rapid growth in the over 85 population. Over the next 15 years, the over 85 population will increase by 80 percent and one out of two will probably need long term care. Decisionmakers are aware that progress against chronic disease adds to this population. Health statistics also tell us that the entire group over 65 are healthier and that only about one out of five of the total retired population need long term care services. This means that financing options for long term care can be explored that have the characteristics of insurance and risk pooling. In addition to demographics, our knowledge of the nursing home population has been augmented by surveys conducted by the NCHS. Ironically, I believe that some of the reluctance of policymakers to deal with the problems of long term care is due to their awareness of the facts and a consequent recoiling from the size of the problem. There is a wariness about consideration of new entitlement programs that stems from the rapid cost escalations of Medicare, Social Security, and other entitlements and the projections of costs into the future--projections based substantially on health statistics.

At the other end of the age spectrum, access to maternal and child health services remains an issue. The recent IOM report on prevention of low birthweight drew heavily on health statistics. Statistical analyses performed for the study indicated the value of prenatal care as an important factor affecting low birthweight. Concern about a leveling off in the decline of infant mortality and possible future reversals has also concerned policymakers, even during a time of constrained resources. As a result, there is political resistance to further cuts in Medicaid and Title V programs, and some states have been increasing their commitments to maternal and child health, reversing earlier cuts.

The monitoring of nutritional status, made possible by the regular surveys of the NCHS, is a factor in the continuing support for food programs targeted to the poor. As a result food programs remain an important part of the social safety net.

A new aspect of the concern about the effectiveness and costs of medical services for the entire population is the attention being given in the last few years to geographic variations in medical practice. Based on the work of Wennberg and others showing wide variations in common medical and surgical procedures by small geographic area, attention of policymakers has been focused on the uncertainties of medical decisions that determine large uses of resources. This attention has been focused almost entirely on the presentation of data. However, this is also an example that illustrates the interaction between the readiness of policymakers to address an issue and their receptivity to available data. Wennberg has been publishing since the early 1970's, and testifying before Congressional Committees for most of that time, yet only in the last three years has his work and similar work by others become widely known among policymakers at the federal and state levels.

The issues of environmental impacts on health continue to attract the attention of policymakers. Disposal of toxic wastes, motor vehicle emissions, radiation safety in a nuclear age, safety of industrial workers exposed to asbestos and other hazardous substances are all examples of health hazards that continue to occupy the attention of policymakers. At a time of record federal deficits, and the realities of an intensely competitive world economy, the costs of controls over environmental hazards force attention to risks and benefits. These analyses are still an uncertain science, with values and technical considerations interwoven. But much of the input into these arguments is based on health statistics.

I believe that these examples, and many others that could be cited, demonstrate that health statistics do make a difference. Furthermore, with decision-makers at all levels of society in both the public and private sectors having wet their feet in data-based arguments about health issues, aided by staff who are more and more trained in the use of data, I believe that the importance and policy relevance of health statistics will continue to increase. We are in an era of choices--choices among competing demands for scarce public dollars in the face of record deficits, choices by consumers among competing health plans, choices by businesses in a world competitive market, choices by labor about the trade-offs between jobs and health benefits and protection against workplace hazards, choices about how to care for those unable to cope for themselves (the dependent elderly, children, and the poor), and finally

choices about how we as a society value health among the many competing values. Values will guide these choices, as they should in a free society, but we can hope and expect that these choices will be defined and informed by health statistics. Thus, in this era of competition, in the health field and elsewhere, and of reconsideration of federal and governmental roles, one of the older governmental functions, the collection and dissemination of objective data, should assume new significance.

The demands for data, arising from the period of policy ferment, constitutes a challenge to our health statistics activities. Our current data sources may be inadequate in the face of these demands. Let me mention several areas where improvements in health statistics are needed to meet the demands of policy relevance from my policy-oriented perspective.

We are undergoing a revolution in the financing and structure of our health delivery system. Changes in both public and private payment methods have introduced new economic incentives that stimulate cost-consciousness and price competition. This is a radical change for a market that has been used to rapid growth and cost reimbursement. DRG's are the most visible symbol of these changes, but private sector actions have also been stimulants for changes. Examples are: the growth of HMO's and PPO's, changes in health insurance benefits to require greater cost-sharing by consumers, introduction of a variety of cost containment measures by business firms that now self insure for health benefits of their employees.

These changes place new demands on our health data sources. The changed incentives raise new questions about the quality of care and improvements in health care outcomes. To monitor quality and outcomes we need better ability to link inputs and outputs in medical care.

The changed economic incentives are also rooting out internal cost subsidies within our health care institutions. The disappearnace of these cost subsidies raises with new force questions about access to health care. Is access being denied? If access is denied, what are the implications for health status?

Stimulated by changes in economic incentives and intensifying competition among health care institutions, many significant structural changes are underway. These include the growth of for profit chains in the provision of health services, the establishment of new relationships between physicians and hospitals, and the rapid growth of vertical and horizontal integration of health care services. Some health analysts have speculated that by the year 2000 most health services will be provided by 20 or 30 major national health care firms, both for profit and not-for-profit. We need to understand more completely the significance of these profound structural changes.

In order to monitor and evaluate the effects of these extraordinarily rapid and profound changes, we need timely data that provides a mechanism for focusing more intense scrutiny. The long time periods between data gathering and availability of analysis that characterizes most of our existing data sources will not suffice for this monitoring purpose.

We still also need small area data that provides a basis to compare local circumstances to state, regional, and national data. Accurate national averages are not enough; these national data may mask local differences that are the most important focus of policy concerns.

Environmental health is another arena that would benefit from modifications in our health statistics. Environmental health effects are often characterized by long times between exposure to hazards and the evident health impact. Our extremely mobile population is a complicating factor in associating exposure to hazards with poor health status. Rapid developments in technologies expose the work force and the general population to new chemicals. These advances in technologies may also alter the work experience, leading to work-related stress. World economic competition further complicates attention to problems of environmental hazards. To address these issues more satisfactorily, we need better sources of data that include longitudinal studies and linkages among data sets.

This era of rapid change also sharpens our concern about the effectiveness of public programs. The growth of entitlement programs in a mature society puts a substantial squeeze on other government expenditures. This leads to intense scrutiny of public expenditures, which will extend into the foreseeable future regardless of which party or political ideology is in power. To respond to this concern about the effectiveness of public programs, we need better links between program data and general health statistics, and we need to be creative in our use of program data to address general problems. We also need to figure out better ways to use privately generated data for public purposes.

All of these limitations of our current health data have been receiving attention, but we need a more intense effort to assure that our health data systems are responsive to perceived needs of decision-makers. In charting new courses, we must strike an appropriate balance between maintaining the value of existing data systems and developing new approaches that can meet more fully current and future needs of the society.

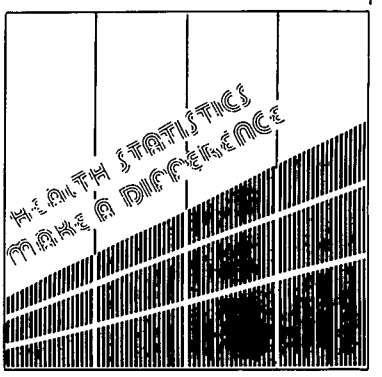
NCHS has met similar challenges during the past 25 years. Policymakers need to be reminded that in this time of policy ferment health statistics are needed more than ever. They also need to be reminded that health statistics are often fragile in the face of these rapid and powerful changes.

I believe that health statistics must remain a cornerstone of public activity, regardless of which political persuasion is dominant at the moment. Those who are vigorous advocates of health care competition as well as those who are strong advocates of public responsibilities in health make frequent use of health data. I am sure that NCHS and the health statistics community will rise to the challenges of this exciting era as they have to the challenges of earlier times. A major test of any successful human organization is whether it can change and adapt appropriately in the face of new circumstances. On the occasion of the fiftieth

anniversary of the NCHS I am confident that the speakers will say that NCHS has once again met the challenges and that health statistics continue to make a difference.

Session A

Accessing Data Bases for Health Monitoring



IMPROVED PERINATAL OUTCOME DATA MANAGEMENT SYSTEM

Jeffrey B. Gould, Ellen Liebman, Carol Nickerson, Frank Many, Miguel Lucero, University of California

OVERVIEW:

As moneys to improve perinatal welfare became increasingly more limited the need for information on which to base health planning and allocation decisions becomes an important priority. Over the last two years we have worked under contract to the State of California, Department of Human Services, Maternal and Child Health Branch, in order to develop the Improved Perinatal Outcome Data Management System (IPODM). The purpose of the system is to enable local health managers to analyze census based sociodemographic data and vital statistics based perinatal data beginning at the census tract and/or zip code level. In designing this system we have tried to meet several key specifications.

1) The system should be interactive and suitable for the health manager with little computer experience.
2) The system should be based on standard elements and operating environments so as to be transportable to a variety of sites.
3) The system should be accessible via modem for "on site" analysis.
4) The system should have the ability to build datasets using census and vital statistics variables and be easily expandable in terms of adding more variables and adding on new datasets.
5) The basic building block should be the census tract and/or the zip code and the system should allow easy aggregation of these units into larger geographic areas.

IPODM was designed for a mainframe computer and an "IBM-CMS" environment. IPODM consists of a library of "SAS" programs and macros that create data bases from the 1980 STF1 and STF3 census tapes and from the 1982 All California Linked perinatal birth-death tape (1983 data will be added this year). Because linked birth death variables (LBD VARS) are reported only by zip code (with the exception of several counties who also include census tract on their birth/death certificates) we had to develop a file of zip code census tract correspondences for all of California's tracted counties. This correspondence file allows census tract sociodemographic data to be aggregated up to the zip code level and merged with zip code based vital statistics data (LBD VARS). In all, six SAS data bases are created - STF1 tract level, STF3 tract level, LBD tract level, and STF1 zip level, STF3, zip level, and LBD zip level. Programs written in CMS REXX (a command interpreter only available with CMS) allow the naive user to access these data bases rapidly, to create user datasets, display data, and perform statistical analysis. The format is interactive requiring no knowledge of SAS or computers. The system also has several help files that can be accessed at any time during a session. For the advanced user the system allows access to the data bases using interactive SAS.

USER DATA BASES:

Data bases can be built using approximately 500 variables. The census variables include information that are commonly used by health planners such as age, race, sex, marital status, household type and relationship, education, labor force status, income, poverty status, housing quality, and overcrowding. The LBD Variables include data on outcomes such as complications of pregnancy, low birth weight, malformations, and mortality; on risks such as age, marital status and ethnicity; and on program indicators such as inadequate prenatal care and short interpregnancy interval.

Two types of user datasets can be built - "regular" and "aggregate". "Regular" datasets list individual tracts or zips for a specific California county or combination of counties. "Aggregate" datasets contain the values for specific aggregates of zips or census tracts. Using the IPODM program "areas", one states the number of aggregates and their composition. These definitions are stored and can be accessed in the future. IPODM3 uses these definitions to perform the aggregations. The ability to easily aggregate is one of IPODM's most important features and allows the health planner to create small area aggregates corresponding to jurisdictions, cities, health planning areas, catchment areas (for visiting nurses, community clinics, etc.), poverty areas, etc. This utility especially when used with IPODM's mapping capability can be used to evaluate the extent, uniqueness, homogeneity, and inclusiveness of traditional and proposed "health service areas."

In building user datasets one is also given the option of restricting the inclusion of an area (zip, tract, or aggregate) by stating a critical value for any system variable. Figure 1 shows the ease of building a dataset. Note that only zips with more than 50 births will be included in this particular dataset.

IPODM UTILITIES-DISPLAY:

After building a dataset IPODM gives one the choice of several interactive options; Pearson correlations, mapping, descriptive statistics, scattergrams of two variables, multiple regression using least squares, interactive SAS, tables, and the ability to select/create another dataset for use.

The two display options are tables and map. Tables can be sorted by any variable and can be immediately hard copied on a parallel printer. Figure 2 lists portions of a typical table that has been sorted on the basis of median family income. The units of observation in this table are census tracts. The map function automatically produces a plot file using SAS Graph's mapping functions. Labels are then added and the file routed to a graphics printer. In the most recent release SAS Graph allows one to label each centroid and to visually separate

aggregates from each other. These features are being incorporated into IPODM3. Figure 3 is an example of our current "map."

ANALYSIS UTILITIES:

Several basic analytic routines have been incorporated into IPODM3 using a menu format. These include basic descriptive statistics, scattergrams, correlation, and multiple regression. A weighting option using the value of any variable is included where appropriate. For more advanced analysis one can enter interactive SAS directly from IPODM3.

DATA SET UTILITIES:

An extremely important feature of IPODM3 is that datasets can be modified. The basic geographic units of a regular dataset (census tracts or zips) can be aggregated to form an aggregated dataset. Aggregated datasets can also be further aggregated. To obtain data for an entire "county" (or combination of counties), one first builds a regular dataset that includes all zips/census tracts for the "county." One then can aggregate these as is illustrated in Figure 4.

For health planning it is often very useful to be able to define geographic areas on the basis of sociodemographic data and to determine certain aspects of health status within these areas. IPODM3 is particularly well suited to this task. The example in Figure 5 shows the output from an "all county" regular dataset that was aggregated into three specific areas - 1) census tracts considered as being extremely poor, 2) census tracts considered as being poor and 3) the remainder of the county.

PERINATAL NEEDS ASSESSMENT INVENTORY:

IPODM3 provides the health manager with several hundred variables from the 1980 Census and the 1982 California Linked Birth/Death File. A major question is how to efficiently utilize the data possibilities in order to obtain a description of perinatal need/outcome in a given county. The Perinatal Needs Assessment Inventory (PNAI) is being developed to provide an efficient starting point for getting an "overall feel" of a county. The PNAI is a listing of 28 variables that are commonly cited in the literature as key sociodemographic indicators and specific indicators of perinatal risk, outcome, and program need. These include: ethnicity variables; sociodemographic factors associated with poor pregnancy outcome such as less than 12th grade education, more than 1.01 persons per room; poverty indicators; perinatal outcome measures such as low birth weight births, perinatal deaths; perinatal risk indicators such as non-marital births, prenatal complications; and program indicators (teen births, poor prenatal care, etc.).

Our first approach was to use a standardization process that would combine these variables and result in a single score summarizing overall perinatal need. From a health planning standpoint however, what is needed is a broadening rather than a narrowing of perspective. With this in mind the single summary score approach was abandoned in favor of a multiple variable inventory organized with respect to the specific usefulness of each group of variables. For these 28 variables sextile values for all California zip codes and tracts have been computed and it is therefore possible to reference specific values of any zip codes/-census tracts against these sextiles. As we gain practical experience with these variables we hope to further refine the scope, format, and usefulness of the inventory.

SUMMARY:

The IPODM system was designed to enable health managers with little computer expertise to access over 500 variables derived from census and vital statistics sources. Using an interactive format one can easily create datasets, small area aggregates, displays, choropleth maps, and perform basic statistical analysis. The system is configured to allow the easy addition of other data bases, and conceptually represents an interactive small area health analysis system with sociodemographic and perinatal outcome modules in place. The next phase of IPODM development involves working on specific analysis projects with several health agencies in order to gain insight into how to further optimize the system - for the purpose of health planning and assessment.

Figure 1: BUILDING A REGULAR DATASET BASED ON ZIP CODES WITH IPODM3

```
Welcome to IPODM3! You can answer any question with HELP to get help or QUIT to
end this session.


Tract or Zipcode level dataset? (T or Z)
z
Create a new dataset? (Y, N, or L where L means list all existing datasets)
y
Make an aggregated dataset? (Y or N)
n
Make the new dataset from the standard datasets or modify one of
your own datasets? (NEW or MODIFY)
new
Name of new data set? (maximum of 8 characters)
all_zips
Do you want to select for particular counties for this dataset? (Y or N)
y
List FIPS codes of counties you want on the next line (dash ok)
19
Select particular observations for this dataset? (Y or N)
y
Select by ZIPCODE number? (Y or N where N means a
different variable will be used for selection criteria)
n
Name of variable to use for selection?
lltlnl
Smallest value allowed in dataset?
50
Do you want variables from STF1? (Y or N)
y
Specify STF1  number variables you want in your dataset. Note that
corresponding percentage variables are automatically carried along.
Separate each specification with at least one space. End by hitting the.
the RETURN key twice.
tlnl t7n2 t19n3

Do you want variables from STF3? (Y or N)
y
Specify STF3  number variables you want in your dataset. Note that
corresponding percentage variables are automatically carried along.
Separate each specification with at least one space. End by hitting the.
the RETURN key twice.
t74nl t9lnl

Do you want variables from LBD1? (Y or N)
y
Specify LBD1  number variables you want in your dataset. Note that
corresponding percentage variables are automatically carried along.
Separate each specification with at least one space. End by hitting the.
the RETURN key twice.
tlnl t7n3 tllnl t15nl

Making dataset. Be patient ....
Make dataset permanent (put on your A disk)? (Y or N)
```

Figure 2: EXAMPLE OF TYPICAL IPODM3 TABLE INCLUDING VARIABLES FROM THREE SOURCES (STF1, STF3, CAL "LINKED" BIRTH/DEATH FILE)

VARNAME	S3T74N1	S1T19N3/P3		L1T2N4/P4		L1T5N3/P3		L1T15N1/P1	
	MEDIAN FAM INC. 1979	1+ PERS FEM. NO HUSB <18YRS		BLACK BIRTHS		TOTAL BIRTHS LT 2500 GR		LT 21 MONTHS SINCE LAST BIRTH	
TRACT	(VALUE)	(NUMBER)	(PCT)	(NUMBER)	(PCT)	(NUMBER)	(PCT)	(NUMBER)	(PCT)
4432.00	50,991	3	1.8	0	0.0	3	7.7	3	7.7
4261.00	48,082	100	11.6	1	0.8	8	6.6	11	9.0
4420.00	43,636	17	4.3	0	0.0	4	5.9	7	10.3
4046.00	41,498	63	12.1	5	3.5	8	5.6	10	7.0
4506.01	40,665	19	9.7	0	0.0	2	5.6	3	8.3
4001.00	39,799	29	14.9	11	11.2	5	5.1	5	5.1
4212.00	39,671	79	15.0	1	1.1	4	4.3	3	3.2
4215.00	39,599	61	15.6	1	1.1	6	6.9	3	3.4
4211.00	39,488	30	11.7	0	0.0	3	5.1	1	1.7
4214.00	37,370	41	17.7	3	6.4	2	4.3	3	6.4
			•						
			•						
			•						
			•						
4020.00	8,750	2	50.0	21	75.0	0	0.0	4	14.3
4021.00	8,367	234	75.5	62	84.9	17	23.3	5	6.8
4018.00	8,217	127	56.7	69	74.2	19	20.4	13	14.0
4028.00	8,194	32	44.4	26	39.4	5	7.6	14	21.2
4022.00	7,635	122	59.8	51	56.0	6	6.6	14	15.4
4025.00	7,325	219	61.9	62	82.7	14	18.7	8	10.7

Figure 3

PERCENT TOTAL BIRTHS LT 2500 GRAMS
1980 ALAMEDA COUNTY ZIP CODES

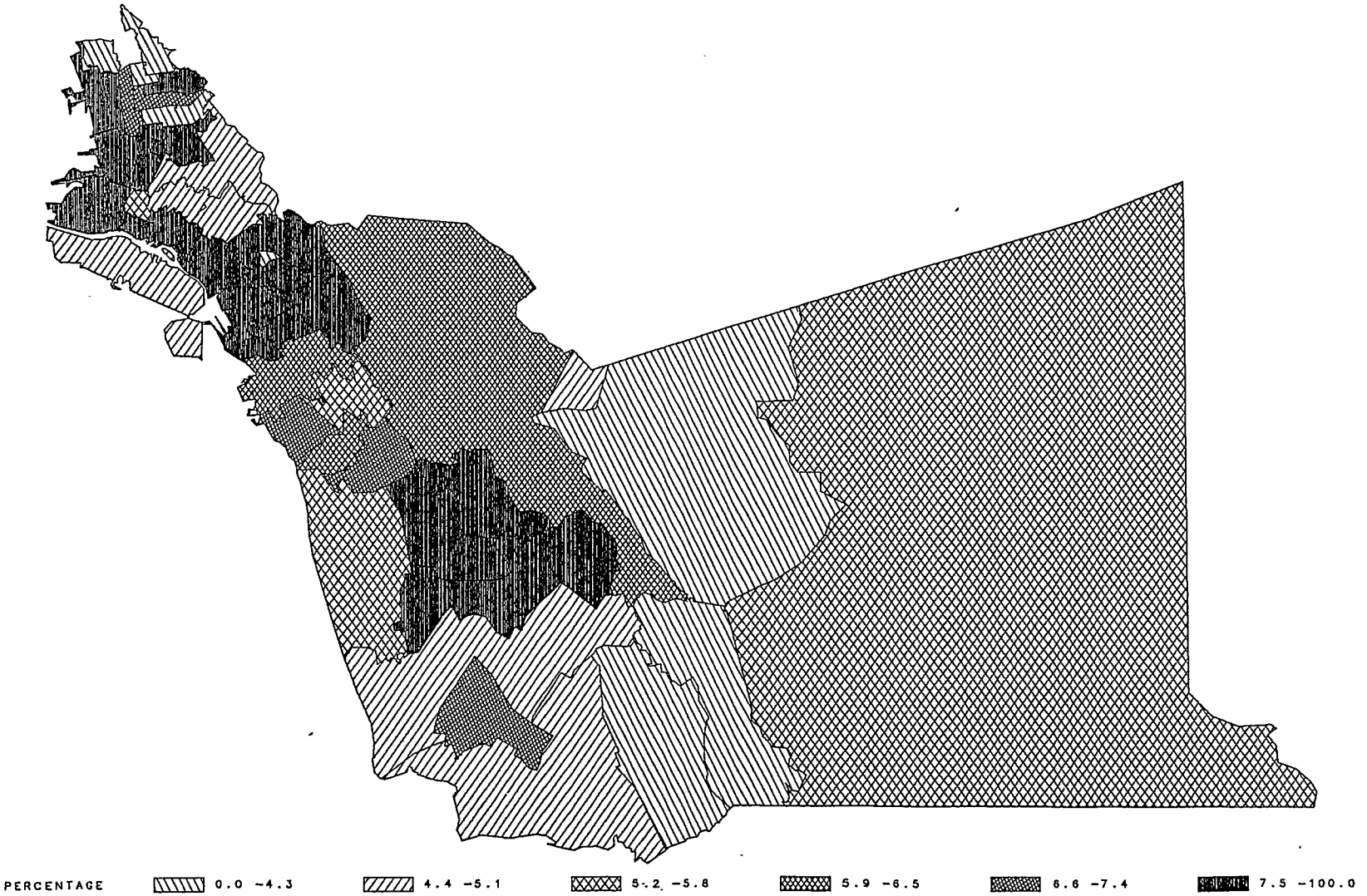


Figure 4A: PROGRAM TO AGGREGATE UP TO THE COUNTY LEVEL

```
DATASET
Tract or Zipcode level dataset? (T or Z)
T
Create a new dataset? (Y, N, or L where L means list all existing datasets)
Y
Make an aggregated dataset? (Y or N)
Y
Make the new dataset from the standard datasets or modify one of
your own datasets? (NEW or MODIFY)
MODIFY
Modify an aggregate or regular dataset? (AGG or REG)
REG
List existing regular datasets? (Y or N)
N
Which regular data set do you want to use?
AL_TRACT

Wait a sec ...

Name of new data set? (maximum of 8 characters)
AL_TOTAL
Do you want to select particular variables from file AL_TRACT? (Y or N)
N
Which county will you be aggregating (use FIPS code)
1
Use predefined areas for county? (Y or N)
NOTE: Answer HELP and look in CTYDEFS to see what these are
N
How many separate aggregations of TRACTs will be formed?  (give number)
1
List TRACTs for aggregate 1 (Last one. Word REST is ok)
REST

Making dataset. Be patient ....
```

Figure 4B: EXAMPLE OF OUTPUT

COUNTY=ALAMEDA

VARNAME	S1T1N1	S1T7N2/P2		S1T19N3/P3		S3T74N1	S3T91N2/P2	
				14 PERS <18YRS		MEDIAN FAM	TOTAL INCOME	
	TOTAL	BLACKS		FEM. NO HUSB		INC. 1979	BELOW POVERTY	
AREA	(NUMBER)	(NUMBER)	(PCT)	(NUMBER)	(PCT)	(VALUE)	(NUMBER)	(PCT)
01	1,103,527	203,323	18.5	36,315	24.3	23,299	121,590	11.3

VARNAME	L1T1N1	L1T11N1/P1		L1T15N1/P1		L1T5N3/P3		L1T7N3/P3	
	TOTAL	AGE OF MOTHER		LT 21 MONTHS		TOTAL BIRTHS		TOTAL	
	BIRTHS	LT 18 YEARS		SINCE LAST BIRTH		LT 2500 GR		PERINATAL DTHS	
AREA	(NUMBER)	(NUMBER)	(PCT)	(NUMBER)	(PCT)	(NUMBER)	(PCT)	(NUMBER)	(RATE)
01	50,728	2,255	4.8	4,530	8.9	3,806	7.5	620	13.8

Figure 5: EXAMPLE OF A COUNTY AGGREGATED INTO THREE SOCIODEMOGRAPHIC AREAS

VARNAME	S1T1N1	S1T7N2/P2		S1T19N3/P3		S3T/4N1	S3T91N2/P2	
	TOTAL POP	BLACKS		1+ PERS <18YRS FEM. NO HUSB		MEDIAN FAM INC. 1979	TOTAL INCOME BELOW POVERTY	
AREA	(NUMBER)	(NUMBER)	(PCT)	(NUMBER)	(PCT)	(VALUE)	(NUMBER)	(PCT)
1	58,162	44,492	76.5	4,201	55.4	11,068	15,960	27.9
2	163,942	75,413	46.0	9,190	37.7	16,306	28,667	17.8
3	881,423	83,418	9.5	22,924	19.5	25,260	76,963	8.9

VARNAME	L1T1N1	L1T11N1/P1		L1T15N1/P1		L1T5N3/P3		L1T7N3/P3	
	TOTAL BIRTHS	AGE OF MOTHER LT 18 YEARS		LT 21 MONTHS SINCE LAST BIRTH		TOTAL BIRTHS LT 2500 GR		TOTAL PERINATAL DTHS	
AREA	(NUMBER)	(NUMBER)	(PCT)	(NUMBER)	(PCT)	(NUMBER)	(PCT)	(NUMBER)	(RATE)
1	3,056	242	7.9	356	11.6	383	12.6	61	21.9
2	10,193	647	6.4	969	9.5	911	8.9	153	15.2
3	37,479	1,366	4.0	3,205	8.6	2,512	6.7	406	12.7

THE EPIDEMIOLOGIC SURVEILLANCE PROJECT
REPORT OF A PILOT PROJECT

Philip L. Graitcer and Anthony H. Burton
Centers for Disease Control

In 1984, the Epidemiologic Surveillance Project (ESP) was initiated by the Centers for Disease Control (CDC) and epidemiologists in six state health departments to:

1. Demonstrate the feasibility of transmitting disease surveillance data electronically, via computers, from state health departments to CDC; and
2. Develop a systematic method by which demographic and epidemiologic characteristics of national disease surveillance data can be analyzed rapidly and comprehensively.

Since these surveillance data include demographic characteristics of the cases as well as information about time of onset and place of residence, detailed analyses of the data can be undertaken which may aid in the rapid identification of disease epidemics and more timely and complete understanding of disease trends.

Three concepts were integral to the design of this system. First, each state health department would use its existing computerized disease surveillance system to transmit data to CDC. Second, specific data concerning each case of a reportable disease -- rather than aggregate case counts -- would be transmitted to CDC, making it possible to do more complete epidemiologic analyses of disease trends. Finally, it was anticipated that these computerized case reports would eventually supplant the telephone reports of 48 reportable diseases made weekly by the states to CDC.

Case Reporting/Disease Conventions

The six state epidemiologists and CDC staff established conventions for the layout of the computerized case record and developed protocols for data transmission. The Medical Information Network (MINET)(1), part of a nationwide public access computer network, was selected as an appropriate system for the transmission of data between CDC and the states. Data and message transmissions are directed from local exchanges to the MINET computer located near Washington, D.C., and are stored in an electronic file until checked by the addressee. Thus, use of MINET to transmit data precludes the need for the addressee to be "on line" since messages are stored in the electronic file until requested.

A record layout was developed for the standardized transmission of data and was coded for several variables: state; case identification number; type of disease; county of residence; patient's age, sex, and race; date of onset; and the date of report. Additional character spaces have been reserved at the end of the record for the coding of variables that may be unique to a particular case report, such as vaccination status or laboratory confirmation of the diagnosis.

Data Entry

Three states were already entering and storing disease surveillance data on high-capacity minicomputers prior to the start of the project. In these states, computer programs were written to abstract the 40-character ESP record from a larger case record that had been prepared previously by the state for its own disease surveillance and reporting needs.

The other three states used popular desk-top microcomputers and commercial data-base management programs for the entry and maintenance of statewide surveillance data. Simple conversion programs were used to change these data into the ESP record format.

Data Transmission

Each state prepared a file of case reports that was transmitted on MINET to an electronic "mailbox" designated by CDC for the receipt of surveillance data. This message was retrieved, stored on a microcomputer, appended to the other state reports, and transferred ("uploaded") to the CDC main computer. Tabulations were made of age, sex, and race/ethnicity frequencies for each disease. The data were then appended to an existing data base of case reports. To assure data quality in the CDC weekly surveillance system during the pilot phase, these data were transmitted in parallel with the weekly telephone reports of disease.

Data Outputs

Various types of data output were generated and returned to the states. A file containing the frequency tabulations of the reported diseases by state was transferred ("downloaded") from the mainframe to the CDC microcomputer and retransmitted on MINET to the states, usually within 4 hours after the case reports were initially received at CDC. Critical to our efforts to maintain the quality of the data base was the quarterly mailing of edit reports to state data clerks. County-specific incidences for various diseases were plotted on national and state maps using the disease reports obtained from the ESP transmissions and population estimates obtained from the 1980 census. These maps were mailed to each state periodically, since detailed graphics cannot be transmitted on the electronic mail system.

Discussion

The ESP is a computer-based system for the transmission of disease surveillance data that is both feasible and efficient in improving the epidemiologic usefulness of surveillance data. With virtually the same personnel resources in the states and at CDC, increased amounts of surveillance data are transmitted to CDC. These data, containing case-specific

characteristics, not only improve CDC's ability to conduct national and regional surveillance of reportable diseases, but they also enhance CDC's efforts to identify changing epidemiologic patterns in these diseases. Changes in age, race, sex, and geographic distributions can be continually monitored and, if necessary, appropriate investigation or intervention may be rapidly undertaken. In addition, by having surveillance data available in a consolidated data base, state health departments are better able to monitor their own intrastate disease trends.

A major advantage of the ESP is that annual state reports of disease morbidity can be prepared rapidly and accurately. At CDC, the ESP data base is queried for the necessary state- and county-specific frequency distributions of diseases by age, sex, and race. Prior to the development of ESP, some of the participating states were using up to three person-months of clerical and statistical effort to produce state reports. Consequently, the MMWR Annual Summary was often delayed in its publication. With ESP, these annual summaries can be generated, edited, and corrected in days rather than months.

The ESP serves as a model for the consolidation of reporting and dissemination of surveillance data to and from state health departments. Presently, program operation and surveillance data from several disease programs are reported separately from the ESP and MMWR systems to CDC by state health departments. Although each of the CDC programs requires disease-specific data elements, the ESP record format can be readily modified so that these data elements can be collected, and the ESP system can be used for reporting of other diseases such as tuberculosis, vaccine preventable diseases, and sexually transmitted diseases. Conversely, CDC programs can use ESP to provide analyses of disease-specific trends on a state, regional, or national basis. The eventual benefit of this consolidation will be the ability of CDC and the states to centralize their data management activities. This consolidated system of disease surveillance will not only permit some reduction in personnel efforts used for data reporting, but will also serve as a centralized source for data on disease prevention and epidemiology activities.

Another advantage of ESP is in the improvement of data analysis. The ESP data can be used to test various forecasting methods and to create statistical limits for the normal incidence of diseases. Using the ESP data base, computer programs that indicate which disease reports are outside these limits can be developed and changes in disease patterns indicative of epidemics or changing epidemiologic patterns can be identified. Disease surveillance can also be made more sensitive through the developement of graphic and tabular outputs of the data. These graphs and tables -- some of which have already been sent to participating states -- can pinpoint "hot-spots" or disease-free counties. These analyses can, of course, be extended on a multi-state or regional basis allowing the identification of disease changes that may be occurring in contiguous counties across state lines.

The ESP makes disease surveillance readily accessible and in a form that is usable by epidemiologists, researchers, and administrators. The delay between reporting, keypunching, and analysis is reduced. The data are more timely. Furthermore, current data are available for analysis and these data can be consistently compared with those data submitted from other states and counties. It is this "currency" of surveillance data that makes the ESP a model for surveillance systems of the future.

THE AFRICA HEALTH TRENDS DATA BASE

Kristine Olsen Powell, Susan P. Enea, and Leo T. Hool
U.S. Bureau of the Census

Background

The U. S. Bureau of the Census International Statistical Programs Center (ISPC) has been involved in assisting developing countries with improvement of their censuses and other data collection activities since the 1940's. The Agency for International Development (AID) has been, and continues to be, the chief source of funding for the international activities of the Census Bureau. However, in recent years the support of other organizations such as the United Nations, the World Bank, and individual country governments has been growing.

ISPC has a training center for participants from developing countries, provides long-term advisors especially for census activities, and provides short-term consultations for data collection and data processing activities such as censuses and household surveys. ISPC provides technical assistance to AID staff in Washington and in field Missions around the world concerning all phases of data collection - identification of data needs, data collection, data processing and analysis. As part of the work with the Washington-based staff in the Africa Bureau of AID, ISPC staff were asked to put together a system for examining trends in health program funding and health status in African countries. The AFTRENDS system is the product of that request.

Types of Data

The AFTRENDS system is unique in bringing together in one easily-accessed system, data concerning socioeconomic and demographic conditions, health and nutritional status of the population, and financial obligations for health. Diverse indicators are useful in decisionmaking for program management and health services planning. The data will be modified to reflect data collection efforts currently underway in the countries included.

Users of the AFTRENDS Data

The primary users of the AFTRENDS data are expected to be AID Washington-based staff, AID Mission Health Officers, and other donors. The system is expected to be used for several purposes: 1) to identify countries in greatest need of health assistance, 2) to examine the relationship between health assistance and health indicators over time; 3) to compare AID's health assistance to that from other donors and to coordinate assistance; 4) to compare AID's assistance in health to assistance in other sectors over time.

Countries

Data are presently available for 50 North African and Sub-Saharan African countries, and the United States has also been included for comparison purposes. Data are available in a time series from 1970 to 1985.

Data Sources

The primary sources of data for the AFTRENDS system are the World Population Reports and the International Data Base of the U.S. Bureau of the Census Center for International Research, United Nations and World Health Organization reports, World Bank reports and data sets, AID's Economic and Social Data Base and Bureau for Program and Policy Coordination reports, and the Organization for Economic Cooperation for Development (OECD). Additional sources of data are currently being reviewed.

Data Base Descriptions

The AFTRENDS system consists of four data files, three of which correspond to the subject areas already described. The fourth is a documentation file which specifies the definition, source of data, and user caveats for each variable included in other files.

The first of the data files contains socioeconomic, demographic, and donor information. This includes GNP and per capita GNP, population, infant mortality rate and life expectancy at birth, and commitments and disbursements of the Official Development Assistance from multilateral and bilateral donors.

The second data file, AIDOBDB, includes information on AID obligations to individual African countries during the time period covered (1970 to 1986 expected obligations for these data). Total amounts for all programs as well as the amounts obligated from each subaccount are shown. For some years certain generalized accounts (such as Sahel Development Fund) can be disaggregated to show the portion that pertains to health activities. Information is also available on AID obligations for other areas, such as population activities, education and human resources activities, and selected development activities.

The third data file, AILMENTS, contains data on the numbers of new cases of certain major communicable diseases which are reported to the World Health Organization. The diseases currently included are those targeted by the Expanded Program of Immunization (EPI)--tuberculosis, diptheria, measles, pertussis, polio, tetanus, and neonatal tetanus.

Data Processing for AFTRENDS

The data are organized in dBASE II, a data base management system which allows ease in data

query and production of reports. (In the next few months the transition will be made to dBASE III.) LOTUS 1-2-3 is used to convert dBASE II files into a spreadsheet format for further manipulation and graphic generation. Data from other large-scale data sets are easily integrated into the AFTRENDS system. LOTUS can convert files from the DATA INTERCHANGE FORMAT (DIF), a common data structure on microcomputers. dBASE can import files from the Standard Data Format (SDF) Structure. Virtually all microcomputer software is capable of generating ASCII text files. dBASE can readily import this type of file as long as the record structure has been defined in a dBASE file. In addition to these possibilities, there is software readily available for mainframe computers which can convert mainframe files to LOTUS, dBASE, or DIF file structures.

Each data file contains a variable called "year/country," a unique identifier which allows rapid searches of the files to locate individual data items, and facilitates linkage between the separate files. Using the data files and dBASE II and LOTUS software, standardized reports, graphs and charts can be generated, and impromptu searches and retrievals also can be made.

Current Status

AFTRENDS is currently operational and available for data access by AID staff. Other users would need to request information through AID's Africa Bureau, Health, Population, and Nutrition Office.

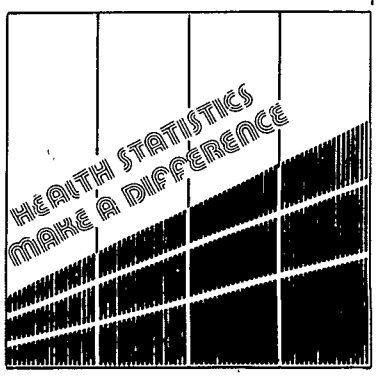
At this time, updates and enhancements are being made to the AFTRENDS system. These include review of additional data sources, documenting data included in the system, performing consistency checks, developing standardized formats for data presentation, and reviewing software options, especially for graphics.

ISPC staff have made several presentations of the system's operation, and have several planned for the next few months. Refinements are being made to make AFTRENDS "user friendly" and to require minimal experience with computers to operate the system. Both dBASE and LOTUS are widely distributed and used, so data extracts from AFTRENDS can easily be sent to distant sites. Presently, modifications will be made to develop the system as menu-driven, further reducing the effort of data entry and retrieval.

To recommend additional data sources or request additional information about the AFTRENDS system, contact Ms. Kristine Olsen Powell, U.S. Bureau of the Census, International Statistical Programs Center, Evaluative Studies Branch, Scuderi 304, Washington, D.C. 20233.

Session B

Statistical Organizations



LOCATING DATA AND DISEASE

Joseph D. Carney, Oregon State Health Division

"The time has come," the Walrus said, "To talk of many things: of shoes -- and ships -- and sealing wax -- of cabbages -- and kings -- and why the sea is boiling hot -- and whether pigs have wings."

That favorite quote of mine from Lewis Carrol, author of "Alice in Wonderland" came to mind as I prepared to put my thoughts together for the topic before us. To talk of today's Centers for Health Statistics is indeed "to talk of many things". I don't know where the ships and sealing wax may fit in, but I'm sure many of you know that the "shoe" with a hole in it is a symbol of the epidemiologist. And since our topic looks at joining health statistics and epidemiology, we most certainly will talk "of shoes"! I suspect that "why the sea is boiling hot" has to do with federal funding of Centers for Health Statistics, so I'll leave that for others to discuss. Lastly, my wife, who is an ex-cop, wouldn't let me touch the question of "whether pigs have wings."!

Basically what this presentation proposes to do is to offer you an organizational configuration which the State of Oregon feels is appropriate for a Center for Health Statistics in our day and time. The organizational configuration unites the statistical and epidemiological areas of expertise normally found in a state health division. To give some order to this "talk of many things" I propose to first review the origins of these two areas of study, that is, vital or health statistics, and epidemiology. Following that I would like to look at several ways these disciplines can and do exist today. We will look in the end at the Oregon organizational cofiguration, some of the early experiences we have had, and what we anticipate in the future.

Statistical Origins

Many of the Centers for Health Statistics which we have today developed from vital statistics offices during the Cooperative Health Statistics Program which was sponsored by the National Center for Health Statistics in the early seventies. Those vital statistics offices had, and many of our Centers for Health Statistics still have, a dual function. They have the legal function of maintaining birth and death records for events occurring in their jurisdiction and the statistical function of using data from those records to produce natality and mortality statistics for their jurisdictions. I am sure most of us are aware that the statistical use of vital records dates back to seventeenth century England when a haberdasher by the name of John Graunt published a book entitled Natural and Political Observations Made Upon the Bills of Mortality. This long history certainly adds to the stature of the profession many of us follow today. Dating back to the seventeenth century certainly gives us substance.

But that's just the history of the statistical function. While researching a paper for presentation at an Iberoamerican conference in Lima some years ago, I discovered the lesser known history of the legal function of vital records. There appear to have been vital registration systems in place as early and as far spread as 1250 BC in Egypt and 720 AD in Japan. It is probable that the Egyptian system did not include all classes of the population, and in both cases it is likely that the systems were ecclesiastically based rather than civil registration systems. Most early systems, indeed, counted baptisms, burials, and weddings (that is, religious ceremonies) rather than births, deaths, and marriages.

As it turns out, however, the earliest recorded civil registration dates back even further than seventeenth century England. It dates back to the Incas of Peru who had developed a civil registration system long before the Spanish conquerors arrived. The Inca system used the intertwining of colored strings and knots to record vital events. This "Peruvian Knot Record" was used because the Incas had no written characters for simple sounds. There was a "Quipucamayu" in each village who kept the knot record, known as a quipu. Today, of course, we have registrars and computers but the basic idea remains unchanged.

Epidemiological Origins

I quote now from an early work on epidemiology: "Whoever wishes to investigate medicine properly should proceed thus: in the first place to consider the seasons of the year, and what effects each of them produces. Then the winds, the hot and the cold, especially such as are common to all countries, and then such as are peculiar to each locality. In the same manner, when one comes into a city to which he is a stranger, he should consider its situation, how it lies as to the winds and the rising of the sun; for its influence is not the same whether it lies to the north or the south, to the rising or to the setting sun. One should consider most attentively the waters which the inhabitants use, whether they be marshy and soft or hard and running from elevated and rocky situations,and then if saltish and unfit for cooking; and the ground, whether it be naked and deficient in water, or wooded and well watered, and whether it lies in a hollow, confined situation, or is elevated and cold; and the mode in which the inhabitants live, and what are their pursuits, whether they are fond of drinking and eating to excess, and given to indolence, or are fond of exercise and labor." (1).

The environmental epidemiologist I just quoted is Hippocrates. The quotation is from

his work On Airs, Waters and Places and dates back almost 2400 years. We have in epidemiology, then, another discipline with the stature of history.

MacMahon et al. in their work, Epidemiology, Principles and Methods define epidemiology as the study of the distribution and determinants of disease frequency in man"(2). In studying the distribution of disease we find epidemiology borrowing much from demography and statistics. A good history of epidemiology will find mention of John Graunt's Natural and Political Observations ... on the Bills of Mortality. Yes, the same John Graunt and the same book mentioned earlier as the foundation of the statistical use of vital records! Histories of epidemiology will also point to the work of William Farr, an early ninteenth century physician, as the root of today's epidemiology.(3). This same William Farr was the first Compiler of Abstracts of the Registrar General's Office and organized the first modern vital statistics system.(4).

MacMahon's definition of epidemiology mentions not only distribution of disease but also determinants of disease. It is in looking at determinants of disease that the epidemiologist becomes involved with causal factors. The classification of causal factors for general use brings into the light of history once again William Farr. In the mid 1800's Farr picked up on the work done by William Cullen and took another major step in the classification of disease which continues to this day in our International Classification of Diseases and Causes of Death.

A lesson in History

This brief review of the historical origins of Centers for Health Statisitcs and of epidemiology shows us that the two disciplines have much in common. John Graunt wove his way into both historical sketches, as also did William Farr. With so much in common in their historical backgrounds, why do the two disciplines seem so far apart today?

Perhaps a part of the answer can be seen if we consider the use of the term "epidemic". MacMahon, et al, point out in their previously cited work that "in the past the term epidemic was used almost exclusively to describe an outbreak of infectious disease. More current definitions stress the concept of excessive prevalence as its basic implication".(5). In the past, and even in some respects today, many think of an epidemic in terms of measles, or flu, or food-borne illness. However, as MacMahon points out in illustration of his revised definition, we have in the United States today two non-infectious diseases which certainly meet the criterion of "excessive prevalence". We can look at coronary heart disease, responsible for a third of U.S. deaths, and we can look at lung cancer which is now some 30 times more common than it was in the 1930s. Both these non-infectious diseases can be classified as outbreaks of excessive prevalence. What is being said here is that an outbreak of disease where frequency is excessive can be classed as epidemic without requiring that the excessive frequency be measured with a time period of merely days or weeks.

What I am suggesting, then, is that thinking of the epidemiologist as one who responds to epidemics of the older, narrower type would certainly relegate that discipline to an existence largely separate and apart from the world of health statistics which is busy gathering data over time and expressing rates over large populations. Think, however, of epidemic with the more current stress on excessive prevalence and the more current denial that the increased frequency need occur during short time periods. If you think of epidemic in such a light as that, then the link to health statistics jumps at you! Not only is the detection of excessive frequency important to the epidemiologist under such an understanding of epidemic, but also routine gathering of data becomes a necessary tool for the epidemiologist to know the norm so that the excessive frequency can be detected.

So, on the first floor of many of our state health divisions the statisticians are busily gleaning data about heart disease and cancer from death certificates while on the fifth floor the epidemiologists in a disease control section concern themselves with the daily emergencies of chemically tainted watermelons, guacamole, and lettuce. I do not mean by this to put down either what the statisticians are doing, or what the epidemiologsts are doing. What I do mean is to underline that by each going a separate way we lose the benefits which can be reaped by what history shows as a definitely symbiotic relationship between the disciplines. Joining the disciplines opens up to the epidemiologist all the rates, population figures, mortality data, and natality data that he will need for surveillance of the currently defined epidemics. The statistician, meanwhile, is rewarded with a new source for morbidity data and medical expertise, and a new array of possible research ventures.

Oregon's Decision

It would be nice to say that after careful consideration of these historical trends, and after studying the patterns of crossovers between statistics and epidemiology through the centuries, and after weighing the implications of the current definition of epidemic that Oregon decided in favor of joining the two disciplines. It would be nice to say that, but it would not be accurate. It would be more accurate to say that like many amateur genealogists we began to check our lineage only after we had become who we are. It was one of those ideas whose time had come. Discussion about reorganization began well in advance of the time that Division budgets would need to be presented for the 1983-85 biennium. Initial discussions took place at quarterly management

meetings and involved all of management.

Another nice thing to be able to say is that the process was smooth and easy. But, again, that would not really reflect the truth. Although, I think it would be fair to say the process was smooth. That, I believe, came from the fact that the Center for Health Statistics staff and the epidemiology staff were far from unfamiliar with each other. A fair amount of interest on the epidemiologists part in environmental epidemiology had already generated numerous contacts and crossovers. The medical doctors in what was then the Disease Control Program had already been tapped by our energetic nosology staff for help in coding odd looking death causes. I'm sure that quirks of the informal organization structure helped the smoothness also. For example, one of the Center's researchers had a background in biology and was called on to aid in the Division reaction to a plague epidemic.

I mention these facts because I am aware from talking individually to members of some state staffs that there is not always a comfortable, open, and communicative relationship existing between the statistical staff and the epidemiology staff. We were fortunate in this regard.

However, although I can allow the word "smooth" for Oregon's transition, I can not allow the modifier "easy". Remember, there was a legislative ways and means committee to be convinced! Many hours of hard work went into the preparation of a budget document which could both accomplish the necessary transfer of funds to allow the reorganization, and at the same time show and explain that transfer simply enough to be understood by county commissioners, legislators, budget analysts, et cetera, et cetera! Somehow it got done, and on July 1, 1983 the Center for Health Statistics was no longer a section of the Office of Staff Services, but instead was a major section in the newly formed Office of Health Status Monitoring. The Office is headed by the State Epidemiologist and contains as other major sections in addition to the Center for Health Statistics, the Epidemiology Services Section, and the Sexually Transmitted Disease Section.

Early Experiences of Union

I would like now to take a few moments to look at some of the activities we have gotten into since July 1, 1983 which are different because of the reorganization.

One of the first activities was organizational in nature. The charge to the newly formed Office of Health Status Monitoring included the development of research projects from ideas generated not only from within the office itself, but also from other offices within the Division, from local health officials, and from other state and local agencies. With such a wide range of places generating project ideas, it was necessary to begin by establishing methods for proposing ideas, and methods for screening and prioritizing the implementation of those project ideas.

A committee composed of myself as manager of the Center for Health Statistics, the supervising research analyst from my research project staff, the chief epidemiologist, a disease epidemiologist, and an environmental epidemiologist was formed to develop proposal and prioritizing methods. In addition to development of these methods, the committee also developed a set of criteria to be used in judging priority of projects.

This same Advisory Committee also took on as one of the new Office's primary tasks, the categorization of data systems easily available to Health Status Monitoring for use in its research projects.

Having set up the project proposal and prioritizing system, one of the first projects to be reviewed was a proposed Trauma Registry. This supplied an early opportunity for the new office to aid in the development of a project in a different office of the Division. The work on the trauma registry, begun in 1983, was brought to fruition with legislative action establishing the program this year.

The OHSM staff was also called on early to work with the legislatively established Agent Orange Committee. This program, although small, has also been re-established with specific legislative goals for continuation during the present biennium.

An early project which got the new office of OHSM involved with programs throughout the Division centered around the 1990 health objectives published by the Surgeon General in Promoting Health/Preventing Disease.(6). Oregon has programs in the Health Division for 13 of the 15 priority areas established in the D.H.H.S. publication. The Office of Health Status Monitoring worked with the program directors to develop objectives for Oregon in each of those thirteen areas. Results of the work were published in a 33 page report which details where Oregon is, and how it anticipates reaching the objectives. The ability to supply a combination of statistical and medical expertise proved exceedingly important in implementing this project.

An area of research currently being worked on is sudden infant death syndrome. The combined forces of epidemiology and statistics are joining in this effort with the Oregon SIDS Institute in pursuing this research area. This is an example of the type project that resulted from the combining of the disciplines serving as a means to broaden the set of outside contact organizations for each of the disciplines.

There are numerous examples of the interaction caused by the reorganization working to the definite advantage of all:

- As a result of a joint meeting with Dr. Sam Milham of the State of Washington, OHSM decided to code occupation on all death certificates beginning in 1984.

- Statistician and epidemiologist combined to develop a system to permit integration of medical examiners files with the death certificate statistical file.

- Vital records was used as a control on an epidemiology developed system of voluntary perinatal morbidity reports.

- Statistical researchers knowledge of software packages proved indispensible in moving forward objectives of the Agent Orange program.

These are just a few specific examples illustrating what really amounts to an overall philosophy change, a mood change, a broadening of perspective.

The Future

Let me close by saying the future for OHSM looks bright indeed. It is a future in which two of the major benefits received from the reorganization will continue. The benefit for the Center for Health Statistics is to grow even further in the change from record keeper to research program; the benefit for epidemiology is to expand from watermelon and lettuce tester to medical researcher responding to the current definition of epidemic with preventive measures to promote public health.

(1) Hippocrates. 1939. The Geniune Works of Hippocrates. Translated from the Greek by Francis Adams. Baltimore; Williams and Wilkins, P 19.

(2) MacMahon, Brian, M.D., Ph.D., D.P.H., Pugh, Thomas F., M.D., M.P.H. 1970. Epidemiology, Principles and Methods. Little, Brown & Co., Boston, P 1.

(3) Ibid., p 6.

(4) Lilienfeld, Abraham M., M.D., M.P.H., Lilienfeld, David E., A.B., M.S. Eng. 1980. Foundations of Epidemiology (Second Edition), New York, Oxford University, P 35.

(5) MacMahon, Brian. Loc. Cit. P 2.

(6) Department of Health and Human Services. Promoting Health/Preventing Disease, Objectives for the Nation. Fall 1980. Washington, D.C.

IMPLEMENTING A STATE CENTER FOR HEALTH STATISTICS--SELECTED FUNDING IMPLICATIONS

Paul D. Gunderson, Minnesota Department of Health

BACKGROUND

State Centers for Health Statistics have emerged in numerous states across the nation. Ranging in size from only a few staff to 80 or more they exhibit a wide diversity of organizational styles and statistical programs. All however have faced significant budgetary constraint as programs have been built.

Several models for funding statistical activities have been used including primary reliance upon federal grant programs, state grant programs, state-level line item, fee-for-service, and private foundation aid. Minnesota's experience involved blending of the above sources...a blending which I will review in order to detail certain managerial implications.

Table 1 indicates the Minnesota Center for Health Statistics funding source pattern for the past decade. The number of funding sources varied as the work load expanded and contracted.

Table 1
Funding Pattern,
Minnesota Center for Health Statistics
1975-1985

1975	7	Funding Sources
1978	8	Funding Sources
1980	11	Funding Sources
1982	17	Funding Sources
1985	15	Funding Sources

The sources of funding are identified in Table 2 below and include seven federal funding sources, two private foundations, and conventional state appropriations.

Table 2
Sources of Funding by Type,
Minnesota Center for Health Statistics
1975-1985

FEDERAL PROGRAMS

Department of Health and Human Services
- NCHS (National Center for Health Statistics)
- CDC (Centers for Disease Control)
- BCP (Bureau of Community Programs)
- NIH (National Institutes of Health)

Other
- EPA (Environmental Protection Agency)
- DOA (Department of Agriculture)
- DOT (Department of Transportation)

PRIVATE FOUNDATION
- Regional Medical Program
- Bush Foundation

STATE GRANT

STATE LINE APPROPRIATION

FEE FOR SERVICE
- Other Public Agencies
- Private Sector

PROJECT COST ACCOUNTING

The Minnesota Center for Health Statistics responds programatically to numerous agencies, both public and private. Each possesses a legitimate interest in insuring that resources are precisely focused upon funded projects. The "blending" of funding sources depicted above required the keeping of audit trails, accurate accounts of "overhead costs," and resort to "pledges," contracts, or letters of agreement between funding sources and the Minnesota Center for Health Statistics. Audit trails were necessary in order to provide a convincing record of both expenditures as well as fund allocations and transfers. Of management significance here is the personnel timekeeping which was instituted to support the audit function. Overhead costs were also tracked very carefully since some funding sources excluded such costs from subsidy awards while others placed strict limits on the proportionate amounts which could be extracted from award levels. The use of pledges as well as contracts and other quasi-contractual devices permitted effective management of cash flows during each fiscal year.

The construction of audit trails required initiation of positive timekeeping. Staff time within the Minnesota Center for Health Statistics (MCHS) is expended in three dimensions:
- direct hours
- unavailable hours (annual leave, sick leave, etc.)
- indirect hours

Costs associated with the three dimensions are allocated across all projects embedded within the annual MCHS workplan. Summary reports are generated monthly and consist of analyses of cost by project (Figure I), cost by project by person (Figure II), time by project by person,

Figure I
Example of Project Cost Analysis by Project

FISCAL YEAR 84
REPORT THRU MONTH OF JUNE

PROJECT NAME
PROJECT NUMBER

	JUL	AUG	SEP
OFFICE, GENERAL			
14001			
PROFESSIONAL	$1057	$229	$52
CLERICAL	$2752	$1755	$2244
PROJECT TOTAL	$3809	$1984	$2296
TIME DISTRIBUTION SYSTE			
14004			
PROFESSIONAL	$247	$230	$390
CLERICAL	$107	$103	$238
PROJECT TOTAL	$354	$333	$628

and time by person by project (Figure III).

The report from which Figure I is excerpted is circulated to program management where appropriate, as well as within MCHS. It serves to highlight temporal expenditure variation within projects as well as provide monthly and final project totals for each fiscal year by type of personnel category.

Figure II
Example of Project Cost Analysis
by Project and Person

FISCAL YEAR 84
REPORT THRU MONTH OF JUNE

PROJECT NAME / PROJECT NUMBER	JUL	AUG	SEP
OFFICE, GENERAL			
14001			
JOSAAS	$1084	$351	$540
BEDARD	$1046	$521	$822
ROSENBAUM			
BRADFORD			
HAMMERSTROM	$622	$880	$864
VARGAS	$1057		
WIGGINTON		$229	
STEINER		$3	
HOLST.			
093640			
SALKOWICZ			
HEIKKILA			
GIBBONS			$52
SWENSON			$18
POLLMAN			
OLSEN			
PROJECT TOTAL	$3809	$1984	$2296

Figure II illustrates expenditure detail for each month by individual MCHS staff person. This report does not circulate outside of MCHS and is used primarily for project control purposes.

Figure III
Example of Project Cost Analysis
by Person by Project

FISCAL YEAR 1984
REPORT THRU MONTH OF JUNE

EMPLOYEE / PROJECT	JUL	AUG	SEP
00011			
ANNUAL LEAVE		10.00	5.00
00022			
SICK LEAVE	10.00	22.00	8.00
00055			
TRAINING			
00056			
MEETING	2.00	1.00	
00066			
HOLIDAY	8.00		8.00
00099			
OTHER			
10003			
CHS - GRANTS(910)		18.00	
10004			
CHS GRANT REVIEW			36.00
14008			
MINN. HEALTH STATISTICS	20.00	14.00	73.00
14012			
GEN. GRAPHICS/TERMINALS			7.00

Figure III shows an excerpt from a report which is used for assessment and control of individual staff commitments to projects. This report is also used when analyzing time proportions associated with unavailable and overhead time (cost).

Mention was made earlier of overhead expense to which some funding sources develop peculiar sensitivities. The blending of funding sources also requires determination of other permissible expenditures. The cost of telephones, office space, furniture, automated equipment, data processing vendors, duplicating and printing, etc. cannot be ignored and must be accommodated in some manner. A decade of statistical activity funding suggests that agreement on these matters is essential prior to incurring expenditures. In particular, we have discovered that such agreement was needed when using resources from within the Department.

In order to reach internal management accord on these matters, we initiated "pledging" within the Department. Figure IV depicts one version which served to solidify expectation and significantly reduce management uncertainty, both within MCHS as well as within separate activity levels of the Department.

Figure IV
Pledging Document

SUBJECT:

Shown below is an estimate of federal funds allocated to MCHS for the provision of services for FY 1978. These estimates were based on actual time and cost factors for services provided to you during FY 1977.

Position Type	Estimated Cost	FTE
Clerical	9,466	1
Data Entry	12,956	1.25
Health Coder	8,056	.75
Statistician	23,319	1
Systems Analyst	30,589	1.85
TOTAL	83,861	

Your signature (below) and indication of the funding source category represents (1) an approval for federal funds to be spent for these services, and (2) the fact that these services are needed.

TO: Accounts and Finance
FROM: MCH
(Signature)
Transfer the above amount from Donald McCampbell
(funding source/category/categories)

MANAGERIAL IMPLICATIONS

There are at least four implications that have emerged from development of the automated system to date, including exterior audit survival, fund allocation and transfer, enhancement of management trust, and system costs.

Surviving HHS audits is a laudable management objective. MCHS has not "failed" any, has been informed by HHS staff and/or staff of the Ninth Federal Reserve (Bank) that our system is exemplary. In keeping with an HHS agreement, hard copy of employee timesheets is kept for one budget year, after which it is shredded and total reliance is placed on electronic image.

Agreement concerning the initial allocation of resources and any transfers of staff between funding sources has often been an issue. Little management doubt, either within MCHS or among grantees, now exists before, during, or after conclusion of a project since each such inquiry

is fielded directly off the automated timekeeping system, and an explicit record of such transfers exists. A reduction in management uncertainty has occurred, permitting a near-total focus on project outcome.

Trust among management levels whose resources are invested in projects developed by MCHS remains high since routine (monthly) accounting is performed on all sources of funding. My experience suggests that, aside from project focus, the next most significant area for disagreement among managers is cost, with staff efficiency a close third. Discussion of both facets is enhanced by information resulting from an explicit cost accounting system.

Managing budget resources costs money. In the nine years MCHS has operated its automated cost accounting system, the annual cost has hovered around one percent of total budget. I submit that the expenditure is almost insignificant, yet is appropriate since it has enhanced our revenue generation capability for the future.

THE SURVIVAL OF THE FITTEST--A COALITION APPROACH TO SELF-SUFFICIENCY AS A DATA BROKER

Elliot M. Stone and Cynthia E. Burghard
Massachusetts Health Data Consortium, Inc.

This presentation is a 20 minute history of the Mass Health Data Consortium which is now in its 8th year of existence.

The Consortium traces its origin to grants from NCHS and its Cooperative Health Statistics System. The CHSS encouraged the 10 demonstration states to select the most appropriate technical and political model for pooling hospital discharge data. The most common models for the holder of a statewide hospital data base are:

- Trade Association
- Insurance Carrier
- Consortium/Coalition
- Regulation/Public Domain

In Massachusetts the consortium/coalition model was chosen for two reasons: 1) The previous successful experience in R.I. and Vermont and 2) a consortium was everyone's second choice. Their first choice was, of course, that their own agency control the statewide data base.

The original CHSS Advisory Committee expanded to 19 members who now pay annual dues in order to set policy for the organization. Every major agency that holds or uses health data is a member/owner of the Consortium, including:

- Hospital Association
- Medical Society
- Federation of Nursing Homes
- Blue Cross
- HIAA
- Statewide PRO
- Four state agencies
  - Public Health
  - Welfare
  - Rate Setting
  - Health Policy
- SHCC
- Six regional planning agencies
- Business Roundtable

The agenda of these groups keeps us involved in issues of access and cost. As you can see, it is a true balance of providers and users--with neither in control.

Case mix and charge data are collected under voluntary contracts with each of the 110 acute care hospitals in Massachusetts. Twenty-five other hospitals from bordering states and V.A. hospitals contribute data so that population-based studies are facilitated. Over 4 million inpatient records are now in the multi-year data base.

Funding for the Consortium has shifted from a reliance on funding from NCHS and HCFA--as high as two thirds in our first year to a greater emphasis on self-sufficiency. We no longer apply for federal funds unless it would complement work in progress. Currently, we are self-sufficient without a dollar of federal funding.

Without federal money, the Consortium has designed its products as an Information Utility.

Our collaborative research with our members and other agencies has resulted in significant applications of the inpatient data base as well as the collection of new data sources. We collaborated with the Greater Boston HSA to publish the first study of variations in surgical procedures in a large industrial state. We collaborated with 5 community hospitals to link nursing data and case mix data. Now for the first time, they can identify nursing hours and nursing costs with any Diagnosis, Procedure or DRG.

We are collaborating with the Mass Business Roundtable to publish a hospital price guide. Collaborative fee-for-service studies are the fastest growing side of our revenue. Findings are the property of the collaborator who finances the study and they decide on public disclosure.

Ten percent of our revenue is derived from education. Conferences are held on the application of data to health policy issues. Our annual meeting attracts over 250 attendees and other successful seminars and workshops have addressed: long term care data, the uses of charge data as well as training physicians to use personal computers. We are currently studying the feasibility of establishing our own Health Information Training Institute.

Independent Research is our way of categorizing products of a sensitive nature that must be approved by the Board. All independent work is financed by the dues and must be disclosed publicly.

These independent products include:

- Patient Origin
- Migration
- Use Rates
- Market Share
- DRG Profiles
- Case Mix Indices
- Charge Profiles by DRG
- ON-LINE Access
- Data Digest

Individual hospitals are identified, but patients are confidential. Our staff is highly product oriented.

Hospitals receive their own reports at no charge in return for pooling data with the Consortium. Market Share is the most popular report, and ON-LINE Access allows users to download the Consortium's data to a user's own personal computer.

Beginning with the 1983 data, we shall publish studies by unique physician numbers; i.e., not identified by name, but with the same number for all the hospitals where the physician attends patients.

These independent products have led to our self-sufficiency as a data broker and now constitute over one-half of our revenue.

The three key ingredients then for our survival as a non-profit data broker are: Politics, the Operation and the Users.

The Chairman of our Board--Dr. Francis Moore--defines politics as "the art of getting someone to agree to your point of view without resorting to physical violence." Our survival is linked to the politics of health care costs in Massachusetts. There are strong lobbies--each pulling in separate directions--all distrusting the others' motives and information! The Consortium supplies impartial numbers to these countervailing forces. Data sharing is a political consensus. Only senior managers participate on our Board and they must make a financial commitment--annually-- to be allowed to sit at the table.

The Operation: The working phrase that characterizes the survival of the Consortium's operation is: single mindedness. We have a full time, core staff that does not divert itself from the principal mission of annual data base building and data pooling.

The Board and staff exert leadership to ensure that an acceptable process is followed so that data providers and users are encouraged to work together. As a non-profit agency, we have always been able to call on our Technical Review Committee for support for the tasks of acquiring and editing data and educating users in order to avoid inappropriate conclusions.

We have a single mindedness about the costs for data base building also. Our first effort in 1979 cost over $600,000: $345,000 and 18 months to acquire the data--since not all hospitals were automated and we manually abstracted at several sites. Logistics were not yet in place for data sharing agreements and staff traveled to numerous meetings with hospital lawyers to finally bring in every hospital in the state.

Over $300,000 was spent in our first effort on data processing. Extensive software development and training was provided for our programmers in coping with hospital data that did not meet specifications. These costs underscore the need to share the burden among many agencies.

By 1983 we had achieved uniformity through the Consortium's feedback as well as the introduction of regulations by the state Rate Setting Commission. Acquisition costs have been lowered to $10,000 and data processing costs were reduced to $40,000 to pool one million patient records that year. The staff uses a time share system and are very cost sensitive on all programming and production runs. Costs increased slightly in 1984 with the introduction of a new data set merging patient charges with the clinical data set.

Talk of self-sufficiency is so much self-congratulation without satisfied clients. I have learned that marketing means more than selling data--it means understanding what clients need.

Our most popular reports were designed for clients who had specific questions for research or managing a facility. Our most active users are mostly within hospital planning departments and they are the ones who know the issues, understand the value of data and are creative.

Paul Densen, of Harvard, former chairman of the CHSS Advisory Committee and the and the Consortium's first President taught my staff early on to avoid creating data tables without first asking: "What is the Question?" We have worked with out clients to address issues such as the following:

- What are the characteristics of patients who leave the area for hospitalization?
- Are there variations among small geographic areas in our state in the rate of inpatient surgery?
- Can you quantify the degree to which a hospital's case mix affects its bed need?
- What is the hospital's market share of orthopedics for patients with private insurance?

Currently our population-based file is being used for the latest round of hospital bed-need hearings. Use rates by community and age group is a standard product which resolves many earlier disputes.

Facility-specific reports on DRG, Charges and Market Share continue to be our mainstay for self-sufficiency. As the data become more sensitive in a highly competitive environment, our reports are more in demand. Naturally, most hospitals wish that we were not in business to disclose their data, but they and the Massachusetts Hospital Association realize that our process is preferable to raw release of the data by public agencies.

In summary--the Consortium has survived because of the creativity of its Board and staff to turn data into useful information.

We are committed to:

- an approach of statewide data reported by region and local community,
- of comparable data which is facility and soon--physician-specific,

- of sharing the large expense among all the major users
- of a dialog between users and data providers, through our Board and Technical Review Committee.

As a result of the Consortium's credibility, the Massachusetts Business Roundtable has not established their own data gathering (as they have in other state coalitions). The membership of the Roundtable ensures the continued involvement of the other major players.

The Consortium's future activities will include the design of ambulatory data bases, but our approach to a creative blend of politics and data pooling will continue.

I have seen the demise of other consortia and data firms. I have heard their complaints that people did not know how to use or ask for the data. In effect they said, "why doesn't anyone love us?" My answer has always been: "Did you make yourself loveable?"

If the Mass Health Data Consortium continues to survive you will know that, through our products, research and services, we have made ourselves loveable to our members and clients.

Session C

Morbidity, Mortality, and Toxic Chemical Exposures

APLASTIC ANEMIA MORTALITY AND OCCUPATIONAL EXPOSURES

Robert Spirtas (National Cancer Institute), Shelia K. Hoar,
Rose Kaminski, Harry Rosenberg

Abstract

To demonstrate the value of vital statistics data in occupational disease surveillance, a NCHS/NIOSH/NCI collaborative study compared the industries mentioned on death certificates of 464 white men dying from aplastic anemia in the U.S. in 1975 with 1459 white men randomly selected from all other causes of death. Excess deaths were observed in the agricultural, forestry, and fisheries industry, lumber and wood products manufacturing, and in the printing and publishing industry. Efforts were made to identify the agents responsible for these excess risks. A job-exposure matrix was used to translate the death certificate job titles and industries into exposure data. The average number of exposures per subject was 25, ranging from 0 to 233. Of particular interest were acetone, adhesives, ammonium chloride, benzene, carbon tetrachloride, ethylene glycol, inks, mineral oil, pesticides, petroleum products, titanium oxide, trinitrotoluene, toluene, wood dust, and wood preservatives. Risk estimates for specific chemical exposures and chemical classes will be presented.

Introduction

The continual introduction of new chemical and physical agents into the workplace argues for ongoing surveillance of the health of occupational groups. Occupational disease surveillance not only serves to identify problems in the workplace but also serves as a sentinel for future problems in the general population which is likely to be exposed to the same agents as workers, but at lower levels and later in time.

This paper describes a NCHS/NIOSH/NCI collaborative study which demonstrates the value of vital statistics data in occupational disease surveillance. Aplastic anemia (ICDA 284), a persistent form of anemia caused by the bone marrow's failure to produce adequate numbers of peripheral blood elements, was chosen for study because of its known and suspected associations with occupational exposures and its rarity. Aplastic anemia has been associated with exposure to benzene (1-2) and trinitrotoluene (3) and has been suspected of being caused by exposure to pesticides (4) and carbon tetrachloride (5). Excess aplastic anemia has been reported in the rubber, printing, and shoe industries (6). The infrequent occurrence of aplastic anemia (U.S. white annual mortality rate: 0.5/100,000 [7]) necessitates a case-control study of nationwide scope to supply enough cases for adequate statistical power.

Methods

All U.S. white men who died in 1975 with aplastic anemia as the underlying cause of death were identified by NCHS, using mortality data provided by individual states as part of the ongoing U.S. vital statistics network. Death certificates were located for 464 men (99%). Approximately 3 controls per case (N=1459) were selected by an age-stratified random sample of all other causes of death.

As part of a special study (8), the U.S. Bureau of Census coded the usual occupation and industry from the death certificates using the 1970 Alphabetical Index of Industries and Occupations (9). All death certificates for persons dieing of four rare causes (aplastic anemia, liver cancer, pneumoconiosis, and dermatitis) and an age-and race-stratified random sample of all causes of death were identified to test the feasibility of coding industry and occupation from death certificates. The results showed that industry and occupation could usually be coded to the three-digit level of the Census code. In the present study we examined the risk of aplastic anemia by usual industry and occupation.

Of particular interest were the chemical exposures sustained in those jobs. Job title and industry separately are often poor surrogates for exposure; e.g., persons with the same job title in different industries can have vastly different exposures. Misclassification of exposure status reduces statistical power and dilutes risk estimates. Job-exposure matrices (JEM), which are cross-classifications of industry-specific job titles with agents to which persons in the jobs are exposed, have been developed to impute exposures from job and industry data (10). Using a JEM developed by Hoar et al. (11), the Census occupation and industry data were converted into the JEM's coding scheme and linked to known or suspect carcinogens. Each occupation-exposure pair included a crude degree of exposure variable: a "3" was assigned to jobs that appeared to involve a heavy degree of exposure to the agent, or that were classified by Hueper and Conway (12) as hazardous because of that exposure; a "2" was assigned to processing occupations in the same industry as other jobs that appeared to entail heavy exposure to the agent, and to occupations classified by Hueper and Conway (12) as suspected of being hazardous; and "1" was assigned to engineers, managers, officials, salespersons, production clerks, or professionals in the same industry as other jobs considered to entail heavy exposure.

The measure of association between aplastic anemia mortality and usual occupation, industry,

or exposure was the odds ratio (OR). When necessary, estimates were adjusted for the effects of age by stratification. Maximum likelihood estimates of the overall risk and 95% confidence intervals (CI) were computed by Gart's method (13).

Results

Table 1 shows the number of deaths among U.S. white men during 1975 due to aplastic anemia and the control group of all other causes, by usual industry. Age-adjusted odds ratios were calculated using three age strata (<= 64 years, 65-74 years, 75+ years) with persons usually employed in the relatively nonexposed public administration category as the referent. Usual industry was listed as "retired" or not reported for 133 cases and 415 controls. Significant excesses of aplastic anemia were seen in association with usual employment in the agriculture, forestry, and fisheries (OR=2.4), construction (OR=2.0), lumber and wood products manufacture (OR=3.7), and printing and publishing (OR=6.2) industries. The agriculture, forestry, and fisheries industry association was attributable to excess risk in agriculture (OR=2.4; 95% CI=1.3,4.7). All but 2 cases and 4 controls were in the agriculture subcategory. Several other industries had non-significant excess with 5 or more deaths due to aplastic anemia: manufacturing of machinery, transportation equipment, and textile and apparel, transportation industry, retail trade, finance, insurance, and real estate, and business and repair services. No significant deficits were observed.

Risk according to usual occupation is presented in Table 2. Using the professional, technical, and kindred workers as a referent, no significant excess of aplastic anemia was seen in any occupational group. Farmers and farm managers had a nonsignificant OR of 1.6 (95% CI=0.9,2.8). Within craftsmen, 21 cases and 32 controls were carpenters (OR=1.8; 95% CI=0.8,3.9). Usual occupation was listed as student or retired or was not reported for 75 cases and 242 controls.

The JEM was used to translate the death certificate job title and industry data into exposure data. The average number of exposures per subject was 25, ranging from 0 to 233. Previous epidemiologic research on aplastic anemia plus the results of this study's industry and occupation analyses focussed attention on 15 chemicals and chemical classes. Table 3 contains the number of subjects ever exposed to the a priori suspect substances and odds ratios. The referent category for each comparison is all subjects not exposed to the substance under consideration. Exposure to wood dust was significantly associated with aplastic anemia (OR=1.7; 95% CI=1.1,2.6). The risks associated with pesticides (OR=1.2; 95% CI=0.9,1.6) and wood preservatives (OR=1.1; 95% CI=0.9,1.4) were of borderline significance.

Most chemicals had either no difference in risk by exposure level or had too few subjects in the low and moderate exposure categories for meaningful evaluation. However, for pesticides, petroleum products, wood dust, and wood preservatives, risks were significantly elevated for the moderately exposed workers (Table 4). Moderate exposure to pesticides carried an OR of 1.5 (95% CI=1.1,2.1). Petroleum products had an OR of 1.3 (95% CI=1.0,1.8) in the moderate group. Both the moderate and high levels of wood dust exposure were significantly elevated. The moderate category, based on small numbers, had an OR of 7.0 (95% CI=1.4,37.7). The heavily exposed subjects, based on larger numbers, had an OR of 1.6 (95% CI=1.0,2.5). The OR for persons exposed to moderate levels of wood preservatives was 1.4 (95% CI=1.1,1.9). When restricted to subjects whose cause of death had been confirmed by an autopsy, almost every risk estimate increased. Moderate exposure to wood preservatives had an almost 3-fold increase in aplastic anemia (OR=2.7; 95% CI=1.0,7.6).

We also examined the aplastic anemia risk associated with other exposures in the JEM that were not a priori suspect chemicals. Elevated risks were seen among persons exposed to beta-naphthylamine (OR=1.3), oil of orange, a flavoring agent and perfume ingredient (OR=1.7), phenol (OR=1.6), carbamates (OR=1.7), coal tar and pitch (OR=1.4), estrogens (OR=1.6), soot (OR=1.1), DDT (OR=1.8), dieldrin (OR=1.8), endrin (OR=1.8), diethylene glycol (OR=1.1), acaroid resin (OR=1.4), calcium cyanide (OR=1.6), thiourea (OR=1.7), arsine (OR=1.5), barium (OR=1.5), molybdenum (OR=1.8), calcium oxide (OR=1.6), phosphorus (OR=1.4), sodium metasilicate (OR=1.9), hydrogen chloride (OR=1.5), ammonia (OR=1.7), and nitrogen oxides (OR=1.5).

Discussion

The project served two purposes. First, it tested and generated hypotheses concerning the etiology of aplastic anemia. Second, it demonstrated the value of vital statistics data in occupational disease surveillance. We used a JEM to quickly and inexpensively supplement the occupation and industry items on the death certificates.

Aplastic anemia was found to be associated with employment in agriculture, construction, lumber and wood products manufacturing, and in printing and publishing. The occupations of farmer and carpenter had nonsignficant excesses of aplastic anemia mortality. Exposures found to be related to aplastic anemia include: wood preservatives, pesticides, carbamates, DDT, dieldrin, endrin, thiourea, calcium cyanide, molybdenum, calcium oxide, sodium metasilicate, phosphorus, phosphorus, and ammonia.

Previous research has linked aplastic anemia to many nonoccupational factors including the drugs chloramphenicol (14), naproxen (15), and phenylbutazone (15), hepatitis infection, infectious mononucleosis, dengue, influenza, high estrogen levels during pregnancy, irradia-

tion, paroxysmal nocturnal hemoglobinuria, leukemia, immunologic disorders, and inherited syndromes such as Fanconi's anemia, dyskeratosis congenita, and the Schwachman-Diamond syndrome (14). Occupational factors related to aplastic anemia are benzene (1,2,6), trinitrotoluene (3), carbon tetrachloride (5), and chlorinated hydrocarbon pesticides, such as DDT, lindane, and chlordane (4,16). However, a case-control study of aplastic anemia deaths in North Carolina found no association with pesticide exposure and concluded that aplastic anemia may be an idiosyncratic reaction to pesticide exposure (16).

Our study examining the role of the workplace in the etiology of aplastic anemia found associations with agriculture and related exposures in all three approaches: industry, job title, and inferred exposures from the JEM. Carbamates, DDT, dieldrin, and endrin are chlorinated hydrocarbon pesticides. Calcium oxide is used in insecticides and fungicides. Calcium cyanide is used in rodenticides and fungicides. Phophorus is used in rodenticides, other pesticides, and fertilizers. Ammonia is also a principal component of agricultural fertilizers. These data did not confirm the previously reported associations with benzene, carbon tetrachloride, or trinitrotoluene. The associations with construction, carpentry, manufacture of lumber and wood products, and wood preservatives form another related group of occupations and exposures that should be evaluated in other data sets.

The findings must be viewed with caution because of the limitations of death certificate occupational data (17) and JEMs (10). Death certificate occupation and industry may be inaccurate (18). The most recent occupation may appear instead of the requested usual occupation. Upgrading occurs with people reporting jobs of higher socioeconomic status than jobs actually held by the decedents. Death certificate information is often thought to be incomplete (18); however, a report on the data on which the present study was based showed that there was usable information for 90% of the death certificates sampled (7).

JEMs are based on exposures inferred from job title and industry, not actual exposure histories for individual study subjects. Workplace variation over time and place can introduce errors in exposure assignment. The JEM used in this analysis had only crude dose measurements. It is conceivable that an entire class of chemicals, such as petroleum products, was indicted when only a few chemicals in the class were hazardous. Also, the JEM was limited to known or suspect carcinogens. There are thousands of other chemicals that were not studied. However, despite the limitations of the JEM, we believe its use enhanced the death certificate-based analysis. It allowed us to go beyond occupation and industry to specific exposures and, thereby, reduced the misclassification of exposure status.

Finally, there were no data on smoking habits or other potential confounding factors, although recent work by Blair et al. (19) suggests that the effects of smoking on occupational associations may be far less than commonly thought.

We believe the data and methods used in this study are particularly useful for surveillance of diseases that are too rare to be studied using living cases identified through hospitals over a short time period or from small geographic areas.

References

1. Snyder R, Kocsis JJ. 1975. Current concepts of benzene toxicity. Critical Reviews in Toxicology 3: 265-288.

2. Mallory TB, Gall EA, Brickley WS. 1939. Chronic exposure to benzene (Benzol), III, The pathologic results. J Indust Hyg Toxicol 21: 355-377.

3. Crawford MAD. 1954. Aplastic anemia due to trinitrotoluene intoxication. Brit Med J 2: 430-437.

4. Sanchez-Medal L, Castenado JP, Garcia-Rojas F. 1963. Insecticides and aplastic anemia. N Engl J Med 269: 1365-1367.

5. Straus B. 1954. Aplastic anemia following exposure to carbon tetrachloride. JAMA 155: 737-739.

6. Vigliana EC. 1976. Leukemia associated with benzene exposure. Ann NY Acad Sci 271: 143-151.

7. Mason TJ, Fraumeni JF Jr, Hoover R, Blot WJ. 1981. AN ATLAS OF MORTALITY FROM SELECTED DISEASES. P. 223. Washington, DC, DHEW (NIH Publ, No 81-2397).

8. Rosenberg HM, Burnham D, Spirtas R, Valdisera V. 1979. Occupation and industry information from the death certificate: Assessment of the completeness of reporting. STATISTICAL USES OF ADMINISTRATIVE RECORDS WITH EMPHASIS ON MORTALITY AND DISABILITY RESEARCH. L Del Bene and F Scheuren, Editors. P. 83-89. Washington, DC, DHEW.

9. U.S. Bureau of Census. 1971. 1970 CENSUS OF POPULATION ALPHABETIC INDEX OF INDUSTRIES AND OCCUPATIONS. Government Printing Office, Washington, D.C.

10. Hoar SK. 1982. Meeting highlights: Job-exposure matrices in occupational epidemiology. JNCI 69: 1419-1420.

11. Hoar SK, Morrison AS, Cole P, Silverman DT. 1980. An occupation and exposure linkage system for the study of occupational carcinogenesis. J Occup Med 22: 722-726.

12. Hueper WC, Conway WD. 1964. CHEMICAL CARCINOGENESIS AND CANCERS. Charles C. Thomas, Springfield.

13. Gart JJ. 1970. Point and interval estimation of the common odds ratio in the combination of 2 X 2 tables with fixed marginals. Biometrika 57: 471-475.

14. Camitta BM, Storb R, Thomas ED. 1982. Aplastic anemia: Pathogenesis, diagnosis, treatment, and prognosis. N Engl J Med 306: 645-652.

15. Arnold R, Heimpel H. 1980. Aplastic anaemia after naproxen? Lancet 8163: 321.

16. Wang HH, Grufferman S. 1981. Aplastic anemia and occupational pesticide exposure: A case-control study. J Occup Med 23: 364-366.

17. Hoar S, Blair A. 1984. Death certificate case-control study of cancers of the prostate and colon and employment in the textile industry. Arch Environ Health 39: 280-283.

18. Buechley R, Dunn JE, Linden G, Breslow L. 1956. Death certificate statement of occupation: Its usefulness in comparing mortalities. Publ Health Rep 71: 1105-1011.

19. Blair A, Hoar SK, Walrath J. In press. Comparison of crude and smoking-adjusted standardized mortality ratios. J Occup Med.

Table 1. Number of deaths due to aplastic anemia among U.S. white men during 1975, controls, and odds ratios by usual industry.

Usual Industry (Census Code)	Cases	Controls	OR (95% CI)*
Public Administration (907-937)	18	96	1.0
Agriculture, forestry, and fisheries (017-028)	73	126	2.4 (1.3,4.7)
Mining (047-057)	4	27	0.6 (0.1,2.3)
Construction (067-077)	50	132	2.0 (1.0,4.0)
Manufacture of:			
Lumber and wood products, except furniture (107-109)	7	9	3.7 (1.0,14.6)
Stone, clay, and glass products (119-138)	1	9	0.4 (0.02,4.7)
Metal industries (139-169)	3	30	0.7 (0.1,3.0)
Machinery (177-209)	5	20	2.2 (0.5,8.6)
Transportation equipment (219-238)	8	29	1.9 (0.6,5.8)
Food and kindred products (268-298)	5	24	0.8 (0.2,2.9)
Textile and apparel (307-327)	5	17	1.3 (0.3,5.0)
Printing and publishing (338-339)	8	14	6.2 (1.6,25.4)
Chemicals (347-369)	2	8	1.9 (0.2,13.4)
Rubber and plastics (379-387)	1	9	0.9 (0.04,9.6)
Other products (118,239-259,298-299, 328-337, 377-378,388-397)	7	18	1.6 (0.5,5.7)
Not otherwise specified (398)	3	12	0.9 (0.2,4.4)
Transportation (407-429)	23	90	1.4 (0.6,2.9)
Communications (447-449)	4	10	2.1 (0.4,10.0)
Utility and sanitary services(467-479)	4	15	2.0 (0.4,9.4)
Wholesale trade (507-588)	7	39	1.0 (0.3,3.0)
Retail trade (607-698)	32	99	1.8 (0.9,3.8)
Finance, insurance, and real estate (707-718)	16	51	1.9 (0.8,4.5)
Business and repair services(727-759)	12	37	1.9 (0.7,4.9)
Personal services (769-798)	2	31	0.3 (0.04,1.5)
Entertainment and recreation services (807-809)	5	10	2.6 (0.6,10.8)
Professional and related services (828-897)	26	82	1.9 (0.9,4.0)

* Odds ratio (95% Confidence Interval), age-adjusted (<= 64 years, 65-74 years, 75+ years).

Table 2. Number of deaths due to aplastic anemia among U.S. white men during 1975, controls, and odds ratios by usual occupation.

Usual Occupation (Census Code)	Cases	Controls	OR (95% CI)*
Professional, technical, and kindred workers (001-195)	36	127	1.0
Managers and administrators, except farms (201-245)	53	174	1.1 (0.6,1.8)
Sales workers (260-280)	19	61	1.0 (0.5,2.0)
Clerical and kindred workers(301-395)	16	58	0.9 (0.4,1.9)
Craftsmen (401-580)	91	295	1.0 (0.6,1.6)
Operatives, except transport(601-695)	36	119	1.0 (0.6,1.8)
Transport equipment operatives (701-715)	12	66	0.7 (0.3,1.6)
Laborers, except farm (740-785)	33	106	1.2 (0.6,2.1)
Farmers and farm managers (801-802)	66	103	1.6 (0.9,2.8)
Farm laborers and farm foremen(821-4)	5	15	1.0 (0.3,3.5)
Service workers (901-965)	22	93	0.8 (0.4,1.5)

* Odds ratio (95% Confidence Interval), age-adjusted (<= 64 years, 65-74 years, 75+ years).

Table 3. Number of deaths due to aplastic anemia among U.S. white men during 1975, controls, and odds ratios according to ever exposure to a priori suspect chemicals or chemical classes.

Exposure	Cases	Controls	OR (95% CI)*
Acetone	6	17	0.9 (0.3,2.5)
Ammonium chloride	2	2	1.9 (0.2,19.4)
Benzene	82	248	1.0 (0.8,1.4)
Ethylene glycol	34	113	0.9 (0.6,1.4)
Carbon tetrachloride	22	60	1.1 (0.6,1.9)
Dyes	0	5	--
Mineral oil	75	340	0.7 (0.5,0.9)
Pesticides	93	210	1.2 (0.9,1.6)
Petroleum products	247	742	1.0 (0.8,1.3)
Naphtha	16	50	0.9 (0.5,1.7)
Titanium oxide	29	91	1.1 (0.7,1.7)
Toluene	17	48	1.1 (0.6,2.0)
Trinitrotoluene	0	7	--
Wood dust	45	85	1.7 (1.1,2.6)
Wood preservatives (Creosote)	178	492	1.1 (0.9,1.4)

* Odds ratio (95% Confidence Interval), age-adjusted (<= 64 years, 65-74 years, 75+ years).

Table 4. Number of deaths due to aplastic anemia, controls, and odds ratios according to ever exposed to selected chemicals, by degree of exposure for all subjects and for autopsied subjects only.

	ALL SUBJECTS			AUTOPSIED SUBJECTS		
Exposure	Cases	Controls	OR (95% CI)*	Cases	Controls	OR (95% CI)
Pesticides						
Never	371	1248	1.0	57	178	1.0
Low	1	4	0.6 (0.03,6.8)	0	0	--
Moderate	76	132	1.5 (1.1,2.1)	4	5	2.5 (0.5,12.4)
High	16	74	0.7 (0.4,1.2)	4	18	0.5 (0.1,1.9)
Petroleum products						
Never	217	716	1.0	33	109	1.0
Low	28	140	0.7 (0.4,1.1)	6	25	0.9 (0.3,2.9)
Moderate	125	270	1.3 (1.0,1.8)	14	21	2.3 (0.9,5.6)
High	94	332	0.9 (0.7,1.2)	12	46	0.9 (0.4,2.0)
Wood dust						
Never	419	1373	1.0	57	190	1.0
Low	0	5	--	0	2	--
Moderate	6	3	7.0 (1.4,37.7)	2	0	--
High	39	77	1.6 (1.0,2.5)	6	9	2.4 (0.7,8.4)
Wood preservatives						
Never	286	966	1.0	42	148	1.0
Low	44	197	0.8 (0.5,1.2)	11	33	1.1 (0.5,2.7)
Moderate -	113	213	1.4 (1.1,1.9)	10	14	2.7 (1.0,7.6)
High	21	82	0.8 (0.5,1.4)	2	6	1.1 (0.1,7.2)

* Odds ratio (95% Confidence Interval), age-adjusted (<= 64 years, 65-74 years, 75+ years).

THE IMPACT OF OCCUPATIONAL EXPOSURE TO TOXIC MATERIAL ON PREVALENCE OF CHRONIC ILLNESS

Richard G. Frank, The Johns Hopkins University, Mark S. Kamlet and Steven Klepper, Carnegie-Mellon University

1. Introduction

This paper reports the results of a preliminary analysis of the effects on health of exposure to occupational pollution.[1] To convey the nature of our analysis, we report the results of a preliminary analysis of one occupational pollutant, carbon monoxide (CO), on the prevalence of one specific chronic condition, cardiovascular illness. These results are representative of the kinds of results we have been getting for other occupational pollutants (we have looked at solvents, lead, cadmium, and benzene) and other chronic conditions.

In order to measure occupational pollution exposure, the literature on occupational epidemiology relies heavily on the classification of individuals into occupational groups, such as by craft or type of job performed. Mortality and morbidity differences across groups are compared and significant group differences are studied to indentify possible toxic agents (for example, see Gamble, Spirtas, and Easter, 1976).

A weakness of this approach is that it does not allow for quantification or direct assessment of the effects of specific exposures on health. An alternative approach is to link occupations to exposures on the basis of expert opinion concerning exposures in various occupations and industries (Hoar et al., 1980). In our analysis we take this type of approach one step further. We link the National Occupational Hazards Survey (OHS) with the 1980 National Health Interview Survey (HIS) and the 1984 Area Resource File (ARF). The OHS provides information on exposure to toxic materials by industry and occupation. The 1980 HIS and its supplements contain information about health, occupational history, health-related habits, and various demographic and economic variables for a large sample of individuals. The ARF provides information by locality for various climatic and economic variables. By linking these information bases together, we are able to construct a data set for a large sample of individuals with both measures of health and occupational pollution exposure as well as many other health relevant variables. While this approach has its own distinctive limitations, as we point out below, it does enable us to estimate a relationship between health and occupational pollution exposures directly, without resort to expert judgments or any other facsimile for direct measurement of pollution exposure. One of the purposes of our analysis is to indicate how our approach can best be implemented and what steps might be taken to make it more useful.

The paper is organized as follows. In Section 2 we briefly describe our analytical framework. This is followed by a discussion of data and measurement issues in Section 3. In Section 4 we present our empirical results. We conclude with a brief summary and caveats in Section 5.

2. Methodology

The theoretical point of departure of our analysis is Grossman's (1972) notion that health is both demanded and produced by individuals. Health is demanded because it provides utility. It is produced in the sense that it is affected by the actions of individuals.

To capture these notions, we first specify a "health-production" function:

$$(1) \qquad y_1 = \sum_{j=2}^{J} \alpha_j y_j + \sum_{i=1}^{K} \beta_i x_i + \epsilon,$$

where y_1 is a measure of health, $y_2, y_3, \ldots, y_J$ are endogenous factors under individual control that affect health, $x_1, x_2, \ldots, x_K$ are exogenous factors that affect health, ϵ is a classical disturbance representing the influence of all other (unobservable) factors that affect health, and $\alpha_2, \alpha_3, \ldots, \alpha_J$ and $\beta_1, \beta_2, \ldots, \beta_K$ are coefficients. Examples of the y-variables are the quantity of medical care, quantity of exercise, and quantity of smoking. Examples of the x-variables are age, race, and sex. Depending on how it is interpreted, occupational pollution exposure might be classified either as an endogenous or exogenous variable. We return to consider this later.

Separate demand equations are specified for each of the inputs $y_2, y_3, \ldots, y_J$ in the health-production function:

$$(2) \qquad y_j = \sum_{l \neq j}^{J} \theta_{jl} y_l + \sum_{n=1}^{N} \gamma_{jn} z_n + \mu_j, \quad j=2,3,\ldots,J,$$

where $z_1, z_2, \ldots, z_N$ are N exogenous variables that affect the quantity chosen by the individual for one or more of the inputs $y_2, y_3, \ldots, y_J$, μ_j is a classical disturbance pertaining to input j, and θ_{jl}, j=2,3,...,J, l=1,2,...,J, and γ_{jn}, j=2,3,...,J, n=1,2,...,N are coefficients. Equation (2) specifies that the amount chosen for each input y_j, j=2,3,...,J, depends upon health, y_1, the amounts of the other inputs, and the N exogenous variables $z_1, z_2, \ldots, z_N$. Examples of the z-variables are income, sex, and the prices of the various inputs (e.g., the price of medical care and the price of cigarettes). Note that the N z-variables do not necessarily enter each equation, in which case some of the γ_{jn} are constrained to equal zero.

Ideally, we would like to estimate the coefficients of equation (1), and in particular the coefficient of the CO variable. One approach to estimate these coefficients is simply to regress y_1 on $y_2, y_3, \ldots, y_J, x_1, x_2, \ldots, x_K$. However, because each of the inputs $y_2, y_3, \ldots, y_J$ not only influences health, y_1, but in turn is determined by health (according to equation (2)), the regression of y_1 on $y_2, y_3, \ldots, y_J, x_1, x_2, \ldots, x_K$ will not yield unbiased estimates of the α_j, j=2,3,...,J or the β_i, i=1,2,...,K.

We proceed by deriving an equation that can be estimated using regression analysis. Using (2) to substitute for $y_2, y_3, \ldots, y_J$ in (1), we can solve out for y_1 in terms of the exogenous variables $x_1, x_2, \ldots, x_K, z_1, z_2, \ldots, z_N$:

$$(3) \qquad y_1 = \sum_{i=1}^{K} \lambda_i x_i + \sum_{i=1}^{N} \kappa_i z_i + \omega,$$

where the λ_i and κ_i are functions of the coefficients of both the health-production function and the various input demand equations, and ω is a classical disturbance.

Equation (3) is called the reduced-form equation for health. Since $x_1,x_2,...,x_K$ and $z_1,z_2,...,z_N$ are all exogenous variables, the regression of y_1 on $x_1,x_2,...,x_K$ and $z_1,z_2,...,z_N$ will yield unbiased estimates of the λ_i and κ_i.

In our empirical work, we assume that besides health, only medical care is endogenous. In effect, we assume that smoking, exercise, diet, and occupational pollution exposure, as well as other factors under individual control, are determined independently of health. Thus, the only y-variables in the model are health, y_1, and medical care, y_2. In this case, it is easy to solve out for the λ_i and κ_i in (3) in terms of the coefficients of (1) and (2):

$$\lambda_i = \beta_i/(1-\alpha_2\theta_1)$$

$$\kappa_i = \alpha_2\gamma_i/(1-\alpha_2\theta_1),$$

where α_2 is the coefficient of y_2 in (1) and θ_1 is the coefficient of y_1 in the y_2 equation as defined by (2) (the j subscript on θ_{1j} in (2) has been deleted since there is only one input equation). Assuming that $\alpha_2 > 0$ (i.e., more medical care contributes to better health) and $\theta_1 < 0$ (i.e., better health leads to less medical care) then the λ_i will be the same sign as the β_i, but smaller in absolute value, while the κ_i will be the same sign as the γ_i, with the relative magnitude of κ_i and γ_i depending on the magnitudes of α_2 and θ_1. The implications for estimating the effect of CO on health is that the regression of y on $x_1,x_2,...,x_K$ and $z_1,z_2,...,z_N$ will yield an attenuated estimate of the "true" effect of CO on health (as defined by the health-production function).

It can be shown that estimates of the coefficients of (3) can be used to generate estimates of the coefficients defined by (1) if there is at least one z-variable in the demand equation for medical care which does not enter the health-production function directly (i.e., which is not also an x-variable). The most promising candidate is the price of medical care, which should affect the quantity of medical care but not directly affect health. For cardiovascular disease, the relevant price would be some composite of the price of specialized physician care and hospital care. The most important determinant of this price for individuals is whether the individual is insured and the nature of the insurance for those who are insured. Unfortunately, we have no information on the latter and nearly all the individuals in our sample are insured. Consequently, estimates of the coefficients of (1) derived from the reduced-form coefficient estimates would not be very reliable. As a result, we concentrate our analysis on the reduced-form coefficients.

3. Data

As indicated above, the data used in the analysis come from three primary sources: the 1980 HIS, the OHS, and the 1984 ARF. The HIS is conducted yearly by the National Center for Health Statistics. It is a stratified, cluster sample of 35,000 households comprised of some 100,000 individuals representing the noninstitutionalized civilian population of the United States. Self-reported and proxy data are collected on a variety of health outcomes and other individual characteristics. Each year different supplemental sets of questions are asked. In 1980, supplemental questions were asked about smoking and work history.

The OHS was conducted between 1972 and 1975 by the National Institute of Occupational Safety and Health with the assistance of the Bureau of Labor Statistics. It is a stratified, cluster sample of nonagricultural businesses. The Bureau of Labor Statistics selected 5,200 facilities in 67 SMSA's covering a wide range of Standard Industrial Classification (SIC) codes. Twenty engineers were hired to measure pollution exposures in the 5,200 facilities. Each engineer went through a nine-week course in fundamental industrial hygiene and then was trained in field-gathering techniques. In inspecting each facility, an engineer observed every plant process and every employee, recording specific exposures. The engineer catalogued all materials utilized for more than 30 minutes per week or three full eight hour days per year. The number of individuals exposed to greater than a threshold level of each pollutant examined was recorded. For carbon monoxide, the number of individuals subject to four hours per day of detectable continuous exposure was reported.

The 1984 ARF is compiled by the Health Resources Administration. It contains county-level information on a variety of health-related economic and weather variables. It was used to measure humidity, temperature, and medical care prices. The HIS, OHS, and ARF data sets were integrated on the basis of the HIS primary sampling unit for the ARF data and the HIS occupational information for the OHS data.

Table 1 gives definitions and acronyms for the explanatory variables in our analysis. With the exception of the CO variable, all the variables are either measured in their natural units (e.g., age in years) or as a dummy variable. The CO variable was constructed as follows. The OHS provides estimates of the total number of workers in the U.S. in each three-digit SIC occupation who were exposed to more than the threshold level of CO. For each occupation we divided this by the total number of individuals in the occupation in the U.S. using the 1970 Census Bureau Subject Reports on Occupational Characteristics. For each occupation, this yields the probability of exposure to above the threshold level of CO for individuals in that occupation.

Results from animal studies of exposure to CO indicate that health abnormalities result after long-term exposure to CO (USEPA, 1979, pp. 10-27). This suggests using a disease model for the effect of CO exposure on cardiovascular illness based on cumulative exposure with little occurrence of repair. Accordingly, a CO exposure measure for each worker was constructed by multiplying the probability of exposure in the occupation held for the longest time by the length of time spent in that occupation.[2]

Our outcome, or dependent, variable in the health-production function is a dichotomous variable assuming the value one if an individual reports a chronic cardiovascular condition and zero otherwise. A cardiovascular condition includes coronary heart disease (ischemic with hypertension and with arteriosclerosis) and hypertensive heart disease.

4. Empirical Results

The empirical analysis was performed on a subsample

of the HIS. The subsample was selected according to the following criteria:

1. Subjects must have been administered the smoking supplement to the HIS. Thus individuals for whom smoking information was not available were eliminated from the sample. Since subjects were chosen randomly for the smoking supplement, the included subjects are a random subsample of the HIS.

2. Subjects must have had some work history outside the home and the agricultural sector. This resulted in a larger fraction of children and housekeepers (most often women) being eliminated from the sample than adult males.

3. Data had to be available for all the relevant variables.

The selection rules resulted in a sample of 10,872 subjects. Table 2 reports a cross tabulation comparing the prevalence of cardiovascular disease of workers who had a nonnegligible probability of being exposed to above the threshold level of CO with those who had no probability of being exposed (in a number of occupations the OHS indicates that no workers were exposed to more than the threshold level of CO). The comparison indicates that 3.2 percent of those exposed experienced chronic cardiovascular disease while 2.9 percent of those who were not exposed experienced chronic cardiovascular disease. While the direction of the association is consistent with the hypothesis that CO exposure leads to cardiovascular disease, the difference in rates of prevalence in the two populations is not large and it is not statistically significant at conventional levels.

To take account of other factors besides CO that may affect health (and that possibly are correlated with CO), we estimated both a linear probability and a logistic regression model (the latter was used because our dependent variable is dichotomous). The explanatory variables in our analyses included: indices of obesity (the ratio of weight to height and the squared value of this ratio), income, sex, race, age, age squared, smoking behavior (whether the individual had ever smoked and whether the individual was currently a smoker), marital status, schooling, the price of healthcare, temperature, geographic region, and humidity. Exposure to CO was entered as both a direct effect and interacted with smoking. The interaction effect allows for the possibility that CO affects the health of smokers more than nonsmokers.

The results for the full model are reported in Table 3. In general the coefficient estimates for the non-CO variables conform to expectations. However, only four coefficient estimates are significant at the 90% confidence level in the logit analysis and two in the linear probability regression. In the logit model the coefficient estimate for income is negative and significant (at the .01 level), implying that higher income individuals have a lower prevalence of cardiovascular disease, holding all other factors constant. However, the income coefficient estimate is positive in the linear probability model, although insignificant (at any conventional significance level). The coefficient estimates for both smoking variables are positive in both the logit and linear probability models, and the estimates are significant at the .05 level in three of four instances. Thus, smokers appear to have a higher prevalence of cardiovascular disease, particularly relative to individuals who have never smoked. The coefficient estimates for age are positive in both the logit and linear probability models and statistically significant at the .10 level in the logit, indicating, as expected, that cardiovascular disease is more prevalent as age increases. Finally, the coefficient estimates for sex differ in sign for the two models. In the one instance in which the estimate is significant at the .05 level, it indicates that males have a higher prevalence of cardiovascular disease.

The coefficient estimate for the CO variable, which measures the direct effect of cumulative exposure to CO, was negative in both models, although insignificant in both instances. The interaction effect of smoking with CO exposure was estimated to contribute toward a higher prevalence of cardiovascular illness, although the estimate was insignificant in both models. Thus, our results suggest no reduced-form direct effect between cardiovascular illness and CO exposure and a weak, if any, reduced-form indirect effect between cardiovascular illness and CO exposure in combination with smoking.

These results are representative of the results we have found in preliminary analyses of the effects of other occupational pollutants on other chronic health problems. For instance, we have studied the effect of lead and solvents exposure on chronic neurological conditions and the effect of benzene exposure on blood disorders. In each case we found some evidence of health effects in simple descriptive analyses (although weak and statistically insignificant) that vanished when additional explanatory variables were added to the analysis.

5. Discussion

Our results concerning the effects of CO exposure are consistent with the epidemiological literature insofar as the epidemiological literature is quite uncertain concerning the effect of CO on cardiovascular illness. Studies by Kuller et al. (1975) and Radford and Weisfeldt (1975) both failed to obtain clear associations between ambient CO levels or long-term CO exposures and heart disease.

However, before reaching any final conclusions, a number of limitations of our analysis should be noted. First, our estimates are only for the coefficients of the reduced-form equation (3). As we noted earlier, these coefficients will be smaller in absolute value than the coefficients of the health-production function, thus understating the true effects of the variables in (1).[3] Second, our pollution measure is *not* a true measure of *individual* CO exposure, but rather of expected or average exposure of individuals in each occupation. As we demonstrate elsewhere in this volume (Kamlet, Klepper, and Frank, 1985), this will bias our estimates of the effects of CO exposure on health. Third, to some extent CO, as well as other pollution exposures, may be endogenous. If in fact individuals who suffer some kind of cardiovascular illness, or who have a greater probability of contracting a cardiovascular illness, choose jobs with less CO exposure, then our estimates will understate the true effects of CO exposure on

health (they might even be the wrong sign). Finally, we are not able to measure very well a number of factors that may affect health that may be correlated with occupation. For example, if the type of individual who is more prone, perhaps genetically or environmentally, to cardiovascular disease is more likely to pursue say office jobs where CO exposure is low, and this is not controlled for, then our estimates will understate the true effect of CO exposure on health. Similar difficulties arise if the level of exposure to CO in settings other than occupation is correlated with the level of occupational exposure.

These problems are in large degree shared by other occupational epidemiology studies. The important thing to realize, however, is that they are not intractable. Consider first the fact that our exposure measure is not really a true measure of individual CO exposure. Kamlet, Klepper, and Frank (1985) indicate that even with the limited information we possess, it may be possible to estimate the effect of *individual* CO exposure on health. While the approach outlined in Kamlet, Klepper, and Frank is not entirely applicable to our analysis here (because the CO variables are composites of CO exposure and other variables--e.g., length of longest job), we are currently working on adapting this approach to our problem.

The other problems with our analysis cannot be solved without additional information. One advantage of our approach is that it indicates the kind of additional information needed to overcome existing estimation difficulties. In particular, information about input prices, early medical and health experiences, and motivations for occupational choices would all be helpful. Such information could conceivably be compiled in subsequent HIS surveys.

While our approach clearly confronts a variety of difficulties in estimating the effects of occupational pollution on health, we feel it remains revealing and valuable. Our current findings suggest the absence of a strong link between occupational exposure to a number of important occupational pollutants and health. Of course, these results may change as certain limitations of our analysis are overcome. But for now we simply note that even using a large sample and actual exposure measures across occupations we are not able to detect much, if any, effect of occupational exposure to CO and the other occupational pollutants we examined on chronic illness.

Notes

[1]We thank Herb Needleman and Usha Sambamoorthi for their helpful suggestions. The work reported here was funded by the Environmental Protection Agency Cooperative Agreement CR811041.

[2]Ideally, the CO exposure measure should sum exposure times length of job for all jobs held by the individual. However, the HIS reports length of job for only the current, last, and longest job held which, for the majority of individuals, were one and the same.

[3]An extreme instance of this is where medical care is used to offset fully the negative health effects of CO exposure. In this case there would be no relationship estimated between health and CO exposure in the reduced-form equation even though CO exposure does lead to negative health effects.

TABLE 1

Descriptive Statistics

Variables	Mean	Standard Deviation
AGE	42.550	17.952
AGE^2	2132.742	1703.094
FAT	2.475	1.440
FAT^2	8.200	23.197
SEX (Male Dummy)	0.478	0.500
MARIT (Nonmarried Dummy)	0.363	0.481
RACE (White Dummy)	0.877	0.329
SCHOOL	11.823	3.386
INCOME	8744.113	8867.376
NWEST (Nonwestern Regional Dummy)	0.815	0.388
AVHUM (Average Humidity)	37.2	30.04
CURSMOKER	0.305	0.461
OCSMOKER (Occasional Smoker Dummy)	0.021	0.143
FORSMOKER (Former Smoker Dummy)	0.186	0.389
UNSMO (Smoking Status Unknown Dummy)	0.0034	0.058
PRICE	4.66	4.78
ATEMP (Average Temperature)	34.17	(27.58)
CO (Probability of CO exposure times length of longest job)	0.0074	0.069
CO x CURSMOKER	0.0032	0.036

N = 10,872

TABLE 2

Cross Tabulation: Carbon Monoxide Exposure with Chronic Cardivascular Illness

	Not Exposed to CO	Exposed to CO
Number of Workers	8,051	2,821
With Chronic Illness	233	91
Rate	2.9%	3.2%

TABLE 3

Multivariate Models Results

Variable	Logit	Linear Prob
CONSTANT	-5.15*	0.0035
	(0.77)	(0.022)
INCOME	-0.00020*	0.10 (10^{-6})
	(0.00001)	(0.17)(10^{-6})
SEX	-0.088	0.0068*
	(0.130)	(0.0033)
FORSMOKER	0.29*	0.0042
	(0.15)	(0.0046)
CURSMOKER	0.333*	0.0056*
	(0.140)	(0.0036)
CO	-0.130	-0.0167
	(2.760)	(0.0190)
CO x CURSMOKER	0.61	0.39
	(2.40)	(0.72)
AGE	0.03*	0.06
	(0.0018)	(0.12)
AGE^2	-0.000043	0.0000090
	(0.00018)	(0.0000060)
AVHUM	0.00069	0.000017
	(0.0058)	(0.00016)
FAT	2.72	0.079.
	(2.73)	(0.077)
FAT^2	-3.70	-0.107
	(93.26)	(0.092)
NWEST	0.034	0.00073
	(0.14)	(0.0042)
MARIT	-0.019	-0.00066
	(0.13)	(0.0037)
RACE	0.057	0.0015
	(0.18)	(0.0050)
OCSMOKER	-19.36	-0.013
	(3510)	(0.011)
UNSMO	-19.4	-0.026
	(3529)	(0.027)
SCHOOL	-0.0099	-0.00030
	(0.017)	(0.00054)
PRICE	-0.018	-0.00051
	(0.017)	(0.00053)
ATEMP	0.00047	0.000016
	(0.0065)	(0.00018)

Standard error in parentheses.
*Significant at a 90% confidence level.

References

Gamble, J.F., Spirtas, R., and Easter, P. (1976). "Applications of a Job Classification System in Occupational Epidemiology," American Journal of Public Health, 66 (8): 768-774.

Grossman, M. (1972). The Demand for Health: A Theoretical and Empirical Investigation. New York: Columbia University Press for NBER.

Hoar, S., et al. (1980). "An Occupation and Exposure Linkage System for the Study of Occupational Carcinogensis," Journal of Occupational Medicine 22 (11): 722-726.

Kamlet, M., Klepper, S., and Frank R. (1985). "Mixing Micro and Macro Data: Statistical Issues and Implications for Data Collection and Reporting," presented at the 1985 Public Health Conference on Records and Statistics.

Kuller, L.H., et al. (1975). "Carbon Monoxide and Heart Attacks," Archives of Environmental Health 30: 477-482.

Radford, E.P., and Weisfeldt, M.L. (1975). Final Report of the Study of the Relationship Between Carboxyhemoglobin on Admission to the Subsequent Hospital Course of Patients Admitted to the Myocardial Infection Unit at the Johns Hopkins Hospital.

U.S. Environmental Protection Agency, Air Quality Criteria for Carbon Monoxide. Washington, D.C.: U.S. Government Printing Office, EPA-600/8-79-022.

EFFECT OF ARSENIC EMISSIONS ON INCIDENCE OF CONGENITAL FACIAL CLEFTS

Mary K. O'Mara
Floyd Frost
Washington State Division of Health

ABSTRACT

Association between incidence of congenital facial clefts and maternal exposure to arsenic emissions from a copper smelter located in Tacoma, Washington, was investigated. The relative risk of facial clefts in live births from 1979 through 1981 was analyzed by level of maternal exposure to arsenic emissions, as determined by proximity of mother's residence to the smelter stack at time of child's birth. Selected maternal and demographic characteristics were examined to assess the similarity of exposed and nonexposed population groups.

No statistically significant relationship between incidence of facial clefts and maternal exposure to arsenic emissions from the smelter was observed. However, significantly more previous fetal deaths among exposed women were observed. This may suggest a relationship between arsenic exposure and fetal death.

INTRODUCTION

In Washington State, the possibility of adverse health effects from the emissions from Tacoma's American Smelting and Refining Company (ASARCO) copper smelter has been a subject of citizen concern for several years. The Washington State Division of Health, Epidemiology Section, has been involved in assessing public health risks associated with smelter emissions since the early 1970's.

ASARCO processes a copper ore known to contain a high level of arsenic. According to recent estimates, the smelter releases approximately 100 tons of arsenic trioxide into the atmosphere each year as a byproduct of their copper smelting activities. (1) In high doses (exposure received by smelter workers having direct contact with arsenic) arsenic is known to produce excess risk of respiratory cancer, skin diseases, and adverse reproductive outcomes. (2-4) At low doses (exposure common to persons living in the area surrounding the smelter) the risk of adverse health effects is less clear.

The purpose of the current study was to determine if an excess risk of oral clefts exists among the offspring of women exposed to arsenic emitted from the ASARCO smelter. It was anticipated that the study findings would provide a basis for determining whether further investigation into arsenic-related malformations was warranted. The decision to examine the incidence and relative risk of oral clefts, as opposed to other possible malformations, was predicated on the findings of epidemiologic studies that examined health effects from a copper smelter located in Sweden. The findings of these studies suggest the presence of a dose-response relationship between arsenic exposure and the incidence of cleft lip/palate conditions in the offspring of smelter workers and community dwellers. (4-7) Data from animal studies have also implicated arsenic as a potentially genotoxic substance. (8,9)

MATERIALS AND METHODS

In this retrospective cohort study, the study population was comprised of all live births to residents of a four county area surrounding the smelter location in Washington State (i.e., Pierce, King, Snohomish, and Kitsap Counties) during the years 1979 through 1981.

Cases were children born alive with a cleft lip and/or cleft palate condition. Cases were identified through birth certificate information filed with the Washington State Office of Vital Statistics and hospital discharge records from the hospitals known to perform cleft lip and palate surgery for residents of the study area (i.e., Children's Orthopedic, Mary Bridge, Harrison Memorial, Providence of Everett, and Madigan Hospitals).

Case identification was accomplished by reviewing birth certificates for children born to residents of the study region during the years 1979 through 1981 where a cleft lip and/or cleft palate (ICD-9 749) was reported on the birth certificate, and by reviewing hospital charts of children born to residents of the study area between 1979 and 1981 who had been admitted for surgical repair of a congenital cleft lip/cleft palate condition during the period 1979 through 1983.

Review of hospital discharge records permitted identification of cases where the cleft lip/palate condition was not reported on the birth certificate. The inclusion of hospital data provided additional assurance of complete case ascertainment because (a) cleft conditions are always repaired except in rare cases when this anomaly is accompanied by other severe malformations which make the prognosis for survival extremely poor, (b) oral clefts are always repaired at one of the five study hospitals, and (c) the vast majority of oral cleft repairs are performed within the first two years following birth.

A computerized data record including all data items reported on the Washington State birth certificate was developed for each member of the study population. Following this step, all births in the study population were assigned to one of three exposure groups, based on proximity of mother's residence to the ASARCO smelter stack at time of child's birth. Census tract of residence, as reported on the birth certificate, was used to determine proximity to the smelter. Exposure group one included births

to mothers having a residence equal to or less than three miles from the smelter stack. Exposure group two included births to mothers residing between three and five miles from the smelter stack. Urine studies and meteorlogic data suggest that low exposure to arsenic is common among persons living within five miles of the smelter (10) and that exposure level diminishes with increasing distance from the smelter stack. (Table 1) Exposure group three included births to mothers having a residence outside of a five mile radius from the smelter stack but inside the county boundaries of Pierce, King, Snohomish or Kitsap Counties. Area residents living outside of a five mile radius from the smelter stack are considered to be unexposed to arsenic from smelter emissions.

The exposed and nonexposed populations and cases were compared with respect to maternal age and parity, previous fetal deaths, mobility (i.e., frequency of change in residence), socioeconomic status, family history of oral clefts, and fertility rate among the age 15 to 44 female population. Maternal characteristics examined were those known to affect the incidence of congenital malformations generally. (11) Demographic characteristics examined (e.g., mobility) were those having potential to produce misclassification of cases. Malformation rates for exposed and nonexposesd populations were compared by determining the relative risk of oral clefts per number of live births among mothers residing in exposed areas to mothers residing in nonexposed areas.

RESULTS

Using birth certificates and hospital discharge records, a total of 117 cases of cleft lip and/or cleft palate were identified. Thirty-five cases (29.9%) were identified through birth certificates only, 38 (32.5%) were identified through hospital records only, and 44 (37.6%) were reported through both birth certificates and hospital records.

Five of the cases were born to women who resided within five miles of the smelter at the time of the child's birth (i.e., exposure areas one and two). The remaining 112 cases were born to women who resided in the nonexposed area (i.e., exposure area 3). (Table 2)

Incidence rates for cleft lip and/or cleft palate conditions for each exposure group were calculated by dividing the total number of cases in each exposure group by the total number of births in each group. In calculating the relative risk by exposure area, exposure groups one and two were combined because of the small number of observed cases.

The relative risk, based on figures shown in Table 2, was 1.19, with a 95 percent confidence interval of 0.490 to 2.931. Using a Chi square test it was found that the excess risk was not statistically significant ($p = .9261$). In order for the excess to have been significant, the relative risk would have needed to be 2.25 or higher. This means that to produce a significant relative risk the observed number of cases among the exposed population would have had to nearly double (i.e., there would have had to have been nine cases instead of five).

Comparison of exposed and nonexposed populations on selected demographic and maternal characteristics showed that the groups were similar with respect to parity. The groups differed with respect to maternal age, previous fetal deaths, fertility rate, and stability of mother's residence. (Table 3) The exposed group tended to have more very young mothers and fewer mothers in the upper end of the age distribution and their length of residence in current home (according to 1980 U.S. census reports) tended to be shorter than their nonexposed counterparts. (12) Both the fertility rate and the number of previous fetal deaths were significantly higher for women in the exposed area.

DISCUSSION

As previously noted, this study was undertaken to provide a basis for determining whether further investigation into arsenic-related malformations was warranted. The study findings do not support a need for further investigation, as there appears to be no significant risk of developing oral clefts as a result of arsenic exposure. The observed relative risk (1.19) was far below what was required for statistical significance (2.25).

The observed differences between exposed and nonexposed populations with respect to maternal age, previous fetal deaths, and fertility rate do not alter the value of the findings since they do not appear to be confounders (i.e., factors associated with the incidence of oral clefts. Of some concern, however, is the observed mobility of the study population and the resulting potential for misclassification. In this study, residence at time of delivery was assumed to be the same as residence at time of conception. If this assumption is in error, the chances of showing an effect are reduced, since it is known that the time of critical exposure for oral clefts is the first eight weeks of pregnancy. (13) It was not possible to account for pregnant women moving out of the exposure area after conception but prior to delivery. U.S. census data for 1980 indicates that women in both the exposed and nonexposed areas move frequently, but that length of residence in the exposed area tends to be shorter. However, in order for mobility to have significantly altered the study findings (i.e., obscured a significant excess risk in the exposed population), nearly 50 percent of the cases in the exposed area would have had to move out of the area, thus being lost to follow-up. Based on the limited evidence that is available, this would not appear likely.

Also of some interest was the observed difference in previous fetal deaths among the exposed and nonexposed populations. Given that women in the exposed area tended to be younger than in the nonexposed area and that the populations in the two areas were similar with

respect to parity, one might expect that, if anything, the population in the exposed area would show somewhat fewer fetal deaths (barring any unexpected exposure). The fact that women in the exposed area had significantly more previous fetal deaths may indicate that arsenic is having an effect on pregnancies, but that the effect cannot be detected by measuring differences among malformation rates for live births. Similarly, the data on fertility rates for exposed and nonexposed women would support the theory that if arsenic is having an effect on reproduction, the effect is occurring after conception since women in the exposed area had a significantly higher fertility rate than their nonexposed counterparts. The observed higher rate of spontaneous abortions in the exposed population is consistent with findings from Swedish studies that examined the incidence of spontaneous abortions among women exposed to arsenic from a copper smelter. (14,15) Exploring fetal deaths and malformations among fetal deaths for women exposed to arsenic might prove a fruitful area of research.

TABLE 1

Mean Annual Arsenic Levels Recorded at Measuring Sites Within Five Miles of ASARCO Smelter Stack, By Proximity of Site to Smelter Stack, 1982

MEASURING SITE	$\bar{X}$ ANNUAL ARSENIC LEVEL (ugm/m3)*	DISTANCE FROM STACK (IN MILES)
Smelter Stack	1.3	0.13
Smelter Parking Lot	0.8	0.23
Site 3	0.6	0.40
Site 4	0.3	0.47
Site 5	0.2	1.50
Site 6	0.2	1.90

*Reported arsenic levels reflect both azimuth and distance from stack in miles, with sites located due north of the smelter reporting higher exposure.

TABLE 2

Relationship Between Arsenic Exposure and Incidence of Cleft Lip/Palate Conditions+

	ARSENIC EXPOSURE		
CLEFT LIP/PALATE CONDITION	EXPOSED*	NONEXPOSED*	TOTAL
Clefts	5 (0.1%)	112 (0.1%)	117 (0.1%)
No Clefts	5,208 (99.9%)	136,156 (99.9%)	141,364 (99.9%)
Total	5,213 (100.0%)	136,268 (100.0%)	141,481 (100.0%)

*Exposure Areas one and two; Exposure Area three is considered not exposed.

+Relative risk = Incidence in exposed group/Incidence in nonexposed group (5/5,213 112/136,268 = 1.19) (95% C.I. = 0.490 - 2.931).

TABLE 3

Comparison of Exposed and Nonexposed Populations on Maternal and Demographic Characteristics

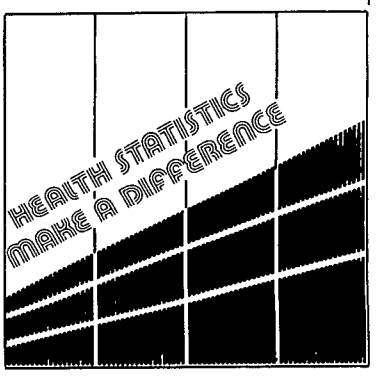
VARIABLE	EXPOSED	NONEXPOSED	SIGNIFICANCE p (Chi Square)
Parity (% with 1+ children)	53.7%	54.5%	0.379
Fertility Rate	95/1000 pop.	86/1000 pop.	0.000
Maternal Age (% under 23 yrs)	33.8%	27.9%	0.000
Fetal Deaths (% with 1+ fetal deaths)	21.0%	19.1%	0.001
Length of Residence (% less than 5 yrs)	57.0%	44.3%	0.000

REFERENCES

1. EPA News Release, October 20, 1983.
2. Pinto, S.S.; Enterline, P.E.; Henderson, V. & Varner, M.O. (1977) Mortality experience in relation to a measured arsenic trioxide exposure. Environ. Health Perspect. 19:127-130.
3. Perry, K.; Bowler, R.G.; Buckell, H.M.; Druett, H.A. & Shilling, R.S.F. (1948) Studies in the incidence of cancer in a factory handling inorganic compounds of arsenic. II. Clinical and environmental investigations. Brit. J. Ind. Med. 5:6-15.
4. Beckman, L. (1978) The Ronnskar smelter. Occupation and environmental effects in and around a polluting industry in northern Sweden. Ambio 7:226-231.
5. Beckman, L.; Myberg, N. (1972) The incidence of cleft lip and palate in northern Sweden. Hum. Hered. 22:417-422.
6. Beckman, L.; Nordstrom, M. (1976) Population studies in northern Sweden. VIII. Frequencies of congenital malformations by region, time, sex, and maternal age. Hereditas 84:35-40.
7. Nordstrom, S.; Beckman, L. & Nordenson, I. (1979) Occupational and environmental risks in and around a smelter in northern Sweden. VI. Congenital malformations. Hereditas 90:297-302.
8. Hood, R.D. (1972) Effects of sodium arsenite on fetal development. Bull. Environ. Contam. Toxicol. 7:216-222.
9. Ferm, V.H. (1977) Arsenic as a teratogenic agent. Environ. Health Perspect., 19:215-217.
10. Milham, S., Jr. & Strong, T. (1974) Human arsenic exposure in relation to a copper smelter. Environ. Res. 7:176-182.
11. Janerich, D.W. & Polednak A.P. (1983) Epidemiology of Birth Defects. Epidemiologic Reviews 5:16-37.
12. U.S. Census Bureau (1983) 1980 Census of Population and Housing; Census Tracts Washington Selected Areas. (U.S. Census Bureau Document Number PHC 80-2-49).
13. Nanda, R. (1975) Teratogenic Effects of Environmental Agents on Embryonic Development. Dent. Clin. North Am. 19(1):181-189.
14. Nordstrom, S.; Beckman, L. & Nordenson I. (1978) Occupational and environmental risks in and around a smelter in northern Sweden. III. Frequencies of spontaneous abortion. Hereditas 88:51-54.
15. Nordstrom, S.; Beckman, L. & Nordenson I. (1979) Occupational and environmental risks in and around a smelter in northern Sweden. V. Spontaneous abortion among female employees and decreased birthweight in offspring. Hereditas 90:291-296.

Session D

Statistical Data Collection, Management and Analysis: A Sample of Microcomputer Applications

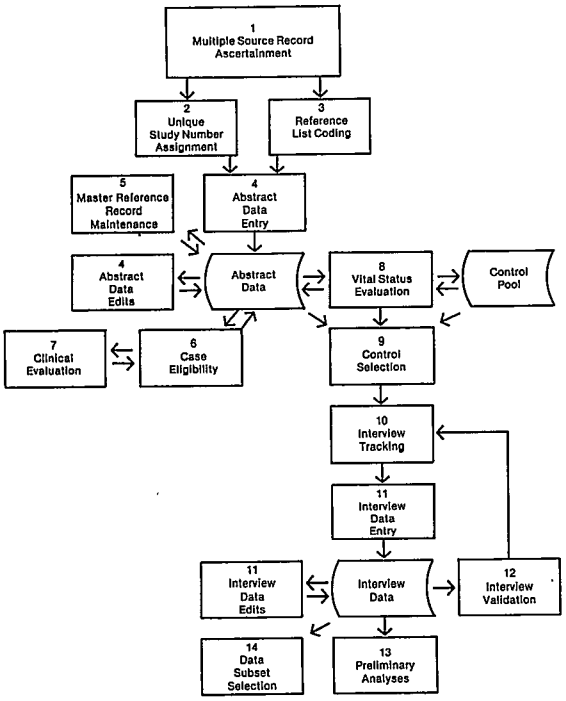


A MICROCOMPUTER-BASED DATA MANAGEMENT SYSTEM FOR CASE-COMPARISON STUDIES

Richard A. Johnson, University of Texas School of Public Health

The case-comparison data system (CCDS) is divided into a series of steps or modules which are common to most epidemiologic studies (Figure 1). These modules may be grouped into four general areas: 1) Case Ascertainment; 2) Control Selection; 3) Interview Data Management and 4) Data Analysis.

FIGURE 1



Multiple Source Record Ascertainment

The first module is Multiple Source Record Ascertainment. The CCDS is designed to handle multiple records from different sources, a common occurrence in epidemiological studies. The ability to track information back to the information collection form is crucial to the correction of data entry errors at later stages. The link is established by the assignment of a sequential document number to each information collection form before entry. This document number is entered as the first field for each computer record created from the document.

Unique Study Number Assignment

The next step in the system is the assignment of unique study numbers to each individual. This is done through an automated system which checks each new subject against all previous subjects on the basis of first and last name or date of birth. All possible matches are displayed on the screen and the data entry person is asked to determine if the subject should get the next consecutive study number or a previously assigned study number.

Possible oversights are later identified through the use of more precise identifiers such as social security number. Duplicated study numbers and multiple study numbers given to the same subject are checked for at that time.

Reference List Coding

Another important step in the management system is the assignment of code numbers to reference lists used in the study. These include lists of physicians, hospitals, counties, etc. Procedures similar to the study number assignment program are used to generate these codes ensuring that no codes are duplicated.

Abstract Data Entry and Edits

The use of commercial software for data entry greatly reduced the amount of time used for program development. Datastar, by Micropro International Corporation, was used for this purpose. Although there are other commercial packages available, no other has the combination of flexible screen formatting, data editing and batch entry provided by Datastar while running on the type of equipment used in this study. Entry screens can be made to look very similar to the data collection forms, thereby increasing the reliability of data entry. The type of data, either numeric or character, may be specified for each item as well as a value range for numeric fields. The most significant data checking is provided through the process of batch reentry. This mode requires the data to be entered twice, comparing the second entry to the first. Only after the two entries match are the data moved to a permanent data file.

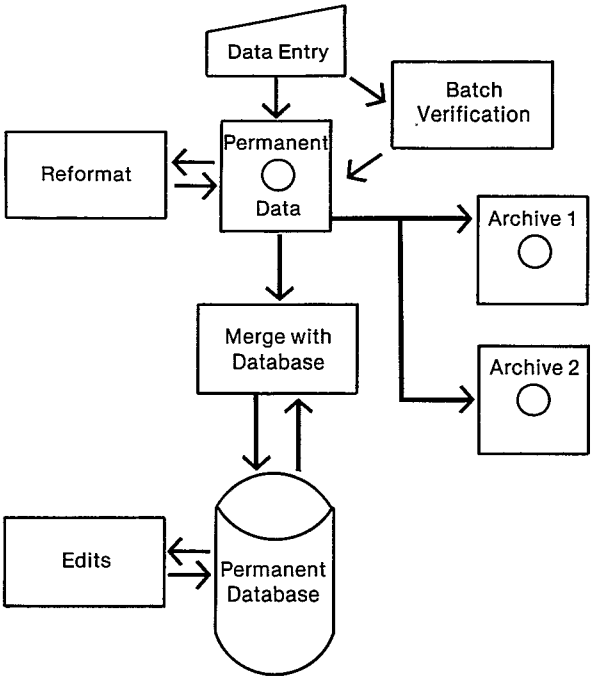
Abstracts are entered in batches which are approximately one-half the capacity of a floppy diskette (Figure 2). After entry, the datafile is reformatted, splitting it into several smaller files. These files are renamed and appended into the main database files which reside on a ten megabyte fixed disk drive. Two archive copies of these files are also made and kept in separate places.

After each batch of abstracts is added, edit programs are run which recheck with greater precision the values allowed for certain variables as well as making logic checks between variables. At this time another check is made for duplicate study numbers which may have occurred during data entry.

Master Reference Record Maintenance

Several of the variables abstracted from various sources need to be represented in the data file only one time. Examples of these are the subject's name, date of birth, vital status and date of last contact. These data are linked to the

FIGURE 2



abstract data and later, to the interview data through the study number. As each batch of abstracts is added to the database, procedures are run which identify discrepancies between the incoming records and any which are already in the Master Reference File. These discrepancies may represent errors or they may simply be values, such as address or vital status, which have changed from one abstract to the other. Many factors may affect the decision of which value is correct. Therefore, discrepancies are listed and appropriate study personnel are consulted to determine which value to use in the master record. A procedure is then used which allows the substitutions to be made.

The creation and maintenance of this master record greatly facilitates routine data processing. This file can be used to obtain a complete list of cases and controls ascertained for the study, whether they have been interviewed or not. The record contains what is considered to be the most reliable information available on a given study subject although the information may have been compiled from several sources.

Case Eligibility

During the editing process, before addition to the main data files, the abstracted records are screened to verify that they meet demographic eligibility requirements such as age, date of diagnosis, residency, sex and other combinations of variables. Ineligible subjects are removed at this point.

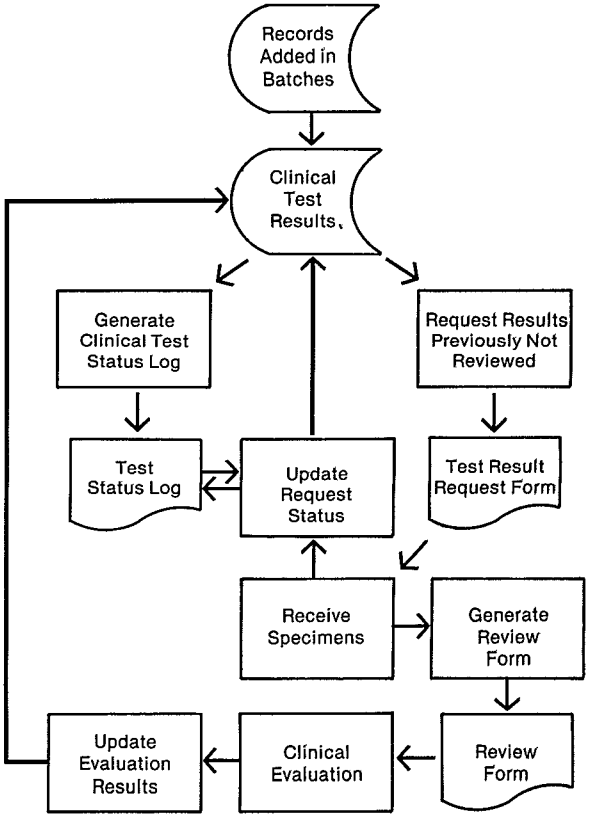
Vital Status Evaluation

As part of the eligibility criteria and as part of control selection, the vital status of potential study subjects must be determined and entered into the CCDS. Procedures are available which list study subjects and all tracing information by year of last contact. Periodically these lists are generated and checked against death certificate records. Changes in vital status as well as additional contact information are entered into the CCDS.

Clinical Evaluation

Various clinical findings are reviewed for each subject before final eligibility status is determined (Figure 3). This process involves the review of pathology reports and the examination of available tissue specimens by a committee of pathologists. The CCDS contains a procedure which identifies and requests test results from participating hospitals, generates a log of clinical test results from each hospital and a clinical review form containing relevant data values. This form is sent to the review committee with the specimens and is returned with their findings.

FIGURE 3



Control Selection

Controls are selected from a pool of eligible subjects derived from driver's license tapes and matched to the cases on several demographic characteristics. Although paired matching can be used if desired, the CCDS is currently utilized to group match controls. Each time new cases are added to the database, a frequency distribution

distribution of the controls is generated. Using the current number of cases in each matching group and the matching ratio, the system calculates the required number of controls per group. The number of new controls needed is determined by subtracting the current number of controls in each group from the number required. A procedure is then used which selects the required number of controls; by group, from the control pool and adds them to the Master file. Once the controls are added they are checked to determine if any have been previously ascertained as cases.

Interview Tracking

The tracking of subjects through the study from ascertainment to completed interview is, by far, the most complex process in the CCDS. Input from several staff members is required at many different points in this process. The tracking system of the CCDS has been implemented as a menu driven procedure which can be used by staff members not familiar with the other data management procedures.

At the heart of the tracking system is a file which is similar to the Master file in that it has a single record for each person selected for the study. This record contains items which indicate the status of key events along the way to the completion of the interview. Examples of items included in the record are: status of the physician's consent letter, the interviewer to whom the subject has been assigned, whether the interview has been sent to the interviewer and the date sent, the date the interview was completed, etc.

Tasks associated with interview tracking primarily involve the periodic updating of status fields as each subject moves through the system. Each task is given its own menu selection. Choosing a task from the menu presents the user with a set of fields which may need to be updated during that task. In this way, the user is prevented from making inadvertent changes to fields which don't apply to that task.

Interview Data Entry and Edits

Entry of the interview data is done in the same manner as described for data from the abstracts. Batch entry with validation by reentry is used. Interview data is subjected to range checking and an intensive set of logic checks. These logic checks edit information within each interview and also check it against information contained in the master record wherever possible.

Data for interview validation follows the same path through the system that is traveled by the interview data. Validation data can be easily compared to interview data regardless of whether all items or only a subset are being compared.

Preliminary Analysis

Preliminary analysis of the data can also be carried out on the microcomputer. The generation of frequency distributions, crosstabulations and descriptive statistics are among the most common of these analyses and are well within the capabilities of a number of packages available for the microcomputer.

The generation of plots and other types of graphs is extremely important when looking at data closely for the first time. The ability to generate these types of tools without having to worry about the cost-performance factor encourages investigators to become familiar with their data before embarking on more complex analyses.

Data Subset Selection

Once the preliminary analyses have been used to determine the key areas of concentration, subsets of the data can be selected. These can then be transferred to a mainframe computer for more complex analyses, such as iterative procedures, which require more speed and accuracy than our microcomputer can provide. This is less of a problem with the current generation of microcomputers and software than in the past.

System Costs

The design, implementation and documentation of the CCDS required approximately one man-year although the system evolves constantly as improvements are added. At present, the system runs under the CP/M operating system on an Apple IIe microcomputer. Peripheral equipment includes a 10 megabyte fixed disk drive, a 1200 baud modem and a dot matrix printer. The total cost (in 1983) of the hardware was $7,000. The software to support the system cost an additional $1,200 and included dBASE II[1] -- the database manager, Datastar[2] -- the data entry package, and Apple Writer IIe[3] -- a word processor. The data entry and data management packages are available on a wide variety of popular MS-DOS and CP/M-based microcomputers including the IBM PC, the Kaypro II and the DEC Rainbow.

The selection of controls from a pool of 2 million potential subjects and transfer of those subjects to the microcomputer cost approximately $600 for computer time. This cost, spanning just a few days, amounts to one-half of what was spent on microcomputer software alone.

Conclusion

In conclusion, the CCDS has been implemented at the University of Texas, Health Science Center at Houston, School of Public Health as part of a large epidemiological study. The system is duplicated at a collaborating center, the Louisiana State University Medical School at New Orleans. Through the use of the system, the data management procedures are uniform between the two centers. The final data set for each institution is expected to include approximately 550 abstracts for 400 cases and their 400 controls. Each completed interview contains over 200 primary responses.

Advances in microcomputer technology over the past few years, coupled with falling prices have brought enormous computing power to a great number of people. Although the initial cost for hardware may seem high, the costs are more easily bud-

geted and are less likely to increase as the size of the database grows. Hardware procurement and system development of the CCDS began over two years ago. We have always believed this project to be feasible even given what must, today, be considered less than average technology. It is certainly more feasible now and becoming more s- every day.

[1]dBASE II is a registered trademark of Ashton-Tate Inc.

[2]DataStar is a registered trademark of MicroPro International Corp.

[3]Apple Writer IIe is a registered trademark of Apple Computer, Inc.

This research is supported through a grant from the National Cancer Institute (Grant No. 501CA32584) and is the result of work done by Dr. Keith Burau, University of Texas School of Public Health.

THE USE OF MICROCOMPUTERS FOR PRIMARY DATA CAPTURE IN A LARGE SCALE FIELD SETTING

Dennis G. Ross-Degnan, International Eye Foundation

INTRODUCTION

The methodology used to collect health data places constraints on its quality and reliability. Dimensions along which methods may vary include: location of the examination or interview; duration of contact with respondents; the experience and skill of the clinical examiner or interviewer; the purpose of the data gathering; and the format of the device or protocol used to collect the data.

In the developing world epidemiologic data are often gathered in circumstances which dictate that contact with subjects is brief, interviewers inexperienced and hastily trained, and examiners rushed. In addition, in order to adequately represent a population, data must often be collected far from any central, controlled environment where adequate supervision of the data collection process is easy to implement.

This paper will describe a computerized system of data capture in such a setting. The Saudi Arabia Eye Survey was a clinical eye examination survey of a statistical sample of the Saudi population carried out in 1984 by the International Eye Foundation and the King Khaled Eye Specialist Hospital to define the national prevalence of blindness and eye disease in Saudi Arabia; to describe regional variations; and to develop detailed plans for targeted interventions to cure and prevent avoidable blindness and eye disease. The 16,810 examinations were conducted over a four-month period by five examining teams operating from a series of 16 camps. Clinical examinations were completed in 75 sample locations that spanned the length and breadth of the Kingdom. The camp was free standing, supplying its own food, water, and electricity, and in addition, its own computer facility.

Although the context is that of a clinical survey, the data collection ideas and techniques are generalizable to other settings. In many ways, the process described can be thought of as a worst-case scenario for the utilization of computers in data gatherings. Having proved feasible in the extremes of the Saudi environment, such a system, if justified, could be implemented in most settings.

WHY COMPUTERS?

Inevitable problems associated with paper-and-pencil data collection techniques include: invalid or out-of-range responses; confusion about intended response when forms are marked ambiguously or changed by the interviewer; erroneously skipped or included items or sections of a protocol; and the turnover time, cost, and errors of post facto keypunching and verification.

Computerization of a data collection process may seem a rather expensive and unwarranted solution to such problems. The issue turns on a number of factors. First, what is the value of accurate data, and what gains in accuracy can be made through computerization? Second, how necessary is it to have access to the data and its contents how rather than later? And third, what are the marginal differences in the costs of collecting information with computers, and are they balanced by benefits that accrue? Although no definitive answers to these questions are possible out of context, consideration of the experience of the Saudi Arabia Eye Survey will highlight some of the relevant parameters.

HARDWARD AND SOFTWARE CONSIDERATIONS

Before describing the particular system used in the Saudi Arabia survey, however, it would be useful to review some issues that must be considered in the development of any system for the capture of health data in the field. First are the characteristics of the machine itself.

Paramount issues in a field application of this sort are the portability and durability of the computers. They must withstand environmental stress including ranges of temperature, dust, and moisture (or the lack of it), and also shock and misuse. In any application that requires portability, the machines must be compact and light, transportable in a single case, and easy to set up for operation.

Although in some settings it might not be necessary, ideally the computers rely on their own internal power supply. Batteries must be rechargeable in a reasonable length of time, have an acceptable life span and charge duration, and be easily replaceable.

Internally the computer must use a processor and architecture fast enough to execute at a speed which does not interfere with the flow of data collection, and for versatility, the operating system should be compatible with one of the major systems on the market. Available memory must be large enough to hold the data capture program and utilities, and lengthy or complicated start-up procedures can be eliminated if the software resides in non-volatile memory.

Display screen adequacy is a weak area in current lap-size portable computers. A screen must be easily readable and adjustable, to suit a variety of situations in which data collection might occur, and must be of adequate size to allow sufficient prompting. For an interview survey, this would be the text of a question and its answers, or better a series of items to allow the operator to keep a sense of the interview context. In some applications, like clinical surveys, prompting is less critical, since the general flow of items and valid responses becomes familiar to examiners in a relatively short period.

The need for a hard copy of results will vary with the particular application. Very few portables have built-in printers, and carrying a portable but separate printer is cumbersome. Built-in printers usually suffer in speed, quality, and noise. For applications where loss of information in machine storage can be tolerated, dispensing with immediate hard copy is preferred.

For lap-size machines, current mass storage alternatives are microcassettes or mini-diskettes. The failure rate of storage media,

especially under operational stress, is an important concern, as is capacity. In most settings, random access to information is essential for correction or updating, so if the storage is to be done on microcassette, its ability to access randomly at a reasonable speed is critical. The final hardware need is the capability of the machine to easily communicate stored data to a larger computer for analysis and archiving.

Software issues of primary concern are the type of program used to develop the data gathering protocol, its friendliness to the operator, and features of program logic which protect against errors.

If the computer operating system is a standard system, then use of commercial database packages for developing the data capture program is possible. Usually this entails sacrifice in sophistication of branching and error checking desirable in many protocols. Speed of developing the application with a commercial package is balanced against the power that programming in a lower level language allows.

A compiled program, as opposed to an interpreted one, runs faster, and so is often more convenient for the operator. Many portable machines have language interpreters, and not compilers, and so run complex programs slowly. Software must also be alterable and able to be reinitialized in the field by the operators to allow recovery in unanticipated circumstances.

Desirable logic features include: immediate range, consistency, and plausibility checks for entries; forced entry of a valid code for required items, and skipping of logically eliminated ones; ability to recall and change previously entered information; ability to tolerate nested errors in data entry with appropriate explanation and handling; automatic logical branching to item subsets; and finally capacity for program interrupt, and temporary storage or restarting of cases.

THE SAUDI ARABIA EYE SURVEY SYSTEM

The decision to use microcomputers in the Saudi Arabia Eye Survey was due to a number of factors. First, in ophthalmic surveys, many items of interest occur rarely, and so the accuracy of each piece of information is important. Previous IEF experience in similar clinical surveys suggested that the level of error was unacceptably high, even with experienced personnel and nightly review for completeness and accuracy. As an example of the possible magnitude of error when control standards are low, in a recent ophthalmic survey in Egypt, 37% of individuals coded as having visual acuity loss had missing diagnoses, the primary data item in the protocol, and 4% of individuals with normal vision received a diagnosis erroneously.

In an attempt to control this type of error, in an eye survey in Kenya in 1982, a prototype "portable" system was developed which used a small industrial control computer and required a bulky twelve-volt battery for power. Despite its shortcomings, the gains in information accuracy were felt to be considerable, and further development with a more adaptable machine justified.

The next factor arguing for the use of a computer-based system was the rapid development of the technology of truly portable machines. By the time the Saudi Survey was being planned in May, 1983, there were on the market a number of potentially usable machines, including the Epson HX-20 actually employed. Many of these are already obsolete compared to current lap-size computers.

Finally resources were available for the development of such a system for this survey, and for political and logistic reasons, it was necessary to report on the data very rapidly. Turnaround time for entry, verification, and complete editing of data gathered on paper forms was felt to be too long to make such rapid reporting possible.

The survey was conducted on a house-to-house basis, to ensure maximum response of eligible participants. Computers needed to be easily transportable, rugged, and quick to set up to keep pace with an average of 60-75 examinations per day. In addition, since many of the homes visited, most notably bedouin tents, would not have electric power, the machines had to be able to operate on their batteries for at least 8 hours at a time.

Each of the five examining teams typically traveled two to four hours daily in vehicles. Vehicles were air-conditioned, but the computers were subjected to many hours of vibration and jarring. In addition, they were exposed to extremes of heat and especially dust, and after a time required periodic removal of the keys and cleaning of the underlying contacts with alcohol swabs to remove the accumulated dust.

Computer operators all had at least a secondary education. None had any previous exposure to computers, however, so the program had to be kept simple to operate and difficult to corrupt. Training time in the operation of the computer system was minimal, and for most operators required two to four days in a pilot setting.

Registration and screening of survey members occurred first in each household, while the ophthalmologist and computer operator prepared themselves to receive screened subjects for examination. A coded demographic and screening information form accompanied each individual to the examination area. While the ophthalmologist began the exam, this information was entered into the computer in response to a sequence of item prompts. Because of the Epson's small screen size, these prompts were brief, and a list of responses could not be fit on the screen to select from. Instead valid codes appeared only on the screening forms or on a reference sheet for the clinical portion of the exam.

Upon completion of the exam, clinical findings, communicated verbally or by notation on the screening card, were entered. The structure of the exam process made this style of data entry feasible. Items were concerned primarily with the existence of pathologies, so positive codes were somewhat infrequent. Only because items and whole sections of the protocol could be skipped rapidly were computers able to keep pace with examinations.

Errors and inconsistencies were identified

at the point of entry, since the program would not allow required information to be skipped or questionable values to be accepted. All codes that had been entered were also printed out for verification by the ophthalmologist before the team proceeded to the next house, and this printout was stapled to the screening form as a permanent record. The final exam results, stored in a random access microcassette file, were therefore as free as possible from recording errors at the point of collection.

A major feature of the system was its ability to retrieve at night summaries of information collected during the day. While the machines were being charged by the portable generators which electrified the camp, these summaries were assembled by a utility program, and were used to monitor geographic, temporal and observer variation. Completed data tapes were transferred later to another computer for management and analysis. Editing time necessary before analysis was minimal+.

BENEFITS OF COMPUTERIZED DATA CAPTURE

The benefits of capturing data directly on microcomputer fall in four general areas. First, downline processing costs are reduced, partially offsetting some of the marginal costs of implementing the system. Such areas of downline costs include data entry and verification, as well as the resources required to validate and edit the resulting data file.

Another major benefit is reduced access time to the data. This allows validation and screening of items as entered, and also opens the possibility for data-based reporting and feedback to local personnel. The time required for producing complex analyses of the information can also be shortened as necessity and staffing allow.

The quality of the data is enhanced significantly. Computers help overcome the inevitable fatigue and routinization that occur in a data gathering operation. Major classes of error, like invalid or illogical codes, and omitted or out of sequence items, are avoided. Other types of error can be reduced, for example by requiring verification of unlikely combinations of items. And finally, intra- and inter-observer variation can be monitored and minimized.

A final area of benefit is the enhanced quality of work for the interviewer/registrars. Morale and hence quality of work are improved by the opportunity to learn a new skill, and by encouraging the feeling of a more important function in the entire survey operation.

AREAS OF ADDITIONAL COST WITH COMPUTERIZATION

The additional marginal costs of computerized data collection as compared to manual methods are difficult to calculate out of a particular context. They depend in part on the use to which hardware and software assembled to implement an application will ultimately be put. Table 1 arrays the equipment costs in dollars and personnel time of the Saudi Arabia Eye Survey, and a theoretical range of costs for different types of application with similar needs.

The major categories of cost are hardware purchase, and software purchase and development. Machine cost, if amortized over a number of years and separate applications, actually reduces to a low cost per application. In a similar way, if generic data collection software is purchased and utilized for other purposes as well, its true cost is reduced considerably.

Marginal increases in personnel costs are also associated with computerized operations (Table 2). Tailoring a program to match the structure of a particular data protocol can be a complex process and full development, testing, and documentation of a single-use protocol could well exceed all the other categories of marginal cost. In Saudi Arabia about ten weeks in person-time was spent in this phase of the operation. A large portion of this was required to develop ways to overcome memory and storage limitations in the machine used, and to implement changes in the algorithms following changes in the examination protocol.

The marginal costs associated with training operators are flexible, and depend upon their prior experience with computer systems. For experienced users, additional training costs can be virtually nil. Similarly, additional clerical time needed to prepare machines and recording media, and to transfer data to a larger machine can be effectively reduced if data is stored on diskette, and if the machine has good communications capabilities.

Field management and supervision of a computerized system is a significant area of concern and cost. Expertise must be available to properly supervise operators; to catalog and organize stored information as it is collected; to modify the system if necessary; to produce and analyze interim reports of results; and to maintain and repair the hardware. If the operation is occurring in a difficult physical environment, away from sources of repair and support, this field management function is vital for the success of the system.

CONCLUSIONS

The feasibility of primary data capture on microcomputer in a large-scale field health survey has been proved by the experience of the Saudi Arabia National Eye Survey. Even under a set of harsh environmental circumstances, it is possible to implement a truly portable system of data collection, operated by personnel unsophisticated in the use of computers, and collect data that are as free as technically possible from avoidable sources of error and available immediately for analysis.

Although feasible, whether such a system is practical or desirable in a particular setting depends upon a number of factors. These include: the magnitude and perceived value of the potential reduction in error, which depends in part upon the skill and experience of the interviewers or data collectors; the complexity of the protocol; the circumstances under which the data collection will be taking place; the plans for long-term utilization of the additional hardware and software that must be assembled to implement the system, the need for rapid access to the data for

monitoring indicators or for feeding back results in an ongoing fashion, and the availability and cost of personnel to develop the programs, operate the system, and most importantly, supervise it in the field.

The availability and performance capability of portable computer equipment will increase dramatically in the near future and the cost of this equipment will decline. In addition, generic software systems will become available that will allow rapid and inexpensive development of data gathering protocols for these machines, protocols which will be flexible in logical and error checking capability and therefore well suited to the collection of complex data. As these trends develop, the benefits of using computers in most data collection settings will soon outweigh declining marginal costs, and the use of systems of computerized data collection will become the norm.

Table 1: Benefits of Computerized Data Capture

1. Processing costs avoided
    a. data entry and verification
    b. data validation and editing
2. Enhanced access to data
    a. immediate correction and validation
    b. immediate capability to produce reports
    c. reduced time to complex analysis
3. Increased quality
    a. prevention of major types of coding error
    b. reduction in observer bias
    c. reduction in other types of error
    d. improved employee satisfaction

* *

Table 2: Additional Costs: Equipment and Materials

	Saudi Arabia Eye Survey	Feasible Range
1. Computer, ROM/RAM. diskette/tape drives, cables, printer	8 @ $900	$400-3500
2. Supplies (diskettes/ tapes, paper, print ribbons)	$1600	$100-2500
3. Maintenance per machine	--	$0-750
4. Software (data entry package or language compiler/interpreter)	$400	$0-1000

* *

Table 3: Additional Costs: Personnel Functions

	Saudi Arabia Eye Survey
1. Applications programming	8 weeks
2. Documentation and training	2 weeks
3. Computer operation	
a. training period	1 week
b. added daily activities	5 hours/day
4. Field supervision and management	3 hours/day
5. Clerical	4 weeks

USING SPREADSHEET SOFTWARE FOR DATA ANALYSIS: AN APPLICATION FOR SUMMARIZING CHART AUDIT DATA

Sarah S. Marter, The Health Data Institute Inc.

INTRODUCTION

For certain analytic tasks, spreadsheet software used on a personal computer offers an efficient alternative to statistical software packages used on a mainframe computer. This paper will give an overview of Lotus 1-2-3 (tm) and show how it was used for tabulating results of hospital chart audits for inappropriate ancillary utilization. The advantages of using spreadsheet software and criteria for evaluating its efficacy will be discussed.

LOTUS 1-2-3 OVERVIEW

A spreadsheet is a columnar pad used in accounting that is larger than a standard piece of paper. An electronic spreadsheet can be much larger, and it offers the flexibility of adding and deleting rows or columns. In Lotus 1-2-3 the spreadsheet, called a worksheet, contains 256 columns and 2048 rows. A cell is defined by its column letter and its row number. Cells can contain numbers, titles or labels, formulas, or macros. (Macros are helpful in automating certain operations. For a good introduction to 1-2-3 macros, see Bingham 1984.) Formulas are created by using the four basic arithmetic operations as well as certain additional functions. Figure 1 shows some of the 1-2-3 functions which are useful in statistical analysis.

Figure 1

SELECTED 1-2-3 FUNCTIONS

@COUNT(List)	Counts the number of all items in list
@SUM(List)	Sums the values of all items in list
@AVG(List)	Averages the values of all items in list
@MIN(List)	Minimum of all items in list
@MAX(List)	Maximum of all items in list
@STD(List)	Standard deviation of all items in list
@VAR(List)	Variance of all items in list
@IF(Cond., X,Y)	The value X if conditon is true, the value Y if condition is false
@ROUND(X,n)	Round a number to n decimal places
@INT(X)	The integer part of X

Other important capabilitites of a spreadsheet include copying, moving, erasing, and formatting cell contents. When formulas or functions are copied to a new location in a worksheet, unless specified otherwise, the cell address (column letter and row number) of each cell readjusts to the new location. This capability is particularly helpful in the case, for example, of several columns of numbers (columns A, B, and C) to be summed. The formula for the first column is written in the cell that will contain the first sum (@SUM A3..A10) and then the formula is copied to the cells that will contain the subsequent sums. The software automatically adjusts the column letters in the formula to the new columns (@SUM B3..B10 and @SUM C3..C10). The most dramatic feature and the one that saves the most time is the automatic recalculation of formulas when new values are substituted. In the past, spreadsheet software was used mainly for financial applications. The classic example is a budget projection. Now spreadsheets are used for almost any application in which straightforward, rapid recalculation is needed.

THE ANALYTIC TASK: SUMMARIZING DATA

The application described here was developed to summarize the results of hospital chart audits, conducted during an Ancillary Services Review Program (ASRP). The audits evaluate the appropriateness of ancillary services ordered, according to clinically proven guidelines. Ancillaries are all the services rendered in a hospital that are not in the categories of room and board or physician services. At The Health Data Institute (HDI) we focus on the use of laboratory tests, EKG's and other diagnostic cardiac tests, respiratory therapy services, antibiotic therapies (pharmacy) and diagnostic radiology. (The ASRP methodology is described in Hughes et al. 1984.)

Before the first spreadsheet was designed for summarizing the laboratory audit data, the data were analyzed on a mainframe computer using a statistical software package. Spreadsheet software is appropriate for this analysis for three reasons. First, the summaries required are straightforward, consisting essentially of sums, averages and percentages. Second, the analysis is repetitive. The data are summarized by diagnosis within a hospital and then overall for all diagnoses at a hospital. Several diagnoses are audited at each hospital, and multiple hospitals are audited for each ancillary. The same summary statistics are calculated for three of the five ancillaries we audit. Variations of the laboratory worksheet are used for the respiratory therapy and cardiac audit data. Additionally, the ASRP is a product offered to multiple clients. Third, the number of charts reviewed for each diagnosis is within a predictable range, usually 15 - 35 charts.

THE SPREADSHEET PROGRAM

The worksheet designed for the laboratory chart audit data is shown in Figure 2. This example shows actual data from six patients'

charts and the summary statistics for all charts audited at Hospital X for gallbladder patients. Three types of laboratory tests were reviewed, electrolytes, enzymes, and urine cultures. For each chart the form number, number of tests reviewed, and number of tests judged inappropriate were entered into the worksheet. Since blanks are treated as zeroes, there was no need to enter all the zero values.

The indicator variable formulas were created using the @IF function and are self-coding variables that depend on the data entered. For example, to code the Presence of Test in Chart (position #7, circled on Figure 2) for electrolytes for the first chart (Form Number 1001), the formula is set up to test the condition of whether or not the value in the Number (of tests) Reviewed (#5) cell for electrolytes is greater than 0. If it is, then the cell is coded 1, if not, the cell is coded 0. Similarly, the Presence of an Inappropriate Test (#8) is coded from the Numbers of Inappropriate Tests (#6), and the indicator variable for the Presence of Any Inappropriateness (#9) is coded from the column directly to its left (#8).

On this template, spaces are provided for data gathered from up to 35 charts. In the example presented in Figure 2, only 24 charts were reviewed. At the bottom of Figure 2, the summary data for this diagnosis are calculated. The formulas use either simple arithmetic or 1-2-3 functions. The Total Electrolytes (below #11) cell is the addition of the values in the cells for each chart which correspond to the number of electrolytes reviewed (+d14+d18+d22+d26 ...including cell addresses for all 35 chart spaces). This formula was typed once and then copied to the three columns (#12, #13 and #14) to the right of the Total # Reviewed column (#11). The Grand Totals were calculated using the @SUM function i.e. @SUM(d145..d147). That formula was copied for the Grand Total of the # Inappr. (#12). The Percentage Tests Inappropriate by Test (#17) was calculated as the ratio of the number of inappropriate tests to the number of tests reviewed by type of test. It was then formatted as a percentage. The remaining summary statistics were generated in a similar manner.

The worksheet for Hospital X contains templates for three diagnoses and a template for a summary table that can be sent to the hospital in an audit report. The worksheet map in Figure 3 illustrates the overall layout of the worksheet. The template for the summary table appears in Figure 4. The Chart Review Results by Diagnosis are automatically copied from the templates for each diagnosis. The Overall Summary of Inappropriate Tests by Type of Test is calculated in the summary table template.

ADVANTAGES OF USING SPREADSHEET SOFTWARE

The traditional analytic method for summarizing these data used a statistical package on a mainframe computer. When comparing the original method to the spreadsheet method, the advantages of the latter method are apparent.

In the traditional mode, all the data for all diagnoses, all hospitals, and all ancillaries were keypunched or entered at a terminal. The large data set was proofread and errors were corrected. A program was written and debugged. The final job was run on the computer, and it produced multiple pages of output. An analyst copied the numbers needed onto a summary table by hand. The table was checked to make sure there were no copying errors. Finally the table was typed and proofread.

In the spreadsheet mode, the data are entered for one diagnosis per hospital per ancillary at a time. The template provides the blanks for up to 35 charts. The data portion of the spreadsheet is printed and proofread. Errors are corrected, and the recalculation feature updates the summary automatically. The other diagnoses for that hospital for that ancilllary are entered and proofread. Then some additional data are added (to fill in the blanks in Figure 4) from an earlier phase of the analysis and the table is ready to be printed. (We have had this table retyped to improve the format, however it is possible to use a wordprocessing program in conjunction with 1-2-3 to eliminate this extra step.)

The sequence of work was much more efficient using the spreadsheet method. The audits were performed on a hospital by hospital basis. We were able to avoid a backlog of data by analyzing it as it arrived in the mail, finishing the summary reports within a few days. We eliminated the possibility of copying errors by having the summary table generated automatically. The portable computer offered more independence. The work could be done either on a desktop computer in the office or on a portable computer outside the office. Unlike the mainframe computer environment, there was no downtime for maintenance and no waiting for a batch job to execute. The operating costs were substantially lower. The cost of the diskettes compared favorably with the traditional charges for connect time and CPU usage. The program was easier to understand. Non-programmers used the templates effectively with minimal understanding of 1-2-3. Recalculation adjusted summary statistics without a second run of the job on a mainframe.

DECIDING TO USE SPREADSHEET SOFTWARE

Answering three questions will help in deciding which analytic tasks could be performed more efficiently with spreadsheet software.

1. Will the spreadsheet software meet my analytic needs?

It is possible to construct very complicated formulas with spreadsheet software; for instance, functions can be nested inside other functions and macros can be designed to automate the worksheet. However, the beginner should probably start with simple applications on relatively small worksheets.

If the formulas needed can be constructed with the functions listed either in Figure 1 or in the manual and the four arithmentic operations, then spreadsheet software may be appropriate. In this application the single template for one diagnosis was developed originally. The template for the summary table followed. Lastly, the combination of templates

Figure 2

SPREADSHEET PROGRAM EXAMPLE

ANCILLARY SERVICES REVIEW PROGRAM
LAB CHART AUDIT RESULTS

HOSPITAL: X DIAGNOSIS: GALLBLADDER

	TYPE OF TEST	FORM NUMBER	DATA: NUMBER REVIEWED (TESTS)	DATA: NUMBER INAPPROP.	INDICATOR VARIABLES (Y=1,N=0): PRESENCE OF TEST IN CHART	INDICATOR VARIABLES (Y=1,N=0): PRESENCE OF INAPPR. (BY TEST)	INDICATOR VARIABLES (Y=1,N=0): PRESENCE OF ANY INAPPR. IN CHART
1	ELECTROLYTES	1001	2	1	1	1	
	ENZYMES		1		1	0	
	URINE CULTURES		1		1	0	1
2	ELECTROLYTES	1002	3	1	1	1	
	ENZYMES		2	1	1	1	
	URINE CULTURES				0	0	1
3	ELECTROLYTES	1003	7	3	1	1	
	ENZYMES		3	2	1	1	
	URINE CULTURES				0	0	1
4	ELECTROLYTES	1004	5	2	1	1	
	ENZYMES		2	1	1	1	
	URINE CULTURES				0	0	1
5	ELECTROLYTES	1005	6	3	1	1	
	ENZYMES		4	3	1	1	
	URINE CULTURES				0	0	1
6	ELECTROLYTES	1006	3	1	1	1	
	ENZYMES		2	1	1	1	
	URINE CULTURES				0	0	1
35.	ELECTROLYTES				0	0	
	ENZYMES				0	0	
	URINE CULTURES				0	0	0

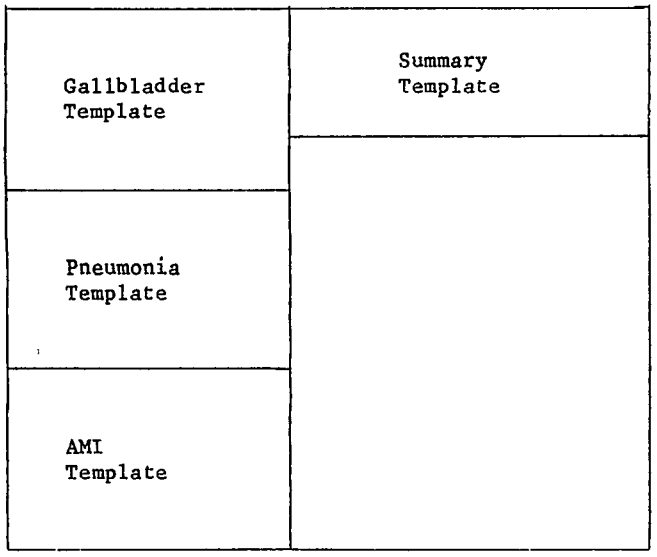
SUMMARY

-------	DIAG.: G.B.	TESTS: TOTAL # REVIEWED	TESTS: TOTAL # INAPPR	TOTAL # CHTS W/TST	TOTAL # CHTS W/INAP	TOTAL # CHTS W/ ANY INAP.	TOTAL # CHARTS REVIEWED
TOTAL ELECTRO		87	24	23	14	16	24
TOTAL ENZYMES		47	20	22	12		
TOTAL URINES		6	1	5	1		
GRAND TOTAL		140	45				

	PERCENTAGE TESTS INAPPROP. BY TEST	PERCENTAGE CHARTS W/ INAP. TEST, BY TEST		PERCENT. CHARTS W/ TEST	MEAN # TESTS/CHT
ELE	27.6%	58.3%		95.83	3.63
ENZ	42.6%	50.0%	PERCENTAGE OF	91.67	1.96
URI	16.7%	4.2%	ALL CHTS W/ ANY INAPPROPRIATENESS	20.83	0.25
TOT	32.1%		66.7%		

Figure 3

1-2-3 Worksheet Map



for multiple diagnoses were linked to the summary table within a single 1-2-3 worksheet (as represented in Figure 3).

2. Is the task repetitive? Is this calculation going to be repeated?

If the same calculation is repeated on a regular basis, then it is probably worth the time required to design the template or spreadsheet. In situations in which data are reported and summarized regularly, the spreadsheet software is appropriate. One example is calculating infection rates by floor or service in a hospital. The infection control nurse has the same data elements to summarize every month. The template approach saves time. A second example is standardization of rates by age and sex. If the same reference population is used every time, the template can be designed to calculate the adjusted rates automatically when the new data are entered.

3. Can I predict how much data I will have to analyze?

This consideration is important for two reasons. First, the amount of data cannot exceed the memory capacity of your personal computer. With advancing technology, this concern will become less of a problem for small data sets. For larger data sets, it will always be necessary to estimate capacity. Second, the structure of the template must fit the data. In the application described, there were only three types of laboratory tests reviewed. The block of data for each patient fit neatly into a space on the spreadsheet. However, for another of the ASRP audits, the radiology audit, a variety of tests may be reviewed. One patient may receive a chest x-ray, while another receives 6 separate studies to determine the cause of abdominal pain. The template would be unduly cumbersome if space were provided for all the possible x-rays for each patient. (A spreadsheet program has been designed for the radiology audit data. It is rather complicated and does not offer many of the advantages of the template approach described in this paper.)

For large data sets or data sets that do not fit the design of a template spreadsheet, the spreadsheet software can be used for data entry. Later, the data can be uploaded to a mainframe computer using inexpensive software that is in the public domain. Other papers presented at this conference have addressed this issue.

CONCLUSION

For appropriate data sets or any repetitive calculations, spreadsheet software offers a fast and efficient alternative to using a statistical package on a mainframe. Accuracy of summary tables can be improved. This simple alternative offers potential for wider gathering and sharing of health statistics.

REFERENCES

Bingham, J. B. 1-2-3 Go. Reading, MA: Addison-Wesley Publishing Co., 1984.

Hughes, R. A., Gertman, P. M., Anderson, J. J., Friedman, N. L., Rosen, M., Ward, A., & Kreger, B. E. The Ancillary Services Review Program in Massachusetts: Experience of the 1982 Pilot Project. Journal of American Medical Association, 1984, 252:1727.

Lotus (tm) 1-2-3 (tm) Users Manual Release 1A. Cambridge, MA: Lotus Development Corporation, 1983.

Figure 4

CHART REVIEW RESULTS
AT
HOSPITAL X

ANCILLARY: LABORATORY

OVERALL SUMMARY OF INAPPROPRIATE TESTS BY TYPE OF TEST

Tests	# Inapprop	# Reviewed	% Inappropriate
Electrolytes:	132	255	37.8%
Enzymes:	64	215	29.8%
Urine Cultures:	45	67	67.2%

CHART REVIEW RESULTS BY DIAGNOSIS

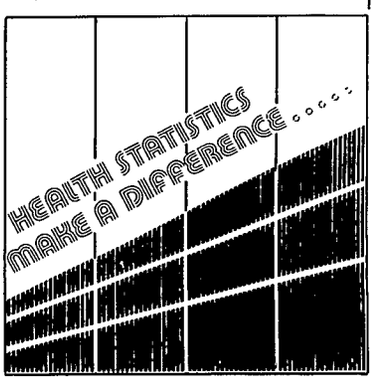
Diagnosis	All Hosp. Avg. Adj. Lab Chg.	HOSP. X Avg. Adj. Lab Chg.	# of Cases	# of Cases Reviewed	Type of Tests Reviewed	# of Tests Reviewed	# of Tests Inapp.	% of Tests Inapp.*	# of Cases w/ Inapp.**
GALLBLADDER					Electrolytes	87	24	27.6%	14
					Enzymes	47	20	42.6%	12
				24	Urine Cultures	6	1	16.7%	1
					TOTAL	140	45	32.1%	16
PNEUMONIA					Electrolytes	124	45	36.3%	19
					Enzymes	76	20	26.3%	14
				26	Urine Cultures	32	21	65.6%	17
					TOTAL	232	86	37.1%	23
ACUTE MYOCARDIAL INFARCTION					Electrolytes	138	63	45.7%	12
					Enzymes	92	24	26.1%	8
				12	Urine Cultures	29	23	79.3%	10
					TOTAL	259	110	42.5%	12

Inappropriate laboratory testing was found in 51 of the 62 cases reviewed.

* % of inappropriate tests = # of inappropriate tests/# of tests reviewed.
** # of cases with inappropriateness.

Session E

Information Systems Based on Linked Records



USING INSURANCE CLAIMS DATA FOR RESEARCH AND COST CONTAINMENT

Richard S. Papel, McGraw-Hill

Introduction

In order to assist corporations and business health coalitions in identifying health care cost and utilization problems, SysteMetrics has analyzed large health care data bases developed from insurance claims files. Although these insurance files were not designed for research, with the proper attention valuable information can either be derived or lifted directly from them. Unfortunately, however, this useful information is often missing, entered erratically or in a highly fragmented way, making accurate analysis extremely difficult.

This paper describes some common characteristics of health insurance claims files including their structures, the kind of data captured, their limitations, and how they may or may not be used. We outline some problems researchers can expect as well as how they can be overcome. And finally, suggestions are provided for the redesign of claims files eo make them more valuable as analytic research tools, without sacrificing their efficiency as accounting tools.

The Claims File

Up until very recently, insurance claim files were designed exclusively for one task, that being the efficient adjudication and paying of insurance claims. Many claims systems, particularly those of smaller insurance companies, are still designed only for this purpose. However, recent efforts at health care cost containment have spurred many carriers to improve their data collection systems and subsequent reporting ability. In addition to containing basic information about charges and amounts paid, they may contain substantial clinical and demographic information.

Generally speaking, these systems are either basic flat files or more complicated segmented or hierarchial systems. In either case, the first portion of the file is devoted to basic administrative information on both the patient and insured. There might be a claim number, processing office number, medical record number, as well as some basic descriptive information such as name, age, sex, and date of birth. This may be followed by detailed charge and clinical information. Segmented files, on the other hand, will be broken down by elements of the actual claim. Administrative and basic descriptive information will be included on the header, followed by a segment containing provider information, a service segment, charge segment, benefit segment, and perhaps a draft segment.

Typically, for a given hospital stay, the insurance company will receive a minimum of one hospital bill, usually with line item charges, a bill from the attending physician, the surgeon, a lab, and perhaps several other professionals or facilities.

As each claim arrives at the insurance company every line item charge is entered into the system individually meaning that a single physician or hospital bill may yield several records. In its simplest form the carrier will capture the date the service was performed, the type of service, the type of facility at which the service was performed, the provider name, and a diagnosis and procedure code. This line item information is then followed by aggregated financial information needed to pay the claim, such as year-to-date deductibles and payments, COB savings, and coinsurance amounts.

Typically, the diagnoses will be taken from the hospital record and procedures from the respective physician records. Most larger carriers use an established coding scheme such as the ICD-9-CM, while many others use their own individual codes. Still others do not use a code at all and instead type in a brief narrative. Each will be received at different times and entered on the claims system without any regard to the hospital stay as an entity. Various claims relating to a single may be scattered throughout the system with no explicit linkage to each other. The problem is to how collect these bills together and sort them into a usable semblance of order.

The preferred format for our research is the "hospital stay record," a chronicle of the hospital episode from admission through discharge. On this "stay record" (sometimes called an episode of care) we want to record as much clinical and charge information as we can accurately obtain from the claims file. Table 1 shows some of the variables we ideally would like to capture.

TABLE 1

Charges	Clinical
Hospital Charges	Principal Diagnosis
- room & board	Secondary Diagnosis
- ancillary	Principal Procedure
Physician Charges	Length of Stay
- attending physician	Patient Age
- surgeon	Patient Sex
Non-Professional Charges	Discharge Status
- labs & X-ray	Admission Date
- drugs	Discharge Date
- equipment	Length of Stay

Other	Amount Paid
Hospital ID	To Hospital
Physician ID	To Physician
Surgeon ID	To Others
Zip Code	
	Savings
	COB
	Coinsurance

Unfortunately most of these desirable fields cannot simply be "lifted" off the file and, worse yet, the available data may be "dirty" - that is, full of inaccuracies, omissions, duplicates, and other harmful distortions. The root of this problem comes from the fact that most carriers are not really interested in anything other than the charge and the amount they are obligated to pay. As long as the charge is not for cosmetic surgery or some other non-covered procedure, they really do not concern themselves with it. Any recording of clinical information is of secondary importance and may be done erratically. Some carriers consider these to be "optional" fields to be entered only at the discretion of the claims auditor or keypunch operator.

Methodology for Producing Episode Files

The basic methodology for producing a stay record is not difficult. In short, identify the admission and discharge dates and collect the records falling within them. The difficulty comes from trying to deal with the preponderance of bad quality data. Even if 95 percent of the data looks good, decision rules are needed to isolate those 5 percent of the records which are bad. Thus there is not one set methods one can use each and every time. Rather, the methods used will vary with each carrier's data. All along the way, data must be examined for its reasonableness. It is important not to assume that once the "code" has been cracked that the results are going to be correct.

The logical first step is to sort the records together by an individual identifier such as a social security number, thus ensuring that all bills for a single individual are placed together. Most of these records will relate to outpatient reimbursements and should therefore be subset from the data. To do this, find an admission and discharge date and keep all records with service dates falling on or between these dates. Chances are there will not be a field explicitly labelled "admission" or "discharge" date but there may be a "start" and "end" date. Look for a "type of service" field and see if there is a code for room & board - the start and end dates for this record will correspond to the admission and discharge dates. It is also possible that, in the case of an interium billing or room change, two separate room and board charges will be listed. Check to see whether the admission date on one is the same as the discharge date (or the discharge date plus one) on a previous bill. Beware of trying to pare the file down on the basis of a "place of service" code, thereby retaining only records specifically labelled as hospital inpatient. From our experience, these codes can be unreliable and may cause problems later when looking for records that should have been retained, but were inappropriately deleted on the basis of a bad code.

Practically all of the variables listed in Table 1 need to be derived. In finding hospital charges, as with admission and discharge dates, you will need to improvise since there will not (or at least it is very unlikely) be a field specifically labelled "hospital charges." The best solution is to retain the provider number of the record previously defind as the "room & board" record and consider all charges having this identification to be hospital charges. All remaining charges will be either those of the physician or outside facility.

Finding accurate clinical information is the most difficult job since this information is often not provided on a consistent basis. Generally speaking, diagnosis information will be found on the hospital record while procedure information will be on the physician record. Assuming you are interested in the principal diagnosis, find the first occurring diagnosis on the hospital record and consider it to be principal, the second occurring being the secondary, etc. Similarly, the principal procedure would be the single most expensive procedure performed by the surgeon. Finding the attending physician's identity may also be tricky since there will likely be multiple physician IDs listed for a single hospital stay and unlikely that there would be a specific code indicating which was the attending physician. A reasonable assumption would be the physician having the greatest cumulative charges for the stay.

Data Quality Issues

Regardless of how much time and effort the researcher expends in obtaining accurate data, the final degree of data quality can be only as good as the data originally keyed into the system by the insurance company. No amount of manipulation or creative improvising can improve bad data. Accordingly, before making an effort to process a particular insurance company's data, carefully assess its quality and determine its potential value in advance. Look carefully at the variable list, and at a dump of the data to ensure that they are consistent. It is equally important to realize that even relatively good quality data can hide numerous ambiguities and distortions that, given the current state of claims systems and hospital billing methods, cannot be completely identified. Any analysis of claims data must point to this fact and be interpreted cautiously.

This issue comes to the forefront of any work dealing with charges in particular, or where data from multiple insurance companies is merged together into a single data base. Identical-sounding variables on different claim systems can mean slightly different things. Similarly, different hospitals bill differently, each having their own peculariities and nuances. Room rates at two different hospitals may vary but one may include certain ancillary items in the room charge where the other may not. It could therefore not be concluded that one hospital is charging more than the other. Physician

charges may also vary in equally subtle ways. Some physicians may list a charge for consultation and a separate charge for a procedure, while another physician lumps all his charges together under one or the other.

The following is a list of the more frequently encountered and serious data quality issues.

A. Duplicate Records - Under certain circumstances it is possible that duplicate entries could be made for the same service. This would likely occur when a bill was rejected by the insuror, later changed in some way, and then resubmitted. For instance, let us say that a bill was sent in without an important piece of information such as the physician's signature. Chances are it will still have been entered on the claims system before being rejected. Two weeks later the same bill is resubmitted with a signature, and paid. How do you know not to double count the charges? True, there would be two identical-looking bills, one showing an amount paid, and one with nothing paid, but this would be insufficient to conclude one was a duplicate and not paid because of this. Ideally, the carrier should eliminate duplicates before producing the tape, but this is rare.

Some carriers, on the other hand, will provide a "pend" table showing why a bill or line item was not paid. Together, with the carrier, go through and pick out various pends most likely to cause duplication. And finally, other carriers may simply count the number of time a particular bill was resubmitted by using a "counter" field. Here, any multiple counter numbers within a claim number might represent a duplicate. Similarly, a sequence number might be applied to each line item, and any duplicated sequence with a line item would be the duplicate. In short, excluding duplicate observations is at best tricky, and will vary radically depending on the carrier.

B. Unreliable "Length of Stay" fields: If you are provided with an explicit "length of stay" field, ignore it and recalculate the length of stay based on the admission and discharge dates. The reason for this is that the insurance company will calculate LOS by claim number and not reconcile it when there are multiple claim numbers and multiple room & board records for a single stay.

C. Insufficient Provider Identification - The most common provider ID is the social security number or Tax ID. Unfortunately, because of errors, administrative pecuciarities, and other non-standard IDs, the provider identification may be unreliable, especially for hospitals. First, some carriers will combine the provider ID with another field such as the address and place it at different positions within the field making it extremely difficult to automatically read. In addition, it is common to find different physicians having the same identification numbers or, conversely, the same physician having multiple numbers. The problem is the same for hospitals, particularly for hospitals owned by the same parent company. And finally, some carriers often use their own codes for provider IDs which are inconsistent with other carriers.

The only solution is to find some basic common identifier that can be applied to all hospitals.

D. Missing Key Variables - Sometimes missing variables can be pieced together with a bit of creativity. So, before concluding that a key variable is missing, speak with the carrier and look over the documentation for other related bits of information that might be used instead. As an example, one carrier we worked with recently didn't bother to capture length of stay, or discharge date. They did; however, have a field that sometimes corresponded to admission date, and an accommodation charge and room-type. Since the data were isolated to a single city, we found the average city-wide daily accommodation rate by room-type which we divided into the total accomodation charge to estimate the discharge date.

While this gave us a "number" we obviously had to interpret it extremely cautiously, and only in comparison with other out-of-area hospitals.

E. Insufficient Clinical Coding - Many insurance companies, including some of the largest, use their own carrier-specific coding schemes, often far-removed from other standardized codes such as RVS or ICD-9. Other carriers, while using the basic codes, may modify them by capturing only a portion of the code or by adding on their own modifiers. Be prepared to spend time converting these carrier-specific codes into one standard format. Worse yet, these codes are frequently entered carelessly or completely omitted. One carrier we worked with counted the number of normal deliveries we derived and insisted that they had paid claims on far more. Upon investigation, we found that the actual CRVS procedure code representing a childbirth was not being keyed in on a consistent basis, and instead only lab procedures were included.

F. Poor documentation - Many claims systems are old and have outdated documentation that no one has bothered to update. Before spending too much time with a file, take a few minutes to check a dump of the tape with the accompanying documentation.

G. Other miscellaneous variables - Some carriers use a whole assortment of cryptic, strange-sounding variables. While they may not appear to mean anything, be sure to check either the documentation or with the carrier first, before discarding them. Often the same basic variable may be called any number of different names.

What Can Insurance Companies Do?

Surprisingly, many carriers have not the faintest idea of what might be done to improve their claims processing systems for reporting purposes. We are keenly aware that data processing systems, developed and fine-tuned over many years, cannot simply be dropped and replaced with something else more

attuned to the needs of health care researchers. But, as quality reporting becomes an increasingly greater competitive asset for insurers, carriers will institute improvements on their own. Understanding the expense and impracticality of altering a claims system, we would like to suggest a few changes which could be made on most systems with a minimum of difficulty.

First, fields need to be utilized more consistently by claim auditors and keypunchers. A field that is used by some and not others is nearly worthless from a reporting point of view. Secondly, audit checks ought to be built into systems to reduce the amount of input error. Something, for instance, that could check codes for validity along with the "reasonableness" of other data would go a long way toward improving overall data quality. This same system could be used to help reduce the number of processing errors, thereby improving its integrity and ultimately saving money. Clinically, there should be checks to ensure the compatibility of sex and diagnosis codes, as well as the overall validity of diagnosis and procedure codes. Missing values could be spotted as well as inconsistencies and contradictions in the codes. Service dates could be checked to make sure they fall between the admission and discharge, while invalid hospital and physician ID's could be spotted and corrected immediately.

Thirdly, much of the difficulty in assigning DRGs could be eliminated if non-standard clinical codes were not used. Further, it does little good to use only a part of an established coding scheme, such as the first three digits of the ICD-9-CM. The problems caused by this short cut are often compounded when zero-fills are used or codes are incorrectly positioned within a larger field.

And finally, it is almost essential that some kind of zip code, preferably the patients, be included in the data set for doing area analysis. Without this, it is difficult, if not impossible to isolate patients living in a given study area.

LINKING MEDICARE PAYMENT RECORDS WITH MEDICAID MANAGEMENT INFORMATION SYSTEM DATA

Frederick Pratter, Abt Associates Inc.

1. Introduction

One fear that has troubled the popular imagination in the last two decades is that monolithic government data systems are being created by linking various agency files, and that these files are used to spy on individuals in the way predicted in Orwell's 1984. Researchers involved in attempting to combine data from different systems know how unfounded these fears are. About four years ago, Abt Associates was assigned a task that at the time seemed quite straightforward. This was to create a patient level database on 1400 New York State Medicaid recipients, containing information on all the public benefits received by these individuals. This presentation describes some of what was discovered in the process, in the hope that it may be useful to those who are currently working on this problem, as well as those who might be considering such an endeavor.

The primary focus of this presentation is on the database constructed for the evaluation of the New York State Long Term Home Health Care Program (LTHHCP) conducted by Abt Associates for the Health Care Financing Administration. This paper will not discuss the findings of the evaluation; those who are interested in the results should consult the project final report, available from the Health Care Financing Administration.[1] In addition, it should be noted that that since the evaluation included only a sample of 1400 clients, it is only partly relevent to those who would attempt to merge Medicare and Medicaid records for an entire state or states beneficiary population.

2. Overview of LTHHCP Data Collection Process

As Figure 1 illustrates, linking Medicare and Medicaid data was only a portion of the database construction effort for the LTHHCP evaluation. The goal of the project was to collect as much information as possible about all of the benefits received by the study individuals. Data were obtained from the New York State Welfare Management System (WMS), the MMIS claims file, the New York City public benefits programs and the Social Security Administration, as well as other primary and secondary sources. In this paper, the focus is on two aspects of the task: the construction of a complete Medicare data set by combining Part A & Part B bills, and the merger of this file with the Medicaid payment record. The rest of this presentation will describe, in some detail, the steps that were required to accomplish this process.

3. Client Identification

The first step was to establish a consistent set of client identifiers. The process was somewhat tedious. Having selected a study sample, a data file was created (that subsequently became known as ID.CENTRAL) containing the name, sex and date of birth of the clients, as well as their Medicaid Client Identification Number (CIN) and Medicare Health Insurance Claim (HIC) number, obtained from the patients' medical record. These numbers were then verified.

To check a HIC number, one prepares a HIPO request (Health Insurance Printout) to the HCFA Master Beneficiary enrollment file. The HIPO printout includes all the case numbers that pertain to a particular client, based on their name, sex and date of birth. It should be noted that the BIC, or claim number suffix for an individual can change over time due to changes in marital status or other conditions. Also, the entire 11 digit claim number can change, so one must also be aware of any cross reference numbers that might occur. The result of a series of HIPO requests was about 2000 pages of printout that were manually reviewed and input to ID.CENTRAL.

The process of verifying CIN numbers was similar, in that each client was looked up in the New York State Welfare Management WINC system, producing one page of printout for each client listing all the case numbers in which that client was involved. It should be noted that New York City was not included in the WMS system, so that these identifiers had to be verified separately. While the CIN number in New York State is unique to each individual, and is not supposed to change over time, this is not true in every state. The result of this process was another 2000 pages of printout to be reviewed, and another set of updates to the identifier master file. Following this step, the project staff was now fairly certain that the ID master file contained correct sets of identifiers for every study client.

4. Data Acquisition

Data were collected from three separate sources: the New York State Medicaid Claims file (MMIS SURCLAIM 8), the Bill History File (BHF) and the Medicare Utilization Bill & Payment File (Claims). The Claims file includes the Part B payment record and is the only available source of physician and outpatient Medicare payments, although it tends to be less complete than the Bill History file, since it is updated later. Figure 2 shows the types of information collected from these three sources

The Bill History file was the primary source for Inpatient, Skilled Nursing Facility (SNF) and Home Health Bills. The Claims file provided another source for inpatient bills (in addition to the BHF), as well as physician bills and other outpatient services. The MMIS includes any portion of inpatient stays in hospitals and SNFs as well as outpatient services that are not covered by Medicare, and intermediate care facility (HRF) bills and prescription drugs.

From these three sources it was possible to compute the total covered utilization and benefits received for all health care services provided to the study population. Data requests were made to New York State and HCFA including the study identifiers collected in the preceding step and the start and end dates of the time periods of interest. For the LTHHCP evaluation, data were collected for twelve months following the study entry date. The resulting data files contained about a quarter of a million bills for the 1400 patient/years. These files were then aggregated to the patient/month level for analysis.

5. Analytic File Construction

The process by which the files were constructed was an elaborate and labor intensive one. In what follows, the procedures used for the inpatient utilization file are described, since (due to the availability of data from three different sources for inpatient episodes) this was the most complicated step.

The first step was the creation of a hierarchical file with two record types: the inpatient episode header record, and the detail record. Figure 3 shows the layout of these two record types. The header record included essentially just the study ID and the start and end dates of the episode. The detail record layout was an attempt to reduce the data from the different sources to a common format. The column locations specified refer to intermediate extract files and not primary data sources; the "X"s indicate where certain items were not available from a particular source. The key to the construction of this file is that the various separate system identifiers have been eliminated; what remains is the sequential study number (AAI ID),

and the utilization, charge and reimbursement amounts from the three sources in a common format.

The existence of the detail record file made it possible to carry out extensive verification and imputation steps. Since all the cases were Medicaid eligible, for every Medicare covered episode there should always be a Medicaid payment record covering the coinsurance and deductible amounts. For any stay of over one day (for Medicare beneficiaries), there should be a Medicare payment (unless the lifetime reserve days were used up in previous inpatient stays). If the corresponding bill was not present, the amount reimbursed and the number of days covered could usually be imputed from the bills that werepresent, using prevailing rates and payment amounts. Prior to imputation, however, cases with incomplete data were listed and manually reviewed for errors or inconsistencies. It is worth noting that while there is a Medicaid involvement indicator on the HCFA payment record, and a corresponding Medicare flag on the MMIS record, these were ignored since there were a substantial number of erroneous values in these fields.

There are a variety of reasons why a bill might be missing. The most frequent was probably lags in submitting reimbursement requests to the fiscal intermediary. Most hospitals billed fairly promptly, but in some cases as much as a year might go by between the date of service and the recorded payment date. It is consequently a good idea to extract bill history data retrospectively, allowing at least a year to elapse after the last date of service required. (Note that the LTHHCP evaluation was conducted prior to PPS, but the principles for file construction described here should still be applicable).

Another source of missing data was that the capture process for Part B claims can only be accomplished prospectively. That is, after a data request to HCFA for a claims file extract, the claims for the selected cases are accumulated starting from the month following the request., Thus some claims were missing from the beginning part of the study period, because the HIC numbers for these individuals had not yet been verified Obviously, this requirement has to be satisfied at the same time as the advice above about waiting a year to allow time for all of the bills to be submitted for payment.

The procedures for linking the other kinds of information (nursing home and outpatient care) were similar to those described, although each had certain unique aspects. For example, almost all of the physician bills were dated the first of the month. As specified in the provider manual, doctors' offices batch bills for submission, and consequently it was not possible to link Part B claims with MMIS information except at the aggregate level.

The final step in the database construction was to aggregate the imputed episode file to the individual patient level. The preferred unit of analysis is the calendar month, since nursing home and (as noted above) physicians offices bill monthly. Inpatient stays were apportioned into calendar months on a percentage basis. Obviously, this ignores the fact that non routine charges are not uniformly distributed across stays, but there was no way in practice to determine the dates on which specific ancillary services were provided.

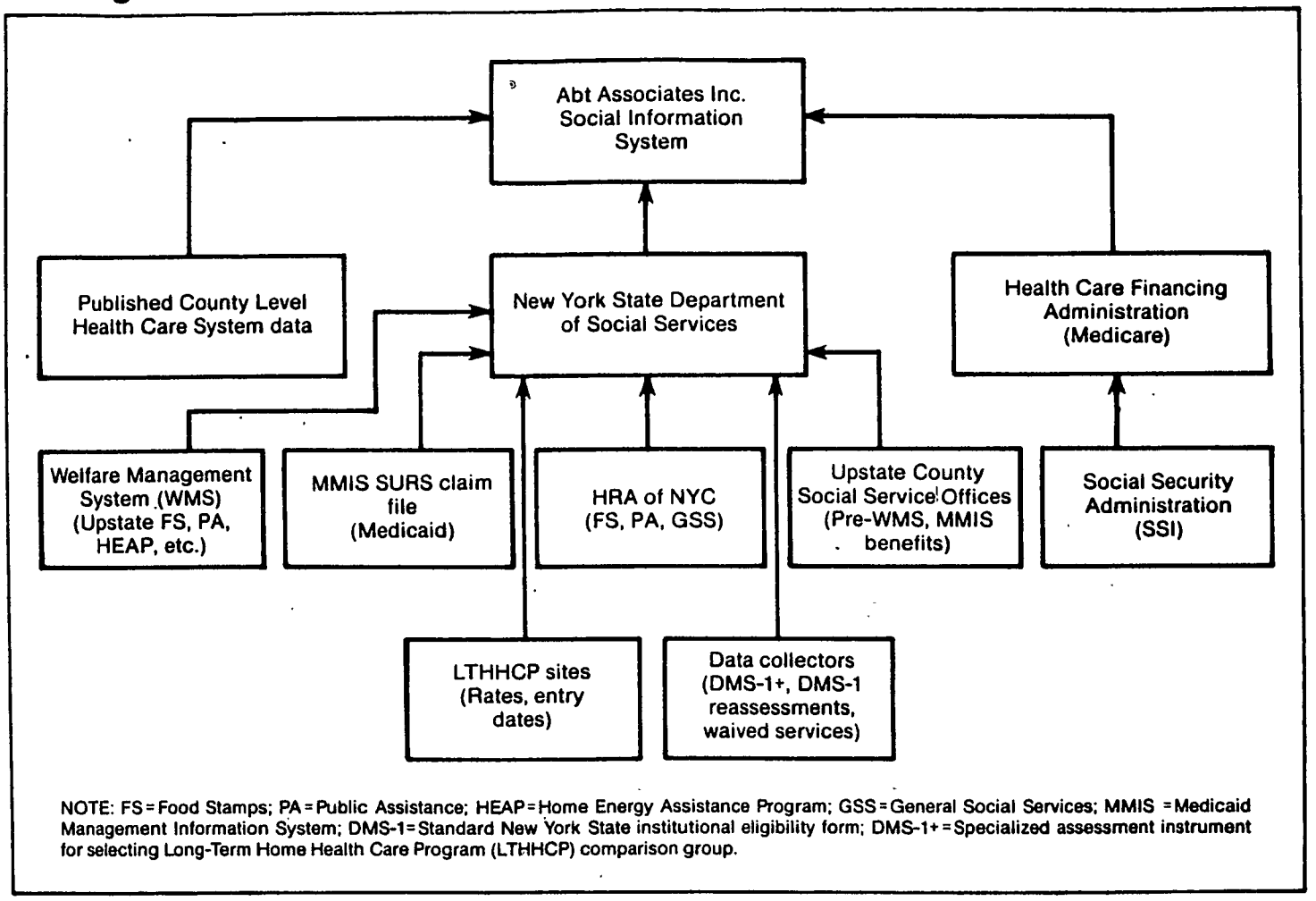
The resultant file included about 13,500 patient/month records. This is slightly less than the number possible (1400 * 12 months) because of censoring due to death. The use of patient/month records made it possible to maintain the file in a relatively convenient format; each record had a study ID followed by a fixed number of fields containing Medicare and Medicaid charges, reimbursements and utilization for each of the types of services, and totals. This file was then merged with primary data taken from interviews and medical records to form the final evaluation analytic file.

6. Conclusions

The most important conclusion to be drawn from this process is that it is possible to link the two data systems at the patient level, but only for a sample of cases. The ID verification and the payment record cross-validation steps are too time consuming and labor intensive to be conducted for a large population. At the same time, they are essential to ensure the validity of the data file. The lesson for those who would attempt to link record systems to detect fraud and abuse is that for the effort to be accurate, the cost may well exceed the savings to be realized. The most important step is the reliable determination of the individual identifiers, and given the current state of health care system record databases, the possibility of error is a matter of real concern. Nevertheless, for planning and evaluation purposes, a carefully drawn sample of individuals can be accurately matched, with results that can be of use for administrators, researchers and policy planners.

Birnbaum, H., G.Gaumer, F.Pratter and K Burke, "Nursing Homes Without Walls: Evaluation of the NY State Long Term Home Health Care Program", Abt Associates (December 1984). This research was conducted under HCFA Contract No. 500 79 0052. The author would like to thank Cathy Ellingson Otto and Richard Yaffe of HCFA for their patience and support in this effort, and to note that the views expressed here are solely his own and not of Abt Associates, HCFA or the NY State DSS.

Figure 1
Long-Term Home Health Care Program Social Information System data sources



NOTE: FS = Food Stamps; PA = Public Assistance; HEAP = Home Energy Assistance Program; GSS = General Social Services; MMIS = Medicaid Management Information System; DMS-1 = Standard New York State institutional eligibility form; DMS-1+ = Specialized assessment instrument for selecting Long-Term Home Health Care Program (LTHHCP) comparison group.

Figure 2

Medicare and Medicaid Covered Health Benefits and Utilization By Category of Service and Data Source

Category	Source	Dependent Measure
Hospital Inpatient Care	MMIS; Medicare Bill History; Medicare Claims	Reimbursement Amt. (Medicaid & Medicare); Covered Days (Medicaid & Medicare); Total length of stay Number of admissions
Skilled Nursing Facility (SNF) Care	MMIS; Bill History	Reimbursement Amt. (Medicaid & Medicare); Covered Days (Medicaid & Medicare);
Health Related Facility (HRF) Care	MMIS	Reimbursement Amt. (Medicaid) Covered Days (Medicaid)
Physician Services	MMIS; Claims	Reimbursement Amt. (Medicaid & Medicare); Number of encounters
Outpatient Services	MMIS; Claims	Reimbursement Amt. (Medicaid & Medicare); Number of encounters
Home Health Care	MMIS; Bill History; HRA/GSS; Waived Services	Reimbursement Amt. (Medicaid, by type of service; Medicare total) Number of encounters (Medicaid, by type of service; Medicare total)
Prescription Medications	MMIS	Reimbursement Amt. (Medicaid) Number of prescriptions
Other Services	MMIS; Claims	Reimbursement Amt. (Medicaid & Medicare); Number of claims

Figure 3

Hospital File
HEADER RECORD LAYOUT

Record Type	1	9	CONSTANT - 0
AAI ID	2 - 5	x(4)	
Site	6 - 7	99	
Date of Study Entry	8 - 13	9(6)	YYMMDD
Admission Date	14 - 18	9(5)	YYDDD
Discharge Date	20 - 24	9(5)	YYDDD
Number of MMIS Bills	25 - 26	99	
Number of Medicare	27 - 28	99	
TOTAL Length of Stay	29 - 31	999	(DIS-AD+1) - or - (END-AD+1) if Flag = 1
Medical Length of Stay	32 - 34	999	

Figure 3 continued

Hospital File
DETAIL RECORD LAYOUT

	MCB.DATA	BH	CLAIMS	MMIS	
AAI ID	2 - 5	307 - 310	169 - 172	23 - 26	X (4)
Site	6 - 7	311 - 312	173 - 174	27 - 28	99
Date of Study Entry	8 - 13	313 - 318	175 - 180	29 - 34	YYMMDD
Admission Date	14 - 18	150 - 154	34 - 38	(251-256)	YYDDD
Discharge Date	20 - 24	156 - 160	144 - 148	(257-262)	YYDDD
From/Service Date	26 - 30	162 - 166	181 0 185	(207-212)	YYDDD
To/End Date	32 - 36	168 - 172	44 - 48	(234-239)	YYDDD
Reimb/Payment Date	38 - 45	133 - 140	26 - 33	78 - 84	9(6)V99
MEDLOS/TOTDAYS	46 - 48	173 - 178 (176-178)	X	278 - 280	999
Deductible	49 - 54	191 - 196	110 - 114	X	9(4)V99
PART A/COINS-DAYS	55 - 60	179 - 184	115 - 117	226 - 227	9(6)
Coinsurance Amount	61 - 67	X	123 - 129	X	9(5)V99
Coinsurance Rate	68 - 72	X	118 - 122	X	9(3)V99
Lifetime Reserve Days	73 - 76	187 - 190	X	X	9(4)
MCAIDDAYS/COVD	77 - 79	X	149 - 151	228 - 229	999
INVOLV.IND	80	X	20 (A)	206 (A)	X
NON-COVDAYS	81 - 84	223 - 226	165 - 168	230 - 231	9(4)
JULIAN-DOE	85 - 89				

COMPUTERIZED LINKAGE OF REGIONAL DEATH RECORDS

Judith Baxter, Lael Gatewood, Gordon Weil, Orlando Gomez, Aaron Folsom, David Jacobs, Shu-Chen Wu, University of Minnesota; Paul Gunderson, Minnesota Department of Health

INTRODUCTION

Many public health or clinical studies can benefit from computerized systems for linking medical records to death records. In such studies, it is frequently desirable to go back to data collected in the past and determine the vital status of individuals last known to be alive at that time.

The availability of the United States National Death Index (NDI), for linkage to death records since 1979, has been hailed, correctly, as the 'removal of an impediment to epidemiologic research' (MacMahon,1983). The Social Security Administration (SSA) can also provide confirmation of death if Social Security Numbers (SSN) are available. It is clear that NDI and SSA can serve a wide variety of epidemiologic needs. Why, then, should anyone develop a regional system?

There are two major reasons. First, not all studies have the required data available. The SSA requires SSN; NDI requires birthdate or SSN. Second, NDI is not available for years prior to 1979.

The Minnesota Heart Survey (Gillum et al., 1982) had both these reasons for developing a regional system. The Minnesota Death Index (MINNDEX) was designed to carry out mortality followup on heart disease patients hospitalized in 1970. MINNDEX is a simple system developed with limited resources on a minicomputer (DEC PDP 11/70). It can provide linkage to Minnesota death certificates since 1960, and can use, for matching, as little data as first and last name, or as much as is available. Our evaluation of the performance of MINNDEX shows that in some situations, a limited system can be used to meet a variety of regional needs for linkage. This experience can help those considering development of a regional system, to understand the elements of linkage systems and the design choices to be considered.

STRUCTURE OF A LINKAGE SYSTEM

As shown in Figure 1, a linkage system deals with two databases, the user database (usually, in death linkage, a set of records for persons whose vital status is unknown), and the search database (the death records). Before linkage can be carried out, both sets of records must be coded and organized for the system. For complex systems, this pre-processing may be quite elaborate. The linkage process can be divided into two functions: the search strategy which determines which death records will be selected for detailed comparison with a particular user record, and the matching strategy which determines how the records will be compared and how the quality of the proposed match will be evaluated and reported.

Computer matching on names is generally carried out using a phonetic coding algorithm such as NYSIIS (Lynch and Arends, 1977) or SOUNDEX (Waters and Murphy, 1979) to compensate for common spelling variations. The linkage system may or may not include automated procedures for making a final decision as to whether the match will be accepted, or procedures for referring borderline cases for manual resolution. The report on records from the search database which match a particular record from the user database may include a probability factor for the likelihood that the match is correct and/or information on the quality of the match for individual fields. Confidentiality laws often determine how the matching on names can be reported.

If either database has limited identifying information or if additional information exists on the original records but not in the computer files, it may be important to evaluate cases manually, despite the risk of error and inconsistencies. Human judgement can recognize similarities of names, such as nicknames, which would be exceedingly complex to computerize. However, manual resolution is not practical for large studies.

Many large death linkage systems have been developed in recent years. An examination of the characteristics of these systems helps clarify the choices that must be made in designing such a system. (Table 1). These systems vary in the size and quality of the databases involved, the user data required, the goals and resources of the studies involved, and the strategies chosen. All the systems shown except NDI can be adapted to deal with user data bases without SSN or birthdate. All except Minnesota have many items in the death records; Minnesota has primarily name, age, and sex prior to 1976, and many items including birth date and SSN after 1975. All these systems except NDI have some way of varying the match criteria used to suit the particular study and user database. NDI now has 12 set combinations of matches on various items that define possible matches. The Canadian system "incorporate(s) virtually any refinement of logic that the human mind finds profitable to employ in a manual search" (Newcombe, 1984). This requires considerable staff resources to tailor the search to the situation. CAMLIS uses a combination of deterministic and probabilistic searches to adjust the search to different situations. MINNDEX allows the user to select a combination of items to define possible matches.

The choice between deterministic and probabilistic search strategies is very important. The CAMLIS and Statistics Canada systems are both large sophisticated systems which include probabilistic procedures to evaluate the likelihood of a correct match given

a particular pattern of matching and discrepant items. Inclusion of probabilistic linkage procedures is more complex, and can add significantly to the performance of the system by linking on the basis of partial agreement and discriminating between rare and common events.

In general, users do not know the quality of their input database and an iterative procedure for determining a strategy or weights for probabilistic procedures can be very useful. Probabilistic procedures require considerably more resources to implement than strictly deterministic ones. On the other hand, NDI can serve many users with a limited staff by providing a standardized, deterministic search.

MINNESOTA DEATH INDEX (MINNDEX)

In contrast to the other systems described above, the Minnesota Death Index (MINNDEX) was developed specifically for an epidemiological study in which very limited data were available in early death records. It allows the user to specify the weights given to matches on various items and the cutoff score that determines which combinations of items will be considered possible matches, but has no probabilistic component. It was developed on a DEC minicomputer with relatively few staff resources. When MINNDEX searches years prior to 1976, manual checking of birth dates and SSNs from death certificates is needed for verification of matches. Manual checking is feasible given the relatively modest Minnesota population size.

Within each year, the MINNDEX death certificate files are divided into 'buckets', according to the NYSIIS code of the last name on each death certificate. Since in NYSIIS, all vowels except the first letter of the name are the same (A), and since many names become "xAy..." where x and y are consonants, each bucket contains only records with the same first and third letters of the coded last name. In order to reduce processing time for records which are unlikely to match, the current version of MINNDEX only searches within the bucket that matches the first and third letters of the NYSIIS-coded last name of the user record.

The user specifies which of the following list of items will be compared in evaluating possible matches.

Sex

Last name, (NYSIIS code)
First name, (NYSIIS code), shorter string
(Ed matches Edward)
First letter of first name

SSN
Age within N years (N can be specified)
Date of Death compared to Date of Last Contact

Month of Birth
Year of Birth
Day of Birth

The user specifies a weight for each item: a positive or negative integer, or zero. These weights have been assigned on an ad hoc basis. The search algorithm is as follows: For each user record, the program finds the appropriate name bucket for the year being searched. For each death record in the bucket, the items or combinations of items with non-zero weights are compared to the corresponding items on the user record. If there is a match, the corresponding weight is added to the score for this pair of records. After all items with non-zero weights are compared, the total score is compared to the 'cutoff' score specified by the user. If the total score is equal to or greater than the cutoff, the death record is considered a 'possible' match, and is reported as such.

The report lists all 'possible' matches for each user record. The user can examine actual (not NYSIIS-coded) names, partial matches on SSN or Birthdate when available, and other items, and decide which possible matches will be considered positive matches.

EVALUATION OF RECORD LINKAGE SYSTEMS

Evaluation of linkage systems is a complex process. Sensitivity and specificity depend not only on the system design but also on the type of study being done. Study-specific characteristics include the quality of the user database, the quality of the search database, the degree to which the search strategy is appropriate for these databases, and the number of false positives or false negatives the user is willing to accept or eliminate by other processes, in order to obtain more true positives.

Many systems, including NDI and MINNDEX, are not designed to give the user a yes/no answer as to whether any potential match is 'true'. They report cases as probable and possible matches, or give a score or probability factor. They assume some further verification procedure, either manual or automated. Sensitivity and specificity can be considerably worse if they are computed before checking possible matches, whether by manual or computer means.

Examples of the performance of MINNDEX show the importance of study specific characteristics (Table 2.) In tests using data on heart disease patients known to have died in the hospital, MINNDEX achieved sensitivities (proportion of dead persons correctly matched to a death certificate) between 92% and 98%. Matches were verified and false positives eliminated by examining the actual death certificates and checking birthdate, SSN (when available), and address. All the user data were collected from 1970 and 1980 hospital records in the same way. The lowest sensitivity corresponded to cases from 1970, when hospital records were presumably less accurate than in later years, and when SSN and birthdate could not be used in searching. The values of 96% and 98% were achieved on one set of 1980 data, with the improvement due to adding SSN as an

additional matching criterion. These variations in performance depend primarily on the quality of the input database and the criteria selected for matching. For comparison, Table 2 also shows the results of submitting the same known deaths (from 1980) to NDI. The differences between the two systems were entirely consistent with the quality of the input data and the matching strategies used by the two systems.

Variations in performance can be seen in another set of tests, in which we varied the match strategy and the number of false positives we were willing to evaluate and eliminate. The test data set used was the set of 459 'known dead' individuals reported as having died in the hospital in 1970. For this test, only names, age, and sex are available in the computerized death file. The results of changing the match strategy to allow larger and larger discrepancies between the records are shown in Table 3. These variations in performance are due both to the changes in the match strategy and the quality of the input database. In this case, where SSN and birthdate could not be used, the limiting factor on performance was the fact that first names differed on the death certificate and the hospital record in about 3% of the cases. Obviously, the strategy of searching on last name and sex alone can only be used in situations where the need for high sensitivity is great.

Specificity of MINNDEX for one study was tested in a population of 2572 known living individuals using the previous year's death records as the search database. This search, which achieved 92.7% specificity without manual evaluation of matches, used the same criteria as the search which achieved 96% sensitivity on known dead cases from 1980. In an actual study, the cases with multiple possible matches could not be resolved without manual examination. With our standard procedures for manual resolution, we expect that all the false matches would have been rejected. Thus the specificity in this test series was probably 100%. These tests show that the performance of MINNDEX depends as much on the input data and the user's strategy as on the design of the system itself.

DESIGN QUESTIONS FOR A REGIONAL DEATH LINKAGE SYSTEM

Before designing a regional death linkage system it is important to think through the following questions.

1. Who will the users be? What kind and quality of user data will the system accept? Are there multiple uses or types of input data? How long is followup? What is the impact of migration out of the region on sensitivity?

2. What data are available in the search database? How large is it? Are there situations in which the users are willing to accept large numbers of false positive matches in order to get the true positives?

3. Will the search and match strategies be deterministic and/or probabilistic? Will the users be able to modify the search and match algorithms for their individual situation?

4. What kind of resources will be available to help users in setting up individualized searches if that is to be possible? Do the user records have information needed for the match criteria?

5. Will the system provide sufficient information to make the final decision in some or all cases? How will manual resolution of borderline cases be accomplished?

RESOURCES REQUIRED FOR MINNDEX

MINNDEX indicates the resources needed to develop a relatively simple regional linkage system. Development took approximately 1.5 person-years for design, programming, testing, and evaluation. In our PDP 11/70 system, one year of the Minnesota death record database (32,000 certificates) occupies roughly 4700 kilobytes of storage, and takes 3/4 hour of CPU time to generate. Searches take an average of 7.4 seconds/input record/year searched, so that running 2200 input records against 2 years of deaths took 9 hours of CPU time.

APPLICATIONS OF MINNDEX

Our experience with MINNDEX shows that a relatively simple system can be very useful in a variety of situations. Over 6000 hospital cases have been followed for 4 years for the Minnesota Heart Survey, as described above. A study of Emergency Medical Services usage after cardiac arrest also employed MINNDEX for follow-up assessment of vital status. In this case, the input data were abstracted from ambulance records for 1972 - 1982, and frequently contained first and last names only. The quality of the user records was poor enough that only 75% of the known deaths were detected by MINNDEX, and some of the names were common enough to produce large numbers of false positives to be checked. However, in this study, the investigators found it more efficient to evaluate the false positives than to use other strategies of follow-up.

Another application highlights the issue of confidentiality in using such systems. The Minnesota Cancer Surveillance Feasibility Study found that one of the institutions providing data would not release full names and did not have SSN's, but would provide initials for last name and first name, and birthdate. MINNDEX was modified to search using initials rather than names. A test using the 1980 known deaths from the Minnesota Heart Survey showed that matching using initials only was slightly less successful (sensitivity 95.4%) than matching on full names (sensitivity 96.2%). However matching using initials generated three times as many false positives to be evaluated manually (with access to the full names) in comparison to matching with full names.

CONCLUSIONS

A simple regional mortality linkage system can be useful in a variety of situations. Design and implementation of such a system must take into account both the performance needs and funding resources available in choosing among complex options. The considerable resources required to implement the most sophisticated systems can be justified by making the system available to as large an audience as possible. If record linkage systems were more available on a regional basis, they could provide more accurate population estimates of disease rates and health outcomes for epidemiological studies.

References:

Arellano M, Petersen G, Petitti D, Smith R: The California Automated Mortality Linkage System (CAMLIS): Am J Public Health 1984; 74:1324-1330.

Gillum R F, Prineas R J, Luepker R V et al.: Decline in coronary deaths: a search for explanations. Minn Med 1982; 65:235-238.

Howe G, Lindsay J: A generalized iterative record linkage computer system for use in medical follow-up studies. Comput Biomed Res 1981;14:327-340.

Lynch B, Arends W: Selection of a surname coding procedure for the SRS record linkage system. Statistical Reporting Service, USDA, Washington, D.C. 1977.

MacMahon B:The National Death Index. Am J Public Health 1983; 73:1247-1248.

Newcombe H: Strategy and art in automated death searches. Am J Public Health 1984; 74:1302-1303.

Rogot E, Feinleib M, Ockay K, Schwartz S, Bilgrad R, Patterson J: On the feasibility of linking Census samples to the National Death Index for epidemiologic studies: a progress report. Am J Public Health 1983; 73:1265-1269.

User's Manual, The National Death Index. US Department of Health and Human Services, Office of Health Research, Statistics, and Technology, National Center for Health Statistics, DHHS Publication No. (PHS) 81-1148

Waters K, Murphy G: Medical Research in Health Information. Aspen Systems Corp., Germantown, MD. 1979.

Supported in part by NIH Grants HL23727, HL25523, RR01632, the Bush Foundation, and the Minnesota Department of Health

Table 1. Characteristics of several record linkage systems.

SYSTEM	USER DATABASE	SEARCH DATABASE	CRITERIA/ STRATEGY	RESOURCES REQUIRED
NDI (NCHS)	Specific	Many items	Fixed/ Deterministic	Large
Statistics Canada (Howe)	Any	Many items	Highly Variable/ Probabilistic	Large
CAMLIS (Arellano)	Any	Many items	Variable/ Probabilistic; Deterministic	Large
MINNDEX	Any	Many items since 1975 Limited before 1975	Variable/ Deterministic	Small

Table 2. Sensitivity matching known deaths with death certificates

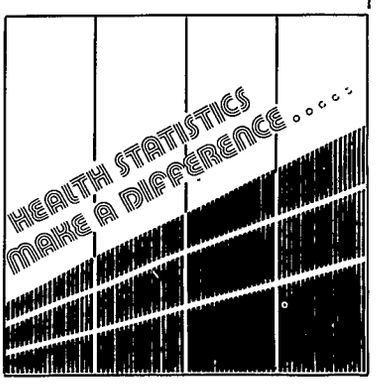
SEARCH YEAR	N	SENSITIVITY	MATCH CRITERIA (MINNDEX unless noted)
1970	459	92.5%	Last Name, First Name, Age +/- 5 Years
1980	371	96.0%	Last Name, First Name, Age +/- 5 Years
1980	371	98.3%	Last Name, First Name, Age +/- 5 Years, or SSN
1980	371	95.4%	NDI (See 1981 User's Manual)

Table 3. Sensitivity and number of 'possible' matches using MINNDEX

SEARCH YEAR	N	SENSITIVITY	MATCH CRITERIA (MINNDEX unless noted)	"POSSIBLE" MATCHES WHICH REQUIRE MANUAL RESOLUTION
1970	459	92.5%	Last Name, First Name, Age +/- 5 Years	493
1970	459	93.5%	Last Name, First Name	812
1970	459	96.3%	Last Name, Sex	15,106

Session F

Evaluation and Revision of the U.S. Standard Certificates and Reports



THE 1988 EVALUATION OF THE UNITED STATES STANDARD VITAL CERTIFICATES AND REPORTS

Mary Anne Freedman, Statistics Vermont

In the United States, national vital statistics are collected through a decentralized cooperative system. Responsibility for the registration of births, deaths, marriages, divorces, fetal deaths, and induced terminations of pregnancy is vested in 57 registration areas. These are the 50 states, the District of Columbia, New York City, Puerto Rico, the Virgin Islands, American Samoa, Guam, and the Trust Territory of the Pacific Islands.

In order to insure the uniformity necessary for national vital statistics, the responsible national agency (originally, the Bureau of the Census and currently the National Center for Health Statistics) periodically issues recommended standards. The registration areas adopt these standards voluntarily. The standards include model laws and regulations, uniform definitions, and reporting forms. The latter are the U.S. Standard Certificates and Reports.

The first standard certificates were developed by the Census Bureau in 1900. They have been revised periodically over the past 85 years. The revision process is a cooperative effort between the States and the federal government. It is accomplished through the work of an "expert" panel appointed by the National Center for Health Statistics.

The present revision panel consists of 30 members and includes state registrars and statisticians, researchers, and representatives of organizations such as the American Medical Association, the American Hospital Association, the American College of Obstetricians and Gynecologists, the National Funeral Directors Association, and the American Bar Association. The Panel works through 6 subgroups, 4 of which are concerned with certificate content, one charged with formatting the final certificates, and one responsible for evaluating the impact of the revised certificates upon the current version of the Model Law. The subgroups report to an umbrella committee called the Parent Group. The Parent Group is charged with reviewing subgroup recommendations and making final recommendations to the National Center for Health Statistics. Ultimately, the National Center will release the Standard Certificates.

At its initial meeting the Panel established objectives for its task. The primary objective was to develop a set of certificates that will meet the health data needs of the 1990's. An equally important objective was to balance statistical considerations with the legal requirements of the vital statistical system. To do this, the Panel developed two criteria that an item must meet in order to be included on the standard certificate.

These are:

1. The item must be needed for personal identification or for establishing the time and place of the event,
   or
   The item must have a high priority among the data needed for scientific or public health program purposes, and
2. There must be a basis for believing that complete and accurate information can be obtained with reasonable effort.

The Panel received input from data providers and users in several ways. Organizations who had special interests were asked to provide written testimony. A limited number of organizations and individuals also came to Panel meetings to present their views and concerns. However, the major means of input was accomplished through a survey process. The Panel developed 6 questionnaires (one for each certificate type) and mailed them to over 1,800 interested parties. The questionnaire responses were tabulated and analyzed. These tabulations provided an important focus for the panel during its evaluation of certificate contents.

The first meeting of the parent group was held in December, 1983. Since that time, the group has held 5 meetings, the most recent being in June, 1985. At the June meeting, the subgroups finalized their recommendations and the Parent Group began deliberations upon those recommendations. The information that I will present today reflects those decisions.

I'm going to direct the rest of my remarks to the proposed revisions of 2 certificates - the live birth certificate and the death certificate. I've chosen these two certificates because the Panel is recommending some radical changes in them. The panel is also recommending major changes to the fetal death certificate, but I won't deal with those changes because they are very similar to the birth certificate changes. The recommended changes to the marriage, divorce, and induced termination of pregnancy records are minimal.

The current U.S. Standard Certificate of Live Birth has several open-ended questions regarding complications, concurrent illnesses, and congenital anomalies. The Panel is recommending that these items be reformatted into checkbox responses. The rationale behind the use of checkboxes is to improve reporting for these important, but often under-reported items.

For example, the two current items, "Complications of Pregnancy" and "Concurrent Illnesses or Conditions Affecting this Pregnancy", have been combined into a single item called "Risk Factors Affecting this Pregnancy". The response list contains 20 checkbox responses - including items such as present conditions, for example, anemia and hypertension - physical attributes of the mother that may impact on pregnancy outcome, such as incompetent cervix, - historical factors, such as previous small for gestational age or large for gestational age babies - and behavioral factors like tobacco and alcohol use. The provider is asked to check all the conditions that apply, or to check "None".

The two other open-ended questions on the current certificate have been reformatted. Complications of labor and delivery include 15 items such as premature rupture of the membranes, placenta previa, precipitous labor, and fetal distress. The most common and/or most important

congenital anomalies are included in the list for "Congenital Anomalies of Child".

In addition to reformatting old items, the panel is recommending the addition of several items to the birth certificate. These include:

Obstetric Procedures: The rationale for the inclusion of "Obstetric Procedures" on the Standard Certificate is to enable the ongoing assessment of the impact of technology and interventions on safety, outcome, and health care costs.

Method of Delivery: There is little national data on delivery practices. Adding this item to the birth certificate will allow for the analysis of birthweight, gestational age, and other outcome indicators in relation to type of delivery. It would also enable us to monitor changing obstetric practices.

Abnormal Conditions of the Newborn: Currently, only birthweight, Apgar Score, and congenital anomalies are available as outcome indicators on the birth certificate. The addition of this question will help to identify other "at risk" conditions of the newborn. This item should be particularly useful in perinatal and public health program planning by identifying high-risk infants who might need special medical and other support services.

The revised birth certificate has several other modifications worth mentioning. These include:

1. A specific question regarding the type of facility in which the birth occurred (e.g., hospital, birthing center, private residence, etc.), and
2. Questions regarding maternal and infant transfers.

In addition, the panel is considering the addition of mother's and father's occupation and industry during the year preceeding birth to the certificate. The final decision regarding this item will be made in October.

I can't show you exactly what the final certificate will look like since the format group has not yet completed its work. It will, of course, be larger than the current certificate which is 8½ by 7½ inches. We expect the new certificate to be 8½ by 14 inches, with a legal section that is 8½ by 7½. The panel felt that size should not be a constraining factor in the determination of what goes on the certificate. We also note that the work done in the area of electronic transmission of birth certificate data in California and other states, will ultimately make size a less important factor.

Now I'd like to give you a brief overview of some anticipated changes in the death certificate. Currently there are 3 standard death certificates - one for use by physicians only, one for use by medical examiners and coroners, and a combined physician-medical examiner certificate. The Panel decided that as of the 1988 revision there should be only one standard certificate of death - a combined physician-medical examiner one.

The Panel's major concerns on the death certificate are the cause of death and the certifier to death.

Figure 1 presents the cause of death section of the current U.S. Standard Certificate. It is the most important item on the death certificate, and arguably, the most important item in the entire vital statistics system. The Death Subgroup spent a lot of time discussing the accuracy and reliability of the information provided by the certifier in this section. They had concerns about current practices in the certification of cause of death and they made several recommendations to address those concerns.

FIGURE 1

Immediate Cause Enter Only One Cause Per Line

PART 1	(a)	Interval between onset and death
	Due to, or as a Consequence of:	
	(b)	Interval between onset and death
	Due to, or as a Consequence of:	
	(c)	Interval between onset and death
PART II	Other Significant Conditions-Conditions contributing to death but not related to cause given in PART 1(a)	

The Panel's first recommendation addresses the need for more physician education in the area of cause of death certification. The Panel recommends that the Department of Health and Human Services, in cooperation with other interested organizations (such as the American Medical Association), develop a program for the ongoing training of physician certifiers and other involved personnel. As a first step, the Panel recommends that DHHS should convene a conference to identify interested parties and begin the development of a plan for this program.

The Panel's second recommendation concerns the format of the cause of death section of the certificate. The Death Subgroup discussed the inversion of the cause of death section so that the underlying causes would follow (Figure 2).

FIGURE 2

UNDERLYING CAUSE - List single most important disease/injury which initiated events resulting in death:

(a)_____

PART I. Resulting conditions in sequence of occurence:

(b)_____

Resulting in:

(c)_____

Immediate cause:

(d)_____

PART II. List other significant conditions contributing to death but not related to cause(s) given in Part I:

The supporters of this proposal argued that this inverted order is similar to that used in the other medical records a physician deals with

and that it conforms to the way physicians are taught to reason. However, since we have no experience with the reverse format we cannot be certain that it will result in improved cause of death reporting.

Therefore, the Panel is recommending that NCHS undertake a study to determine whether this inverted cause of death format, together with physician training, will improve the quality of cause of death data. The Panel recommends that the study be timed so that if the results warrant it, the revised format could be implemented concurrently with ICD10 (i.e., in 1993).

As an interim measure, the Panel recommends that the cause of death section be modified to provide additional instructions and clarification, to the certifier. The revision also provides additional space for multiple causes. Note that the study the Panel is recommending will evaluate improvement in cause of death reporting between this interim format and the inverted format.

The final item I'd like to discuss is the "Certifier to Death". Current practice in many hospitals requires that the death certificate be completed before the body is released to the funeral director. This often means that the certificate, including the cause of death section, is filled out by a staff physician rather than the attending physician. The staff physician's association with the case may have been (and often is) minimal. In order to improve the reporting of cause of death in these instances, the Panel has recommended that the death certificate provide for 2 certifiers. In instances where the attending physician is not immediately available, the first certifier (e.g., the staff physician) could certify to the time and place of death only. This would allow the body to be released for burial preparation. The attending physician would then certify to the cause and manner of death at a later time. The Panel has approved this recommendation in concept, although a format has not yet been finalized.

In closing, I think that the Panel is conducting a complete and careful review of the current certificates that will result in an improved vital statistics system in the 1990's. Thank you.

PROPOSED NEW ITEMS ON THE 1988 STANDARD VITAL RECORDS DOCUMENTS

George Van Amburg, Michigan Department of Public Health

In this presentation, I will focus on major changes proposed for the certificates of live birth and death. Reports of Fetal Death will most likely be patterned after the standard for live births. The Induced Termination of Pregnancy form will have few changes. Time will not permit a discussion of changes proposed for the Marriage and Divorce records. While these are important legal and statistical documents, there are relatively few changes.

The revision of 1948 was the first to include a confidential section on a vital record for the purpose of collecting additional, sensitive medical information. Three items were added in a confidential section on the standard certificate of live birth. It was not until the revision of 1968 that the medical section of birth records was expanded. The 1968 revision was considered to be quite radical. Nine new statistical items were added. In addition, some items were shifted from the non-confidential section to the confidential section. The 1968 standard certificates reflected the increased interest and demand for statistical information concerning births and other vital events that occurred during the 1960s.

The 1978 revisions were, for the most part, a simple refinement of the 1968 records with few additions. For example, the only major statistical addition to the standard certificate of birth was the Apgar score. This brings us to the 1988 revisions. The proposals for the 1988 revisions appear, after 20 years of stability, to be a radical departure from the current certificates. Not only are there new items, but major changes in document formats and methods of collecting information are being considered. It is being recommended that some items be renamed and check boxes used to provide greater clarity in communicating what information is being sought.

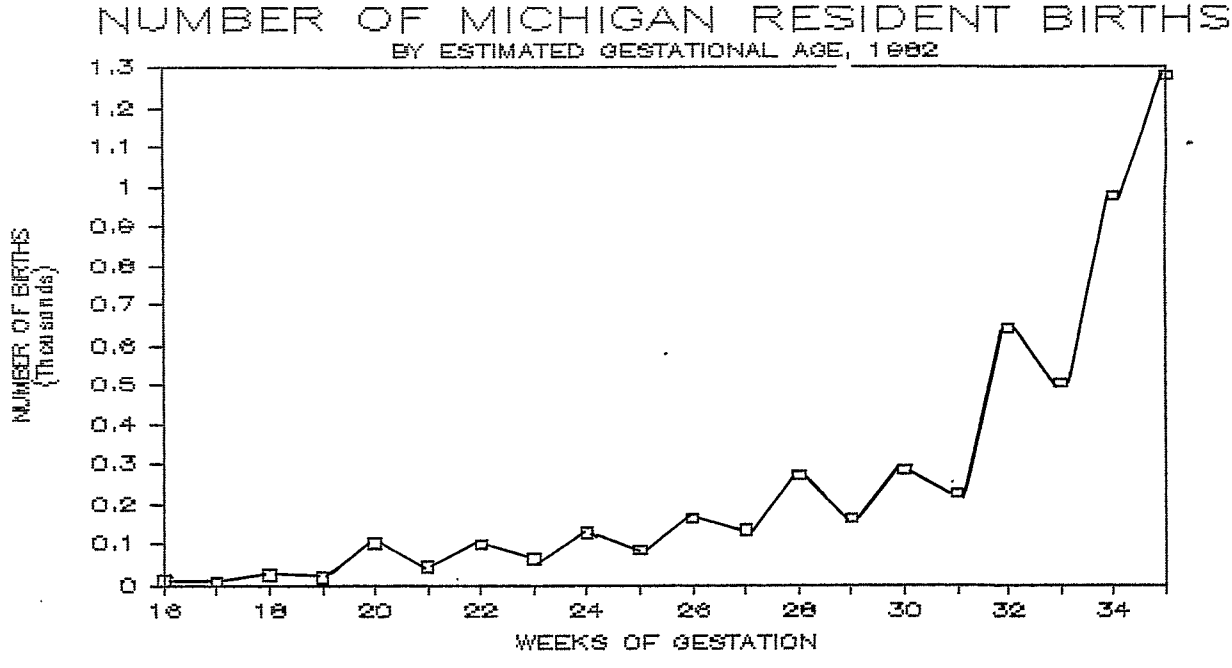
There are three types of changes being considered for statistical items on the standard certificates.

1. Changes in format and terminology to clarify what data are being requested.
2. Additional items to augment and clarify data currently being obtained.
3. Completely new data items.

1. CHANGES IN FORMAT AND TERMINOLOGY TO CLARIFY WHAT DATA ARE BEING SOUGHT

The use of revised terminology to describe the data being sought should result in more complete and better quality data for researchers and policy makers. For example, on the current standard certificate of birth the items "Complications of Pregnancy" and " Concurrent Conditions and Illnesses Affecting this Pregnancy". These items attempt to obtain information on potential risk factors that would aid in evaluating the outcome of the pregnancy. The data obtained are so incomplete and poor that not many states bother to code it.

FIGURE 1



The proposal for the 1988 revision combines these two items with a new title; "Risk Factors Affecting This Pregnancy". Further clarification of what is being sought is provided by check boxes for 20 specific risk factors. The risk factors specifically identified on the form will likely be both medical and health practice or life style. Researchers will find these data particularly helpful in evaluating adverse outcomes and identifying potentially preventable health problems.

2. CHANGES TO AUGMENT DATA CURRENTLY BEING COLLECTED TO IMPROVE COMPLETENESS AND QUALITY

The two most important items on the certificate of live birth related to pregnancy outcome are birthweight and length of gestation. Birthweight is easily obtained and accurately measured and recorded. On the other hand, length of gestation has been a problem for years. Prior to the 1968 revision the standard certificate included the item "Length of Pregnancy in Weeks". Two problems were observed with data collected in this manner.. First, is the even number syndrome. Person reporting these data tended to report even weeks of gestation. Second was the fact that the distribution was too peaked around forty weeks gestation because of the tendency to automatically report 40 weeks gestation for a term pregnancy. Fifty-three percent of the records indicated an estimated 40 weeks gestation whereas using the LMP date to calculate a gestational age resulted in 20 percent at 40 weeks gestation.

TABLE 1

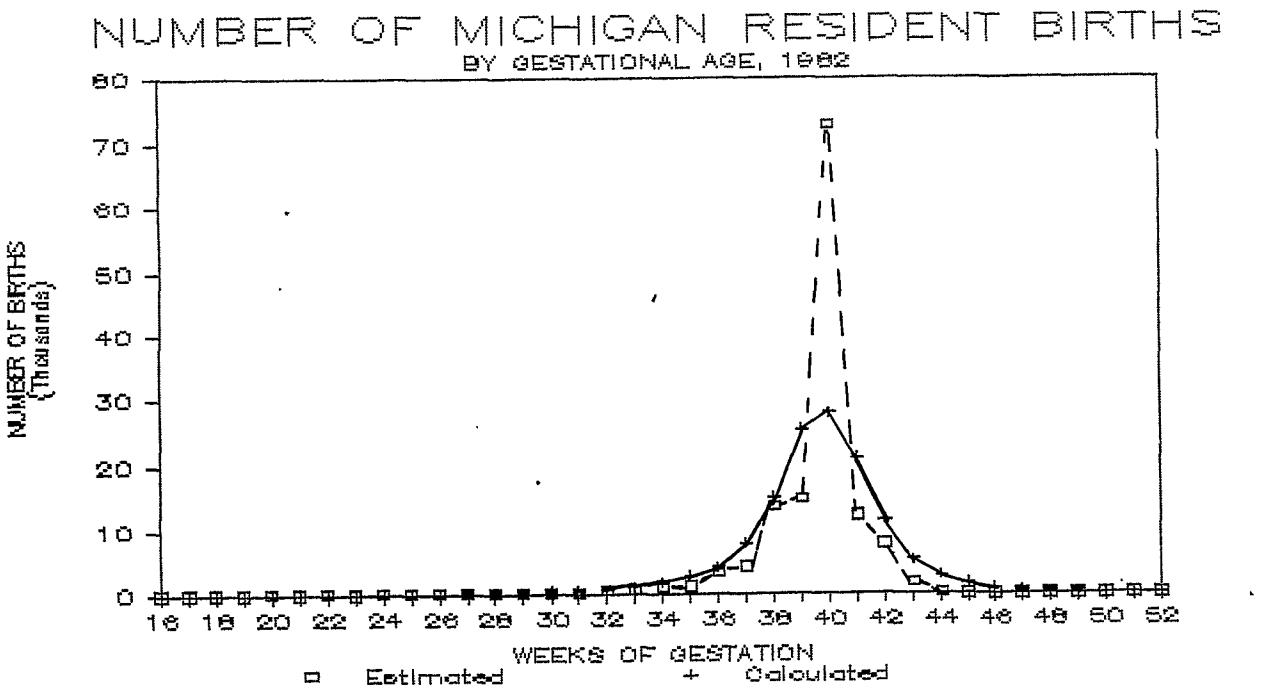
COMPARISON OF CALCULATED WEEKS OF GESTATION
vs
PHYSICIAN ESTIMATED WEEKS OF GESTATION
MICHIGAN RESIDENTS, 1982

Date of Last Menses	Estimated Wks Gestation		TOTAL
	Present	Missing Invalid	
All Complete	101,335	1,148	102,290
Day of Month Missing	32,109	323	32,432
Missing /Invalid	3,718	132	3,850
TOTAL	136,969	1,603	138,572

In 1968 the item was changed to "Date of Last Normal Menses" It was thought that a better distribution of gestational age could be obtained. As can be seen by the green line on the slide this was correct. However, the number of certificates with these data missing increased. In addition there is a bias in reporting the day. The fifteenth of the month followed by the 1st, 10th, 20th, and 25th were most popular. Interpolation methods were devised to handle these cases and those where part of the date was missing. In Michigan, we added the LMP date in 1968 but then added back a question on estimated weeks gestation in 1978. As a result from 1978 to present, both LMP date and estimated weeks gestation were requested on the Michigan certificate. This is what is being proposed for the 1988 revision. The reason can be seen in Table 1.

Slightly over 2.5 percent (3,718) of the certificates have a non-useful "Date of Last

FIGURE 2



Normal Menses" but only a total or 138 certificates have both the LMP date and "Estimated Weeks" as not useful or missing. In approximately 1 percent of the cases where the LMP date is technically usable, we substitute the estimated gestation for the calculated because the calculated results in an unreasonable length of gestation (outliers). With both items, researchers will have the opportunity of obtaining good statistics on length of gestation for essentially the complete cohort of births.

Statistics on the attendant at birth have, in the past, been obtained by coding the written title as provided by the attendant. Data obtained in this manner are not complete. Recently researchers and policy makers have become more interested in monitoring changing patterns of attendants at birth. As a result, a check box item for the title of the attendant at birth is recommended for the 1988 standard birth certificate. The Michigan certificate has included a check box item for "attendant at birth" since 1978.

TABLE 2

LIVE BIRTHS BY TYPE OF ATTENDANT AT BIRTH
MICHIGAN OCCURRENCES, 1979 and 1983

Attendant	1983	1979	1983%	1979%
M.D.	112486	121590	85.3%	84.9%
D.O.	18428	20967	14.0%	14.6%
Nurse	104	137	0.1%	0.1%
Midwife	171	32	0.1%	.0%
Nurse-Midwife	148	12	0.1%	.0%
Physicians Assistant	12	8	.0%	.0%
Husband	259	257	0.2%	0.2%
Other	147	170	0.1%	.0%
No Attend/ Unknown	83	67	0.1%	.0%
TOTAL	131838	143240	100.0%	100.0

As can be seen with these data, the vast majority of deliveries are reported to be attended by an MD or DO. However, these data do show that the number of deliveries reported to be attended by midwives and nurse-midwives has increased in the last four years. The addition of this item to the standard certificate will give researchers information on the geographic variability of attendants. The data can also be used in analyses of pregnancy outcome.

3. POTENTIAL NEW ITEMS

DEATH RECORD

Three new statistical items are being considered for the death record.
1. Hispanic Origin or Descent
2. Education of Deceased
3. Place of Death if Other Than Hospital

An item on Hispanic origin or descent is being considered for all standard certificates or reports (birth, death, fetal death, Itop, marriage, and divorce). The research implications for these data are obvious. The data can be used by researchers and policy makers for identifying health issues in this population subgroup and to evaluate the impact of various programs. It will also be possible to identify cohorts of Hispanics dying of certain diseases for retrospective epidemiological studies. As with statistics on Native Americans obtained from vital records, there will be some concern over the completeness, quality, and thus, usefulness of the data collected. An item on the education of the deceased is being considered in order to provide, in combination with occupation, a reasonable surrogate for economic status of the deceased. Research has indicated that these two items combined into a single index provide a better estimate of economic status than either item alone. The result of having this addition will be the availability of cause of death information by a reasonable surrogate for economic status.

Place of death, if other than a hospital, is being proposed in order to provide statistics on changing patterns of medical care especially for the terminally ill. This item has been on the Michigan record of death since 1978.

TABLE 3

ACTUAL PLACE OF DEATH
MICHIGAN OCCURRENCES
1979 and 1983

Actual Place	1983	1979	1983%	1979%
Hospital	49658	47077	65.8%	64.9%
Nursing Home	12040	11390	16.0%	15.7%
Home	11629	10549	15.4%	14.6%
Ambulance	112	164	0.1%	0.2%
Other Institution	36	41	.0%	0.1%
Extended Care Facility	392	39	0.5%	0.1%
Other	1433	1973	1.9%	2.7%
Unknown	143	1267	0.2%	1.7%
TOTAL	75443	72500	100.0%	100.0%

As you can see from Table 3, the vast majority of deaths occur in hospitals and for the last few years the pattern has remained essentially constant. Researchers and policy makers expect the pattern to change over the next decade. The availability of these data from the death record will allow easy monitoring of any changes in the pattern.

BIRTH RECORD

Six new items are being considered for the standard certificate of live birth.
1. Hispanic Origin or Descent
2. Occupation of Mother and Father in the Previous Year

3. Mother or Infant Transferred
4. Obstetric Procedures
5. Method of Delivery
6. Abnormal Conditions of the Newborn

Consideration is being given to including the occupation of both parents on the birth record and the fetal death report. Proponents for including this item argue that it would provide information on possible occupational hazards and poor outcomes. If added, the item in combination with education would also provide a surrogate measure of economic status.

Two additional items concerning medical procedures are proposed: obstetric procedures; and method of delivery. Clearly the literature indicates that the distribution of types of deliveries has changed significantly over the last ten years especially with respect to C Sections. It is expected that the pattern may continue to change as the result of medical research and cost containment efforts. The availability of information on the method of delivery for the complete cohort of births will provide researchers with a readily available source of information upon which to study geographic variations in methods of delivery as well as examining this variable with respect to adverse outcomes.

It is proposed that data for six specific types of obstetric procedures be captured on birth records. Since the last certificate revision, many new procedures for the evaluation of the unborn child have seen widespread use. However, there are no national data on the use of these procedures with respect to health risks and pregnancy outcome. These data from birth records combined with the other new and revised data elements on the standard certificates would allow for such an analysis.

The inclusion of a data element with respect to the transfer of an infant or mother will allow for the evaluation of facility performance with respect to the transfer of high risk mothers and infants to facilities equipped to handle these exceptional cases.

The standard certificate of live birth has classically contained little information on the outcome of the pregnancy. It is proposed that an item titled "Abnormal Conditions of the Newborn" be added to increase information available on the condition of newborns. This would be in addition to the Apgar score and Congenital Anomaly items. These three items along with the information on risk factors, obstetric procedures, method of delivery, and demographic characteristics of the parents will allow for a much more complete analysis of births than has ever been possible before without resorting to costly follow back studies.

SUMMARY

In a nutshell, the use of check boxes and revised terminology should allow easier completion with responses focused on what is considered important for health statistics and research. Several states using check boxes report increased responses. Many of the new statistical items being proposed will be a welcome addition to state and national vital statistics data bases.

However, there are also some drawbacks to the proposals that need to be recognized particularly with respect to the format changes. First, it will be easy to check the wrong box. As a result, the accuracy of the data may be questioned. Second, it is likely that if the information is not identified as a check box it will not be recorded no matter how important it might be in evaluating the outcome of the event. Third, persons using these data for health policy analysis may find discontinuities in time trends for those items that are important for their studies.

Fourth, epidemiologists, public health program persons and other health researchers that like to look at the medical terminology on records will find check boxes unsatisfactory. For example, the proposed standard birth certificate calls for the item cleft lip/palate under the heading congenital anomalies. Many researchers consider these completely separate anomalies. Collecting data in the proposed manner will not permit the analysis and tabulation of these two distinct anomalies.

On balance, it appears the changes being considered will deliver more accurate, complete and useful data resulting in improved state and national data bases.

PROCEDURES FOR STATE IMPLEMENTATION OF STANDARD CERTIFICATES

Irvin G. Franzen, Kansas Vital Statistics

INTRODUCTORY NOTES

In discussing the subject of the Implementation of Revised Standard Certificates we are of course referring to the revised forms of the birth, death, fetal death, marriage, divorce and induced termination of pregnancy certificates or reports that are scheduled for adoption on January 1, 1988. The execution of all the necessary details between now and the effective date of the revisions, as well as the essential follow-up activities, represent a truly awesome administrative responsibility. It is therefore important that we alert ourselves to all that has to happen, what steps of our respective responsibilities have to be completed when, and what kinds of questions must we be ready to answer along the way in order to assure a uniform and efficient revision process.

Perhaps some of the details may seem elementary to the more experienced state executives but I'm sure they'll gladly bear with those that are less experienced. On the basis of the length of service, we apparently do represent a wide range of experience in administering the implementation of revised certificates. Some of us, I'm sure, have never been involved in such a transition, some of us were probably on the scene when it happened but were not so deeply involved as to bear much of the responsibility and some of us have had the responsibility but may have forgotten about some of the procedures or may have some good ideas on how techniques should be modified in keeping pace with other changes. At any rate, whichever of these categories might apply, it will now be a matter of learning, of reviewing or of being the experts that answer all the questions or that fill in the gaps that others of us may overlook.

Since my assigned subject is "Procedures for State Implementation of Standard Certificates," I have prepared a schedule of some of the important steps that will have to be carried out state by state and have also included some historical notes concerning certificate revisions. I have tried to incorporate the significance of uniformity and I have noted some basic principles that I believe are important to bear in mind during the transition process. Finally, I have recommended certain measures for prompt action. I am not proposing that my recommendations or suggested schedules represent the final or the best answers to an effective implementations of the revised standard certificates, but hopefully their provocative purpose may help to point us in the right direction as we pursue the necessary implementation steps.

HISTORICAL NOTES

Historical accounts indicate that ever since the turn of the century standard certificates have served as the principal means for gaining uniformity in the minimum content of the documents used to collect information on vital events.

Throughout the span of the development of the registration system, the periodic revisions of standard certificates have been a careful and democratic process conducted by the national vital statistics agency in consultation with State health officers and registrars; with federal agencies concerned with vital statistics; with national, state and county medical societies; with representatives of the legal profession and with others working in the fields of public health, social welfare, demography, sociology and insurance. This revision process has assured careful evaluation of each item in terms of its current and future usefulness for registration, identification, legal, medical and research purposes.

In reviewing historical accounts, it's also noted that the Association of Registration Executives was actively involved in the various stages of the development of the registration system and particularly in the endeavors to bring about uniformity in procedures and records forms. This organization of course is currently the Association for Vital Records and Health Statistics and most of us are fully aware of its deep concern about the needs for improvements in standard forms and in uniformity of the whole registration process, as evidenced by the time and financial resources given to these endeavors approximately at decennial intervals.

The attached pages list some of the particular years or span of years that mark some of the important developments relative to the adoption of standard certificates and attempts to achieve uniformity in vital statistics procedures.

HISTORICAL NOTES RELATIVE TO THE DEVELOPMENT OF STANDARD CERTIFICATES AND UNIFORM REGISTRATION PROCEDURES

Span of Time	Noted Activity
1850	Collection of the first national vital statistics thru the mechanism of the Seventh Federal Census.
1850 - 1900	National vital statistics were collected on a decennial basis along with other census data.
1880	Bureau of the Census established a national registration area for deaths, including two states, the District of Columbia and a number of cities.
1880 - 1900	The Census Office consistently advocated national uniformity in State supervision, in basic procedures, and in the forms used for the registration of deaths.

1887 Congress passed an act directing the Commissioner of Labor to collect statistics on marriages and divorces for the years 1867 to 1886.

1900 Effective year of the use of the first standard death certificate adopted in full by 12 states and in part by 6 states and the District of Columbia.

1900 The beginning of annual compilations of death statistics representing 10 states, the District of Columbia and a number of cities located in non-registration states.

1902 The Bureau of the Census was made a permanent full time agency by an act of Congress which also authorized the Director of the Bureau to obtain annually copies of records filed in the vital statistics offices of those states and cities having adequate death registration systems.

1905 The President recommended in a special message to Congress that the Director of the Census be authorized by appropriate legislation to collect and publish statistics pertaining to marriages and divorces covering the period of years from 1886 to 1906. Thereafter estimates and surveys were made for limited annual compilations of marriage and divorce statistics.

1915 The Bureau of the Census established the national birth registration area, including 10 states and the District of Columbia.

1900 - 1933 The fundamental task of the Bureau of the Census in the field of vital statistics was to extend the birth and death registration areas in accordance with established standards.

1933 The birth and death registration areas were completed to include all States.

1934 The initial bringing together of state registrars for work conferences to exchange viewpoints and unify registration practices by cooperative agreements.

1939 The first official recommendation of the three revised standard certificates of live birth, death and stillbirth for simultaneous adoption by all States.

1940 Beginning of collection of marriage and divorce transcripts from States able to provide them from their State offices of Vital Statistics.

1941 The Division of Vital Statistics of the Census Bureau was called upon by State registrars to aid in the development of acceptable standards for the filing of delayed birth certificates, and subsequent conferences between federal agencies and state representatives resulted in a set of recommendations that were incorporated in a Manual of Uniform Procedures for the Delayed Registration of Birth.

1943 A recommendation agreed to by the Budget Division, the Association of State and Territorial Health Officers and a special Commission on Vital Records, emphasized the need for a cooperative vital records system with the coordinating responsibility placed in a single national agency.

1946 The National Office of Vital Statistics was established in the U.S.P.H.S. with explicit responsibility for federal functions in vital statistics.

1949 The Public Health Conference on Records and Statistics was officially launched, having been conceived as a working arrangement with a broader scope than previous conferences in that its activities would now be expanded to embrace the whole field of public health statistics in addition to that of vital records and vital statistics.

1949 Effective year of the first revisions of the standard certificates of live birth, death and stillbirth after becoming part of the Public Health Services. Some of the most significant changes included the establishment of a medical and health section on the standard birth certificate, the revision of the death certificate's medical certification in accordance with the form recommended by the World Health Organization for use with the Sixth Revision of International List of Diseases and Causes of Death, and adjusting items on the stillbirth certificate to correspond to information being collected on birth certificates.

1954 Standard Records of Marriage and Divorce or Annulment were approved by the Public Health Service and the Public Health Conference on Records and Statistics and were recommended for adoption by all States.

1956 Effective year of revision of the standard certificates of birth, death and fetal death. The principal revisions included the addition of the questions "Is residence inside city limits?" and "Is residence on a farm?"; on the death certificate the wording of the cause-of-death item was changed to improve its clarity and the questions on operations were omitted and on the fetal death certificate the cause item was revised to conform to the wording used on the death certificate.

1957 The MRA was established with 32 States and 2 other areas participating.

1958 The DRA was established with 16 States and 1 other area participating.

1968 Effective year of revision of standard certificates. One of the most controversial changes was the removal from the death certificate of the item "Was Decedent Ever in U.S. Armed Forces?"

1971 The beginning of providing NCHS with State-coded computer data tapes through the Cooperative Health Statistics System.

1978 Effective year of revision of standard certificates. Birth certificate revisions included the addition of items on 1 and 5 minute Apgar scores, the deletion of the item on birth injuries and changes in wording on legitimacy status and previous pregnancies. The item "Was Decedent Ever in U.S. Armed Forces?" was restored to the death certificate and the question on "whether or not autopsy findings were considered in determining the cause of death" was dropped. The marriage record was revised to include a section providing the license to marry and items on the title of whoever performed the ceremony and whether it was a religious or civil ceremony. The divorce record was revised to delete legal grounds and to whom decree was granted and change children involved to children born alive of this marriage and the number under 18 years old at the time of dissolution.

1980 The MRA had grown to 42 States and 3 other areas and the DRA to 30 States and 1 other area. Machine readable data tapes were being provided to NCHS, via VSCP contracts, for both marriages and divorces by 8 States and for marriages only by 4 States.

1985 All States are under VSCP contract to provide data tapes to NCHS on births and deaths (demographic).

CONTINUING NEEDS FOR GREATER UNIFORMITY

In reviewing the development of the registration system since the first standard certificates were adopted at the turn of the century, it is obvious that thru considerable difficulties some progress has been made. It is also obvious that we have not yet achieved the optimum in uniformity of records and procedures and that there is a definite need to continue to do something about it.

Some of the evidence of lacking uniformity has been vividly portrayed during the standard certificate revision process over the past couple of years. Considerable differences have been noted in the color, size and item arrangement of standard certificates as well as in its content and in the wording of certificate items; wide variations exist in the type and extent of instructions provided for hospitals, funeral directors, and local officials involved in the registration process; differences in laws and regulations pertaining to registration practices continue to represent a problem, and important differences exist in state office procedures in querying incomplete and inconsistently reported data. Further evidence of lacking uniformity is reflected by the fact that there are still 8 States not in the Marriage Registration Area and 20 States not in the Divorce Registration Area.

To consider some of the evidence of basic needs for uniformity, the respective standard certificates must be regarded in their dual role of serving as legal instruments and as the source of valuable statistical data.

From the standpoint of vital records as legal documents I believe we should especially be concerned about uniformity in the following respects:

1. To create and maintain a favorable public image of the total registration system. Standard certificates and uniform procedures across state lines become increasingly important in this respect with the increasing mobility of the population. At the present time approximately 2.6% of all registered births and about 3.7% of the registered deaths occur outside the registrant's state of residence. Numerically this represents approximate annual totals of 95,000 births and 75,000 deaths. It is also noted that in 1980 nearly one-third of the U.S. natives were living in a different state than their state of birth.

2. To strengthen the whole registration system and the respective positions of the individual states for a united stand when challenged by law suits or confronted with proposed unfriendly legislation relative to any standard certificates or registration procedures. In our experience whenever we've been faced with law suits or legislative proposals regarding particular certificate items, our position has certainly been enhanced by the evidence that the terminology or data item in question is uniformly used by the various states, by the NCHS and by the Bureau of the Census.

3. To achieve universal acceptance of the respective vital records as valid documentary evidence. When another governmental agency gives delayed birth certificates a third rate priority as documents on proof of age, all is not well.

We have had a Manual of Uniform Procedures for the Delayed Registration of Births and model laws pertaining thereto since 1941 but apparently the intended goal of uniform principles has not yet been achieved. At least, the variations among the States in establishing delayed birth certificates has been referred to as a major difficulty in determining what kind of documentation represents adequate proof-of-age.

From the standpoint of vital records as statistical documents, I believe we should be particularly concerned with uniformity in the following respects:

1. To achieve greater efficiency in processing statistical data at the national level for the maximum usefulness of the collected data. The history of the registration system shows that statistics collected through vital records have through the years served as the basis of many useful health indexes and that there have also been significant advancements in data processing and analytical methods. With these increasingly sophisticated research activities and with the ever increasing mobility of the population, well coordinated regional and national studies in many cases will most efficiently yield health and demographic indicators that are valuable to the individual states as well as the nation.

2. To be most effective in keeping a current account of health problems within the individual state and its communities, while at the same time accumulating useful information for local, regional and national comparisons.

3. To maintain consistency with established classification systems that represent denominators for basic vital statistics rates. One of the principal values of vital statistics data depends upon the ability to compute meaningful rates in which the vital events of a given class are related to the population of a like defined class. Vital statistics and population statistics must therefore be consistently classified for valid tabulations of comparable groups. Hence it behooves us to always pay attention to what the Census Bureau is doing and to maintain an effective working relationship and line of communication with any agencies whose activities significantly impact upon the value and usefulness of vital statistics.

PROPOSED SCHEDULE OF PROCEDURES FOR STATE IMPLEMENTATION OF STANDARD CERTIFICATES

Approximate Time	Specific Activities
July-Dec., 1986	Prepare 1988 fiscal year budget to provide for printing, postage and extra field work expenses.
July-Dec., 1986	Become completely familiar with all revisions being made and the reasons therefor, and finalize any state additions to the standard items.
Jan.-Mar., 1987	Submit revised certificates to State Health Officer or other Chief Administrator and to the official Board or Commission that needs to approve any changes.
Jan.-Mar., 1987	Submit well substantiated proposals for statutory changes to state legislatures where this is necessary for the revision of certificate items, and follow any such changes with any necessary pursuant regulatory changes.
Apr.-June, 1987	Send copies of revised forms to the respective local officials, hospitals and funeral directors involved in the registration process, and include summary explanations of the revision process, referring to previous communications concerning this and informing that this is the final result of the deliberations that have involved representatives of their professions and interests as well as state and national registration executives, and further informing of the remaining steps in completing the transition.
Apr.-June, 1987	Send copies of revised forms to other agencies and to professional societies to inform of the revisions being made, and to solicit their interest and cooperation in calling any special problems to our attention or otherwise aiding in the implementing of the new standard certificates.
Apr.-June, 1987	Prepare Instruction Manuals or at least Instruction Sheets to accompany the new certificates when they are distributed.
Apr.-June, 1987	Send notifications of scheduled revisions to printing plants, particularly in areas where local units handle their own printing of certificate forms.

July-Sep., 1987	Submit orders for printing an adequate initial supply of the revised certificates, of hospital and other worksheets and of instruction manuals.
Oct.-Nov., 1987	Type or print address labels and prepare for mailing of forms and manuals.
July-Dec., 1987	Field contacts, utilizing all available personnel to visit with local officials concerning problems of implementing the new certificates.
Nov.-early Dec., 1987	Mail standard certificates and instruction manuals to hospitals, local registrars, funeral directors, coroners, and to county officials involved in marriage and divorce registration.
Dec.(last week) 1987	Send final reminder that at midnight December 31 the use of new certificates is to begin and that all old certificates are to be disposed of promptly thereafter.
Jan.-Mar., 1988	Follow back by telephone contacts, field visits and correspondence on any incoming old certificate forms.
April, 1988	Field visits to any local officials still using old certificates to personally exchange new forms for the old ones and checking out any problems relative to the transition.
Apr.-June, 1988	Letter to all local officials, hospitals and funeral directors to express gratitude for their cooperation in the transition process and to invite their questions and comments regarding any particular problems with the use of the new forms.
1988 and thereafter	Monitor and evaluate the accuracy and completeness of data collected, keep in touch with data providers, be sure that new records personnel at the collecting points are appropriately trained and provided with instruction manuals, and insist that data analyzers and users understand the meaning and limitations of what has been made available via the vital statistics system.

ELEMENTS OF A SUCCESSFUL TRANSITION

In planning and executing the necessary details to implement the transition to new standard certificates, it is also important that the process be approached with the kind of attitude and fortitude that will yield effective results. I believe the following represent some important basic elements in achieving a coordinated and successful transition.

S-cheduling - There must be timeliness in planning a step by step procedure to effect the transition process, allowing ample time to accomplish each phase of the total process.

U-niformity - To efficiently achieve our goal it's important to use a standard and well synchronized approach from state to state and across the nation so that local officials, hospitals and funeral directors will be informed and their cooperative action simultaneously enlisted.

C-larity - We must clearly explain the changes to be made in terms that are applicable and meaningful to the respective officials and organizations involved.

C-onviction - Our attitude must show that we believe in what we ask hospitals, funeral directors and local officials to do in accomplishing the transition and that we firmly believe in the value of revising standard certificates.

E-ffort - We must be willing to expend considerable energy in carrying out the step by step process, in dispensing the applicable information to all people involved and in answering all questions pertaining to their respective responsibilities.

S-ubstantiation- We must become so well informed ourselves that we will be ready to support with convincing and complete information any of the changes we are proposing.

S-agaciousness - We must maintain a keen sense of perception regarding the motives of those that challenge us or propose deviations from committed standards, or misuse the data being collected. Along with that, we must be able to exercise calm and sound judgment in handling criticisms of the process by officials at any

level, even in legislative chambers.

If each of us includes all of the above in our approach, it can't help but spell SUCCESS in effectively implementing the revised standard certificates.

RECOMMENDATIONS FOR SPECIFIC ACTION TO ENHANCE THE PROCESS OF STATE IMPLEMENTATION OF STANDARD CERTIFICATES

1. That the Executive Committee of the AVRHS confer with representatives of the Public Health Conference on Records and Statistics to plan for the preparation of standard letters of introduction to local officials regarding the revised items on the respective certificates and the reasons for the revisions, along with a brief summary of past revisions and the types of valuable information that have resulted.

2. That the AVRHS work with NCHS staff in developing a schedule of activities designed for a smooth transition.

3. That a committee of experts be designated to be available to assist or advise any states that encounter any special transition problems.

4. That a committee or study group be assigned the task of compiling a manual on standard procedures, including the Uniform Vital Statistics Act, Uniform Regulations, standard certificates, and references to other pertinent information and published papers concerning uniform registration practices and principles; and that the entire compilation be prepared in the form of a loose leaf manual so it can be kept up to date in serving as an introduction to the standard system when new state registrars arrive on the scene as well as being a handy reference when questions about reasons for certain items on certificates are posed by physicians, hospitals, local officials, legislators, budget analysts and others.

5. That the AVRHS and NCHS collaborate in developing some well prepared budget attachments that can be adapted to the use of the various states in seeking increased budget allocations for communications, printing, field activities and other categories involved in the effective implementation of standard certificates.

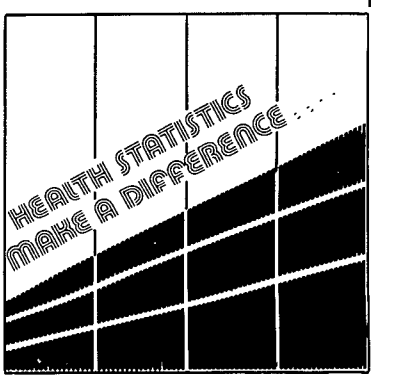
6. That the model handbooks on registration of vital events, as published by NCHS in 1978, be updated along with the certificate revisions and the AVRHS Executive Committee collaborate with NCHS in effecting a generous and timely distribution thereof to all states.

7. That in view of the fact that an important reason for certificate revisions is to improve the quality and type of statistical data, and since such data are needed for the planning and evaluation of many public health activities at the state and national level, every effort should be made to make grants available to all registration areas to support the costs of implementing the revised standard certificates.

8. That all executives in state and city vital statistics offices strive for maximum uniformity. Since lack of uniformity has sometimes been responsible for unpleasant encounters with other governmental agencies and with the public, and since we are concerned about producing valuable statistical data, it behooves us to use our individualities wisely. In other words, let's don't deviate from a standard procedure or a proposed standard certificate unless we have substantial reason for doing so.

Session G

Evaluating the Cost and Use of Health Services: Examples from Medicaid and Medicare Populations



MONITORING POLICY DECISIONS FOR QUALITY OF CARE IMPACTS: THE CASE OF MEDICAID COPAYMENTS

Timothy J. Tyson, Wisconsin Department of Health and Social Services

State level programmatic decisions are routinely implemented as policymakers attempt to mold the health care system so that it operates in a more effective and efficient manner. Today concern over rapidly escalating health care costs has moved the debate over the allocation of health care resources into an arena where in many instances, evaluators and policy analysts can have much greater input than in the past. Many current health care reform activities and proposals entail the rationing of medical care either directly or indirectly. Consumer cost-sharing, capitated payment plans, reimbursement by diagnostic grouping and changes in the tax treatment of health insurance premiums are only a few examples.

The issue raised by economic solutions to health care problems as aptly noted by Schwartz (1983) is that "treatment will not be provided to everyone who might benefit from it." For the evaluator, this highlights a particular problem; namely, the need to encompass in their evaluation designs, a means for assessing patient care effects of policy decisions that are chiefly viewed as economic issues.

In this paper, an evaluation is presented that not only evalutes the cost impacts of a policy decision to implement recipient copayments for Medicaid recipients, but also assesses the quality of care impact of the policy. The argument is advanced that a simple and easily implemented quality assessment tool can enhance the integrity and utility of an evaluation and, in turn, improve decision-making.

PROGRAM EVALUATION AND QUALITY ASSESSMENT

Program evaluation tends to be a much broader term than quality assurance. Program evaluation concerns itself with activities of health providers and the performance of the health care system as well as the direct provision of services to patients (Donabedian, 1978). Program evaluation entails an aggregate level analysis of performance to determine if it meets organizational or societal goals. Quality assessment, on the other hand, has been the domain of physicians and has focused on individual patient needs. Questions of efficiency and equity were left to others. Actual quality assessment methods have followed a similar vein with physicians resisting assessment approaches that lacked a clear clinical foundation (Brook, Lohr, Chassin et al., 1984). Donabedian (1978) argues that the distinction between program evaluation and quality assessment has become blurred and that changes in the organization and financing of health care have resulted in a greater need for quality assessment to concern itself with collective issues.

Geographic variations in the use and cost of services have been noted by a number of investigators and these findings have heightened interest in the use of indicators for assessing the performance of the health care system. Significant variations between geographic areas have been documented and attributed to a variety of factors such as financing mechanisms, the supply of providers and characteristics of the delivery system (Roos and Roos, 1982; and Wennberg and Gittlesohn, 1975, 1982). These wide variations have raised questions regarding the impact of more services on the costs and quality of care. Brooks, Lohr, Chassin et al. (1984) in a recent article discuss the missing clinical link between indicators of use of service and actual use and admonish physicians to become involved in providing insight to efficacy questions. If they do not become involved, "Unwise cuts (in services) will be made; wise cuts will not. Physicians may lose the battle to those who can at least count the costs of medical care accurately."

To the evaluator this can be viewed as an opportunity. An opportunity, with sufficient clinical input, to incorporate a patient care dimension to many evaluations that tend to be predominantly fiscally directed. Difficult to measure attributes such as the quality of medical care present special problems for the evaluator and as a result are often omitted from the analysis (Loveland, 1980). This is likely to contribute to the under-utilization of the evaluation; a situation that Guttentag (1977) has gone so far as referring to as a failure on the part of the evaluator. The evaluator desiring to provide a relevant product and at the same time maintain scientific rigor can only follow Campbell's (1969) advice to "do the best we can with what is available to us." This is the stance of this paper as well.

Before presenting the particular example of copayments for Medicaid services, a brief review of quality assessment approaches will be presented. The case will be made that the complexity and expense involved in applying traditional quality assessment methods are so great that they are not likely to be used in routine evaluations of policy decisions. Rather, a more modest quality monitoring or screening measure can be successfully applied consistent with the dictum of doing the best with what is available.

REVIEW OF QUALITY ASSESSMENT METHODS

Most observers would agree that the assessment of the quality of health care is a nascent art and is a long way from providing an answer to the question of what constitutes quality of care. Donabedian (1978) has pointed out that variations in the quality of care are not random phenomena but "are highly patterned and responsive to causative factors that we need to identify and understand if the quality of care is to be successfully safeguarded." This line of reasoning suggests

quality assessment is indeed a crucial activity and can potentially provide an answer to questions surrounding quality. Donabedian refers to the inquiry aimed at supplying the answers as the epidemiology of quality.

Quality of care assessments have generally been categorized as relating to the structure, process or outcome of medical services (Brook, 1974, Donabedian, 1978 and Williamson, 1978). Assessments of structure, i.e., facilities, personnel, etc., have the longest history growing out of licensure and accreditation requirements. Process assessments, perhaps the most commonly used form of assessment, developed as medical knowledge more clearly identified the type of intervention necessary for a particular medical problem. Standards for treatment became more common providing the foundation for the process assessment methods of chart review and medical audit (Williamson, 1978). The notion of outcomes of medical intervention was the last to emerge partly due to the strong belief that structure and process were directly related to outcome and thus were good proxies for outcome. This underlying assumption that process is causally linked to outcome has come under investigation in a number of studies. Brook and Appel (1973) report the results of reviews of the care of 296 patients with urinary tract infections, hypertension or stomach ulcers using five peer review methods based on process and outcome measures. They found weak correlations between process judgments and outcome measures. Romm, Hulka and Mayo (1976) studying outpatient services for patients with congestive heart failure found no significant relationships between process measures and outcome indicators taken six months later. Other studies have produced similar findings (Tompkins, Burner and Cable, 1977; Martin, Donaldson and Landon et al., 1974 and Schroeder, Schliftman and Piemme, 1974).

Most quality assessments are highly resource consuming since they customarily focus on care at the patient and provider level. A description and analysis of diagnostic and therapeutic patient management is generally necessary for quality assurance and the information is usually obtained through literature searches, case record reviews and from expert judgment. Costs and benefits are sometimes assigned to all possible strategies and outcomes. Although published well over ten years ago, a study by Schwartz, Gorry and Kassirer et al. (1973) illustrate the complexities involved in modeling clinical practice. These realities coupled with questions surrounding measurement of outcomes and the unproven relationship between process assessment and outcome, undermine the application of traditional quality assessment methods to routine policy decisions.

THE CASE OF MEDICAID COPAYMENTS

As was discussed earlier, the future of health care is being decided more and more on the basis of economic considerations, and clinical input into this arena has been sorely lacking. As a result, resource allocation decisions that generally involve cost-quality tradeoffs may not be made with full information as to the likely impact of a policy decision on patient care.

Medicaid recipients sharing in the cost of their health care is one such example of a policy decision motivated by economic considerations yet may well have patient care implications, potentially adverse. This section of the paper presents an evaluation of the impact of a copayment requirement on Medicaid recipients with a particular emphasis on the patient care impact. The intent was not to include the type of assessment that traditionally comes to mind when speaking of quality assessment, but rather an approach that can be simple and inexpensively applied. It is the argument of this analysis that quality assessments are not common occurrences in routine state level evaluations of policy changes that are primarily economically motivated. It is the further contention of this paper that the inclusion of an, albeit, simple quality of care indicator can go a long way toward enhancing the acceptability, utilization and impact of an evaluation.

THE COPAYMENT PROVISION

Although requiring copayments on the part of Medicaid recipients is not a new idea and has been previously proposed in Wisconsin, it was only with the passage of Wisconsin Laws of 1981 that copayment requirements were first implemented. Copayments were implemented for only those services that were not federally mandated, i.e., optional services. They include such services as dental, vision care, therapy, chiropractic, transportation and drugs. Copayment amounts range from $.50 to $3.00.

Nursing home residents, Medicare covered services provided to Medicare eligibles, persons enrolled in health maintenance organizations and children in susidized adoption and foster care placements were exempted from the copayment requirement. Copayments were to be collected by the provider and automatically deducted from the provider's payment by the fiscal intermediary. The effective date of the copayment provision was November 1, 1981.

At the time copayments were enacted, federal legislation did not permit charging copayments on mandatory services to most Medicaid recipients. However, with the passage of the Tax Equity and Fiscal Responsibility act of 1982 (TEFRA), states were allowed to charge copayments on mandatory services in addition to optional services.

STUDY OBJECTIVES

This study represents an effort to assess the various impacts that copayments may have on the Medicaid program. The study was designed to be completed within a relatively short time period, about six months, in order to be timely for budget and legislative deliberations. The purpose of the study was to provide information and background on the potential effects of copayments, summarize what the early,

although limited, data suggest, and raise relevant issues concerning copayments in need for further research.

More specifically, the objectives of the study were twofold:

1. Assess the effect of copayments on the cost of Medicaid services, and
2. Identify possible effects of copayments on the appropriateness of care received by Medicaid recipients.

THE EVALUATION DESIGN

In most instances, policy and program changes are made in a manner not conducive to rigorous evaluation or policy analysis. In the case of copayments, one would ideally prefer an experimental approach whereby Medicaid recipients are randomly assigned to one group required to make copayments and a second group not required to make copayments, i.e., a control group. Of course, this was not the situation and thus the methodology chosen can only approximate the effects of copayments.

The methodological approach used in this study was twofold. First, a time series model was used to estimate the change in expenditures and utilization for Medicaid services after copayments took effect. Secondly, a case review of a sample of Medicaid recipients' utilization of specific medical services before and after copayment was conducted using a specially designed quality assessment screen to determine, among other things, if the appropriateness of care had changed.

Time Series Design

The time series design was one that basically sought to test for a discontinuity in a series of observations. The discontinuity is presumably the result of an intervention or interruption in the series of observations. In this evaluation, the observations were expenditure and utilization data on Medicaid services and the intervention was the implementation of the copayment requirement. A regression analysis procedure was employed to test the magnitude and significance of the discontinuity in the data series. The model utilized a monthly data base extending from November, 1979 through October, 1982.

Separate models were developed for each service for expenditures and utilization. Variables reflecting the number of Medicaid recipients and overall historical trends in Medicaid, as well as the introducton of copayments and other program changes, were included in the model. All data in this study have been adjusted for lag in the billing of claims. The specification of the model was as follows:

$$Y = a + B_1 t + B_2 COPAY + B \quad RECIPIENT + e$$

where:

Y = monthly expenditure and utilization observations for each Medicaid service in the analysis. The data ran from November, 1979 through October, 1982.

t = monthly time variable; 1 = November, 1979, 2 = December, 1979, etc.

COPAY = a dichotomous variable representing the introduction of copayments; 1 = November, 1981 through October, 1982; 0 = prior to November, 1981.

RECIPIENT = the monthly number of Medicaid recipients.

e = random error term.

Quality Assessment Screen

To determine in greater detail the effect of copayments on Medicaid recipients, a sample of recipients was selected and their utilization patterns reviewed. A record of all services utilized by any Medicaid recipient is available through what is referred to as a recipient history. Recipient histories were reviewed to determine changes in the volume and mix of services consumed. Medicaid recipients with a minimum of 24 continuous months of eligibility (12 months prior to copayments and 12 months after copayments) were identified. The utilization patterns before and after the introduction of copayments were then compared and changes noted.

To ascertain whether the quality of care received by Medicaid recipients deteriorated as a result of copayments, a quality assessment monitoring method was developed. As presented earlier, a method was desired that could be easily applied using data that could be collected at a reasonable cost. Traditional quality assessment methods cannot meet these criteria and thus an alternative was sought.

One factor in designing and selecting a quality assurance method is the definition of the term quality. Definitions of quality health care range from very narrow to very broad as typified by the World Health Organization (1960) definition that "Health is a state of complete physical, mental and social well-being and not merely the absence of disease and infirmity." Quality is also often thought to consist of a number of dimensions such as quantity, availability, accessibility, cost and timeliness (Kerr and Trantow, 1969). Depending on how one wishes to define quality, it can call for very different assessment methods.

Although not explicitly stated, the goal of the Medicaid program is thought to provide access to a basic level of medical care for all persons in need (Holahan, 1975). Quality in this instance likely refers to the receipt of services approriate to the patient's needs. Thus the quality of care screen developed in this paper has been termed an appropriateness of care measure to distinguish it from a more technical quality of care measure.

Index Construction

To develop the index, the physician and dental consultants to the Medicaid program were asked to identify services that if decreased in quantity would likely suggest a drop in the quality or appropriateness of care.

They were also asked to identify services that if decreased would not likely cause concern over appropriateness or quality of care. The services that fall into the former category were summed to form the "appropriateness of care" index. The services were dental exams and cleanings, vision care exams and heart disease and blood pressure, gastrointestinal, epilepsy and diabetes medications. The services in the latter group were termed secondary services and included dental fillings, eyeglasses, sedatives, tranquilizers and painkillers and antibiotics. Table 1 details these services. In addition, services to which copayments did not apply were monitored as a type of control or comparison group.

Recipient use patterns for these selected services 12 months before and 12 months after the introduction of copayments were then compared. Two samples were selected, one a random sample of 88 Medicaid recipients and the other a sample of 30 Medicaid recipients who had been identified as users of a large quantity of services. In all, over 2,000 months of services were reviewed. All data for the analysis were obtained from the state Department of Health and Social Services' Medicaid Management Information System.

EVALUATION FINDINGS

Findings of the evaluation will be presented in two sections relating to the effect of copayments on Medicaid expenditures and secondly the impact of copayments on patient care.

Cost Findings

The findings relative to the cost containment potential of copayments will be reviewed only briefly. A report entitled "Preliminary Findings of a Review of the Medicaid Copayment Policy," that more thoroughly analyzes the cost impact of copayments can be obtained from the Wisconsin Department of Health and Social Services.

Table 2 summarizes the results of the cost savings attributable to copayments. Expenditures for services to which copayments applied averaged $7.9 million per month for the 12 month period immediately preceding the introduction of copayments. The statistical models described earlier indicated a reduction in expenditures of 5.5% or about $5.2 million annually over what would have been expected in the absence of copayments. Most, about 60%, of the copayment savings resulted from the direct effect of recipients paying for a portion of their care while the remainder was due to decreased utilization. In terms of services, the greatest percentage savings were for chiropractic and transportation; but in absolute dollars, drugs were responsible for the greatest savings.

Patient Care Impact

Table 3 presents the impact of copayments on the utilization of services and Table 4 the impact on the appropriateness of care index. Table 3 indicates that utilization has dropped for both services to which copay applies as well as services to which copay did not apply. However, the drop for the copay services was about twice that as non-copay services. High users of services had utilization drops almost twice that of the random sample suggesting that copay may be an effective method for reducing excessive utilization. It should be pointed out again, that the recipients whose utilization patterns were analyzed in this section of the analysis were required to have 24 months of continuous eligibility. While the random sample was chosen to represent various medical status groups (aged, blind and disabled) and geographic areas within the state, it may not be representative of all Medicaid recipients.

In terms of the appropriateness of care index, the findings were different for the high users and the random sample. The random sample experienced a reduction in the index of 24% while the high users increased slightly. This suggests that copayments may have a greater adverse effect on the average Medicaid recipient as opposed to those using many services. Those services identified as less indicative of appropriate care, i.e., secondary services, followed the opposite pattern. Secondary services decreased 10% in the random sample but 31% among high users. This finding coupled with the results in Table 3 might indicate that while on the one hand copayments may curb excess utilization, they may also result in less appropriate utilization patterns in the typical Medicaid recipient. The finding of a reduction in the appropriateness of care suggests that there may be value to exempting from copayments, basic, point-of-entry services such as periodic dental exams, vision screening or treatment where patients generally are not experiencing pain such as anti-hypertensive drugs. In this way, any tendency for copayments to discourage initial encounters with the medical care system would be minimized.

SUMMARY

This paper has presented a simple and easily implemented method for screening for quality care impacts of policy decisions. Policy decisions are often made in the absence of information on the effect of a program or policy change on patient care and instead rely primarily on fiscal data. This is especially true, today, where economic realities are more and more outweighing other concerns. The decision to require Medicaid recipients to share in a portion of the cost of their care was selected as an example of a decision viewed basically as an economic issue. The decision was evaluated for not only its cost impact, but also its impact on the quality of care.

An appropriateness of care index was developed that sought to achieve as its purpose, simplicity in development and ease of use while maintaining at least face validity. The thrust of the argument in this paper is that the evaluator faced with analyzing

a policy decision must do the best they can with what is available to them. In assessing quality of medical care, traditional approaches are often too specific and too resource demanding for application to routine evaluation activities. As a result, quality of care is often an omitted variable. Although dominated by clinicians and separated from fiscal analyses, the distinction between quality assessment and program evaluation is becoming blurred. The evaluator wishing to be responsive to the demands of policy-makers must balance rigor and relevance in deciding upon an appropriate evaluation strategy.

In the illustration presented in this paper, copayments by Medicaid recipients were found to lower Medicaid expenditures and utilization, but were also found to have adverse patient care effects as evidenced through the use of selected medical services. The findings suggest that it may be important to target copayments on certain services and exempt from copayments basic, point-of-entry services such a periodic dental exams, vision screening and specific drugs such as anti-hypertensive medications.

This evaluation, it is hoped, has demonstrated that a form of quality of care assessment can be combined with routine program evaluation to produce an evaluation that is more responsive to the needs of policy-makers. The evaluation is not intended to answer the question of what constitutes quality, but rather provide a screen for determining whether quality of care issues are potential problems in need of further assessment.

References

Brook, R. Quality of Care Assessment: A Comparison of Five Methods of Peer Review. Maryland: Department of Health, Education and Welfare, 1974.

Brook, R. et al. Quality of Care Assessment: Choosing A Method for Peer Review. New England Journal of Medicine, 1973, 288:1323.

Brook, R. et al. Geographic Variations in the Use of Services: Do They Have Clinical Significance? Health Affairs, 1984, 3:63.

Campbell, D. Reforms as Experiments. American Psychologist, 1969, 24:409.

Donabedian, A. Needed Research in the Assessment and Monitoring of the Quality of Medical Care. Maryland: Department of Health, Education and Welfare, 1978.

Guttentag, M. Evaluation and Society. In M. Guttentag (Ed.), Evaluation Studies Review Annual, No. 2. Beverly Hills, California: Sage, 1977.

Holahan, J. Financing Health Care for the Poor. Massachusetts: Lexington, 1975.

Kerr, M. et al. Defining, Measuring and Assessing the Quality of Health Services. Public Health Reports, 1969, 84:415.

Loveland, E. Measuring the Hard to Measure. San Francisco: Jossey-Bass, Inc., 1980.

Martin, S. et al. Inputs into Coronary Care During 30 Years, A Cost Effectiveness Study. Annals of Internal Medicine, 1974, 81:289.

Romm, F. et al. Correlates of Outcomes in Patients with Congestive Heart Failure. Medical Care, 1976, 14:765.

Schroeder, S. et al. Variation Among Physicians in Use of Laboratory Tests: Relation to Quality of Care. Medical Care, 1974, 12:709.

Roos, N. et al. Surgical Rate Variations: Do They Reflect the Health or Socio-economic Characteristics of the Population. Medical Care, 1982, 20.945.

Schwartz, W. The Competitive Strategy: Will It Affect Quality of Care. In J. Meyer (Ed.), Reforms in Health Care: Current Issues, New Directions, Strategic Decisions. Washington, D.C.: American Enterprise Institute, 1983.

Schwartz, W. et al. Decision Analysis and Clinical Judgment. The American Journal of Medicine, 1973, 55:459.

Tompkins, R. et al. An Analysis of Cost-Effectiveness of Pharyngitis Management and Acute Rheumatic Fever Prevention. Annals of Internal Medicine, 1977, 86:481.

Wennberg, J. et al. Health Care Delivery in Maine: Patterns of Use of Common Surgical Procedures. Journal of the Maine Medical Association, 1975, 6:123.

Wennberg, J. et al. Small Area Variations in Health Care Delivery. Scientific American, 1982, 246:120.

Williamson, J. Assessing and Improving Health Care Outcomes. Massachusetts: Ballinger, 1978.

World Health Organization. Constitution. Geneva: World Health Organization, 1960.

Table 1
Services and Procedures Used to Construct the
Appropriateness of Care Index

Appropriateness of Care Index Services

Dental
- o Initial oral examination
- o Periodic oral examination
- o Prophylaxis

Vision
- o Basic screening exam
- o Comprehensive vision exam

Drugs
- o Heart/Blood Pressure - Inderal, Dyazide, Lasix Oral, Lanaxin, Aldomet, Hydrochlorothiazide, HydroDiuril, Slow-K, Isordil, Hygroton, Lopressor, Aldoril, Digoxin, Nitroglycerin, Aldactazide, Persantine, Minipress, Ser-Ap-Es, Nitro-Bid, Apresoline, Potassium Chloride, Diuril Oral.
- o Gastrointestinal - Tagamet, Donnatol, Lomotil.
- o Diabetes - Diabenese, Insulin.

Secondary Services

Dental
- o Fillings

Vision
- o Eyeglasses

Drugs
- o Sedatives/Tranquilizers/Painkillers - Valium, Tylenol/Codeine, Dalmane, Darvocet-N-100, Empirin/Codeine, Tranxene, Elavil, Librium, Fiorinal, Ativan, Mellaril, Zomax, Atarax, Percodan, Triavil, Sinequan, Darvon, Meprobamate, Haldol, Serax.
- o Antibiotics - Ampicillin, Tetracycline, Penicillin, VK, Keflex, Amoxicillin, Erythromycin, V-Cillin K, Amoxil E-Mycin, Vibramycin.

Table 2
Cost Impact of Copayments by Service Category

	Copay Implementation		Copay Effect			
Service	Before	After	Direct	Indirect	Total	%
Chiropractic	$ 94,999	$ 69,940	$ 4,594	$ 21,512	$ 26,106	27.5%
Dental	1,782,731	996,208	66,090	39,892	105,982	5.9
Drugs-Legend	3,083,886	2,868,543	92,212	69,009	161,221	5.9
Drugs-NonLegend	363,934	199,255	5,654	8,144	13,798	3.8
Equip. & Supplies	181,423	195,301	2,167	NS	2,167	1.2
Psych. Hosp.	993,418	1,491,652	436	NS	436	.0
Therapies	420,067	202,871	7,032	9,400	16,432	3.9
Transp.	373,659	379,283	68,958	NS	68,958	18.5
Vision Care	603,133	406,335	26,622	13,496	40,118	6.7
TOTAL	7,897,250	6,809,335	273,765	161,453	435,218	5.5

NS indicates no significant effect. All other significant at p .05.

Table 3
Utilization Before and After the Implementation of Copayments, Random Sample and High Users

	Random Sample			High Users		
Services	Before	After	% Change	Before	After	% Change
Copayment	532	449	-15.6%	337	243	-27.9%
No Copayment	342	316	- 7.6	153	130	-15.0

Table 4
Appropriateness of Care Utilization Before and After Copayments, Random Sample and High Users

	Random Sample			High Users		
Services	Before	After	% Change	Before	After	% Change
Appropriateness Index	348	265	-23.9%	110	115	- 4.5%
Secondary Services	312	280	-10.3	340	235	-30.9
No Copayment	301	354	- 9.5	195	171	-12.3

EPISODES OF CARE FOR MEDICARE BENEFICIARIES

Donald M. Steinwachs, Richard Frank, David Salkever, The Johns Hopkins University, Balto., MD.

The introduction of new payment methods for hospital care under the Medicare program has directed attention to the importance of being able to measure all costs associated with any cost containment initiative. While it is expected that costs savings will result from DRGs, there is considerable uncertainty regarding the magnitude of such savings and the impact on the patient's total care. In exploring these policy evaluation issues, methodological questions arise regarding the appropriate unit of analysis for studying cost savings. This is particularly important if one is concerned with the possibility that hospitals may shift responsibility for patient care to other settings and providers as a method of containing costs. The purpose of this research is to examine one approach to using Medicare claims data to construct episodes of care and to link these data to uniform hospital discharge abstract data.

It has long been recognized that the episode of care is a basic measure of the use of health services, particularly for acute disorders (Solon 1967). The construct also can be applied to chronic disorders if one accepts periodic visits for medical management as individual episodes. The advantages of the episode of care as a measure of use are reasonably clear; (1) all the services, including initial, follow-up, and consultation visits as well as tests and procedures, are incorporated into the episode for a given problem or diagnosis, and (2) measurement of outcomes at points in time after the onset of illness/injury are consistent with this framework. The relatively infrequent application of the episode framework, however, appears to be related to uncertainties in measurement: (1) the beginning and end of episodes may be ill-defined, and (2) the episode may involve two or more problems or diagnoses and their treatment. Even so, the episode of care represents an important construct for examining patterns of care, provider practices, costs and outcomes of care. In this paper, Medicare episodes will be used to examine the relationship of physician inpatient fees to the characteristics of the patient and hospital in selected DRGs. The question is whether hospital and physician charges are related, and thus asks whether it is feasible to pay physicians using the same DRGs as are used to pay hospitals.

Construction of Episodes. The sources of data for the analysis are Medicare claims data for Parts A and B that are processed by Maryland Blue Cross as the intermediary, and Maryland hospital discharge abstract data submitted by all hospitals to the Health Services Cost Review Commission. With the cooperation of Maryland Blue Cross, all Medicare beneficiaries hospitalized between January 1 and June 30, 1980 were identified. For each person, all Part A and B claims were obtained for a period of 90 days before admission to 90 days after admission. If an individual had more than one hospital admission during the six month selection period, the first admission was selected as the index. All personal identifiers were removed from the claims.

Ideally, episodes of care in which a hospitalization occurs would include all related services and exclude unrelated care. Unfortunately, Part B claims for professional services do not include diagnoses, making it infeasible to screen pre-and post hospital care for relatedness. As a result, all care received during the arbitrarily defined pre and post periods were included in the episode. A summary file for each individual was developed in which pre-index admission services were summarized by treatment setting (acute inpatient, nursing home, home health and ambulatory) and by type of charge (facility or physician). Post-index admission were summarized in a similar fashion, and characteristics of the index admission were included in the file. This file includes admissions to 39 of the 46 general acute care hospitals in Maryland, with 44,750 individuals having been hospitalized one or more times in the six month period.

Preliminary analysis of the data indicated that approximately six per cent of the beneficiaries had no Part B claims. This is higher than expected. In part, it reflects a problem with using one Medicare intermediary as the source of data. In the Maryland suburbs of Washington, D.C., hospitals file claims through the Maryland intermediary while many of the physicians file claims through the Washington intermediary. Thus, those without Part B claims include those who have not enrolled in Part B and those with missing data. As a result, it was decided to exclude all individuals having no Part B claims from the analysis file.

The decision to match discharge abstract data to the index admission of the Medicare beneficiary was undertaken due (1) to the limited data available on the claim for properly assigning a DRG classification, and (2) for the additional information available on type of admission, secondary diagnoses, and disposition status. It was assumed that the discharge abstract and the index hospital claim matched if the hospital identifier, dates of admission, and discharge and date of birth matched. The successful match rate was 91%; the ones not matched either had missing data on the Medicare claim or an exact match was not achieved. The resulting file was passed through a DRG grouper to assign DRGs to each of the index admissions.

Characteristics of Episodes. In Table 1, the characteristics of a 10% random sample of episodes are shown. Almost 66% are 75 years or older and the distribution of principal diagnoses submitted to Medicare reflect a high proportion of circulatory, digestive, and infectious disorders. In Table 2, the utilization characteristics are shown. Almost seven per cent had a prior hospital admission in the 90 days preceding the index admission and 19% had a readmission within the 90 day post-index admission period. Other characteristics of ambulatory physician and hospital outpatient department (OPD) use are shown. In Table 3, the charges

TABLE 1

CHARACTERISTICS OF MEDICARE EPISODES IN MARYLAND
(Per Cent Distribution)

TOTAL NUMBER	1980
Total Number*	4,059
Total Per Cent	100.0
Age	
Under 65	7.3
65-74	26.8
75-84	58.3
85 +	7.6
Sex	
Female	54.1
Principal Diagnosis	
Infections	12.7
Neoplasms	8.8
Endocrine	1.2
Mental	1.9
Nervous-sense	7.2
Circulatory	25.4
Respiratory	4.3
Digestive	11.7
Genitourinary	4.8
Dermatology	1.2
Musculoskeletal	4.4
Symptoms	5.3
Injury	6.1
Blood	1.2
Other	0.9
No Diagnosis	0.9

* Based on Systematic 10% sample, excluding cases in which there were no Part B charges.

TABLE 2

CHARACTERISTICS OF MEDICARE EPISODES IN MARYLAND

	1980
Total Number	4,059
Pre-Index Admission	
(3 Months)	
Per Cent Hospital Admission	6.7
Per Cent Nursing Home	0.1
Per Cent Physician Visit	59.9
Per Cent Hospital OPD Visit	12.6
Index Admission	
Average LOS	13.63
Post-Index Admission	
Per Cent Readmitted	19.0
Per Cent Nursing Home Discharge	1.0
Per Cent Physician Visit	61.1
Per Cent Hospital OPD Visit	12.7

TABLE 3

CHARACTERISTICS OF MEDICARE EPISODES IN MARYLAND

	1980
Total Number	4,059
Total Episode Charges	$ 5,065.96
Pre-Index Admission	413.14
Index Admission	
Total Hospital	2,894.87
Routine	1,536.29
Ancillary	1,358.58
Total Physician*	1,074.19
Post-Index Admission	683.76

* Includes any post-index inpatient physician charges.

for services received are shown. The average episode generated, $5,066 in charges, 21% was accounted for by inpatient physician charges and 57% is for hospital charges. The remaining 22% includes all other services before and after the index admission. If this distribution were compared to the distribution of Medicare payments it would be somewhat different. In Maryland, Medicare pays full hospital charges less a small discount, but pays usual, customary, and reasonable charges for physician services. As a result, the percentage of total Medicare payments going to the hospital is understated in these percentages.

Physician Payment and DRGs. One application for this data base is to examine the relationship of physician payment to DRGs. Payment of physicians for inpatient care based on DRGs has been under consideration by the Health Care Financing Administration. For purposes of this analysis, four categories of DRGs were selected that included digestive, respiratory and neurological conditions and neoplasms. In total, eighteen DRGs and 4,471 episodes are included in the analysis. The multivariate logistic model is applied to examine the ratio of inptient physician charges to total hospital and physician charges. The decision to use the ratio of charges as the dependent variable reflects our hypothesis that the ratio would be constant within a DRG, but might vary across DRGs reflecting variations in intensity and type of treatment. Another advantage of the ratio of charges as the dependent variable is that it reduces the effects of variations in local area wages and prices. In general, it is expected that the number and intensity of the hospital services will be directly related to the duration and intensity of physician services within the selected DRGs (DRG categories include 10-19, 85-90, 96-97 and 172-173). Further, it should be noted that the DRGs selected for analysis exclude any surgical DRGs, which we will be analyzing separately. The mathematical form of the dependent variable used in the analysis is the logarithm of the odds ratio. This circumvents the truncation of the ratio at one when estimated using ordinary least squares.

The model estimated includes characteristics of the hospital and patient, plus dichotomous dummy variables for each hospital and each DRG. The hospital dummy variables are designed to capture unmeasured constant differences among hospitals, e.g., effects of overall casemix on charges. Hospital characteristics included in the model are teaching status and the logarithm of number of beds. The race, age, sex, logarithm of length of stay and of special care days, whether or not there had been a prior admission in the previous 90 days, and whether or not the patient was discharged dead are specified for each patient.

In Table 4 the results of the analysis are shown. The model explains 27% of the variance in the odds ratio. The results indicate that a 10% change in length of stay reduces the odds ratio of 3%. Since the average of the odds ratio for these DRGs is .27, the 3% reduction in the odds ratio leads to approximately a 3% reduction in physician charges. This result in consistent with physicians generating most of their charges early in the stay, and relatively less as patients stay longer. The variable for teaching status indicates that physician charges are a lower proportion of totalcharges in teaching as compared to non-teaching hospitals. Any hospital with at least one approved residency program was included in the teaching hospital category. This result could be explained by physician charges being similar across hospitals, but teaching hospitals have higher average charges, and/or it could be due to having physicians on salary, specifically housestaff, and not charging a separate professional fee. The policy implications of these alternatives are quite different and the relative contribution of each explanation needs to be estimated. Note that hospital size is marginally significant and positive suggesting that in larger hospitals the physician fee accounts for a higher proportion of total inpatient episode charges.

Patient characteristics are not significant in this group of DRGs, except for race and recent prior hospitalization. Whites, on the average, have lower physician charges than non-whites in the selected DRGs. This is an unexpected finding and the explanation is uncertain. Prior hospitalization within 90 days of this current admission was included in the model as an indicator of severity. It is significant but has a negative effect on physician fees. This suggests that although these individuals may be in poorer health, the physician care is less intense. One possible explanation is that the patient's condition is known and there is little need for diagnostic and consultative services. Furthermore, since these are medical DRGs, the care being received at this admission does not include major surgery which would probably lead to a higher physician to total charge ratio.

TABLE 4

RELATIONSHIP OF HOSPITAL AND PATIENT CHARACTERISTS TO THE LOGARITHM OF THE RATIO OF PHYSICIAN TO TOTAL INPATIENT CHARGES

VARIABLES	COEFFICIENT	t-STATISTIC
Ln Age	-.0457	- .51
Race (White=1)	-.2395	- 6.66
Sex (Male =1)	.0269	1.09
Previous Hospital	-.0213	- 5.31
Ln LOS	-.2712	-14.56
Ln Special Care Days	-.4884	- 6.80
Discharged Dead	.0188	.41
Teaching Hospital	-1.1838	-5.02
Ln Beds	.2608	1.88
Hospital Dummy Variable	Included	
DRG Dummy Variable	Included	
Constant	-1.6212	- 2.38

n = 4471
Overall F = 27.28
R-Squared = .2740

Possibly the most important results concern the contribution of the four categories of variables to the total explanatory. In Table 5, the per cent variance explained is shown for each category. Hospital characteristics explain little of the variation while patient characteristics, the DRG and the hospital dummy variables have approximately equal explanatory power. This suggests that there are important hospital related differences in practice and/or charging patterns that affect the ratio of physician to total charges. The importance of patient characteristics suggest that the DRGs are not capturing aspects of patient mix that are explanatory of variations in physician charges.

TABLE 5

EXPLANATORY POWER OF VARIABLE CLASSES

VARIABLE CLASS VARIABLE CLASS	PER CENT OF VARIANCE EXPLAINED
Hospital Characteristics	1.5
Patient Characteristics	5.0
DRGs	6.0
Hospital Dummies	7.3

Discussion. The results of the analysis raise several interesting policy issues. One concerns the recent trends toward lower lengths of stay that have been associated with the implementation of DRGs. Our results suggest that these trends are likely to have relatively little impact on physician income. A ten per cent reduction in length of stay contributes to an average reduction in physician fees of roughly three per cent. The impact of teaching and housestaff on professional fees could have considerable importance. If higher hospital costs are being offset, in part, by lower physician fees on the teaching hospital, this would strengthen the rationale for special treatment of teaching hospitals under Medicare DRG payment. Alternatively, the physician fee component may not be substantially less than in non-teaching hospitals, but the hospital's charges are, on the average, higher for the same DRGs. Even if this is the case, it would suggest that teaching hospitals that generally have a more severe case mix, do not generate higher physician fees for higher intensity care. Further probing into these relationships, as well as extending the analysis to other DRG categories, will be important in assessing the contribution of teaching status to total inpatient episode costs.

The purpose of the analysis presented was to illustrate one of the many potential applications of claims data aggregated into an episode of care framework. Other applications include an examination of factors related to readmissions, the extent of substitution of pre or post hospital services for inpatient services, as well as the evaluation of cost containment initiatives. In the context of evaluating hospital cost containment initiatives, the episode framework facilitates the analysis of substitution effects and can provide insights into changes in the patterns of care that might suggest better or poorer outcomes of care (e.g., readmission).

The principal limitations to the methodology employed are the arbitrary nature of the duration of the episode and the lack of diagnostic data to relate pre and post inpatient care to the reason for hospitalization. There is, however, procedural information available on the physician claim that might be used to refine the episode structure, although uncertainty regarding its relationship to the hospitalization will persist in many cases. On the positive side, many of the episodes analyzed are for individuals who are very ill and the argument can be made that all the care is related to maintaining or improving the patient's status.

Episode based analyses can be expected to be particularly important in efforts to evaluate the impact of outpatient surgery and diagnostic procedures on total costs and indicators of outcome. Instead of using the admission as the critical event, the occurrence of the procedure would be the event, and the care received before and after would be integrated into the episode. As this example illustrates, the structure of episodes will vary with the issue being examined. Even so, episode analyses based on claims data represent powerful analytic tools as evidenced by recent research (Roos, 1983; Steinwachs 1985). Further applications and testing of episode based strategies should be encouraged.

REFERENCES

1. Solon JA, Feeney JJ, Jones SH, et al. Delineating episodes of medical care. Am J. Public Health 1967; 57:401.

2. Roos LL, Jr., Roos NP. Assessing Existing Technologies: The Manitoba Study of Common Surgical Procedures. Medical Care 1983; 21 (4):454-462.

3. Steinwachs DM, Salkever D, Rupp A. Impact of Per Case versus Per Service Hospital Payment in Maryland. Final Report submitted to the National Center for Health Services Research, DHHS, May 1985 (Available at NTIS No. PB 85-164739/AS and PB 85-169747/AS).

Supported in part by The Robert Wood Johnson Foundation Grant No. 09026.

ASSESSING THE IMPACT OF AN HMO ON THE HOSPITALIZATION RATE OF MEDICARE ENROLLEES

Jan Ouren

Introduction

Little is known about health maintenance organization (HMO) performance with an aged population. With younger populations, HMOs have consistently demonstrated lower rates of hospital admissions, lower overall costs and sometimes briefer lengths of stay than the fee-for-service (FFS) system. If these findings result from HMO service delivery per se, and not from selection bias in enrollment, we might expect that reductions in HMO hospital use may be particularly visible in an aged population where a large volume of service is involved. On the other hand, because aged adults generally have a less discretionary need for health services, the type of delivery system may insignificantly alter the population's demand.

The research at hand investigated the relationship between inpatient hospital admissions and HMO enrollment in an aged cohort of Medicare beneficiaries. It attempted to answer how the hospital use of the aged changed after joining an HMO. This study--part of a larger research project--attempted to define the relationship by applying the time series methodology of Box and Jenkins, 1976.

This particular time series technique models the dynamic, underlying process between variables of interest, not just the overall relationship. The usefulness of this approach is that gradual changes, tied to the occurrence of some event in time, can be evaluated. The group's hospital admissions, a set of observations that were spaced over equal time periods, and the occurrence of enrollment in the HMO, with its known onset and duration within the series, permitted the application of this technique.

The methodology of Box and Jenkins has found many successful applications in marketing and social science research. Its use in health services research was thought to be unprecedented at the time of this study, yet well suited to the longitudinal data bases being developed by the National Center for Health Statistics and others.

Methodology

Research model. The following research model was developed from a review of HMO and hospital utilization literature. Two impacts on the hospital admissions of the enrollees were hypothesized. First, pent-up demand was predicted. This factor would cause a brief upswing in admissions after enrollment in the HMO. A pent-up demand, or deferred utilization, hypothesis assumes that people joining HMOs are able to postpone needed medical treatment until they are covered by a prepaid, comprehensive system. Their temporary treatment deferral results in increased demand that surfaces soon after HMO enrollment, when the out-of-pocket price drops to the consumer. This effect should be brief, due to the inability to predict and to forestall necessary treatment for long. Next, a reduction in the level of inpatient hospital utilization was predicted. This decrease occurs as the HMO's incentives to control utilization take hold, constraining expensive hospital use by the new enrollees.

Data and design. The Health Care Financing Administration provided longitudinal, inpatient utilization and eligibility data on Medicare beneficiaries from the Puget Sound area of the State of Washington for this study. This was the same data base that Paul Eggers (1980) used in his comparison of FFS and HMO Medicare enrollees prior to HMO enrollment. The time frame of this

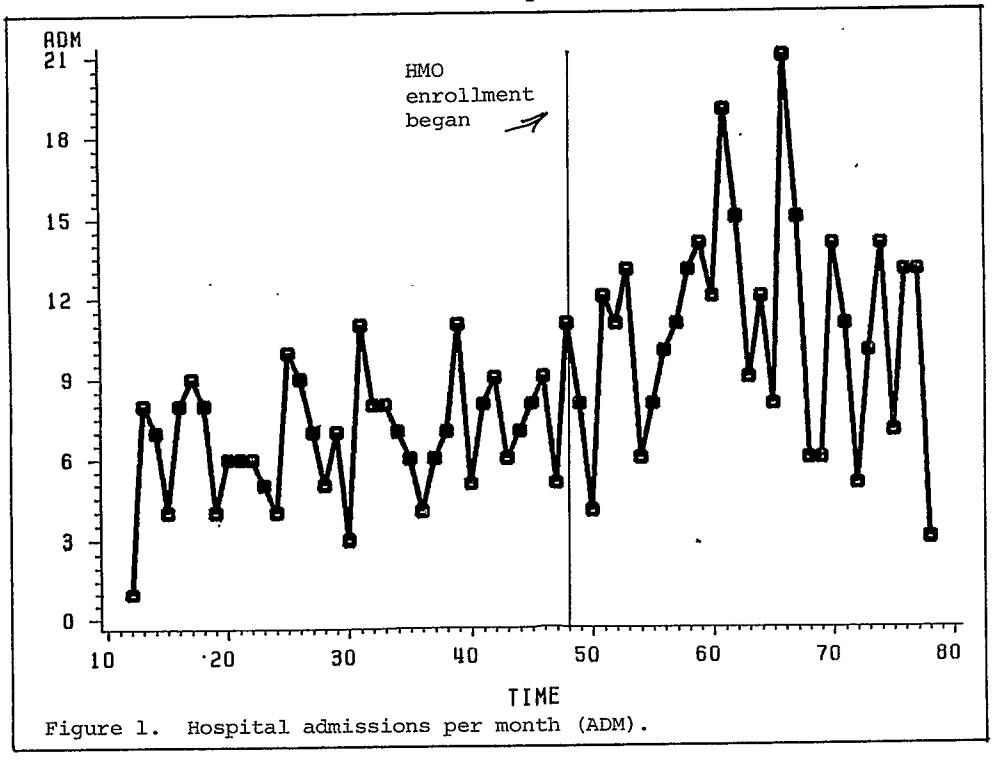


Figure 1. Hospital admissions per month (ADM).

study included the 68 months from December 1973 to July 1979.

The cohort consisted of individuals who voluntarily joined the Group Health Cooperative of Puget Sound (GHC) during an open enrollment period that began in August 1976 and lasted throughout the study period. The GHC had a risk contract with Medicare and was paid a per beneficiary/enrollee amount based on its costs in relationship to 95% of the Average Adjusted Per Capita Cost--the AAPCC--in the area. The data base included each individual's use of hospital services in the FFS system for at least two years prior to HMO enrollment, as well as any post-enrollment utilization in either the HMO or the FFS if the member disenrolled. The enrollment of members thus "interrupted" the time series of admissions, a quasi-experimental design that permitted assessment of the HMO's impact on the hospital experience of the group.

Analytic plan. The first step in creating the necessary analytic variables was to change the day-to-day hospital experiences of individuals to monthly group frequencies. The files contained the beneficiary-specific admission date, discharge date and length of stay. Admissions were converted to a time series measuring the group's hospital use by tallying the variable's frequency across monthly time periods. This admission variable, called ADM, is seen in Figure 1.

ADM reflected hospital use in both HMO and FFS settings because it covered pre- and post-enrollment periods for each individual and because individuals joined the HMO at different times. As more people enrolled and were "locked in" in terms of Medicare, fewer hospitalizations took place in the FFS system. Toward the end of the study, nearly all people from this cohort were established as HMO members.

While health care data are most often expressed and contrasted as rates, a rate format was not necessary with the methodology selected, which requires only that the raw series, such as ADM, be stationary before the parameters are estimated. This eliminates the need to adjust for the changing numbers of people contributing their observations to the process. Avoiding the use of rates can be a considerable advantage when denominators are unknown or costly to obtain. Two different analyses will be presented here. In the second analysis, admission frequencies were changed, however, to a rate format to allow the data to be transformed.

Box and Jenkins provide a method to describe the stochastic or random component of a time series of observations, such as ADM, as an equation consisting of three structural parameters, p, d and q. The p parameter represents the autoregressive process--the relationship between adjacent observations in the series. The d parameter represents the process of trend, the systematic increase or decrease in the level of the series. D is the number of differences required to make the integrated process stationary or without trend. The q parameter represents the moving average process in the series or the extent to which the series is dependent on previous random shock. These (p,d,q) models can be used to describe any lengthy series of observations over equally spaced time periods, as long as the series is stationary or can be made stationary. The set of equations with its Autoregressive, Integrated and Moving Average parameters is referred to as an ARIMA (p,d,q) model.

One identifies a potential ARIMA (p,d,q) model through examination of the autocorrelation function and the partical autocorrelation function, both functions relating the series to itself at successive time lags. Once a possible model is identified, the parameters p and q are then estimated with nonlinear programs.

After this process, the adequacy of the model is diagnosed by examining the series' residuals. Residuals that appear to be unsystematic, i.e., white noise, are evenly distributed around a zero mean with constant variance, a configuration that indicates when an adequate model has been found. This strategy of identifying, estimating and diagnosing is repeated until an adequate model is found. The results of the application of this technique are presented in the following section.

Results

Descriptive statistics. The HMO enrolled 885 Medicare beneficiaries age 65 or older during the open enrollment period. Of this number, 882 had enrolled in the HMO by the end of the study, 40 had disenrolled, and 19 had died. Three more people joined the HMO after the end of the study period in July 1979 and thus contributed to the earlier observations in the FFS.

By the first month of the study, 532 persons in the cohort had become Medicare-eligible. This was 60.1% of the cohort's final population. All hospital admissions were in the FFS. During the next 67 months, this cohort grew to its final size of 885 persons. Over the study period, they experienced a total of 576 hospital admissions, an average of 8.6 admissions a month.

Impact assessment. The admission series ADM can be thought of as one realization of an underlying theoretical process. To test the impact of the HMO on this process, a dummy variable--a time series--was created to represent the HMO enrollment period which began during the 48th month. This variable was called STEP. Its values were 0, prior to enrollment, and 1, during the enrollment period. It was correlated with the univariate ARIMA model of ADM, which had been diagnosed as a (0,1,1), a first-differenced first-order moving average model.

The cross-correlation function between STEP and ADM had no significant spikes, and other than a small, consistently negative relationship when enrollment led admissions in time, there was no evidence of any significant relationship. Thus the data did not initially support the hypothesis that the HMO altered the group's use of inpatient hospital services.

Two interpretations were still possible. Indeed, the HMO may have had no influence on this cohort's use of inpatient care. Or, the effect may have been hidden by feedback between the two variables. Earlier, a cross-correlation function between the prewhitened series ADM and a stochastic enrollment series (the actual monthly enrollment figures) had demonstrated that being hospitalized dampened HMO enrollment two to three

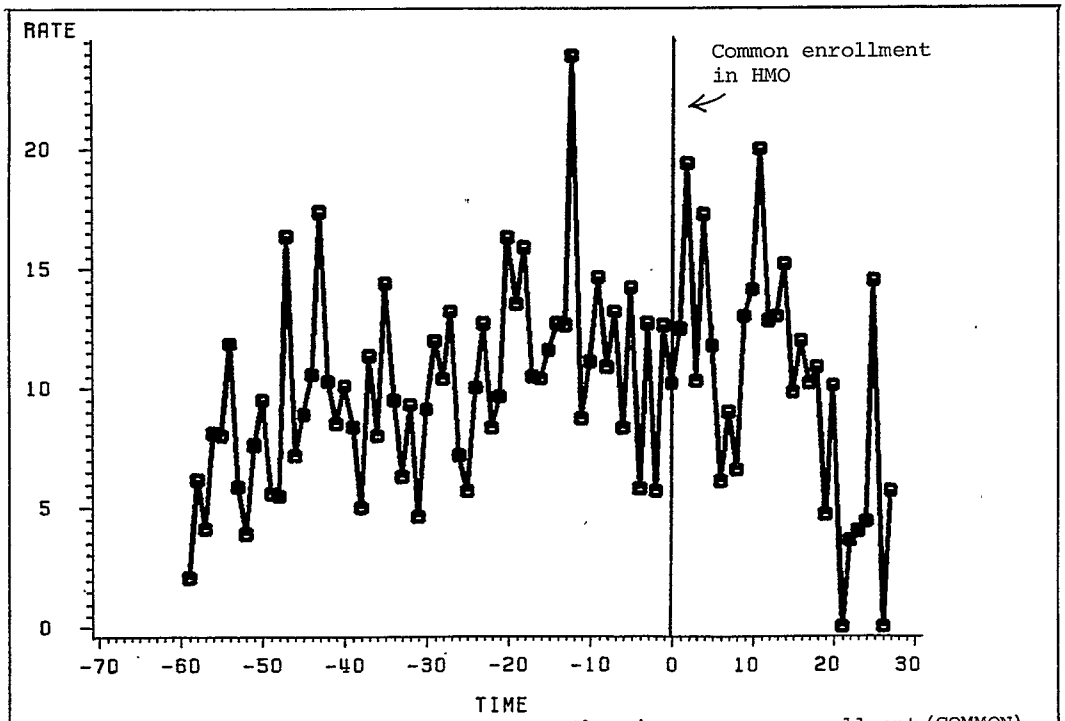


Figure 2. Admission rate per 1000 per month, given common enrollment (COMMON).

months later. Thus, it was possible that feedback was occurring between hospital use and enrollment, wherein the strong relationship between prior use and enrollment was overshadowing gradual changes in use that resulted from HMO enrollment.

The idea that an appropriate transformation of the data might eliminate this feedback loop and uncover the relationship between HMO enrollment and hospital use developed. In this next analysis, the time of enrollment was transformed so as to eliminate the possibility of feedback.

Each individual's enrollment date was adjusted so that the group had a common time of enrollment, E. This adjustment moved individual histories forwards or backwards in time to align each enrollment date with the common axis. Thus not only enrollment dates shifted. So did entitlement, hospitalizations, disenrollment and mortality, with all dates maintaining their position relative to the person's time of enrollment.

After this realignment, the hospital admissions of the group were calculated on this new time scale, centered on E. This series provided the monthly numerator for calculating the admission rate. The denominator, which was the population covered by Medicare during each time period, was constructed by counting the people entitled to benefits (i.e., on the data base) each month before enrollment and each month after enrollment. Admissions were then divided by the number of people covered in each time period and multiplied by 1000 to provide the rate per month per 1000 population. The resulting series was restricted to a length of 87 continuous time periods in which the population base always exceeded 170 persons. This time series, called COMMON, is shown in Figure 2. A vertical line marks the period of enrollment (E) at time=0.

A statistical test, reported elsewhere (Ouren, 1983), demonstrated that a significant decrease in the admission rate occurred after HMO enrollment. In the next analysis, the dynamics of this HMO intervention effect are measured, using Box-Jenkins methodology to provide an estimate of the change.

First, an ARIMA model was built for COMMON. Like the real admission series ADM, COMMON was also diagnosed as an ARIMA (0,1,1), a univariate moving average model that required first-differencing. Its parameters and mean square forecast error (MSFE) for four one-step forecasts are given in Table 1.

The test of forecast accuracy, the mean square forecast error (McCleary and Hay, 1980) compares the forecasts or the conditional expectations of a process with the observed values. As the best forecasts generate the lowest MSFE, the best model is defined as the one with the lowest MSFE. Using this criteria, we will be able to select the better model of the process that underlies hospital use in an aged cohort.

Several intervention components for COMMON were considered. First, a binary pulse variable representing the initial enrollment of the group in the HMO was created to test the deferred utilization effect. This variable, called PULSE, took the value of 1 during the month of enrollment and carried zero values at all other times. Since deferred utilization refers to a temporary effect occurring shortly after enrollment, the first six lags of PULSE were evaluated in a cross-correlation function with the prewhitened COMMON series. The only significant correlation was at lag +2, when COMMON led enrollment by two months and was positively associated with it. An oscillating effect was also visually present.

Next, a second binary variable called STEP was created to represent the presence or absence of the group in the HMO (before enrollment=0;

after enrollment=1). This variable was used to test for a constant change in the admission rate after enrollment. Because HMOs have financial incentives to substitute less expensive services for hospital admissions, this variable was expected to show a negative, constant change after enrollment, reflecting the alteration of incentives from the FFS sector. Indeed, the cross-correlation function between COMMON and STEP was marked by consistently negative correlations, all less than significant but never dying out. This pattern suggested a constant, small negative relationship over time.

Table 1. COMMON: Models, Parameters and MSFE

Variable	Model	Parameter	Estimate	se	t-Test	RMS	MSFE
COMMON	(0,1,1)	MA 1	0.77	.08	9.26*	15.1	0.89
COMMON	(0,1,1)	MA 1	0.89	.05	16.79*	15.7	0.80
	STEP	Uporder 1	-0.36	.15	-2.40*		
	PULSE	Uporder 2	7.16	3.16	2.26*		
	PULSE	Sporder 1	-0.66	.35	-1.89		

*p<.05

To see whether the data significantly supported a compound intervention component representing these two different impacts, a zero-order ω_1 STEP component and a second-order $\omega_1/(1-\delta B)^2$ PULSE component were added to COMMON (0,1,1). This intervention component tested an abrupt, constant change in admissions by the first month, followed by a brief, temporary effect starting the second month. This intervention or impact structure was in keeping with the cross-correlation results and with the hypotheses. The lagged timing of the temporary effect (with PULSE) could be explained as an orientation period during which enrollees familiarized themselves with the new health care system and HMO physicians accomplished the necessary casefinding, diagnosing and scheduling of inpatient hospital admissions.

This model's parameters, estimated by conditional least squares regression (Liu, 1981) are seen in Table 1. Both HMO intervention variables, STEP representing constant change and PULSE representing temporary effects, had significant parameters. An abrupt, constant decrease in the admission rate was present one month after enrollment. This decrease in the monthly admission rate per 1000, estimated as -0.36, was interpreted as the slope of constant change. The temporary increase in the admission rate, estimated as +7.16, oscillated from the second month on, dying out at a moderate rate (d=0.66).

Using B as the notation for the backshift operator, the equation for the inpatient hospital admission rate of the cohort was:

$$(1-B)Y_t = -.36\ \text{STEP} + (7.16B^2/1-.66B)\ \text{PULSE} + (1-.89B)a_t$$

This equation expressed the admission rate as a function of its own past, a constant HMO intervention, a temporary HMO intervention and noise.

Let us now compare the forecast accuracy of this intervention model with the univariate model, COMMON(0,1,1), using the MSFE as the criterion for picking the better model. Remember that the last four observations had been excluded from parameter estimation in both univariate and intervention models of COMMON. One-step forecasts were now made for these time periods using both models. The MSFE was calculated for each one.

As seen in Table 1, the intervention model with an MSFE of 0.80 outperformed the univariate model with an MSFE of 0.89. Thus, the better model was the compound intervention model, the one that included two exogenous HMO "disturbances" to explain the admission rate. The temporal information about HMO enrollment improved the admission rate forecasts beyond the conditional expectations derived from the past admission rate alone. In this sense, HMO enrollment caused the admission rate to change. The way in which it changed was consistent with theories of both deferred utilization and HMO incentives to control utilization.

This concludes the impact assessment of an HMO on the inpatient hospital use of aged enrollees. The methodology will undoubtedly find a useful place as payors attempt to set capitation rates for HMOs and others want a way to measure the dynamics of change in health care settings.

References

Box, G.E.P. and G.M. Jenkins. Time Series Analysis: Forecasting and Control, Revised Edition. San Francisco: Holden Day, 1976.

Eggers, P.W. "Risk Differential Between Medicare Beneficiaries Enrolled and Not Enrolled in an HMO," Health Care Financing Review,

Liu, L.M. "Box-Jenkins Time Series Analysis," BMDP Statistical Software. Los Angeles: University of California Press, 1981:639-660.

McCleary, R. and R.A. Hay, Jr. Applied Time Series Analysis. Beverly Hills: Sage Publications, 1980.

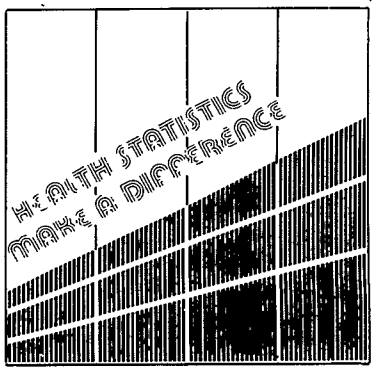
Ouren, J. HMO Performance with Medicare Beneficiaries: Time Series Analysis. Dissertation, University of California, Los Angeles, 1983.

Acknowledgments

I wish to thank Professor Stuart O. Schweitzer, UCLA School of Public Health, and Professor Dominique M. Hanssens, UCLA Graduate School of Management, for their many valuable contributions to my dissertation research, of which the preceding work was part.

Session H

Microcomputer and CATI Applications



NISS: THE NORC INTEGRATED SURVEY SYSTEM

Reginald P. Baker, NORC
Christine Beard, NORC
Joseph Taylor, NORC

Introduction

NORC, A Social Science Research Center, was founded in 1941 and has been affiliated with the University of Chicago since 1947. Among the many activities of NORC is the ongoing collection of data on a variety of important topics including health care, education, the labor force, and the family. This data collection is achieved through a number of survey projects, large and small, that in past years have included the National Ambulatory Medical Care Survey (NAMCS), High School and Beyond, the Air Force Health Study, the General Social Survey (GSS), and the National Longitudinal Study of Labor Force Behavior. NORC collects these data and prepares them for use by researchers in government, academia, and the private sector.

In four decades of survey research activity, NORC has developed a variety of procedures and automated systems to support its data collection, preparation, and research efforts. These procedures and systems include management support tools to control and monitor survey progress as well as data preparation treatments that produce high quality, error-free datasets.

The automated systems currently in use at NORC are the product of an evolutionary development process. As computer technology has changed, existing systems have been modified and manual procedures automated to exploit the advantages of new technologies.

Today NORC confronts another technological challenge. We are meeting that challenge by undertaking a full-scale redesign of existing systems. That redesign emphasizes system integration, shared databases, and microcomputers. The overall goal is a full scale, state-of-the-art automated capability to support survey research. The system is called the NORC Integrated Survey Systems (NISS). Its full development and implementation is expected to take three years or more.

Overview of System Requirements

Our analysis of existing procedures and systems indicated that the proposed system must have four key capabilities.

1. A questionnaire processor to assist in the development of survey instruments. This processor must also perform an analysis of the instrument to enforce appropriate rules of syntax and style and to provide important data to other processors in the system about question text, allowable responses, and skip patterns. Item banking and production of camera-ready copy are other highly desireable features of the questionnaire processor.

2. A data capture capability that operates in three distinct modes: computer-assisted data entry (CADE), computer-assisted telephone interviewing (CATI), and computer-assisted personal interviewing (CAPI). Automated coding of verbatims and open-ended responses should be integrated into all three data captures modes. Extensive error checking at the point of data entry must be provided. Further, all three of these modes must be driven by data generated by the questionnaire processor.

3. A survey control capability to track key survey events for each respondent including both field events -- successful interviews, refusals, etc. -- and in-house events -- completed data entry, retrieval required, editing completed, etc. The survey control capability also tracks survey costs and staff productivity as well as key quality control measures. It provides a shared source of information for both survey managers and field personnel about such important issues as overall survey progress, interviewer assignments, and location of respondents.

4. A capability to manipulate very large datasets in order to produce analysis files for researchers, compute important

statistical measures such as weights and standard errors. and perform tabulations for clients.

The system requirements further specify that all of these capabilities, to the extent possible, must be fully integrated with one another. The data capture processors, for example, must be able to access a questionnaire database generated by the questionnaire processor to determine the appropriate data entry rules. The survey control component must provide online access to survey managers, field personnel, and data capture activities.

System Hardware

In our view, the overall system requirements did not mandate selection of a particular hardware configuration. Rather, there was considerable flexibility that would allow us to take advantage of different hardware capabilities and costs. While the file manipulation required to produce analysis files, deliver large-scale datasets to clients, and produce statistical or tabular summaries seemed to require a mainframe capability, it also seemed to us that other system requirements could be met with minicomputer or networked microcomputer hardware. We chose the latter.

In addition to the previously-described system requirements, two other factors influenced the decision to select microcomputer hardware. First, NORC had already made a commitment to microcomputers as part of an office automation strategy featuring word-processing, spreadsheets, and electronic mail. Second, the lower cost of a microcomputer solution, both acquisition and maintenance, made it more attractive.

The decision to use microcomputers further mandated some form of networking in order to achieve the goal of system integration. We eventually chose an IBM-compatible Novell local area network (LAN). The network architecture is based on a star typology in which every workstation has its own cable to and from the file server, a 68000-based microprocessor. We chose the star because of its ability to manage a large number of workstations with only minimal degradation in response time. A single network can consist of a maximum of 24 devices, any one of which can be linked to another network or networks.

As of this writing, NORC has installed three Novell LANs. Each has a file server with 120 megabytes of disk. The three LANs serve a total of about 50 microcomputers. Some of these stations are dedicated to production activities such as data entry or survey control updating, while others are used by survey managers for a variety of management-related tasks.

In the near future, we expect to add additional LANs with a goal of installing a file server which has only other file servers as its workstations. This so-called "superstar" manages communications between networks in the same way that individual file servers manage communications among the workstations on their networks.

In addition to the workstations and file servers, the superstar LAN has a number of important peripherals. They include high-speed dot matrix and laser printers, a nine track, mainframe compatible tape drive, and a high speed streaming tape unit for system backup and archiving. Several of the workstations are also equipped with modems for remote communications. Future plans call for a high-speed gateway to the mainframe. A schematic representation of the hardware configuration is shown in Figure I.

The NISS Software Design

The key feature of the system design is the sharing of data through three integrated databases. These databases, portrayed graphically in Figure II, contain all of the information necessary to collect the data, convert them to machine-readable form, perform a machine-edit, and produce a final, fully documented dataset. The databases also contain data that permit a survey manager to monitor a survey's progress and make decisions about management initiatives that may be necessary to ensure the survey is concluded on time and within budget.

The Questionnaire Database

This database contains a complete description of the survey instrument. Complete question and answer text, legitimate response values, skip patterns, missing values to be used for each question, coding tables for any open-ended items, and output positions in the final data record are all stored in this database. The database is served by both a pre- and a post-processor.

The questionnaire processor or pre-processor is a design tool for

FIGURE I

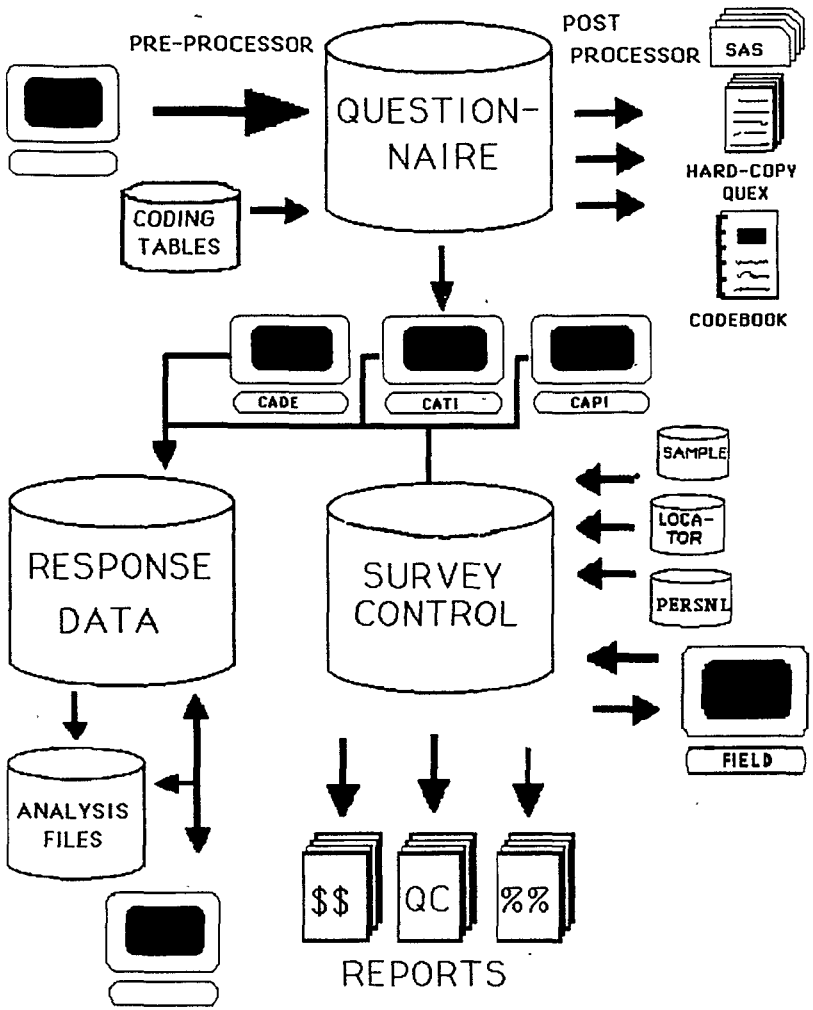
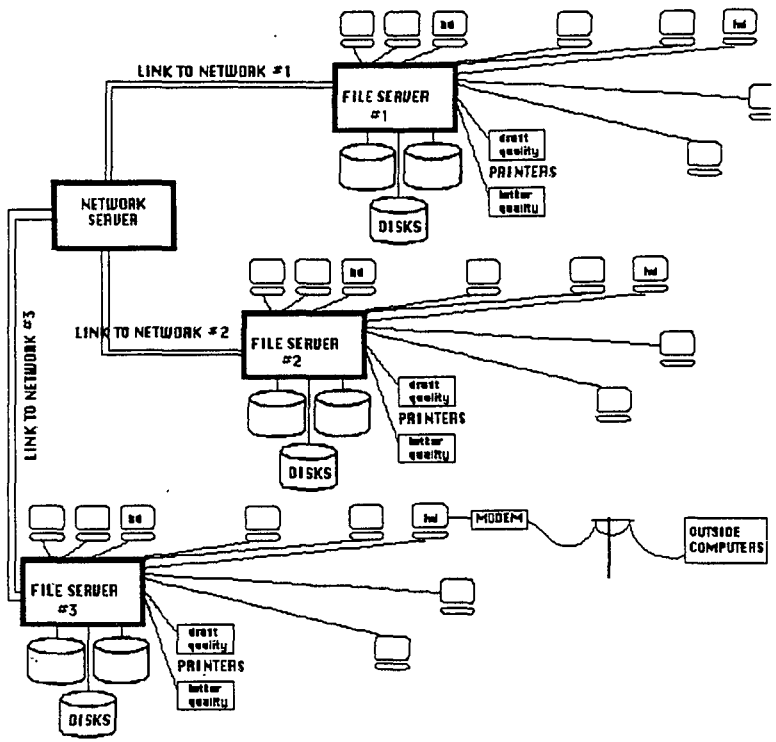


FIGURE II



GENERIC SUPERSTAR NETWORK
-- 3 STAR NETWORKS CONNECTED
BY A SINGLE NETWORK SERVER

instrument development. It functions much like a wordprocessor. Questionnaires can be written and formatted interactively. When design of the instrument is complete, the pre-processor creates the Questionnaire Database.

The post-processor takes data from the Questionnaire Database and creates output files for use by other programs and systems. For example, control cards for statistical systems such as SAS or SPSS (mainframe and PC versions) can be generated. A hard-copy version of the questionnaire can be generated on a letter-quality device such as a laser printer. Questionnaire text can be exported for input to a program that produces a final codebook with record layout, frequencies, etc.

The Questionnaire Database also provides input to one of the most important components of the overall system, data capture. The data capture subsystem can operate in any of three modes: computer-assisted data entry (CADE), computer-assisted telephone interviewing (CATI), and computer-assisted personal interviewing (CAPI). Each of these three modes share common features that include the following:

* Complete checking of all entries so that they conform to valid ranges defined for the particular data item.

* Automatic routing through the instrument to enforce skip patterns.

* Inter-item consistency checks as appropriate for a particular instrument or survey.

* Linkage to a computer-assisted coding (CAC) subsystem that allows for coding from predefined coding tables using a keyword-in-context search technique.

* Automatic linkage to the Survey Control Database.

* In the case of CATI and CAPI, display of full question and answer text.

* Output of respondent data to the Response Database.

The design of the data capture subsystem allows for a maximum of flexibility so that applications may be tailored to meet the varying requirements of individual surveys.

The Response Database

Responses to questions in the survey instrument are entered to the data capture subsystem and then output to the Response Database. Thus, this database contains all of the data collected by the survey and by all instruments used in the survey. This database can have a simple structure, i.e., a simple rectangular file, or, in the case of surveys with a variety of instruments, a very complex structure that may be expressed in a relational or network model. Files containing subsets of the Database may be created and exported for analysis by researchers, delivery to clients, etc.

The Survey Control Database

The Survey Control Database is principally a management tool, although it also provides data to support other processing tasks such as weighting. This database is setup at the very beginning of the project before any interviewing takes place. Among the kinds of information loaded are the following:

* The sample to be surveyed. Each survey respondent will be identified as an entity in the database. Applicable subsample information or stratification indicators may also be loaded.

* Locating information to help an interviewer contact a respondent.

* Assignment data indicating which interviewers and interviewer supervisors have been assigned to each case.

Once loaded, the Survey Control subsystem tracks all events associated with a particular case. It also notes the date of the event and the relevant actor. The former include activities such as a successful interview, receipt of an instrument in-house, completion of data entry, and other such events which are of interest to the survey manager. Actors can include interviewers, coders, data entry personnel, etc. Updates to the database may be done by in-house staff or remotely from the field. In-house updates from other subsystems such as data capture are automatic at the time the event occurs.

The Survey Control Database may also contains cost data. These data may be both in-house staff and processing costs as well as field costs on such items as travel,

interviewer salaries, telephone charges, or other costs which the survey managers wishes to monitor.

Reports on the database inform the survey manager about overall survey progress and costs as well as the status of individual cases. Reports also can be used to measure staff productivity and calculate important cost per case measures that allow the projection of likely costs to complete a survey.

The Survey Control Database has two other important functions. First, it provides a means of communication about the survey between the survey manager and the field interviewing staff. Changes in interviewer assignments, new locating information on respondents, and updates about survey documents either sent to or received at the central office can all be communicated almost automatically to key central office and field personnel. Second, if properly constructed with the appropriate sample-defining data, the Survey Control Database can be the principal source by which a survey statistician calculates weights and non-response adjustments.

Mainframe Functions

As noted above, our design assumes that there are a variety of tasks, particularly at the end of the data preparation process, that require the computing power only a mainframe computer can provide. Post-cleaning or machine-editing of data is one such activity. Others include production of a final file for delivery to the survey sponsor, production of analysis files for researchers, and generation of detailed tabulations or statistical analyses.

A link from the LAN to a mainframe is clearly required. This link will be achieved over a high-speed telecommunications gateway between the LAN and NORC's mainframe at the University of Chicago. Creation of mainframe-compatible tapes on the LAN provides another means for the transfer of data.

A Pilot Project: The NAMCS Integrated Survey Processing and Control System

The general outlines of the overall NISS design are now in place. Actual development of the system began in late 1984 as part of a pilot project to support the 1985 National Ambulatory Medical Care Survey (NAMCS). This project surveys a national sample of approximately 5,000 physicians. Participating physicians complete a brief data collection form (called the Patient Record Form) for a sampled set of individual patients seen by the physician during one week in 1985. The system developed to support this data collection is called the NAMCS Integrated Survey Processing and Control System.

The goal of the pilot project was to integrate the three principal data processing activities of the 1985 NAMCS: survey control, data entry, and coding. The resulting system provides for the rapid and natural flow of work with stringent error control.

Prior rounds of NAMCS were processed in the traditional manner. Documents were collected and bundled into batches as they arrived from the field. Processing occurred in a series of shops: mail receipt and batch ticketing, coding, data entry, and editing. Tracking of documents through the shops was done largely with manual systems. Problems with misplaced documents, duplicate data entry, errors in batch tickets, etc. were often not discovered until the final data preparation was fully underway.

The system developed for the 1985 round of NAMCS is a single system in a single location. All of the above activities, around which the labor intensive serially-dependent shops were organized, are functions of the current system. When a document arrives at NORC's central office from the field it is logged into the system. It then is data entered, verified, coded, and cleaned, all in this single location using one system. At the same time, the document is tracked throughout each of the processing stages automatically.

As work progresses all entries and updates to the NAMCS databases are checked to verify that the document is being processed in the proper sequence and that the entry or update is valid. For example, documents cannot be data entered until they have been fully edited and logged to the system's Survey Control Database. Document identifiers are fully checked at the time of data entry to be sure that the identifier is valid, belongs to the physician for whom data entry is currently being done, and has not been previously entered.

The system has a full CADE capability that supports data entry with valid range checking and missing value substitutions. A data entry verification capability is also built in.

Computer-assisted coding is another key feature of the system.

NAMCS requires coding of open-ended questions about the reasons for patient visits, the physician's diagnoses, and the names of any medications that have been prescribed. The coding subsystem uses a keyword-in-context approach to search the appropriate coding tables (including ICD-9) and then present a menu of potential codes to a coder who selects one or requests additional searching. Like the CADE component, coding also has built-in, independent verification.

Information from both CADE and coding is stored in the Response Database as soon as it is entered and verified. The data are then uploaded to a mainframe computer for additional editing and final data preparation. The Survey Control Database is also update automatically whenever any action is taken from CADE or coding.

Reports from the system are based on complete, accurate, and up to the minute data. They include information about current completion rates, field status, cases with problems, and productivity reports for data entry operators and coders. Thus, survey managers and supervisors are able to maintain close control simultaneously over both data collection and data processing activities.

The system runs on one of NORC's microcomputer-based LANs with a multi-user data base management system. Each microcomputer on the LAN is physically connected to a 120 megabyte disk. The disk contains software that enables each person at his/her microcomputer (or work station) to read and write files on the 120 megabyte disk and to communicate with other stations. A multi-user data base management system on the network provides the data structure and protection necessary when more than one user is accessing and writing to a single database.

Conclusion

We feel that the system design described here offers survey managers a wide range of automated tools with which to conduct a survey and process the resulting data. It relies heavily on state-of-the-art microcomputer technology. Its full implementation awaits additional technical development, particularly in the area of micro-mainframe interfaces.

Our initial evaluation of the NAMCS pilot suggests that the overall system design is a feasible one. While there have been problems with response time and system performance on some tasks such as report generation these problems appear to have solutions in modified software designs.

The experience of NAMCS is invaluable as we work on the development of new NISS modules for use in a number of large surveys to be fielded in late 1985 and 1986. The generalized survey control capability and data capture subsystems are the top priorities. As each is developed, tested, and placed in production we move a step closer to the overall goal of a state-of-the-art integrated system. At the same time, we are making incremental contributions to NORC's long-standing capability to produce survey datasets of the highest possible quality.

AN AUTOMATED SYSTEM FOR ASSESSING PHYSICAL FITNESS IN SCHOOL CHILDREN

Debra G. Weber, Harold W. Kohl, Marilu D. Meredith, and Steven N. Blair
Institute for Aerobics Research

The 1980 Surgeon General's report, Promoting Health/Preventing Disease: Objectives for the Nation (1), identifies exercise and physical fitness objectives for the U.S. population for the year 1990. Five of these objectives relate to exercise and physical fitness in young people 10 - 17 years old, and one objective specifically states that 70% of 10 - 17 year olds should participate in systematic physical fitness assessment.

Current research in youth fitness and health shows low levels of physical fitness and health status in many American children. A temporal decline in youth fitness has been shown over the past 10-15 years (2). For example, the National Children and Youth Fitness Study (NCYFS), conducted for the Department of Health and Human Services, has documented a substantial increase in youth skinfold measures, indicating a significant increase in body fat in children and adolescents aged 10 to 15 years (2). Decrements in other fitness measures were also observed. Compared to nationally established norms, mean mile walk/run scores were higher and mean sit-up scores were lower for both males and females in the NCYFS study. The study found that only 47% of American children participate in appropriate amounts of exercise year-round that may lead to lifetime participation. Physical activity that is regular, vigorous and prolonged is accepted as appropriate. In the NCYFS study this refers to activities utilizing large muscle groups in a dynamic fashion for a period of 20 minutes or longer, three or more times per week, and at an intensity of 60 percent or more of the individuals aerobic capacity (2). Low physical fitness and sedentary living contribute to future risk of chronic disease (3-5) and an inverse association between levels of physical activity and serum lipid values in children has been shown by Thorland, et al (6). Moreover, higher levels of physical fitness as measured by maximal oxygen uptake have been correlated with better overall coronary risk profiles in children and adolescents (7).

In response to low fitness levels in youth, wide spread physical fitness testing of school age children began in 1958 with the introduction of the American Alliance For Health, Physical Education, Recreation and Dance (AAHPERD) Youth Fitness Test (8). The Youth Fitness Test includes: 1) muscular strength and endurance (pullups or modified pullups and situps); 2) power, speed (standing long jump, shuttle run, and 50-yard dash); and 4) aerobic power, speed (distance walk/run). Several test items in this battery might more properly be classified as motor performance rather than physical fitness variables. Performance is primarily determined by genetic potential rather than exercise training. In more recent years physical education professionals recognized the need to develop a fitness test focused on health rather than motor performance or athletic ability, and the AAHPERD Health-Related Physical Fitness Test was developed (9). The Health-Related Physical Fitness Test includes: 1) aerobic power (distance run); 2) strength and endurance of the abdominal wall musculature (situps); 3) body composition (skinfolds); and 4) flexibility of the lower back and posterior thigh muscles and connective tissue (sit and reach).

The two AAHPERD tests are widely used but the data are not systematically processed, nor are the test administrators subject to quality control procedures as data are collected. There have been national surveys for both the Youth Fitness Test (8) and the Health- Related Physical Fitness Test (9), but the purpose has been to develop norms, rather than to establish a wide-spread physical fitness testing system. In order to assess progress towards the physical fitness testing objective, it will be necessary to establish a nationwide system of data collection, verification, processing and dissemination.

The purpose of this report is to describe the process developed by the Institute for Aerobics Research (IAR) for mass testing and monitoring of youth fitness and to present implementation plans for nationwide physical fitness testing of children and youth.

The FITNESSGRAM program is a nationwide computerized system assessing physical fitness levels of school children in grades kindergarten through high school. The FITNESSGRAM evaluation is designed to inform both student and parents about the student's physical fitness status. The results of the student's performance are documented on a report card (FITNESSGRAM) which is sent to the parents.

In addition to providing students with direct feedback regarding physical fitness status and the teacher with information about status of the total group, the FITNESSGRAM report card can be a beneficial tool in increasing public awareness regarding the physical fitness of children and youth in general. The report card communicates to parents the level of physical fitness in their child. Areas in which the child needs improvement are indicated along with suggested activities for improving the performance. A desired outcome is that with this increased awareness of their child's fitness status, the parents will take an active interest in encouraging the development, improvement and maintenance of physical fitness through appropriate exercise and physical activity programs.

METHODOLOGY

The FITNESSGRAM process is delivered by a series of four steps designed to maintain the quality and integrity of the data. In addition, the delivery process is designed to be implemented in a manner that is expedient and conducive to implementation in a school setting.

The initial step, data collection, is performed by the physical education staff in the participating schools. Data collection involves assessing the physical fitness levels of the students with the Youth Fitness Test or the Health-Related Physical Fitness Test. Data collection includes a visual check of test scores for quality control by the physical education instructor. School districts generally administer tests annually with assessments scheduled for the fall or spring semester. Tests are administered in mass and require three to four class meetings to complete. Scores are recorded during testing on a field assessment form and subsequently prepared for data entry.

The second step, data verification for quality control follows data entry. The computer program has stringent range and consistency check routines as part of the verification process. Moreover, a verification report is produced for a manual check of test data. After reviewing the verification report for possible errors, each instructor makes necessary corrections.

The next step is data processing, which includes production of individual student FITNESSGRAMS. A semester report, an awards report, and a statistical report are provided for the teacher and the school district to aid in program evaluation. The semester report consists of an alphabetical listing of all students, including demographic information; physical fitness scores; total fitness scores; and percentile rankings. The awards report is a listing of students who have qualified for awards available from AAHPERD and the President's Council on Physical Fitness and Sports. This report includes: test date(s), sex, age, height, weight, grade level, and class period. In addition, the raw score and percentile ranking for each test item is reported.

The statistical report is a group (class) analysis of physical fitness performance on each variable. The analysis is age/sex specific and provides for visual comparison of the group measures of central tendency and variability nationally established norms.

The basic product of the data processing stage is the youth fitness report card. In addition to the individual student's demographic characteristics, each FITNESSSGRAM displays raw physical fitness scores, percentile rankings according to his/her respective age and sex group based on nationally established norms, histograms showing percentile rankings for each test item, raw scores and percentile rankings of the student's performance on previous tests (up to three previous recordings), test date, total fitness score (a weighted sum of the student's normalized individual test scores). This equation is used with each test score and then summed.), and exercise recommendations. Each student receives the individual report card which also contains a note for the parents, explaining the purpose and implications of the FITNESSGRAM program.

IMPLEMENTATION

IAR Delivery:

During the 1984 - 85 school year IAR delivery of FITNESSGRAM involved 168,800 students in 122 school districts throughout the 48 contiguous states. Each student was assessed using the Youth Fitness Test or the Health-Related Physical Fitness Test. Participating school districts administered tests annually with assessments scheduled for the fall or spring semester. The delivery schedule for FITNESSGRAM testing and reporting was created by the IAR with a strict time schedule due to the large number of tests. The IAR delivery of FITNESSGRAM for 1985 - 86 will involve 225,000 students in 178 school districts in 50 states. Participating school districts will be scheduled for annual FITNESSGRAM production according to a fall or spring delivery schedule. School districts may again choose to administer either the Health-Related Physical Fitness Test or the Youth Fitness Test .

After completing the test administration and data recording in the data collection phase, the teachers involved in IAR delivery of FITNESSGRAM transfer student and school identification information, student demographic and fitness test results to optically scanned entry cards. A visual check of the data is performed by each teacher. The score cards are then batched and sent to the IAR.

The data verification step by the IAR begins with optical scanning of the score cards and production of the verification report. This report is a listing of each student's data and is priority sorted according to the following variables: by school; by teacher; by grade; by period; by sex; and by alpha. In addition, at the end of the verification report is an error listing that includes a listing of students with suspected errors, a description of where the possible errors exist, and instructions with the appropriate action necessary to correct the error(s). The verification report and optically scanned correction cards are returned to the school district. The data verification stage identifies range and consistency violations in the student fitness data. The purpose of the verification report is to afford each teacher the opportunity to correct errors prior to the production of the FITNESSGRAMS.

The physical education teacher reviews the verification report (and accompanying error report) for possible reporting errors. Any corrections that need to be made are recorded on the optically scanned correction cards and returned to the IAR for processing. After receiving the correction cards, the IAR begins data processing and produces the individual FITNESSGRAMS, the semester report, the awards report, and a statistical report. Records with incorrect or incomplete data will be produced as

such. All materials are returned to the school district and each child receives a FITNESSGRAM to take home and share with his or her parents. The process spans approximately a 12 week period, usually allowing completion within a given semester's time.

Microcomputer Delivery:

The 1984 - 85 school year marked the pilot year of FITNESSGRAM delivery through a microcomputer procedure. The FITNESSGRAM software is designed to operate on an Apple IIe, double disk drive microcomputer, with 64K random access memory. The Health-Related Physical Fitness Test or the Youth Fitness Test may be implemented by the school district. The pilot year of FITNESSGRAM through microcomputer delivery involved 12 school districts in 12 different states. A maximum of 2,100 students per school district could be tested. The pilot delivery consisted of extensive pretesting phases involving four alpha test sites, with the school districts administering FITNESSGRAM in both the fall and spring semesters. All alpha sites administered the Youth Fitness Test. There were eight beta test sites, with the school districts administering FITNESSGRAM in the spring only. The Health-Related Physical Fitness Test was administered to students in two beta sites and the Youth Fitness Test was administered to students in the other six beta sites.

Three floppy diskettes are included in the FITNESSGRAM software package: the report card program diskette, the statistical analysis diskette, and a blank data diskette. Each data diskette is encrypted with school district name, and this name is printed on each of the student report cards. This system discourages inter-district trading of copied diskettes, while encouraging intra-district dissemination of the diskettes. The data diskette does have limited storage capacity; a maximum of five-hundred individual records can be stored on a data diskette.

The initial step of the microcomputer implementation of FITNESSGRAM is data collection. On the microcomputer level the actual data entry of student information serves as the first step in the data verification, range and consistency routines on the program diskettes prohibit entry of invalid data values. Errors that may exist within the ranges/qualifications of the variables may still be reviewed with the verification report.

The next step is data processing and production of the individual FITNESSGRAMS, the semester report, awards report, and the statistical report. The FITNESSGRAM software is designed to operate on three popular dot matrix printers, but virtually any printer may be used with minor alterations.

During the 1985 - 86 school year 5,000 FITNESSGRAM software units will be distributed to school districts in 50 states. The mass capability of the FITNESSGRAM software affords the opportunity to involve at least 2.5 million students (kindergarten through high school).

Consultation Delivery

The third delivery mode of FITNESSGRAM is the consultation model involving local mainframe delivery of the system. IAR systems programmers provide consultation and assistance in implementation. Those school districts interested in implementing FITNESSGRAM on their own mainframe system will coordinate all aspects with the IAR. The consultation service includes assistance in the following areas: system flow, file layouts, equipment, storage and personnel requirements, data transfer and exchange methods (including percentile tables and statistical analysis routines) and programming and production techniques. The consultation service is also available to assist a district using the microcomputer delivery system to develop the technology to upload and download FITNESSGRAM data to and from a local mainframe to perform district-wide analyses.

The four stage delivery system (collection, verification, processing and dissemination is theoretically identical to the IAR delivery system. The major advantage of having a local operating system is that the school district is free to create its own timelines for delivery of the program, and hence be less dependent on the IAR.

CONCLUSION

The FITNESSGRAM process is designed to measure physical fitness levels; enhance the awareness of students and parents about physical fitness; and concurrently increase the ability of teachers, administrators, and researchers to track and evaluate fitness performance. Cross-sectional and longitudinal research can be conducted within the project by using the current (through Spring 1985) data base of 226,700 student fitness scores. Approximately 35 percent of the fitness scores are repeat tests, affording an opportunity to relate changes in fitness to the growth curve of children and adolescents. The increased exposure of FITNESSGRAM through the three delivery modes will yield one of the largest data bases of student fitness scores in the United States.

Ancillary research projects can be added to the FITNESSGRAM project. Data obtained from a survey of 400 teachers who participated in the program in the state of Oklahoma are currently being analyzed. The purpose of this study is to examine possible associations between teacher characteristics (demographic factors, education, experience, and personal health habits), physical education program characteristics (type of program, extent of program), and students' physical fitness performance. Future studies are planned with parents and school administrators.

Data captured in microcomputer delivery of FITNESSGRAM can also be used for research purposes. Selection of school districts prior to actual delivery will allow districts involved to duplicate data diskettes and send them directly to IAR to be uploaded and analyzed.

In addition to the impact of FITNESSGRAM on youth fitness research, there are practical applications of the program. The computer assisted program decreases labor intensivness of physical fitness test reporting (data dissemination). This results in affording the physical education staff increased "time on task" both in planning and working with the students. The opportunity to involve the student's parents or guardians improves their awareness of the child's physical fitness status. The communication between the school physical education staff and the family members provides greater involvement in the student's growth and development from a health and fitness perspective. Systematic data collection, verification, processing, and dissemination is important for further evaluation and development of physical fitness programs. The administrative task of evaluating the curriculum for effectiveness is enhanced through the ready availability of data analyses of physical fitness information. These data analyses may be processed for the individual school, the school district, for a given region, or on a statewide level.

In 1985-86, schools participating in the IAR delivery of FITNESSGRAM needed only to pay return postage for sending data input cards to IAR for processing and any reproduction costs of training materials for teachers (manuals, etc). Microcomputer delivery schools receive the software and blank report cards for 75% of the students to be tested at no charge. Districts participating in the consultation delivery are not required to pay for the services of the IAR programming staff or any information received from the IAR.

This mass distribution system has substantial public health implications. FITNESSGRAM may help maintain progress towards achieving high levels of physical fitness in America's young people. A populace that maintains a healthy and active lifestyle will benefit greatly with a higher level of physical fitness and improved health status. This program may lead to a healthier, more physically fit nation in the future.

ACKNOWLEDGMENTS

We thank Mike Smith, Suzan Lewis, Alan McQuiston, Cindy Woods, Connie Fitzgerald, Paula Kahn, and Liz Kohl of the Institute for Aerobics Research for program development and data set maintenance. In addition, we thank the American Alliance for Health, Physical Education, Recreation, and Dance and the President's Council for Physical Fitness and Sports for their assistance in the continued success of FITNESSGRAM.

This research is supported by Campbell Soup Company.

More information on the FITNESSGRAM project may be obtained from the IAR, Youth Fitness, 12200 Preston Road, Dallas, Texas 75230.

REFERENCES

1. United States Department of Health and Human Services. Promoting Health/Preventing Disease. Objectives For The Nation. U.S. Government Printing Office, Washington, D.C. 20402, pp. 79-81, 1980.
2. Ross JG and Gilbert GG. A summary of findings. Journal of Physical Education, Recreation and Dance, 56(1): 3-8, 1985.
3. Blair SN, Goodyear NN, Gibbons LW, and Cooper KH. Physical fitness and incidence of hypertension in healthy normotensive men and women. Journal of the American Medical Association, 252(4): 487-490, 1984.
4. Paffenbarger RS, Wing AL, and Hyde RT. Physical activity as an index of heart attack risk in college alumni. American Journal of Epidemiology, 108(3): 161-175, 1978.
5. Morris JN, Everitt MG, Pollard R, Chave SPW, and Semmence AM. Vigorous exercise in leisure-time: protection against coronary heart disease. Lancet, 2(8206)(Dec 6):1207-1210, 1980.
6. Thorland WG and Gilliam TB. Comparison of serum lipids between habitually high and low active pre-adolescent males. Medicine and Science in Sport and Exercise, 13(5): 316-321, 1981.
7. Fripp RR, Hodgson JL, Kwiterovich PO, Werner JC, Schuler GH, and Whitman V. Aerobic capacity, obesity and atherosclerotic risk factors in male adolescents. Pediatrics, 75(5): 813-818, 1985.
8. American Alliance for Health, Physical Education, Recreation and Dance. Youth Fitness Test Manual. AAHPERD 1900 Association Drive, Reston, VA 22091, 1976.
9. American Alliance for Health, Physical Education, Recreation and Dance. Health-Related Test Manual. AAHPERD 1900 Association Drive, Reston, VA 22091, 1980.

A DISTRIBUTED DATA PROCESSING SYSTEM IN A MULTICENTER CLINICAL TRIAL

Barry R. Davis, Donald J. Slymen, Charles J. Cooper, Lynne Mutchler
The University of Texas School of Public Health

INTRODUCTION

Multicenter clinical trials typically involve a number of clinical and special activity centers and a coordinating center. The coordinating center is responsible for instituting, coordinating and monitoring the data-gathering activities of the trial as a whole and processing, storing and analyzing the large volume of data that is collected.

Until recently, the entry and management of the data in multicenter clinical trials was solely accomplished at the coordinating center. Technological advances have now led to the use of distributed data processing (DDP) systems wherein the collection, entry and some management of the data are done locally at each clinical center. Several clinical trials have used or are now using DDP systems. These include the Coronary Artery Surgery Study (CASS) (1), the Hypertension Prevention Trial (HPT) (2), and the pilot study for the Systolic Hypertension in the Elderly Program (SHEP) (3).

In July 1984 the planning for a full scale SHEP began. The trial started in March 1985. SHEP is a multicenter, randomized, double-blind, placebo-controlled clinical trial involving 17 clinical centers, a coordinating center, a central lab, a CT (computed tomography) scan reading center, and an ECG (electrocardiogram) reading center. The purpose of the trial is to determine whether long-term administration of antihypertensive medications will reduce the five year incidence of fatal and nonfatal stroke in people 60 years of age or over with isolated systolic hypertension. For the purposes of the trial, isolated systolic hypertension is defined as systolic blood pressure $\geq$160 mm Hg and diastolic blood pressure <90 mm Hg.

This paper examines several areas of clinical trial management that were considered important in designing the DDP system for SHEP. These areas include data collection and entry, data quality and quality control, data management, personnel training, implementation time and effort, and mid-trial modifications.

ELEMENTS OF THE SYSTEM

Overall Organization

The distributed data processing system consists of a DEC VAX 11/750 and four DEC Rainbow microcomputers residing at the coordinating center and 19 DEC Rainbow microcomputers distributed at each of the remote sites--the 17 SHEP clinics, the Project Office at the National Heart, Lung and Blood Institute, and the office of the Chairman of the SHEP Steering Committee. The microcomputers are used for data entry, transmission and reception. Communication occurs between the remote site and the coordinating center via ordinary telephone lines. A schedule of weekly transmission times was set up by the coordinating center, but alternate times and free times are also allowed.

All data for participants in SHEP are collected on two-part paper forms designed specifically for the trial. After the data from a form have been entered into the clinic computer and verified by blind re-entry, the original of the paper form is sent to the coordinating center by mail, and the electronic image of the data record produced from the form is sent to the coordinating center over the telephone lines.

At the end of a transmission session, when data files have been received from the clinic, files containing messages, memos or error reports of previously processed data are sent over the phone line to the clinic. In this way, errors found in the data are sent back to the clinic for action on a timely basis. Corrections to the data are made by the clinic on a paper form designed for making changes to transmitted data. This form is treated just like any other form--it is entered into the microcomputer and sent to the coordinating center, with one copy kept in the clinic files.

Paper flows in only one direction--from the clinics to the coordinating center. Other messages or requests for further information are handled on the two-way electronic circuit (Figure 1).

When the transmission from the clinic is complete, processing is initiated in the central computer to check the gross quality of the transmitted data. Later, many further checks are made to assure that the electronic image is as error-free as possible.

Data Collection

The SHEP study has 36 forms ranging from one to twelve pages in length. The number of variables per form ranges from 14 to 250. All data pertaining to SHEP are collected on forms, except the results of analysis of blood samples by the central laboratory, and the record of randomization. All but two types of forms are printed on two-part no carbon required paper, which creates one copy automatically as the original is created. This enables the clinic to keep a copy while sending a duplicate to the coordinating center. The rules for making any corrections to a form are such that the two copies will always remain identical. The coordinating center copy is the official one.

There are several reasons for using paper forms in the trial as opposed to using direct data entry. First, paper forms allow for simultaneous data collection on several patients at several sites within a clinical center. Direct data entry would require several microcomputers. Secondly, participants may fill out forms. (In SHEP, two forms are filled out by the participants.) Direct data entry would require participants to use the microcomputers. Finally, a form requires the signature of the person completing it. Diskettes are not legal copy, and an official audit of the data could not be performed.

The SHEP forms were designed to optimize the collection of complete, accurate information, with considerations of data entry secondary. However, the data entry process is relatively easy.

Data Entry

A large multicenter study such as SHEP requires the accumulation of a large body of data. We estimate that on the average 18,000 items will have been collected on each participant by the end of the five-year follow-up period for SHEP. In a DDP system, major responsibility for the quality of the data is placed at the source of data collection. The error detection and correction process is carried out at the clinical site rather than at the coordinating center, through the use of microcomputers and specialized software.

Commercially available data entry and data management software packages were evaluated for the SHEP study. The following points were considered in selecting the package:

1) Ease of use;
2) Ability to handle entire hard copy forms without breaks in the continuity of data entry;
3) Inclusion of range checks, date checks, conditional field checks with conditional skips, response checks and required entry fields; and
4) Provision of a "blind" data re-entry mechanism for verification.

Based on these and other considerations, the package developed by Viking software was chosen out of twenty packages reviewed. With the Viking package, the hard copy SHEP forms were translated into data entry formats on the microcomputers to provide for speed and accuracy of data entry.

Exact screen copies of paper forms are not used in the SHEP DDP system. Since data are entered from the SHEP forms, only item numbers appearing on the data entry screen are needed to uniquely specify questions on any particular form.

There were two reasons for using skeleton screen versions of the paper forms. First, exact copies of the paper forms would require several screens. Skeleton versions require just one screen for most forms. Secondly, initial testing with exact copies proved cumbersome and boring to data entry operators as they gained experience.

Response and range checking are a vital part of the system. Entry values are matched against a list of possible values (e.g., yes/no responses) for validity. Ranges are specified for measured variables, such as blood pressures. In this case, the range is checked, and if an out-of-range value has been entered, an error message is displayed and the keyboard will not respond until a reset key is touched. Then, the data entry operator can ascertain what the value is and enter it.

An operator can override an out-of-range value and allow it to be accepted. The range for a particular item is designed to capture most of the expected values. However, a participant may actually have an out-of-range value. If it were not possible to defeat this range check, then the form for the participant could not be entered. Thus the suitability of the range check depends on the judgment of the data entry operator.

Many sequences of questions on the SHEP forms are conditional. If a certain condition applies in response to an initial question, a following set of question must be answered. If the condition does not apply, the questions are skipped. The software provides automatic skipping when the appropriate entry has been made.

The data entry package is supplied to the clinics on two diskettes. To enter the data, the clinic operators only need to boot a diskette to execute the program. A SHEP form is then selected from a menu. Once a set of forms are entered, the forms need to be re-entered for verification. New data are compared with previously entered data. If the current data do not match the previously entered data, the clinic operator is immediately asked to make the necessary corrections.

The transfer process employs two programs, the first of which allows for the selection of forms to be transferred and the second, a commercial software package, performs the actual file transfer over the telephone lines. One day a week, during a scheduled half hour time period, the clinic operator boots the transfer program disks on the microcomputer and the first program is automatically executed. Ordinarily, all of the data entry files for the previous week will be selected for transfer. However, should the need arise, the program allows for selection of specific files for transfer.

Once the files have been selected for transfer, the operator is requested to dial up the coordinating center computer. The modems used for the study allow for automatic dial-up. After the operator logs on to the coordinating center computer, the actual file transfer automatically begins. Each data entry file for each form is transferred twice and labeled appropriately for the identification of the transmission. A weekly transfer from a clinic requires from 3 to 15 minutes of phone connect time.

Once a clinic has transferred data to the coordinating center, the two copies of each data entry file are automatically compared. If there are any discrepancies between copies, the clinic staff is notified that their transmission was not successful and is asked to retransmit the files after any necessary corrections.

After each file has been successfully transferred to the coordinating center computer, it is renamed on the clinic diskette and identified as a previously transmitted file. Once the transfer process is complete, the operator is instructed to remove the data entry diskette and insert a diskette upon which files can be received from the coordinating center computer. When all of these files are received, the clinic operator can then examine them. This provides a method of communicating messages or instructions to the clinics from the coordinating center.

The clinic is required to maintain the data entry diskettes for the previous six weeks, thus ensuring a six week backup of previously

transferred forms. This allows the coordinating center sufficient time to determine that all forms scheduled to be received have been received. If forms are missing, the clinic operator is notified and the data entry files for the appropriate week are examined.

Should clinic staff fail to transfer their files on the assigned day and time, they are allowed to transfer the files during the same time period on the next day. A computerized log is kept of the day, time and length of time that each clinic transferred their data. After the scheduled transfer days of the clinics have passed, this log is checked to see if all clinics have transferred their data. If a clinic fails to transmit during their scheduled time periods of either the primary or secondary days, the clinic is called by the coordinating center to ascertain the problem.

Data Management

Traditionally, a database management system facilitates the orderly collection of large volumes of data, often from many sources, and assists the investigators in the preparation of data for storage and subsequent analysis by a separate statistical package. Most database systems allow for error checking and other quality control checks at the time of data entry. A distributed data processing system organizes the collection of data, error checking, and other logical checks directly at the clinical center and at the time of or shortly following the patient visit.

In a DDP system, participant data are received on a timely basis. A primary drawback of a paper system that sends forms from the clinics to the coordinating center is the length of time that must elapse before data enters the central computer. In some trials using paper systems, it may take from two to three months to as long as six months before forms are entered into the main computer (1).

Initial results from the first five months of the SHEP DDP system demonstrate the timely receipt of the data. Table 1 presents the distribution of number of weeks that elapsed between the time two baseline visit paper forms were filled out and the time the edited, electronic revisions were received at the coordinating center computer. Most clinics transmitted most of their forms within one to two weeks. The overall median time from the initiation of a form to its receipt at the coordinating center was 1.14 weeks for the baseline visit 1 form and 1.00 weeks for the baseline visit 2 form.

After the electronic transmission of forms from a clinic is completed, they are written on to a file as new data records and a summary report is generated. This report is sent to the clinic at the end of the next week's transmission. Each data record is equivalent to one SHEP form. These formlength records are appended to a file of previously received forms, which are waiting to be added to the SHEP master database, called the SHEP masterfile.

The electronic transmission of files on a weekly basis allows, in turn, frequent updates of small special-purpose databases. For example, an important function of the coordinating center is to provide timely reports on recruitment and randomization. To facilitate this task, a "minidatabase" is maintained. Each record contains items from baseline visits concerning birthdate, blood pressure, exclusion criteria, etc., plus eligibility information concerning randomization. The minidatabase is updated with the weekly transmission. This file, once written, is never changed. To prevent drift between this file and the SHEP masterfile, which does undergo correction processes, the minidatabase will be created anew from the masterfile after each masterfile update.

As part of the randomization procedure, eligibility criteria obtained at the first baseline visit are verified at the time of telephone contact. A small data file is maintained which contains records for all participants who are still eligible at the end of the first baseline visit. This file is automatically updated with the weekly transmission of new files. At the time of randomization (10-28 days after the first baseline visit), this file is searched for a participant's record and the eligibility information is displayed on the screen and verified with the caller.

All analyses and clinic monitoring reports will be based on data obtained from the masterfile. The software used to update the masterfile and to retrieve from it were written and are maintained by coordinating center personnel. This software has been in development for over 15 years. It now is quite general, and is used for all our multicenter collaborative clinical trials.

The update and retrieval software is table-driven. That is, study specific information is coded into a set of tables which can then be accessed. To add a form to the study, or to add a new data item to an old form, takes only a few hours. No software needs to be altered. Only some tables need to be regenerated with new information put in its proper place.

Once the data from all clinics are received, the data on the masterfile are updated and error checking across forms is done to ensure that appropriate forms have been entered and transferred. If errors are found, appropriate error messages are written to a file which can be transferred back to the clinics at the next scheduled transfer. The clinic can either (1) complete a special form and transmit this form to the coordinating center where the actual correcting of the masterfile takes place, or (2) re-enter a form to be used as a replacement for a previously transmitted form. Also, status reports are transferred back to the clinics to provide information on the number of forms transferred and number of errors detected.

Quality Control

Quality control is a major concern in clinical trials. During all phases of the study, sufficient effort should be spent to ensure that all key data are of high quality. A major part of data quality control consists of the detection, review, and correction of errors in the collected data. A variety of manual and computer procedures have been used in clinical trials for error detection and correction (4,5).

In the previous section, several quality control procedures used to ensure the receipt of accurate data were discussed. This section will elaborate on these procedures and, in addition, describe other quality control procedures of the DDP system.

At the coordinating center, identification information from each paper form received is entered on a file and checked against the electronically transmitted forms on the masterfile. If no masterfile record exists, the form will be entered at the coordinating center. A masterfile form for which no paper form has been received is identified and a request is made to the clinic to send the paper form to the coordinating center.

Range and consistency checks are performed for each data item obtained in SHEP. Any invalid entries are detected at the time of entry and the operator is requested to correct the data before proceeding. In the event that the required information cannot be validated at that time, the clinic operator is requested to document an explanation and submit this report with a copy of the form to the coordinating center.

The system software keeps track of the number of times data have been entered. The system requires double data entry before files can be transferred. However, even blind re-entry does not ensure that data re-entered at the clinic will be correct. In order to monitor this potential problem, a random sample of paper forms will be selected for re-entry at the coordinating center to verify that the data have been entered correctly.

Response and range checks that were performed on participant data on the microcomputers are repeated on the central computer. In addition, consistency checks across forms are made. Failures of any of these checks result in edit reports which are transmitted back to the clinical sites at the next scheduled transmission. Each item in the report is reviewed and appropriate action taken.

Weekly monitoring reports are produced using the SHEP minidatabase, described previously. Information is generated on participant recruitment at the two baseline visits, including the number interviewed, the number scheduled to be interviewed, reasons for ineligibility and breakdowns by medication status at the initial contact and by Clinical Center. A report on the number of randomized participants is broken down by Clinical Center and displays (1) the number interviewed at initial contact, (2) the percent randomized of the number expected, (3) the percent of the total number randomized, and (4) the number randomized during the past seven days (Table 2). At the end of the weekly transmission period, this report is transmitted to each clinic.

Personnel Training

Local entry of data is accomplished by personnel who are generally not full-time computer programmers or data managers. As such, the software for the microcomputers of the DDP system must be made more user-friendly than that used in centralized systems. There must be interactive programs that users with little or no computer experience can easily use.

Two-day training sessions were held at the coordinating center for clinic personnel to learn how to enter, transmit and receive data. A comprehensive instruction manual was prepared to assist in this training.

Each clinical center had a training account. A successful practice transmission was required before any center was allowed to transmit actual trial data.

Implementation Time and Effort

Full time work on the DDP system began in July 1984. Several coordinating center personnel were involved in its design and implementation. Statisticians, systems analysts, computer programmers, forms designers, data entry operators and secretaries all made substantial contributions.

The main bottleneck in the implementation of the system was the completion of the SHEP forms. Once this was accomplished, it took three months for the system to become operational. Time was needed to (a) put the forms in and test the system, (b) develop software for transferring data, and (c) prepare an instructional manual and training sessions for clinical center data entry operators. The SHEP DDP system was operational by January 1985.

Mid-trial Modifications

During the study, forms may be changed. If this happens, each center will receive new software which includes these form revisions. This software can be downloaded from the central computer to the microcomputer or can be sent through the mail on a diskette.

DISCUSSION

The SHEP DDP system has been in full operation since March 25, 1985. Since that time, all clinics have successfully transmitted data. Of the 25 possible form types in the study, 22 have been used in the trial to date. There have been no problems in the data entry and transmission of these form types.

Initially, unverified incomplete records were being sent to the coordinating center. To prevent this from occurring, a special program was written and sent to the clinical sites to detect unverified and incomplete records before transmission. Also, software at the coordinating center was modified to disallow acceptance of unverified and incomplete records.

A DDP system provides several opportunities for enhancement. We are presently planning to include an electronic message center. This will allow notices concerning conference calls, meetings or other reminders to be transmitted to the clinics from the coordinating center. Clinical centers will also be able to transmit messages to each other via the coordinating center computer.

In addition, a DDP system can promote protocol adherence. Presently, we are using the system to aid in randomizing only protocol-eligible patients. If patients have missed required visits or lab work, the clinical center will be notified via the edit reports.

A further possibility is using the microcomputers for local patient management, scheduling, and data analyses. Such a system would allow each clinic to maintain the trial records on each patient but also allow the clinics to maintain other pertinent patient information, e.g., data for an ancillary study. The performance of local data analyses can be accomplished, but such an undertaking must be considered from the viewpoint of the greater cooperative trial. Such local analyses have the potential for compromising the larger study.

A possibility for DDP systems used in a future multicenter clinical trial is local randomization. Patient eligibility could be determined at the clinical site by having the microcomputer access the central computer. Once eligibility is established, randomization could be performed by the central computer.

In summary, distributed data processing systems are now in use in large multicenter clinical trials. They increase the quality of data collected in the trial and greatly lessen the time required to update information and determine ongoing problems.

REFERENCES

1. Kronmal RA, Davis K, Fisher LD, Jones RA and Gillespie MJ: Data management for a large collaborative clinical trial (CASS: Coronary Artery Surgery Study). Computer Biomed Res 11:553-566, 1978.
2. J Jeffreys for the HPT Investigation Group: Performance characteristics of the Hypertension Prevention Trial distributed data system. (Abstract) Controlled Clinical Trials 4:148, 1983.
3. Molvig K, Fox C, Bagniewska A, Black D, Edlavitch S, Hulley SB: Planning a distributed data processing system for a multicenter clinical trial. (Abstract) Controlled Clinical Trials 4:158, 1983.
4. Knatterud GL: Methods of quality control and of continuous audit procedures for controlled clinical trials. Controlled Clinical Trials 1:327-332, 1981.
5. Karrison T: Data editing in a clinical trial. Controlled Clinical Trials 2:15-29, 1981.

ACKNOWLEDGMENTS

The research upon which this publication is based was performed pursuant to Contract No. NO1-HC-48052 with the National Heart, Lung, and Blood Institute, and the National Institute on Aging.

SHEP DDP SYSTEM

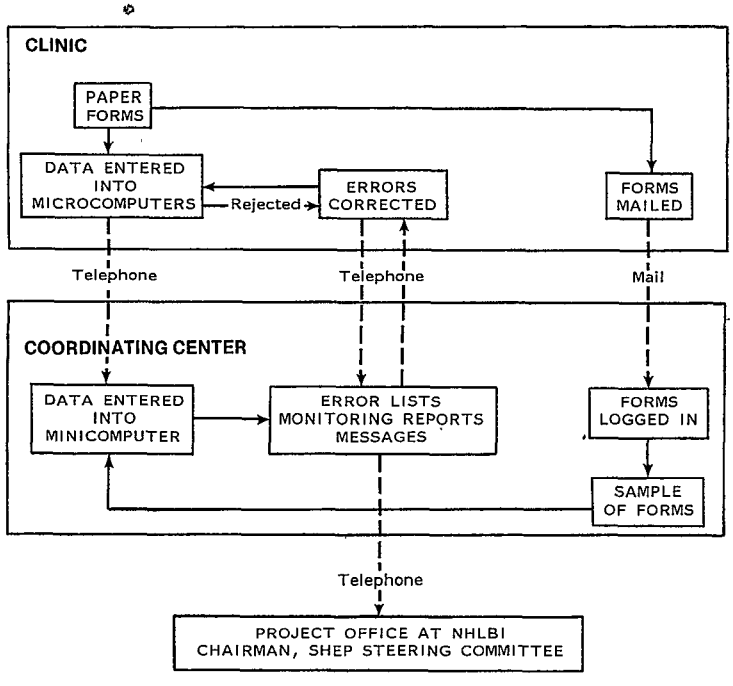


FIGURE 1

TABLE 1

Time from initiation of baseline visit forms to receipt at coordinating center.

Form	Number of Forms	0-1 Week (%)	1-2 Weeks (%)	>2 Weeks (%)	Median Number of Weeks
BV 1	1626	47.4	28.8	23.8	1.14
BV 2	887	53.9	29.3	16.8	1.00

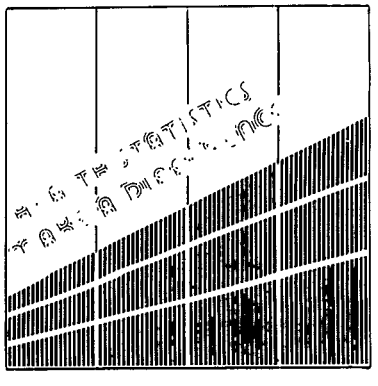
TABLE 2

Randomized SHEP Participants by Initial Medication Status and Clinical Center

Clinical Center	Initial Medication Status		Number Randomized in Clinic	% Randomized of Number Expected+	% of Total Randomized in Study	Number Randomized in Past 7 Days
	On Meds	Off Meds				
1						
.						
.						
.						
17						

Second Plenary Session

Statistics Make a Difference—The State Perspective



THE ROLE OF STATISTICS IN STATE HEALTH POLICY DECISIONS

Bailus Walker, Jr., Health, Massachusetts

INTRODUCTION

At all government levels, experience has shown that making rational choices among alternative health policy decisions depends heavily on careful interpretation of quantitative information provided by health statistics. Even though the tasks of collecting and analyzing health statistics may be costly, difficult, and time-consuming, these tasks are precursors to the formulation and implementation of health policy, and of programming consistent with established policy.

In the past, minimal registration and reporting were apt to be the only health statistics activities conducted by state health departments. Registration of births and deaths often was relegated to clerical personnel who lacked formal training in public health and who accepted, transcribed, and filed documents after only cursory examination. These routine procedures rendered health statistics "dead" in a literal sense.

Now, when all available health statistics are collected, correlated, and analyzed by a health department, they become of fourfold value in problem-solving. They make possible: 1) problem definition; 2) the development of logical programming for problem-solving; 3) planning of procedures and records, for administration and analysis of the programs as they progress; and 4) evaluation of program results.

GENERAL TRENDS

To illustrate current practice, I will first describe some general trends in the use of statistics for development of health policy and programs, and then report a few specific examples from our experience in Massachusetts.

For a number of years there existed an administrative separation between health professionals who collect, analyze, and present statistical data, and those who work in direct health service areas such as environmental health, chronic disease prevention and control, maternal and child health, and health system planning. Strong arguments were made for not including statisticians in policy-making deliberations; it was implied that the statistician lacked requisite expertise and experience. Now health service providers are increasingly recognizing how important statisticians' contributions can be to the decision- and policy-making processes. Cochran has stated the case well:[1]

> There are strong arguments for the presence on policy-making boards of statisticians experienced in collaborating with health workers. They can help to explain the meaning of data, to draw attention to the subject, to avoid statistical fallacies, and to discuss the implications of the inaccuracies inevitable in all data.

Technological advances during recent years, such as those in computers and in statistical sampling techniques, have considerably enriched our information sources.

Public health planners and program managers are beneficiaries of the National Health and Nutrition Examination Survey, the National Health Interview Survey, and other ongoing efforts to build an adequate health data base for the United States.

These efforts mean that we are much better equipped than were our predecessors to develop and establish health policy and programs. Particularly in periods of resource constraints, data are needed to sharpen our focus on the areas of greatest need and on the outcomes of interventions we undertake. Having better information available does not necessarily mean that our decisions will always be wiser, but it does mean that we cannot claim to be ill-informed.

As the health care system attempts to respond to the conflicting pressures of economic constraint and budget limitation on one hand, and newly-emerging health care needs and new technological capabilities on the other, there is increasing demand for reliable data. We need information that will help us identify groups at risk of needing health services and translating their needs into demand for use of medical care resources. In this context, Pollack has shown[2] how data from various national sources can be put together in a number of ways to predict demands on the medical care system, to indicate the size of the groups affected by specific proposed changes in the system, and to evaluate the impact of such changes over time.

In another area of public health responsibility, that of environmental health, health statistics development is proceeding apace as the importance of exposure to environmental toxicants is increasingly appreciated. Much of the credit for this shift in the importance of environmental health vis-à-vis health statistics must go to the National Center for Health Statistics and its establishment in the late 1970s of a Division of Environmental Epidemiology. Public Law 95-623 expanded the new division's program by directing the Center to prepare guidelines for determining the effects of employment and environmental conditions on the public's health.

The National Center also is supporting the states in their coding of occupational information on death certificates, a development which helps identify occupational differentials in mortality. Of great importance to environmental health programs is the vast array of data collected through the National Center's major data systems - including the already-noted National Health Interview Survey and National Health and Nutrition Examination Survey which developed data on pesticide residues and metabolites in blood and urine. These data have been used to identify and assign priorities for research on the health effects of pesticides to which large segments of the United States population are being exposed.

It is reasonable to anticipate that statistics on environmental factors affecting health will continue to improve. These new data

are essential to enhance our epidemiological knowledge so that we can translate experimentally-established principles into practical prevention programs.

In summary, the increasing overall importance accorded health statistics generally reflects improvement in systematic approaches to problem-identification and problem-solving essential for attainment of measurable health goals and objectives.

In Massachusetts we are increasingly using health statistics to delineate health problems in a fashion that will attract attention, achieve understanding, and gain appropriate responses on the part of health administrators and community and governmental leaders. Some specific examples illustrate this approach.

STATISTICS AND ADDICTIVE DISEASE CONTROL

In the fall of 1984, the Massachusetts Department of Public Health conducted a survey of drug and alcohol use among secondary school students in the state. Sixty percent of those surveyed had used one or more illicit drugs during their lifetime and 31 percent had used one or more illicit drugs in the month prior to the survey. Analysis of lifetime illicit drug use indicates that marijuana is the drug of choice; marijuana was used by 51 percent of the respondents, amphetamines by 24 percent, and cocaine by 17 percent.

These statistics and several other pieces of data defined in a clear and concise way the problem of drug abuse among Massachusetts adolescents. Reviewing these data, Governor Michael Dukakis expressed personal interest in the problem. He chaired a number of statewide meetings of community leaders and state and local officials, to glean ideas for a comprehensive multifaceted state-supported program. Program strategy was worked out based on a statistical analysis of where the state was and is, and what possibilities there are for future progress. Since the formulation and official pronouncement of state policy, and the commitment of necessary resources, this evaluation process has been continued, using epidemiological analyses of trends in teenage drug abuse.

Here is a clear example of using statistical data to adequately bring understanding of a modern health issue to the general public and to special groups that need to comprehend addictive disease issues - legislators, educators, primary and secondary school students, religious and social leaders, ethnic minority groups, and criminal justice students, to name a few of those most affected.

STATISTICS AND CHRONIC DISEASE PREVENTION

A second example concerns the three major chronic diseases - heart disease, cancer, and stroke.

In 1981, integrated statistical data from several sources indicated that these so-called degenerative diseases had reached epidemic levels, accounting for more than 65 percent of all deaths in Massachusetts. Approximately one-half of these deaths occurred before age 75, the average life expectancy in the United States, and therefore could be classified as premature. We calculated that the economic burden on the state resulting from these deaths was $1.5 billion, including health care costs of $500 million.

Combining our findings with existing knowledge of these diseases' etiology, we concluded that the evidence was compelling that many deaths from heart disease, cancer, and cerebrovascular disease could be prevented through the reduction of well-known risk factors; what was needed was a coordinated statewide effort focused on reduction of multiple risk factors.

Armed with a statistical analysis and a comprehensive proposal to address the critical issues, we held a series of discussions with committees of the Massachusetts Legislature. Here we translated statistical data into economic, social, and political impact statements. We then spelled out what the state might expect to achieve by instituting a comprehensive chronic disease prevention program: a decline in the age-adjusted mortality rate for heart disease of 3.7 percent annually; simultaneously at least a 7.0 percent annual decrease in the age-adjusted mortality rate for cardiovascular disease; and after ten years a 5.0 percent per year decline in the age-adjusted mortality rates for cancer.

The Legislature responded to both these facts and to our proposal for action by appropriating $2.2 million for the establishment of the Massachusetts Center for Health Promotion and Environmental Disease Prevention within the Massachusetts Department of Public Health. The Center is now implementing an aggressive statewide program using proven methods of risk reduction and of intervention, including modification of physical and chemical hazards in occupational and nonoccupational settings, and of individuals' behavior patterns.

In this pluralistic approach involving both public and private participants, another data source, the Massachusetts Cancer Registry, is being utilized to study risk factors in cancers. Using Registry data, we have prepared various compilations and tabulations which answer to our administrative and programmatic needs. In particular, geographic clusters and unusual sporadic associations have been the object of rigorous analyses. Findings of unexpected associations such as rare tumors in towns and cities with environmental pollution problems have prompted special studies: of leukemia in the town of Woburn; of pancreatic cancer in Peabody; and of kidney cancer in several Merrimack Valley communities. Our experience so far indicates that there are limitations even in well-established registries. Although inherent limitations should be taken into account in evaluating findings, they by no means invalidate use of registry data for epidemiologic studies and program planning.

STATISTICS AND INFANT MORTALITY

A third example of use of health statistics to shed light on current health problems and what to do about them relates to infant mortality.

In Massachusetts the infant mortality rate increased from 9.6 infant deaths for every 1,000 live births in 1981 to 10.1 in 1982. This was

the first increase in the state in nine years and the largest in seventeen years. Additionally, comparable to trends across the country, the infant mortality rate for blacks in Massachusetts was more than double that for whites. Also, a significant geographic variation in rates has persisted in Massachustts and descriptive statistics have identified discrepancies in available access to prenatal care in certain components of the health care system.

Recognizing that solving problems relating to infant mortality would require efforts by both public and private sector groups, we convened a nineteen-member Task Force on Prevention of Low Birthweight and Infant Mortality. The group's membership represented a wide range of expertise, experience, knowledge, and perspectives on health issues. Its members were asked to address low birthweight as well as infant mortality, and to advise the Commissioner of Public Health on strategies which could be implemented to address both problems. The Task Force deliberated for eight months and submitted its report in May 1985. Its recommendations, summarized within five broad strategy areas, note strong imperatives for a comprehensive plan of action and propose initiatives designed to reverse current trends.

These recommendations are providing a blueprint to move Massachusetts forward from the level of progress already achieved. Additional resources to implement the recommendations were included in the Fiscal Year 1986 state budget, not only for the Department of Public Health, but also for other agencies such as the Department of Public Welfare, where over $1 million was added to provide maternity care for low-income uninsured teenagers, thus reducing financial barriers to health care for one of the target high risk groups identified in the task force report.

STATISTICS AND THE HEALTH CARE SYSTEM

Some of the most difficult decisions public health departments now are having to make relate to the revolutionary changes that are sweeping the entire health care system. Factors such as competition, new technology, and new payment mechanisms have come together so rapidly, we have been hard pressed to respond speedily and appropriately. Having health statistics readily available has been crucial to formulation of constructive policy and program determinations.

After many years of growth and high occupancy levels, hospitals in Massachusetts, as elsewhere, are experiencing a downturn in their occupancy rates. When this trend first emerged, some in the hospital field viewed it as a short-term slowdown in utilization. But it is now clear that significant long-term change is taking place. Causal factors in Massachusetts include: increasing competition among health care facilities and for-profit chains; more market penetration by health maintenance organizations and various free-standing special-service clinics and centers; overall changes in forms of medical practice; and greater emphasis on and awareness of wellness. Because of a non-DRG reimbursement system established in Massachusetts three years ago, hospitals in the state have had a waiver from the Medicare-DRG system. This expires October 1, 1985 and we expect then to be included in the Medicare system which has affected hospital utilization rates so dramatically in other parts of the country.

While debate continues about which factors have had the greatest impact, it is evident that the health care system must be downsized in ways consistent with the kinds of system changes that are occurring. To decide on the type and amount of downsizing necessary, to define "appropriate hospital capacity," or to state the minimum size at which a health care facility can operate effectively and efficiently, agencies charged with making such determinations must have accurate data from a number of sources.

And we must know specifics about the administrative structure necessary to maintain adequate medical records as well as physical plant and grounds. We must be able to determine the possible impact of closing an entire section or department in a facility. If neurological services are reduced, what will be the effect on existing surgery, radiology, pathology, and physical therapy services? To address these sorts of questions, now being asked frequently by hospital boards and administrators, health planners, and regulatory agencies, we are turning increasingly to the models and methodologies of operations research and other management sciences. We are recognizing that informed use of quantitative models and techniques can lead to better policy planning and utimately to greater benefit for both the community and the health care industry.

A useful approach to analyzing the practicality of reduction in hospital capacity is a model that takes into account the state's census data, particularly on the size of various age groups in the population; fertility rates; the number of Medicaid patients awaiting long-term care services; the number of enrollees projected by various health insurance plans; and expected numbers of inpatient days. Using such a model, we have experimented with changing various parameters. Based on the results, an analyst can infer how different facility or system configurations - i.e., different numbers of beds and different kinds and quantities of services - will behave under given sets of circumstances. These sorts of complex analyses/evaluations require the most complete and accurate health-related data obtainable. Fortunately, in Massachusetts we have, and are continuing to develop, both governmental and private sector sources of data that can enhance decision-making at this level and can allow us to take into account various social and political as well as economic dimensions of the health care system.

Concommitantly with decisions about inpatient facilities in the health care system, state public health and health planning agencies are being called upon to decide how many and what kinds of non-hospital health service facilities are needed, in what geographic locations. For example, the Massachusetts Department of Public Health recently was confronted with two separate applications for permission to establish birth centers in one of the state's health service areas. In the absence of an as

yet established formal method for determining need for this new type of facility, we began by determining the need for additional maternity services in the area.

The Department's Division of Health Statistics and Research displayed statistics which showed the number of births in Massachusetts to have increased since 1976. The data also indicated that the number of births occurring outside of hospitals had increased, from 217 in 1970 to 634 in 1982. Using 1990 population projections, normative use rates including those for Medicaid obstetrical utilization, and patient origin and case mix data, we developed a projection of obstetrical bed need. Factored into this equation were the infant mortality rate and the need for comprehensive prenatal care readily accessible to the population at risk in the communities under consideration.

We further analyzed the financial feasibility of establishing one or both birth centers, including proposed charges per delivery and the projected operating costs for the first full year.

Based on these analyses, we concluded that: birth centers are appropriate alternatives to inhospital delivery facilities; a birth center should be a regional resource and centrally located to serve the entire region; and the initiation of a birth center can dramatically decrease costs and charges for obstetrical delivery. However, since a birth center in this instance was an addition of resources rather than a substitute, the total cost to the health care system would increase rather than decrease.

It was decided that approval should be granted for establishment of one centrally located birth center, to serve the population with the greatest need for the services made available. We chose to err in favor of protecting the health of mothers and infants, rather than protecting money.

CONCLUSIONS

These few examples of the use of health statistical data in Massachusetts indicate how difficult, if not impossible, it would be for a modern state health agency to develop policy, plan programs, provide services, and achieve its disease prevention and health promotion goals without such information. Also obvious is the fact that state health statistics alone are insufficient for meeting existing needs.

Over recent years, development of a well-designed national health data system comprising a family of interrelated health surveys, and improved vital statistics systems, provide reliable and aggregated sources of information about a vast array of health conditions, health problems, health service resources, and costs of health care. This veritable gold mine must be utilized in conjunction with available regional, state, and local data.

The widespread use of available data gives clear indication that there no longer can be any debate over the critical role of statistical information in development of health policy and direct health service programs. Equally crucial is the role data systems are playing in bringing about a clearer understanding of health issues on the part of the general public and various special groups with particular needs. Our legislators, local and national leaders of industry and business, labor leaders and their constituencies, ethnic minority groups and their leaders, all need and are entitled to full and accurate information.

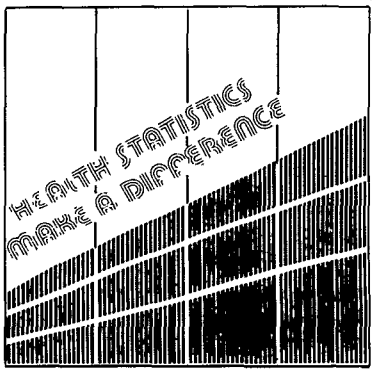
Massachusetts' experience in translating health statistics into policy, programs, and services suggests that building health service excellence requires that managers and administrators go beyond understanding the issues and acting on their understanding. We also must communicate our vision to all in our own agencies and to the community at large, so that together we can keep pace with change and meet the challenges it presents to all of us.

REFERENCES

1. Cochran, William G. The role of statistics in national health policy. American J. of Epidemiology 104:375, 1984.
2. Pollack, Earl S. The epidemiology of cancer and the delivery of medical care services. Public Health Reports 99:476-482, 1976.

Session I

Automated Systems for Collecting Client and Services Data



USE OF A SINGLE AUTOMATED PATIENT MANAGEMENT SYSTEM FOR PROGRAM EVALUATION, BILLING, AND HEALTH RESEARCH

Terri Combs-Orme, Northwestern University; James Masterson, Chicago Department of Health; Janet Reis, Northwestern University; Glenn Good, Northwestern University; Brenda Davis, Northwestern University

Recent years have seen decreasing amounts of Federal money available for many health care and related projects, and the result has been increasing pressure to document the need for and effectiveness of the programs competing for that money. As part of this process, publicly supported ambulatory health care centers have been increasingly called upon to document effective and efficient service delivery efforts. In response to this call, and partly because of burgeoning technologies, there has been an increasing demand for sophisticated automated data collection systems in the health care field (c.f. Densen, 1973; Freeborn & Greenlick, 1973).

The Chicago Department of Health (CDOH) faced the need for better methods of data collection in the early 1970's and responded with the development of a uniform single reporting system called the Patient Registry System or PRS. This presentation describes the history and development of that system, how it works to document the services delivered to about 250,000 patients per year, current uses, and future plans for the PRS.

History and Development of PRS

CDOH, like most other departments of health throughout the country, exerienced great growth through the late 1960's and early 1970's due to the large number of categorical grants obtained. The extensive reporting requirements associated with these grants resulted in a great deal of data collection to generate many monthly, quarterly, and annual reports. The volume of these documents, each generated from independently maintained data sources, was one of the pressures to improve the data collection procedure. In addition, as Federal monies began to decline, competition arose among the various health care agencies in the city, and it became necessary to justify the agency's continued existence and funding level.

Development of the PRS began in 1976 with the formation of a team composed of personnel from administration, planning, Maternal and Child Health, Child and Youth Care, Nursing, and neighborhood health center directors. This team surveyed administrative and service personnel in the various facilities to determine their patient load, the extent of their needs for data and data collection procedures at the time, and other details.

In addition during this time the design specifications for the system were conceptualized. Customary data collection forms were reviewed to determine which items needed to be retained and which eliminated. The decision was made to keypunch data in the larger facilities and to send data collection forms from the smaller facilities to a central site for processing. The design specifications for the system required approximately 18 months for final approval.

Due to the need to retrain all personnel from data entry clerks all the way up to clinic managers, implementation of the PRS was accomplished gradually. In 1979 the system was implemented in one test site, and revisions were made in the system based upon what was learned. Full implementation required three and one-half to four years and was just recently completed.

How the Patient Registry System Works

The Patient Registry System is an online interactive communication system using a hierarchical data base for storing and retrieving data. It requires approximately 30 remote terminal operators and a professional support staff of three. Once completed, data are entered directly from the forms by remote terminal operators working on-site at most of the facilities where the services are provided. The smaller facilities, representing about one fourth of the total service volume, forward their data collection forms to the central office in downtown Chicago for processing. The system runs on an IBM 4381 in an OS/MVS environment. IBM 3276's and 3278's are used for data entry.

With its installment, the PRS replaced six or seven independent systems which were all "batch mode" systems. Under this old system, completed documents were forwarded from the service delivery sites to a downtown location where they were keypunched. A computer edit detected errors, and forms were sent back to the facility for correction in this difficult and time-consuming process. Moving the point of data entry closer to the point of data creation has greatly improved the timeliness and accuracy of the health statistics now available.

Data Collection Forms

The health care services provided in the Department of Health's 19 facilities are recorded using two PRS data collection instruments. These instruments are the registration/enrollment form and the visit/encounter (V/E) form.

The registration/enrollment form (Figure 1) collects basic demographic, program of care, and billing data on registered patients. With the exception of one-time-only and urgency visits on patients for whom a medical record will not be constructed, all patients must be registered into the PRS. A unique patient identification number is created from a complex algorithm based on selected patient characteristics, and this number stays with the patient throughout his or her history with the department.

The V/E History Form (Figure 2) is completed upon each patient visit and collects a wide variety of items, including demographic variables such as race, age, income, family size, and geographic area of residence. This four page document also records information on 13 programs of care in which the patient may be enrolled, such as maternal, family planning, and adult care. Information is collected on the number and kinds of services received by the patient enrolled within a program of care. Specific

services include dental, social service, laboratory, etc. Type of provider providing the service (for example, physician or nurse) is also recorded on this form.

As is clear from the slide representation of the V/E form, individual treatments and procedures are also collected on this form. Within the maternal clinic, for example, there are codes for medical history, physical exam, postpartum exam, and others.

Uses of the Patient Registry System

The PRS allows a great deal of flexibility in the development of both routine and ad-hoc reports for the purposes of program planning and evaluation, compliance with existing reporting requirements, and the generation of bills.

One important use of the PRS is the production of utilization statistics for program evaluation and program planning. There are four basic statistics that are available for analysis by program of care, by facility, by community area, or by any combination thereof. These four utilization statistics are unduplicated patient counts, visits, units of service, and encounters. "Units of service" refer to visits to a specific service area, such as the dental clinic. Since patients will usually be seen in more than one service area on a single visit, each visit will usually involve multiple units of service. By contrast, an encounter is defined as a face-to-face contact between a patient and a provider, so each unit of service might involve encounters with both a nurse and a physician.

Table 1 shows an example of a regular utilization report for one clinic for one month. This report shows unduplicated numbers of patients, number of clinic visits, and visits per patient. The other columns show services and providers for each of the many service areas. This information is provided for each program of care and for all registered vs. unregistered patients. More detailed reports in this same job stream provide specific information by facility, program of care and cluster, so that facility managers can determine the demand on the various service areas.

Billing is also an important function of the PRS. Table 2 is a summary report of sources of income for the same facility used in the previous example. The great majority of services at this facility are provided to medically indigent patients. The Chicago Department of Health does ask that patients who can pay do so on a sliding fee scale, however patients are not refused service if they do not pay.

The PRS system is currently used to bill Medicare, through the creation of forms, such as the currently generated Medicare 1500 form. Future plans call for providing this information on tape. Beginning in October we will be using the PRS to generate tapes containing all the necessary data to obtain reimbursement from the Illinois Department of Public Aid for services provided.

The PRS is also useful for program planning. Currently we are beginning the process of merging data from the PRS with data from a Budget Expenditure System to create a management information data set. The data set allows us to analyze costs in order to examine the efficiency of service provision in the various facilities. Facility managers may want to know the use of various levels of professionals (e.g., physicians vs. nurses) for certain procedures in order to project personnel and budgetary needs, and to use this information to project costs and productivity. Such reports are generated on an as-needed basis. Thus we can map each of our 19 facilities and compare their cost per unit of service, whether unit of services is conceptualized as a visit to the facility or as a visit to individual clusters within a facility.

In regard to program evaluation, the PRS is being used in a collaborative study between CDOH and Northwestern University to evaluate a Federal Maternal and Child Health Block Grant demonstration project in one high-risk area of the south side of Chicago. PRS provides information on whether outreach services are bringing in more patients, and especially high-risk patients, in the target area; whether the mandate for comprehensive services is indeed being met; and whether the timing and the content of prenatal care are consistent with accepted standards. Vital statistics data for the area may then be examined to determine if changes in service utilization correspond to improvements in infant mortality and low birthweight.

Future Plans for the Patient Registry System

Since implementation of the PRS has recently been completed in all of the department's facilities, we now can proceed with improvements in the system. Current discussions on changes are focussing on elimination of items which are no longer required by grantors, and on clarification of the treatment and procedure codes to eliminate ambiguous and duplicate codes.

Recently a preliminary audit of the system was conducted using a representative sample of cases, in order to determine the accuracy and completeness of selected items of data on some system documents (Davis, 1985). Studied were items which are especially important for the administrative and billing components of the system, including patient name, address, program of care, dates of registration and next appointment, etc. The results of the audit indicate that recorded data were highly consistent and complete on the studied documents. For example, data on the Visit/Encounter form were found to be 90% complete and consistent when compared to other sources. We intend to conduct ongoing audits for data quality and to correct the problems found, since any system must maintain constant vigilance to ensure high-quality data.

Another important goal for the PRS is to link it to birth and death records in order to assess outcomes for pregnancies serviced by the Department. This procedure is more complex than was originally anticipated, and currently matches are being achieved in slightly over 50 percent of cases. However, we intend to pursue this problem of linking PRS with Vital Statistics in order to create an outcome data file for births. Such a system will be useful in the ongoing, and thus far encouragingly successful, effort of the Chicago Department of Health to serve the health needs of Chicago.

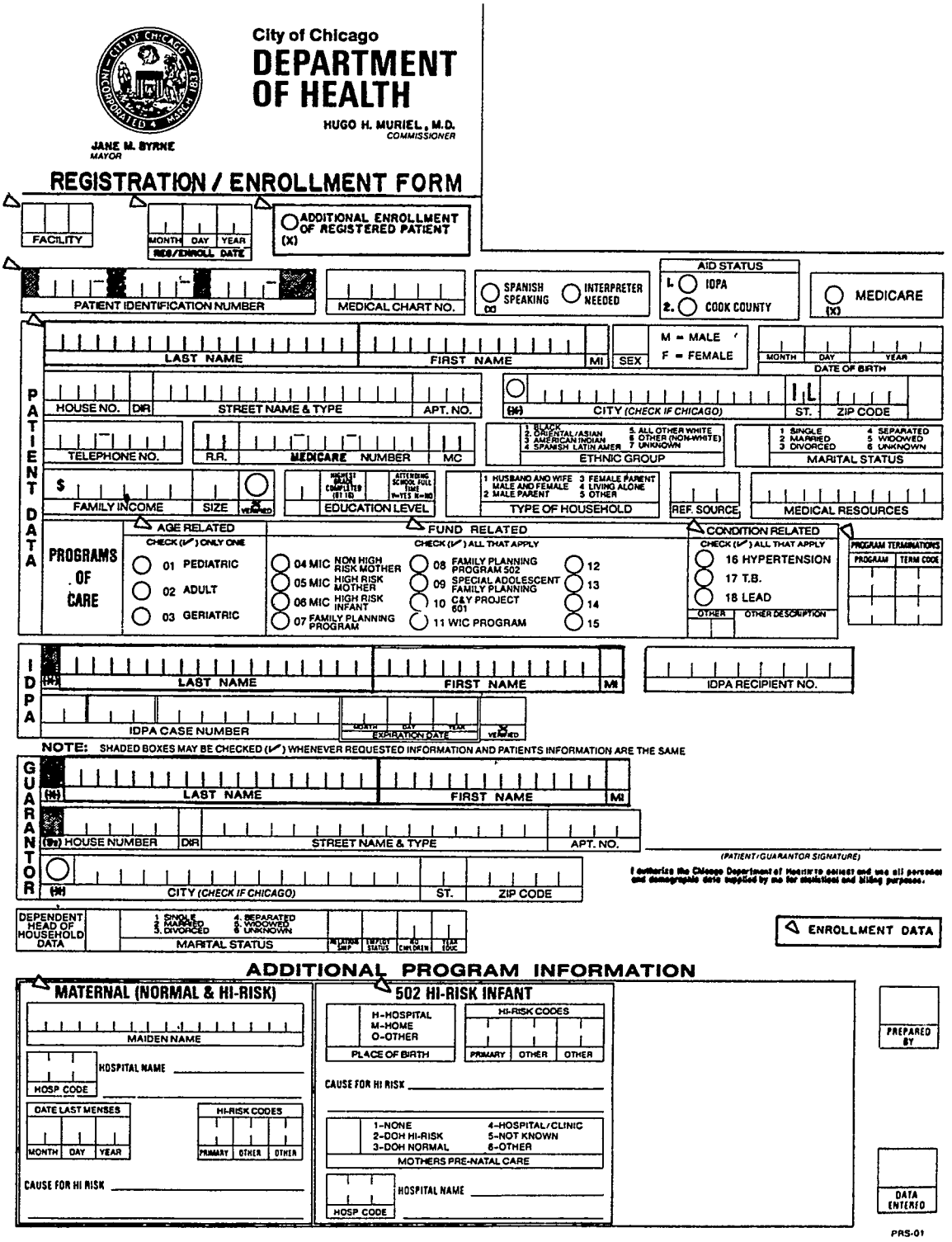
References

Davis, Brenda. Evaluation of the Chicago Department of Health's Patient Registry System: An automated data management system. Unpublished research project, University of Illinois School of Public Health, 1985.

Densen, Paul M. The design of reporting systems for health care programs. Medical Care, March-April 1973, XI (2), Supplement, 145-157.

Freeborn, Donald K. and Greenlick, Merwyn R. Evaluation of the performance of ambulatory care systems: Research requirements and opportunities. Medical Care, March-April 1973, XI (2), Supplement, 68-75.

FIGURE 1



City of Chicago
DEPARTMENT OF HEALTH
HUGO H. MURIEL, M.D.
COMMISSIONER

JANE M. BYRNE
MAYOR

REGISTRATION / ENROLLMENT FORM

FACILITY
MONTH DAY YEAR REG/ENROLL DATE
ADDITIONAL ENROLLMENT OF REGISTERED PATIENT (X)
PATIENT IDENTIFICATION NUMBER
MEDICAL CHART NO.
SPANISH SPEAKING
INTERPRETER NEEDED
AID STATUS 1. IDPA 2. COOK COUNTY
MEDICARE (X)

PATIENT DATA
LAST NAME
FIRST NAME
MI
SEX M = MALE F = FEMALE
DATE OF BIRTH MONTH DAY YEAR
HOUSE NO.
DIR
STREET NAME & TYPE
APT. NO.
CITY (CHECK IF CHICAGO)
ST.
ZIP CODE
TELEPHONE NO.
R.R.
MEDICARE NUMBER
MC
ETHNIC GROUP 1 BLACK 2 ORIENTAL/ASIAN 3 AMERICAN INDIAN 4 SPANISH LATIN AMER 5 ALL OTHER WHITE 6 OTHER (NON-WHITE) 7 UNKNOWN
MARITAL STATUS 1 SINGLE 2 MARRIED 3 DIVORCED 4 SEPARATED 5 WIDOWED 6 UNKNOWN
FAMILY INCOME
SIZE
VERIFIED
EDUCATION LEVEL
TYPE OF HOUSEHOLD 1 HUSBAND AND WIFE 2 MALE PARENT 3 FEMALE PARENT 4 LIVING ALONE 5 OTHER
REF. SOURCE
MEDICAL RESOURCES

PROGRAMS OF CARE
AGE RELATED CHECK (✓) ONLY ONE: 01 PEDIATRIC; 02 ADULT; 03 GERIATRIC
FUND RELATED CHECK (✓) ALL THAT APPLY: 04 MIC NON HIGH RISK MOTHER; 05 MIC HIGH RISK MOTHER; 06 MIC HIGH RISK INFANT; 07 FAMILY PLANNING PROGRAM; 08 FAMILY PLANNING PROGRAM 502; 09 SPECIAL ADOLESCENT FAMILY PLANNING; 10 C&Y PROJECT 601; 11 WIC PROGRAM; 12; 13; 14; 15
CONDITION RELATED CHECK (✓) ALL THAT APPLY: 16 HYPERTENSION; 17 T.B.; 18 LEAD; OTHER; OTHER DESCRIPTION
PROGRAM TERMINATIONS: PROGRAM; TERM CODE

IDPA
LAST NAME
FIRST NAME
MI
IDPA RECIPIENT NO.
IDPA CASE NUMBER
EXPIRATION DATE MONTH DAY YEAR
VERIFIED

NOTE: SHADED BOXES MAY BE CHECKED (✓) WHENEVER REQUESTED INFORMATION AND PATIENTS INFORMATION ARE THE SAME

GUARANTOR
LAST NAME
FIRST NAME
MI
HOUSE NUMBER
DIR
STREET NAME & TYPE
APT. NO.
CITY (CHECK IF CHICAGO)
ST.
ZIP CODE
(PATIENT/GUARANTOR SIGNATURE)
I authorize the Chicago Department of Health to correct and use all personal and demographic data supplied by me for statistical and billing purposes.

DEPENDENT HEAD OF HOUSEHOLD DATA
MARITAL STATUS 1 SINGLE 2 MARRIED 3 DIVORCED 4 SEPARATED 5 WIDOWED 6 UNKNOWN
RELATIONSHIP
EMPLOYMENT STATUS
NO. CHILDREN
YEAR EDUC.

ENROLLMENT DATA

ADDITIONAL PROGRAM INFORMATION

MATERNAL (NORMAL & HI-RISK)
MAIDEN NAME
HOSP CODE
HOSPITAL NAME
DATE LAST MENSES MONTH DAY YEAR
HI-RISK CODES PRIMARY OTHER OTHER
CAUSE FOR HI RISK

502 HI-RISK INFANT
PLACE OF BIRTH H-HOSPITAL M-HOME O-OTHER
HI-RISK CODES PRIMARY OTHER OTHER
CAUSE FOR HI RISK
MOTHERS PRE-NATAL CARE 1-NONE 2-DOH HI-RISK 3-DOH NORMAL 4-HOSPITAL/CLINIC 5-NOT KNOWN 6-OTHER
HOSP CODE
HOSPITAL NAME

PREPARED BY
DATA ENTERED

PRS-01

FIGURE 2

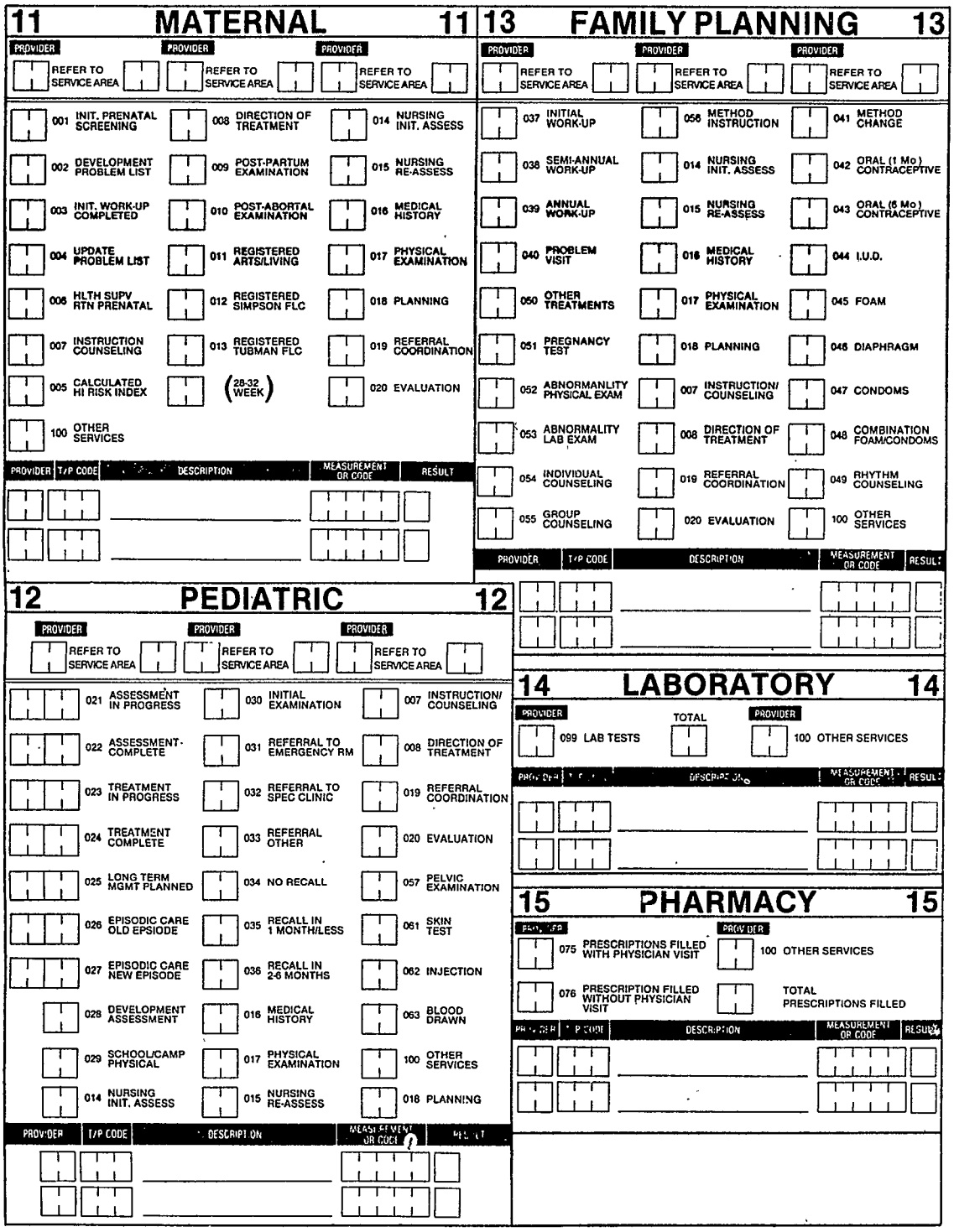

11 MATERNAL 11
PROVIDER
REFER TO SERVICE AREA
001 INIT. PRENATAL SCREENING
002 DEVELOPMENT PROBLEM LIST
003 INIT. WORK-UP COMPLETED
004 UPDATE PROBLEM LIST
006 HLTH SUPV RTN PRENATAL
007 INSTRUCTION COUNSELING
005 CALCULATED HI RISK INDEX
100 OTHER SERVICES
008 DIRECTION OF TREATMENT
009 POST-PARTUM EXAMINATION
010 POST-ABORTAL EXAMINATION
011 REGISTERED ARTS/LIVING
012 REGISTERED SIMPSON FLC
013 REGISTERED TUBMAN FLC
(28-32 WEEK)
014 NURSING INIT. ASSESS
015 NURSING RE-ASSESS
016 MEDICAL HISTORY
017 PHYSICAL EXAMINATION
018 PLANNING
019 REFERRAL COORDINATION
020 EVALUATION
PROVIDER
T/P CODE
DESCRIPTION
MEASUREMENT OR CODE
RESULT
13 FAMILY PLANNING 13
037 INITIAL WORK-UP
038 SEMI-ANNUAL WORK-UP
039 ANNUAL WORK-UP
040 PROBLEM VISIT
050 OTHER TREATMENTS
051 PREGNANCY TEST
052 ABNORMANLITY PHYSICAL EXAM
053 ABNORMALITY LAB EXAM
054 INDIVIDUAL COUNSELING
055 GROUP COUNSELING
056 METHOD INSTRUCTION
014 NURSING INIT. ASSESS
015 NURSING RE-ASSESS
016 MEDICAL HISTORY
017 PHYSICAL EXAMINATION
018 PLANNING
007 INSTRUCTION/ COUNSELING
008 DIRECTION OF TREATMENT
019 REFERRAL COORDINATION
020 EVALUATION
041 METHOD CHANGE
042 ORAL (1 Mo.) CONTRACEPTIVE
043 ORAL (6 Mo.) CONTRACEPTIVE
044 I.U.D.
045 FOAM
046 DIAPHRAGM
047 CONDOMS
048 COMBINATION FOAM/CONDOMS
049 RHYTHM COUNSELING
100 OTHER SERVICES
12 PEDIATRIC 12
021 ASSESSMENT IN PROGRESS
022 ASSESSMENT COMPLETE
023 TREATMENT IN PROGRESS
024 TREATMENT COMPLETE
025 LONG TERM MGMT PLANNED
026 EPISODIC CARE OLD EPSIODE
027 EPISODIC CARE NEW EPISODE
028 DEVELOPMENT ASSESSMENT
029 SCHOOL/CAMP PHYSICAL
014 NURSING INIT. ASSESS
030 INITIAL EXAMINATION
031 REFERRAL TO EMERGENCY RM
032 REFERRAL TO SPEC CLINIC
033 REFERRAL OTHER
034 NO RECALL
035 RECALL IN 1 MONTH/LESS
036 RECALL IN 2-6 MONTHS
016 MEDICAL HISTORY
017 PHYSICAL EXAMINATION
015 NURSING RE-ASSESS
007 INSTRUCTION/ COUNSELING
008 DIRECTION OF TREATMENT
019 REFERRAL COORDINATION
020 EVALUATION
057 PELVIC EXAMINATION
061 SKIN TEST
062 INJECTION
063 BLOOD DRAWN
100 OTHER SERVICES
016 PLANNING
14 LABORATORY 14
099 LAB TESTS
TOTAL
100 OTHER SERVICES
15 PHARMACY 15
075 PRESCRIPTIONS FILLED WITH PHYSICIAN VISIT
076 PRESCRIPTION FILLED WITHOUT PHYSICIAN VISIT
TOTAL PRESCRIPTIONS FILLED
100 OTHER SERVICES

TABLE 1

HLTH3124-01 REPORT DATE 08/20/85 * DEP'T. OF HEALTH PATIENT REGISTRY SYSTEM * FROM 7/01/85 TO 7/31/85 PAGE 18

C L I N I C A C T I V I T Y A N A L Y S I S B Y P R O G R A M A N D R E G I S T R A T I O N S T A T U S

CLINIC: 42

PROGRAMS AND REGISTRATION STATUS	PATIENTS	CLINIC VISITS	VISITS PER PATIENT	SERVICE AREA VISITS	SERVICE AREA ENCOUNTERS	TREATMENT PROCEDURE UNITS	UNITS PER ENCOUNTER
PEDIATRIC	864	1020	1.1	2237	3411	7829	2.2
ADULT	933	1109	1.1	2341	3079	7609	2.4
GERIATRIC	431	514	1.1	1096	1404	3439	2.4
MIC(NHR MOTHER)	115	159	1.3	418	548	1288	2.3
MIC(HR MOTHER)	12	18	1.5	40	53	164	3.0
MIC(HR INFANT)	16	18	1.1	28	51	114	2.2
712 FAM PLN PRG	247	262	1.0	546	731	2054	2.8
502 FAM PLN PRG	97	104	1.0	228	300	842	2.8
WIC PROGRAM	148	150	1.0	197	278	890	3.2
LEAD *	17	22	1.2	41	61	121	1.9
REGISTRANTS	2228	2643	1.1	5674	7894	18877	2.3
NON-REGISTRANTS	491	530	1.0	910	1231	3163	2.5
* * TOTALS * *	2719	3173	1.1	6584	9125	22040	2.4

TABLE 2

CHICAGO DEPARTMENT OF HEALTH 6:35 THURSDAY, AUGUST 8, 1985 177
BUREAU OF OPERATIONS RESEARCH

• • • PATIENT REGISTRY SYSTEM • • •

TREATMENT AND PROCEDURE CODES BY NURSE PROVIDER CITYWIDE
REPORT PERIOD - 04/01/85 THROUGH 06/30/85

TABLE OF TRT_PROC BY PROVIDER

TRT_PROC TREATMENT/PROCEDURE CODE PROVIDER

TRT_PROC	FREQUENCY / ROW PCT / COL PCT	NURSE PRACTINR	NURSE MIDWIFE	CLINICAL NURSE	PUB HLTH NURSE	LIC PRAC NURSE	NURSE'S AIDE	TOTAL
Init.Prenatal Screening	1	21 4.17 0.20	2 0.40 2.94	263 52.18 0.40	0 0.00 0.00	218 43.25 0.35	0 0.00 0.00	504
Development Problem List	2	0 0.00 0.00	2 0.36 2.94	522 94.22 0.80	0 0.00 0.00	30 5.42 0.05	0 0.00 0.00	554
Init.Workup Completed	3	14 12.07 0.14	3 2.59 4.41	71 61.21 0.11	0 0.00 0.00	28 24.14 0.04	0 0.00 0.00	116
Update Problem List	4	1 0.05 0.01	1 0.05 1.47	1254 63.21 1.92	0 0.00 0.00	151 7.61 0.24	577 29.08 1.51	1984
Calculated Hi-risk Index	5	1 0.65 0.01	0 0.00 0.00	149 96.13 0.23	0 0.00 0.00	3 1.94 0.00	2 1.29 0.01	155
Hlth.Supv. Rtn.Prenatal	6	238 32.60 2.30	8 1.10 11.76	468 64.11 0.72	0 0.00 0.00	13 1.78 0.02	3 0.41 0.01	730
Instruction Counseling	7	1635 10.91 15.83	13 0.09 19.12	7106 47.42 10.86	0 0.00 0.00	5881 39.24 9.45	351 2.34 0.92	14986
Direction of Treatment	8	653 23.56 6.32	3 0.11 4.41	1125 40.58 1.72	0 0.00 0.00	991 35.75 1.59	0 0.00 0.00	2772
Post-Partum Examination	9	1 33.33 0.01	0 0.00 0.00	2 66.67 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	3
Post-Abortal Examination	10	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	1 100.00 0.00	1
	TOTAL	10331	68	65423	27	62240	38186	176276

(CONTINUED)

THE MEDICARE/MEDICAID AUTOMATED CERTIFICATION SYSTEM: APPLICATIONS TO LONG-TERM CARE

Alan S. Friedlob and Elizabeth S. Cornelius,
Health Care Financing Administration
Office of Research and Demonstrations

Robert Dickerson
Health Care Financing Administration
Bureau of Data Management and Strategy

What is MMACS?

In the early 1970s, the Health Care Financing Administration (HCFA) created the Medicare/Medicaid Automated Certification System (MMACS) to verify the certification status of providers and to track provider deficiencies. MMACS contains information on over 51,000 Medicare and Medicaid participating providers including hospitals, intermediate and skilled nursing home facilities, home health agencies, independent clinical laboratories, and rural health clinics.

The information contained in MMACS derives from the Medicare and Medicaid certification process conducted by a state licensure/certification agency. Section 1864 of the Social Security Act requires the Secretary of the Department of Health and Human Services to enter into agreements with state health agencies under which these agencies determine whether various types of health care facilities meet prescribed regulations to assure the welfare of Medicare patients (i.e., Conditions of Participation). Section 1902 (A)(9)(a) requires the state Medicaid agency to use this same health agency for surveying and certifying providers serving Medicaid eligible patients.

Certification is a recommendation made by the state health agency to the Health Care Financing Administration on the degree of compliance of providers with the Conditions of Participation and standards. Certifying recommendations are based on data gathered through on-site facility surveys.

Transforming MMACS from a management to a research data base

This paper describes the use of the institutional long-term care component of MMACS in health services research and policy analysis. In the context of this paper, it is important to observe that MMACS was not originally conceived as an analytical data base. Its sole purpose was as an automated management tool to monitor the certification status of Medicare and Medicaid providers, to efficiently schedule facility recertification surveys, and to record and track facility deficiencies resulting from these surveys.

In 1983, HCFA and the Assistant Secretary for Planning and Evaluation/Health (ASPE/H) realized that by making relatively simple changes in the manner in which MMACS data is entered into the data base, MMACS could serve as a useful source of data for long-term care services planning and policy research. To accomplish this objective, HCFA and ASPE/H contracted with SysteMetrics, Inc. to review the problems in using MMACS analytically and to create a model research data base using the last complete year of MMACS data, 1981.

Federal long-term care policy analysts and state certification personnel had already recognized many of MMACS' limitations. The principal problem in using MMACS long-term care data analytically rather than administratively was the duplicate counting of long-term care facilities, beds, and staffing levels. Duplicate counting was caused by maintaining separate files for skilled nursing facilities (SNFs) and intermediate care facilities (ICFs) when a single facility had dual certification (i.e., provided both SNF and ICF level of care).

In verifying the MMACS data by comparing the 1981 data base maintained by HCFA with licensing and certification documents maintained by state health agencies, SysteMetrics, Inc. identified the following limitations to its analytic use:

- Duplication due to multiple levels of care in a single facility. For example, if a single long-term care facility has two levels of care (i.e., SNF and ICF), state certification personnel were required to submit duplicate information for the facility (i.e., recognizing it as two separate facilities). Similarly, beds that are dually certified under Medicaid as SNF and ICF (Medicare does not recognize the ICF level of care for reimbursement purposes) were counted twice.

- Clerical errors. A single facility may be entered into MMACS more than once because of errors in the facility name (i.e., different abbreviations or spelling of facility name) or address (i.e., same facility name but different address and/or different zipcode).

- Undercounting of facilities. When the HCFA Regional Office received the certification report from the state health agency, edit checks were performed. If the facility failed the edits, it was placed in an orbit file and did not appear in the MMACS masterfile until corrections were made. This problem has since been corrected by eliminating the orbit file and creating a "transaction" file within the MMACS masterfile to accommodate facilities that failed the Regional Office edit.

- Undercounting of bed supply. Non-certified beds were often not reported on certification forms. In verifying bed counts, SysteMetrics' observed that while the number of certified beds appears accurate, discrepancies were noted in the reporting of non-certified beds. These discrepancies led to undercounting the total number of long-term care beds in certified facilities. This error has important policy implications since the mix of certified and non-certified beds can be related to obtaining an accurate picture of the mix of Medicare, Medicaid, and private pay patients cared for in nursing homes.

By taking these factors into account in designing a systematic process for unduplicating and verifying the 1981 MMACS file, SysteMetrics was able to reduce the 18,421 certified long term care facilities included in the system to 13,391 facilities. This process entailed a labor intensive effort of manually comparing entries in the MMACS file with source documents. Table 1 shows the variables included on the 1981 MMACS research file.

Table 1

1981 MMACS SNF/ICF RESEARCH FILE VARIABLES

1. **Facility Identification**

   Facility Name
   Facility Street Address
   Facility city, state and zip code
   Provider number
   Type of facility (as identified on HCFA Form- 1516)
   (identifies the facility as hospital-based, SNF, or ICF but does not recognize the possibility of a facility being a combined SNF/ICF part of a hospital)
   Type of ownership (voluntary non-profit, proprietary, government, other)
   Type of facility (created variable-- Medicare/Medicaid SNF only, Medicaid SNF only, Medicare/Medicaid SNF/ICF Distinct Part, Medicaid SNF/ICF Distinct Part, Medicare/Medicaid SNF/ICF Dual, Medicaid SNF/ICF Dual, ICF Only)
   Hospital based (created variable)

2. **Beds**

   Number of certified beds (as identified on HCFA Form- 1539)
   Total certified beds (created variable)
   Non-participating beds (created variable)
   Total beds for facility (computer created)

3. **Staffing** (as identified on HCFA Form-1516)

   Number of FTE RNs
   Number of FTE LPNs
   Number of FTE Physical Therapists
   Number of FTE Occupational Therapists
   Number of FTE Speech Pathologists
   Number of FTE Licensed Pharmacists
   Number of FTE Social Workers
   Number of FTE Dieticians
   Identification of 16 services whether provided by staff or arrangement
   Total FTE RNs and LPNs for the entire facility (created variable)
   Ratio of combined RN and LPN staff to number of beds (created variable)
   Staffing index based on federal regulations (a created variable taking into account nurse staffing levels necessary to meet conditions of participation and the availability of rehabilitation services ranging from 1 (very low staffing) to 9 (high staffing).
   Staffing index based on nurse;bed ratios (a created variable taking into account normative licensed nurse staffing:bed ratios and the availability of rehabilitation services. SNFs and ICFs are expected to have different nurse staffing ratios. This variable also ranges from 1-9.

4. **Deficiencies**

   Number of standard-level nurse staffing deficiencies (created variable by assigning each facility a number ranging from 1-4 based on the number of deficiencies cited for five standards related to nursing staff available and rehabilitative services)
   Presence of any rehabilitation deficiency
   Identification of all deficiencies cited on HCFA Form- 1539 between 1971-1981
   Composition of survey team, most recent survey

5. **Medicare claims data**

   Volume of Medicare inpatient SNF claims
   Volume of Medicare inpatient SNF claims with reimbursement
   Volume of Medicare inpatient SNF claims without reimbursement
   Amount of provider payment for inpatient SNF claims
   Volume of Medicare inpatient medical services claims with reimbursement
   Volume of Medicare inpatient medical services claims without reimbursement
   Amount of provider payment for outpatient physical therapy with reimbursement

In addition to developing a reliable data base on nursing homes participating in the Medicare and Medicaid programs, a major reason for the Departmental review of MMACS was to determine the feasibility of examining relationships between nursing home supply (i.e. number of facilities, beds, and staffing) and the extent and type of deficiencies cited by state surveyors. A major component of the survey and certification process is to review a facility's compliance with 431 elements in the Conditions of Participation for SNFs and 258 elements for ICFs.

In conducting its review, SysteMetrics concluded that the data related to facility deficiencies should not be used for many analytical purposes. It reached this conclusion based on a number of considerations involving inter-state variation in

TABLE 2

State Variation in Nursing Homes-- Selected Variables

	Total Certified Facilities	Total Beds in Certified Facilities	Beds in Certified Facilities/ 1000 persons 65+	Beds in Certified Facilities/ 1000 persons 85+	Percent Facilities Medicare Certified	Percent Facilities ICF-only Certified	Beds per Licensed Nurse
REGION I							
Connecticut	231	24783	66.8	638.5	75	11	6.8
Maine	145	9140	63.2	592.9	11	88	8.5
Massachusetts	513	45005	62.2	553.4	21	48	7.5
New Hampshire	74	6740	63.7	642.5	34	66	7.1
Rhode Island	106	8545	67.5	622	56	44	7.3
Vermont	44	2982	50.3	466.7	41	48	7
REGION I TOTAL	1113	97195	63.5	587.9	36	47	7.3
REGION II							
New Jersey	233	32232	37.2	404.7	54	9	8.1
New York	570	94124	44	424.3	91	8	5.8
REGION II TOTAL	803	126356	42.1	419.1	80	8	6.2
REGION III							
Delaware	26	2789	45.9	491	58	38	7.1
Dist. of Col.	6	1166	16.6	144.7	50	50	6.5
Maryland	174	20909	53.5	583.6	53	47	9.2
Pennsylvania	556	68969	44.7	488	52	6	7.6
Virginia	163	20428	40.6	490.8	32	68	7.1
West Virginia	74	5721	24.2	275.1	46	54	7.8
REGION III TOTAL	999	119982	42.8	466.5	54	28	7.7
REGION IV							
Alabama	206	20742	47.5	546.4	89	9	8.4
Florida	306	34705	21.3	270.7	67	2	11
Georgia	301	30649	59.9	715	23	24	9
Kentucky	204	20304	49.8	540.7	47	53	12.6
Mississippi	143	12294	43.3	447.7	9	17	8.2
North Carolina	202	21722	35.8	440.8	67	28	8.2
South Carolina	123	10880	37.8	489	75	25	6.9
Tennessee	229	24540	47.9	552	25	75	9.6
REGION IV TOTAL	1714	175836	37.6	450.8	49	29	9
REGION V							
Illinois	687	90107	71.7	725.5	32	44	12.7
Indiana	424	41604	70.5	731.6	28	69	16.9
Michigan	421	46275	49.4	524.1	65	31	10.1
Minnesota	454	46335	95	811.8	18	31	10.1
Ohio	856	70799	59.8	610.7	43	57	9
Wisconsin	438	53617	92.7	895	16	26	10.4
REGION V TOTAL	3280	348737	69.8	694.3	33	45	11
REGION VI							
Arkansas	207	19574	63.8	698.6	2	60	11
Louisiana	225	24648	63.9	659.4	5	94	11.3
New Mexico	43	3565	30.1	352.4	10	91	9
Oklahoma	363	28330	77.3	794.2	3	97	18.8
Texas	976	100059	74.4	841	4	79	12.9
REGION VI TOTAL	1814	176176	69.8	765.5	4	83	13
REGION VII							
Iowa	427	34118	87.1	721.5	6	94	11.8
Kansas	368	25694	83.6	726.3	7	85	16.9
Missouri	237	26243	40.7	398	22	63	12.2
Nebraska	217	17425	84.1	694.7	7	86	13.8
REGION VII TOTAL	1249	103480	66.7	595.8	9	84	13.2
REGION VIII							
Colorado	173	18936	75.7	726.2	33	17	10.8
Montana	94	6334	72.2	686	70	12	8.2
North Dakota	83	6570	79.3	746	68	31	9.7
South Dakota	114	7880	84.9	705.6	6	49	11.5
Utah	80	5214	46.5	538	32	50	8.9
Wyoming	26	1904	49.2	506.6	4	35	9.7
REGION VIII TOTAL	570	46838	70.5	680.8	37	30	10
REGION IX							
Arizona	25	3217	10.4	147.3	100	0	16.2
California	1184	114468	47.7	487.1	82	3	9.3
Hawaii	34	2516	32.4	414.4	76	24	4.8
Nevada	26	2269	32.6	534.4	89	11	6.6
REGION IX TOTAL	1269	122470	42.9	458.4	82	3	9.1
REGION X							
Alaska	13	644	54.6	825.6	31	23	4.5
Idaho	62	4769	48.8	532.4	71	10	8.3
Oregon	178	14868	48.1	488.7	26	71	11.6
Washington	262	24872	56.7	561	32	15	8.5
REGION X TOTAL	515	45153	52.7	534.4	34	34	9.2
NATIONAL TOTAL	13326	1362223	53.4	558.2	39	43	9.4

the way surveys are scheduled (i.e., unannounced versus pre-scheduled site visits and the frequency in which these visits occur); the composition of survey teams (e.g., a survey team containing a pharmacist may more rigorously scrutinize pharmacy and medication practices than one that has no pharmacists); differences in the content and duration of surveyor training; and the way in which citations are made (e.g., a deficiency of rehabilitation goals and progess not documented on patient records may be cited under rehabilitation, patient records, or both).

In addition to these inter-state variations, SysteMetrics, Inc. observed that different survey teams within the same state reviewing the same facility may use different elements and standards to cite the same deficiencies. For example, were two different survey teams to find the charge nurse on leave and the director of nursing serving as charge nurse, one team might cite this deficiency under the charge nurse standard, the other team under the director of nursing standard.

To the extent that cited deficiencies are indicators of poor structural or process measures of the quality of nursing home care, MMACS' use in longitudinal or inter-state quality of care studies should be particularly cautioned.

Having described MMACS' origin and purpose and the steps taken to expand its management-oriented functions to include research applications, the next part of this paper describes how the data contained in MMACS has been and can be applied to problems in long-term care delivery and policy analysis.

MMACS Applications

The primary value of MMACS to long-term care policy analysts and health services researchers lies in the data base's ability to describe the structural characteristics of nursing home industry participation in Medicare and Medicaid. The relevance of such data to long-term care policy analysis can be illustrated by three examples---analyzing health manpower needs in nursing homes, assessing changes in Medicare skilled nursing facility reimbursement policy, and examining the relationship between structural characteristics of nursing homes and institutionalized patients' quality of life.

An application to health manpower planning

Policy analysts and long-term care planners agree that the supply of nursing home beds and licensed nursing staff is not primarily influenced by the federal Medicare program. Long-term care reimbursement policies of state Medicaid programs and local demand of private pay patients are the dominant factors affecting nursing home supply. The Medicare skilled nursing facility benefit accounts for only 2 percent of nursing homes' revenue while state Medicaid programs account for 50 percent of revenue.

With the creation of an unduplicated research file, MMACS data accurately documents the widespread variation among states in the number of nursing home beds available to the Medicare and Medicaid population, the types of beds available (i.e., skilled or intermediate care facility), and the availability of nursing and rehabilitation personnel to staff these beds. For example, in 1981 beds per 1000 population 85 years of age or older varies from 895 in Wisconsin to 145 in Washington, D.C. with a median of 553 beds; beds per nurse from 18.8 in Oklahoma to 4.5 in Alaska with a median of one nurse for every 9 beds; and the percent of beds certified as ICF-only from 97 percent in Oklahoma to 1 percent in Florida. Table 2 details this inter-state variation.

The inter-state variation in available nursing home resources raises important health manpower planning issues regarding professional staffing standards. The essential difference in classifying facilities as skilled nursing or intermediate care rests upon the amount of licensed nursing care available. Unlike SNFs, ICFs need not provide licensed nursing care on a 24 hour basis.

MMACS can be used to examine the question of what nursing resources are needed to provide an optimal standard of care to the nation's institutionalized elderly. A recent Institute of Medicine survey of state certification personnel found that 68 percent of respondents favored adoption of specific minimum nursing staff to patient ratios in federal regulations (Institute of Medicine, 1985).

For example, in a recent report, the Department of Health and Human Services' Bureau of Health Professions, Division of Nursing concludes that a "lower bound" of nursing personnel requirements for nursing homes for the year 1990 is 10.2 RNs, 10.2 LPNs and 40.5 nursing aides per 100 patients (Moses, 1985). These projections are based on deliberations of an expert panel. The "lower bound" is defined as a standard that the panel believed all states could meet.

As of 1981, MMACS data indicates that 53 percent of facilities would have been able to meet a standard of a licensed nurse for every ten beds but only 13 percent of certified long-term care facilities had bed:registered nurse ratios less than or equal to 10:1. To meet a standard of one licensed nurse for every ten beds would require an additional 26,542 nurses. To meet the Division of Nursing standard of a registered nurse for every 10 beds would require an additional 80,336 registered nurses.

Alternatively, to provide 24 hour licensed nursing coverage, a situation that 63 percent of certified ICFs could not meet in 1981, would require 4,717 additional nurses.

Applying MMACS nurse staffing data in this manner provides projections of the quantity of nursing manpower required to meet a particular normative objective. Our preliminary assessment of the 1984 MMACS file indicates a movement toward increases in licensed nurse staffing since

1981. However, this approach fails to provide insight into the dynamics and determinants of the demand for nursing services in nursing homes. As a national inventory of nursing home resources, MMACS is limited in the amount of information it provides about how the amount of nursing services available varies depending on the health status and disability levels of nursing home patients. Such information is necessary to develop nurse staffing criteria that takes into account the heterogeneous care needs of the Medicare and Medicaid patient populations.

An application to assessing Medicare reimbursement policy

In addition to examining state variation in nursing home supply, policy analysts and health services researchers can use MMACS to compare various types of certified long-term care facilities. The research file classifies SNF and ICF into seven types based upon whether the facility is certified to serve Medicare and/or Medicaid beneficiaries and the level of care provided (i.e. skilled nursing and/or intermediate care). These seven classifications are Medicare/Medicaid SNF only; Medicaid SNF only; Medicare/Medicaid SNF/ICF Distinct Part; Medicare/Medicaid SNF/ICF Dual; Medicaid SNF/ICF Dual; and ICF only). In addition, the file identifies whether the facility is hospital-based or free-standing.

In support of a recently completed Congressionally mandated study to examine the status of the Medicare skilled nursing facility benefit, researchers at the Urban Institute (Sulvetta and Holahan, 1984) merged the 1981 MMACS data with 1980 Medicare cost report data to examine if structural characteristics explain the cost differences between hospital-based versus free-standing Medicare certified skilled nursing facilities. The merging of MMACS with cost report data produced a data base which included 3,492 of the 4900 Medicare certified skilled nursing facilities filing 1980 cost reports.

Sulvetta and Holahan found that the 761 hospital-based Medicare SNFs accounted for 20 percent of Medicare SNF patient days although these facilities comprised only 10 percent of certified beds and 14 percent of facilities. Of the 3,492 nursing homes submitting cost reports, approximately 10 percent provided 40 percent of total Medicare SNF days. Approximately 20 percent of urban and 16 percent of rural free-standing facilities have licensed nurse to bed ratios that are below one nurse per 14 beds. By contrast, 3 percent of the urban and less than 2 percent of the rural hospital-based facilities are below this level. The data also confirm differences in rehabilitation personnel staffing with approximately 35 percent of hospital-based Medicare SNFs providing two or more rehabilitation services compared with 15 percent of free-standing homes.

This Urban Institute study also relied on a second 1981 MMACS file that contains facility level case-mix data for 1584 Medicare SNF facilities (i.e., 1373 free-standing and 211 hospital-based) in 20 states. Table 3 shows the additional variables contained in this file. This data is gathered by state surveyors on patients resident in the facility or skilled nursing unit on the day the survey occurs. Sulvetta and Holahan found that hospital-based SNF patients were receiving intravenous or blood transfusions, special skin care, and bowel/bladder training more frequently than free-standing SNF patients. The percentage of hospital-based patients that were bedfast and/or had indwelling catheters was also greater than among free-standing facility patients.

After controlling for structural characteristics and facility case-mix, the cost differences between hospital-based and free-standing facilities remained large--$26.11 for total costs and $18.51 for routine operating costs. Case-mix and structural characteristics, however, explained only 43 percent of the observed difference in costs between hospital-based and free-standing Medicare SNF facilities. While MMACS data indicates that hospital-based SNFs have more nursing hours, more licensed nurses, and a greater orientation toward rehabilitation than free-standing SNFs suggesting a different case-mix between these facility types, the authors conclude that "More than half the difference remains unaccounted for, either attributable to unmeasured differences in quality or to inefficiency."

Table 3

MMACS Patient Characteristics Variables

Patient Census Day of Survey (Medicare, Medicaid, Other)
Number of Completely Bedfast Patients
Number of Patients Requiring No Assistance with Ambulation
Number of Patients Requiring Assistance with Ambulation (i.e., wheelchair, cane, walker)
Number of Patients Requiring Full Assistance in Eating
Number of Patients Requiring Some Assistance in Eating
Number of Patients with Indwelling Catheters
Number of Incontinent Patients (Bowel and/or Bladder)
Number of Patients with Decubiti
Number of Patients on Individually Written Bowel and Bladder Retraining Programs
Number of Patients Receiving Special Skin Care
Number of Confused or Disoriented Patients
Number of Patients Receiving Intravenous Therapy or Blood Transfusions
Number of Bed-to-Chair Patients

Application to assessing quality of care

MMACS tells us much about facility characteristics (e.g., number of beds, staffing patterns, levels of care available, ownership patterns) but relatively little about the actual patterns of the long-term care stay, the variation in resource requirements related to patients' health and functional status, and the relationship between available resources and patients' quality of care. The work of Morris, Sherwood, and Bernstein (1985) at the Hebrew Rehabilitation Center on Aging in Boston in developing a nursing home patient classification system suggests an application of MMACS data to examining the relationship between structural characteristics of nursing homes and process and outcome measures of how facilities treat their patients.

Morris, et al.'s thesis is that nursing home facilities need to be evaluated in terms of their impact on patients' quality of life, ideally measured in terms of changes in functional status over the duration of a patient's stay. Having classified nursing home patients into groupings that have similar nursing care needs, it may then be possible to develop facility-level parameters regarding patient outcomes along physical, emotional, and mental domains of functioning. Facilities whose patients have clinical outcomes that are better than average can be considered above average facilities; conversely, facilities where the average patient outcomes are below standards for a particular clinical domain can be considered as less adequate facilities.

Morris, et al. applied their classification schema to 23,481 nursing home residents in 107 facilities in 11 states and the District of Columbia. As part of their study, these researchers examined four process measures of how a facility treats its patients-- (1) the percentage of patients in isolation; (2) the percentage of patients that are tube-fed; (3) the percentage of patients on IVs; and (4) the percentage of patients physically restrained. For example, while overall, 0.6 percent of existing nursing home patients in the study were receiving IVs, 3.4 percent of Medicare patients with heavy incontinence were receiving IVs and 1.4 percent of Medicaid patients with heavy incontinence had IVs.

By using MMACS it may be feasible to similarly examine how the structural characteristics of nursing homes may influence process and outcome-related quality of life standards. for example, a 1984 MMACS data file that includes patient characteristics data for 2855 Medicare SNF certified facilities in 30 states indicates that 0.8 percent of patients had IVs. However, the number of patients with IVs varied from 6.1 percent in Kentucky to zero in Nebraska, Mississippi, and Nevada.

MMACS: Future Directions

An unduplicated MMACS research file for all Medicare and Medicaid certified long-term care facilities exists for calendar year 1981 only. As indicated above, 1981 and 1984 files consisting of non-random samples of Medicare SNF facilities including patient characteristics data also exists. HCFA is in the process of unduplicating the 1984 and 1985 MMACS management files to create research files comparable to the 1981 file. In its routine use as an automated management tool, the degree of duplicate entries in MMACS has significantly declined since 1981. Nevertheless, without special efforts to monitor and correct duplicate entries it is unlikely that the MMACS management data base can serve as a reliable longitudinal research data base.

With the creation of the 1984 and 1985 research files, it will be possible to reliably assess changes in the nursing home industry's participation in the Medicare and Medicaid programs that may have occurred as a result of the introduction of the Medicare prospective payment system for hospitals and the implementation of community-based alternatives to institutionalization to serve Medicaid eligible beneficiaries in their homes. Further consideration will need to be given to how MMACS data can aid in designing Medicare and Medicaid prospective payment demonstrations for nursing home care and how patient-centered assessment data collected as part of the certification survey can be used to examine the relationship between statewide and facility-type variation in structural characteristics, patient case-mix, and quality of life standards.

References

Institute of Medicine, National Academy of Sciences, Committee on Nursing Home Regulations, Data from Survey of State Licensure and Certification Agency Directors, 1985.

Moses, E.: The 1984 evaluation and update of the staffing criteria for the criteria-based model. Rockville,MD. Division of Nursing, Health Resources and Services Administration, Public Health Service, April 1985.

Morris,J.N., Sherwood,S., and Bernstein,E.: Quality of Life Standards in Long-Term Care Institutions. Boston,Mass. Hebrew Rehabilitation Center for the Aged, January 1985.

Sulvetta,M.B., and Holahan,J.: Cost and Casemix Differences Between Hospital-Based and Freestanding Nursing Homes. Washington,D.C. The Urban Institute, 1984.

MEDICARE AUTOMATED DATA RETRIEVAL SYSTEM (MADRS)

By Paul Lichtenstein, M.H.S., Malcolm Sneen, B.S., and Richard Yaffe, M.S.
Health Care Financing Administration

Background

The Health Care Financing Administration (HCFA) administers the Medicare and Medicaid programs. To carry out this function, HCFA maintains a large data collection and processing system to support management, daily operations and research efforts. This paper describes a new database being developed called the Medicare Automated Data Retrieval System (MADRS) that will support the research and evaluation functions of the Office of Research and Demonstrations (ORD) at HCFA.

Because of the large volume of Medicare data, approximately 200 million claims records per year, and the lack of complete diagnostic information on those claims, HCFA has in the past, maintained sample files bill and payment record files for research purposes. These sample files are adequate for a majority of research activities, however, ORD conducts numerous demonstration project to test the efficacy of changes in the Medicare program and in these demonstration projects, sample data is not adequate. The reason that sample files are not adequate is that the demonstration projects usually involve a small geographic area or a small number of Medicare benficiaries. A couple of recent examples of such a changes were allowing the State of New Jersey to reimburse hospitals by DRGs and allowing Medicare benficiaries in southern California to utilize the services of Clinical Socialworkers. Because of the small samples size in these demonstration projects, it is necessary to have data on every individual in the geographic region or on every participant in the demonstration.

When it is possible to identify in advance the geographic regions and/or the beneficiaries involved in the demonstration, one can place a prospective tap on data as it comes into HCFA. In most instances, it is not possible to identify in advance which Medicare beneficiaries will participate in the demonstration. Even when it is possible to identify regions or beneficiaries in advance, researchers often require data for a period of time before the demonstration began in order to detect changes caused by the implementation of the demonstration.

When data is required on a retrospective basis, HCFA has had to go back and search through its massive 100% bill and payment record files. Since the data in these files is now only organized by the date the bill or payment record was received at HCFA, it is neccessary to search through the entire file for long periods of time to find these small amounts of data.

The purpose of MADRS is to reorganize the massive 100% bill and payment record file and to index them so that it will be easier and less expensive to retrieve data for research and demonstration purposes.

Content and Organization of the MADRS file

The MADRS files will contain 100% of all Medicare claims data. Specifically, MADRS will have all hospital bills, outpatient bills, skilled nursing facility bills, home health bills, and physician and supplier payment records. Often claims records are separated into part A and part B depending on which Medicare insurance program covers them. MADRS will contain both types of claims. The more important data elements available in these claims are patient health insurance number, provider number, HMO enrollment, reason for entitlement, dates of service, types of services, diagnosis, charges, reimbursed amounts and coinsurance and deductables.

A number of issues had to be resolved in the design of the MADRS file, the most important of which was what was the best way to organize the file. As noted above, the principal purpose of the MADRS file will be to support the data needs of ORD demonstration projects. In examining the entire range of ORD demonstration projects, it was clear that most often data was needed by either geographic region or by Medicare health insurance number (HIC#) and that this data was needed for specific time periods. Due to these needs, it was determined that the best way to organize the file was to sort it first by year (using the date of service of the claim), then within year by geographic region (using state/county codes) and finally within geographic region by HIC#. In order to facilitate the retrieval of data from the files, indices to the files will be created. The indices will give the file locations by state/county code, by HIC# and by Medicare provider number (only for institutional providers not for individual physicians). Using the indices, researchers will be able to know exactly where needed data in the MADRS files are located and be able to retrieve the data without having to search through the entire file.

Because of the need for data for specific time periods, it was decided that the file

should be based on date of service versus date of reciept of the claim. Due to this fact, a decision had to be made as to how long beyond the end of the year to wait for outstanding data. In making this decision, one has to balance the desire to have the files be as complete as possible with the need to have access to data in a timely manner. A review of the percentage of completeness of the data at various time periods after the end of the year revealed that at 3 months data was only 93% complete, at 6 months data was 97% complete and that at one year the data was over 99% complete. The final decision was to create a primary MADRS file for each year with data available at the 6 month cut-off point and to have a supplimentary MADRS file for data that came in between 6 months and 12 months after the end of the year. The data in the supplimentary MADRS file for each year will be processed into the MADRS format. Data that comes in after 12 months will be retained but not processed into the MADRS format.

A number of decisions also had to be made with respect to the beneficiary records. It had been determined that all records for a particular beneficiary should be located together for ease in retrieval of the data. The first issue was did we want to try and have a record in each yearly MADRS file for every Medicare beneficiary even if they did not have any use during the year. Some 25% of all Medicare beneficiaries do not have a use in a particular year. If this were done, one could then either select comparison/control groups and calculate use rates directly from MADRS files. This could be accomplished by getting header records for non-users from the quarterly Health Insurance Master (HIMA) file and then merging them into the MADRS file. The HIMA file is the master beneficiary record for every Medicare beneficiary and it contains their enrollment status and other demographic information. Due to the difficulty and cost of carrying out this merge, it was determined that MADRS will only have records for those Medicare beneficiaries who have a use during the year and that research will have to separately run through the HIMA file to obtain nonuser information.

The second issue raised was what to do with beneficiaries who moved during the year. For these individuals, a convention was adopted that the county recorded in the first use of the year would be the county all their records for the year would be located.

Finally, in the design of MADRS a number of decisions were made regarding the format of the data in the final MADRS files. In October of 1983, HCFA changed the method of reimbursing hospital service from cost-based to the prospective payment system using Diagnostically Related Groups (DRG). This change is outside the scope of this paper but suffice it to say that major changes were made in the amount and type of data collected after October 1983. In order to make pre and post October 1983 data as comparable as possible, a single format for MADRS data is utilized. The formats for the data are given in the appendices of the paper. In these formats, one should note two things in particular. One is that in many instances pre-October 1983 data is not available for many of the data elements. The other thing to note is that prior to October 1983, diagnostic information was only required for a sample of records and that after October 1983 the samples changed.

Basic summary statistics on cost and utilization by county will be generated from the MADRS files. A list of the statistics to be generated is provided in the appendices of this paper. The statistics will be generated for the primary file, the supplimentary file and for both combined.

Creation of MADRS

Complete documentation of the system design and the programs that create MADRS are available from ORD. For the purposes of this paper, a brief description of the major phases in the creation of MADRS is provided and a diagram of these phases is provided in the appendix. MADRS is created from the weekly part A and B claims files. These files are organized by date of receipt of the claims and within the weekly files it is sorted by HIC#. Due to the size of the files to be sorted, the first step is the separation of the files into HIC# ranges. This is accomplished in phase I. Also accomplished in phase I is the assignment of records into the appropriate yearly files. In phase II, the HIC# range are processed into person records and a county designation is assigned. Finally, in phase III, the person records are sorted in state/county files. During this step, the indices and summary statistics are created.

Using the MADRS Files

Data will be retrieved from the MADRS files by using the indices and programs designed to search the files. Researchers can specify a geographic region or a list of either Medicare beneficiary number or Medicare provider numbers. Because the geographic region indices specifies a list of 3000 state/county codes, it will be able to be manually searched. The HIC# and Medicare provider number indices will be larger and thus maintained as online data sets which can be computer searched. The search of the indices will in turn produce a list of tape addresses where the needed data is located. The tape addresses along with the orginal researcher lists are then input into programs that will automatically search the file and retrieve the data. When data is

requested for Medicare providers, MADRS can either retrieve just the records for the Medicare provider or it can be set to retrieve all data for any Medicare beneficiary who was seen by that provider.

In a typical ORD demonstration project, a researcher might request all data from seven counties or on 500 Medicare patients known to have a particular disease to study the cost of their treatment. Another important use of the MADRS file will be to allow linkage of Medicare data will other database such as the National Ambulatory Health Care Survey, National Mortality Survey, etc. Using MADRS with the HIMA file, researcher can select control or comparison groups for studies and develop rates of use and cost by county.

Availability of the MADRS Files

MADRS data will be created for each year begining with 1980. Each year in June, a new MADRS file will be created for the previous year. In December of each year, the supplimentary MADRS file for each year will be created.

The system and programs to create MADRS are in the final stages of development and HCFA expects to produce the first MADRS file for year 1980 by December 1985. Subsequent years of data will be produced as soon as possible thereafter until the MADRS file are up to date. It is expected that MADRS will be up to date by June of 1986.

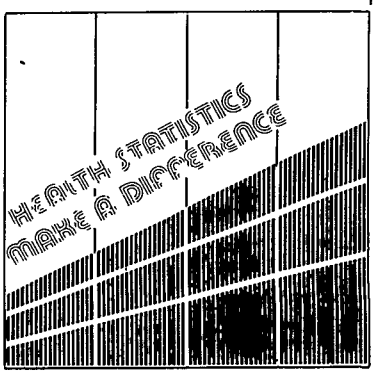
Because of the confidential nature of the data in MADRS, access to MADRS data will necessarily be controlled. Although the process for requesting MADRS data has not yet been established, it is likely to be analogous to applying to HCFA for a grant. Researchers will be expected to describe their proposed studies, how they plan to utilize the MADRS data and how they will maintain the confidentiality of the data.

Appendices

Due to the length of the appendices, they are not included in conference procedings. Copies may be obtained from the authors at the following address: Office of Research and Demonstrations, Room 2306, Oak Meadows Building, 6325 Security Boulevard, Baltimore, MD 21207.

Session J

Administrative Systems and Planning



DATA SETS AND INVENTORY ACTIVITIES OF THE AMERICAN HOSPITAL ASSOCIATION

Ross Mullner, American Hospital Association

Introduction

The increasing number of changes in, and the complexity of the health care industry have created an urgent need for more detailed, up-to-date, and accurate data about its activities on the part of health care researchers, government agencies, and the hospital industry itself. To meet these needs, existing organizations have expanded their data-collection activities, while new ones have sprung up to create their own data bases. As a result, the suppliers of these data -- primarily hospitals -- have been inundated with requests for data, many of them redundant and many of them costly to comply with. At the same time, much of the wealth of data about hospitals and other health care providers has not been used as extensively or intensively as might be possible.

The American Hospital Association (AHA) is doubly concerned with these problems. It is not only the national representative and spokesman of the nation's hospitals, more than 98 percent of which hold membership in it, but is also the principal collector and source of national hospital data.

The objective of this paper is to describe the Association's four major data collection mechanisms. They include: 1) the Annual Survey of Hospitals, 2) the National Hospital Panel Survey, 3) various Special Surveys, and 4) the Inventory of U.S. Health Care Data Bases.

The Annual Survey of Hospitals

The Annual Survey of Hospitals is the principal data collection mechanism of the AHA. Conducted by the Association since 1946, the main purpose of the survey is to provide a cross-sectional view of the hospital industry each year and to make it possible to monitor hospital performance over time. The data that it gathers from its universe of over 7,000 hospitals concern primarily the availability of services, utilization, personnel, finances, and governance.

The most recent Annual Survey (1984) gathers data on eight major areas:

1. Reporting Period. In this section, the beginning and ending dates of the reporting period are requested, as well as information about the hospital's current fiscal year. Although respondents are asked to provide data for a 12-month period beginning October 1 and ending September 30 of the following year, they have the option of reporting data for any consecutive 12-month period.

2. Classification. Includes questions about governance and the principal medical service provided.

3. Facilities and Services. Includes questions about services available.

4. Beds and Utilization by Inpatient Service. Includes questions about beds set up and staffed within distinct inpatient service areas of the hospital and about the utilization of these units in terms of discharges and patient days.

5. Total Facility Beds and Utilization. Includes questions about the total number of beds set up and staffed, admissions, discharges, patient days, discharge days, outpatient utilization, and surgical operations for the entire reporting period.

6. Financial Data. Includes questions about total patient and non-patient revenue, payroll and nonpayroll expenses, restricted and unrestricted assets and liabilities.

7. Hospital Personnel. Includes questions about full and part-time staff divided into occupational categories.

8. Hospital Medical Staff. Includes questions about the number of practitioners on the active and associate medical staff for various specialty groups.

In October of each year, the survey is mailed to all hospitals in the U.S. and its territories. The mailing universe consists of both AHA registered hospitals and nonregistered institutions. Registered hospitals comprise approximately 98 percent of the mailing universe. Identification of nonregistered hospitals is provided by sources such as the National Center for Health Statistics and other national organizations such as the Federation of American Hospitals.

The overall response rate averages approximately 90 percent each year. The response varies, however, between groups of hospitals categorized by size, ownership, service, geographical location, and membership status. The response rate of community hospitals, defined as all nonfederal, short-term general and other special hospitals, is generally higher than that of noncommunity hospitals. The response rate of registered hospitals averages approximately 90 percent, while that of nonregistered hospitals averages less than 60 percent. The response rate of hospitals with more than 100 beds averages over 92 percent; that of hospitals with fewer than 100 beds, 82 percent.

When questionnaires are returned partially completed, or are not returned at all and contacts with the hospitals do not yield completed questionnaires, estimates for most missing data items are generated on the basis of their values in the previous year, whether they were actual or estimated, and on the basis of data reported by hospitals similar to the nonrespondents in size, type of control, principal medical service provided, and length of stay(long- or short-term).

Because of the importance of the Annual Survey, information reported is carefully edited.

A major component of editing involves testing the reliability of information reported in the current survey against data reported by the same hospital in previous years. Unusual changes from one year to the next may indicate data problems. Additional tests include comparing data from a responding hospital with average values for data reported by similar hospitals and testing each response for consistency and agreement with the other information reported on the questionnaire. Once all additions and corrections to these data are completed, aggregate totals for geographical areas, hospital types, and hospital size are compiled for each item. The aggregates are then compared with those obtained in the past. If the changes in aggregate levels are inconsistent with historical trends, the individual case data are re-evaluated until either the findings are confirmed or a specific problem is identified.

The individual hospital is contacted for clarification and confirmation of specific responses that fail the editing tests. As a result of this contact, the data are modified if necessary. On the average, 3,000 hospitals have been contacted each year for resolution of problems.

Data from the Annual Survey are compiled each year in two publications: the Guide to the Health Care Field and Hospital Statistics. The Guide lists all AHA registered hospitals and presents general descriptive information about each, such as location, governance, primary service provided, available facilities and services, total beds, and total utilization, expense, and personnel indicators. Hospital Statistics contains aggregates of most data items for hospitals grouped by location in U.S. Census Divisions and by state. Within these geographical categories, the data are disaggregated for groups of hospitals classified by control and service, length of stay, and size. Other tables show data for selected metropolitan areas and for other special hospital groups.

The complete data set is available through the AHA in magnetic tape format. Information about the revenue, assets, and liabilities of individual hospitals is confidential, however, and is not released. Essentially, two versions of the data tape are available. The nonestimated file contains only reported data, with nonrespondents and missing items recorded as blanks. The estimated file, which represents the entire universe of hospitals, contains both reported and estimated data, the latter identified by a special coding scheme.

The National Hospital Panel Survey

The National Hospital Panel Survey is the only source of monthly data about the finances and utilization of U.S. community hospitals. Conducted by the Association since 1963, the main purpose of the survey is to obtain a limited set of data, each month, from a representative sample of community hospitals. These data are used in longitudinal analysis and monitoring seasonal variations in the utilization, finances, and staffing of all community hospital throughout the country. From the data the Panel Survey collects, estimates of hospital performance indicators are derived for a series of hospital bed-size groups at national and regional levels.

The questionnaire, asking for information from the previous month, is sent to a national sample of 2,000 hospitals. Approximately 1,700 (70 percent) of them respond to the survey in any given month.

The survey's one-page questionnaire is divided into the following sections:

1. Bed and Bassinets. Includes questions about the number of adult and pediatric beds and the number of newborn beds set up and staffed for use.

2. Utilization. Includes questions about the number of admissions and inpatient days for adult and pediatric inpatients, the number of births and newborn days, the number of outpatient visits by type, and the total number of surgical operations.

3. Finances. Includes questions about revenue, expenses, current assets, and current liabilities.

4. Personnel. Includes questions about the number of full- and part-time regularly employed personnel.

5. Utilization: Age 65 and Over. Includes questions about the number of admissions and inpatient days for elderly patients.

From the 39 data items gathered by the questionnaire, estimates are made of more than 100 hospital indicators. These include:

Beds: staffed beds; staffed bassinets.

Utilization: total admissions; 65-and-over admissions; 65-and-under admissions; surgeries; length of stay; 65-and-over length of stay; 65-and-under length of stay; patient census; outpatient visits; occupancy rate.

Finances: total expenses; expenses per adjusted patient day; expenses per adjusted admission; labor expenses; payroll expenses per FTE employee; non-labor expenses; interest expenses; depreciation expenses; supply expenses.

Personnel: FTE personnel; staffing ratio (FTEs per patient day).

Data reported on each hospital's survey are checked by computer edit for internal consistency with data reported on earlier surveys, or for consistency with responses given on the Annual Survey. Hospitals are contacted for clarification and confirmation of responses that fail the edit.

Estimates of the indicators of the universe of community hospitals are made using the hospital bed as the basic unit of computation, since the number of hospital beds is the variable most

highly correlayed with the other indicators. Each month, indicators are estimated separately for 72 strata reflecting eight hospital bed-size groups (6-24, 25-49, 50-99, 100-199, 200-299, 300-399, 400-499, and 500 or more beds) within each of the nine U.S. Census Divisions. These indicators are then summed over the appropriate strata to produce national and regional indicators for the eight bed-size groups.

Data from the Panel Survey are available in a variety of forms. Data pertaining to individual hospitals, however, are considered confidential and are not released. The data from each month's survey are processed within 30 days of the end of the reporting period, and the results are published within 75 days. A regular subcription series of Panel reports is available from the AHA's Hospital Data Center. The reports contain national, regional, and bed-size group estimates for 76 performance indicators, and compare current month, year-to-date, and quarterly figures to figures for the corresponding periods of the previous year.

Data are also available from the Hospital Data Center in magnetic tape format. Two types of tapes are normally available: monthly national and regional estimates for all Panel indicators classified by eight bed-size groups which cover 1976 to the present; and monthly national, regional, and some state estimates for all panel indicators classified by five bed-size groups (6-49, 50-99, 100-199, 200-399, 400 or more beds), which cover 1976 to the present.

Special Surveys

In addition to the Annual and Panel Surveys, approximately ten Special Surveys are conducted each year. Most of these surveys are proposed by units of the AHA, but some are conducted in cooperation with outside organizations. Most surveys are done for a single time only, although some are done repeatedly.

A list of surveys presently being conducted includes:

1. Hospital Supply Survey - 1985. The purpose of this survey is to collect data on hospital expenditures for supplies in order to make national estimates taking into account variations by geographic region and bedsize. The survey questionnaire was mailed to a national sample of 2,165 community hospitals.

2. Survey of Clinical Clerkships/Externships - 1985. The purpose of this survey is to obtain a basic set of data on the nature and extent to which hospitals provide clinical clerkship/externship experiences for students currently enrolled in foreign medical schools. A postcard mailing was first sent to all hospitals in the U.S. to determine which of them provided clinical clerkship/externship experiences. Then a questionnaire was sent to those hospitals that answered in the affirmative, asking about the number and type of foreign medical students accepted.

3. Survey of Hospital Governance - 1985. Sent to all short-term, non-federal hospitals in the U.S., this survey asks for information about: the by-laws and other requirements applying to the governing board; the composition of the board; the executive and other committees; the organizational relationships of the board; and the demographic characteristics, reviews of performance, and compensation of board members.

4. Survey of Medical Care for the Poor and Hospitals' Financial Status - 1985. This survey, conducted in cooperation with the Urban Institute, is the third in the series. The survey was sent to 1,800 short-term, non-federal, nonprofit hospitals. The questionnaire asked for information about the amount of care hospitals provide to low-income and uninsured persons; about the amount of their charity care, bad debts, and financial status; and about how hospitals are coping with changes in payment methods and increased competition.

5. Survey of American Society for Nursing Service Administrators (ASNSA) Members - 1984. Sent to a representative sample of 500 members, the survey questionnaire asks for information about: present and past job positions; the academic degrees received; membership in other professional organizations; and current salary.

6. Capital Finance Survey - 1984. The purpose of this survey, the second in the series, is to obtain current information on hospital capital finances, hospital construction, and hospital capital investing activities. The survey questionnaire was sent to all U.S. community hospitals.

7. Inventory of General Hospital Mental Health Services - 1984. The survey is being conducted in cooperation with the National Institute of Mental Health. It was mailed to all of the approximately 3,500 general hospitals whose responses to the AHA's Annual Survey of Hospitals or to a special screener questionnaire indicated that they provide these services. It collects information on: the kinds of facilities and services into which patients are admitted for diagnosis and treatment of mental disorders or alcoholism, and their relationship to other hospital departments and services; the beds and utilization of separate inpatient psychiatric or alcoholism services; the utilization of separate outpatient services; other types of psychiatric or alcoholism services; and the

operating expenses for the psychiatric services.

As the foregoing list indicates, the Special Surveys conducted by the AHA differ not only in their subject matter, but also in their universe and in the number of hospitals surveyed.

Overall response rates to Special Surveys are almost always high, ranging between 60 and 85 percent. Variations among surveys in response rates are largely due to the differences in the subjects covered by them. It is also generally the case that response rates to any particular survey vary among the various types and categories of hospitals.

Data reported in the surveys submitted by each hospital are checked for internal consistency by computer edit; hospitals are contacted for clarification and confirmation of responses that fail the edit.

Results of special surveys are usually published in a variety of journals. The complete data sets for most Special Surveys are available from the AHA in magnetic tape format. However, information considered to be confidential, e.g., data pertaining to the finances of individual hospitals, is not released. Specially prepared listings and tabulations are also available on request.

Inventory of U.S. Health Care Data Bases

Another important ongoing data activity undertaken by the AHA is the updating and publishing of its Inventory of U.S. Health Care Data Bases. For the last several years the Association has attempted to identify all nonbibliographic, computer-readable data bases containing national health care information that have been collected by public and private sector organizations and agencies throughout the U.S. and that are available to researchers outside the sponsoring organization. The first version of the Inventory, covering the years 1976-1982, was published in the Review of Public Data Use 11 (1983):85-192. The most recent version of the Inventory was published by the Bureau of Health Professions of the Health Resources and Services Administration (Inventory of U.S. Health Care Data Bases, 1976-1983, DHHS Publication NO. (HRSA) HRS-P-OD 84-5).

Currently the Association is attempting to compile another version of the Inventory in which each abstract will contain a standardized bibliographic citation of the data base and will include more information about data collection methodology, geographic and time coverage, subject coverage (Medical Subject Headings (MeSH) will be used as subject descriptors), technical characteristics (e.g. file structure and size), and availability. The Association is planning to publish the abstracts in book form, with the abstracts arranged and indexed so that they can be searched by data base producer, by title, and by Medical Subject Headings.

The compilation of this expanded and updated version of the Inventory will, it is hoped, provide health care researchers with a unique and comprehensive source of information about available U.S. health care data bases; and will promote the use of these data bases in secondary analysis by improving access to them.

THE HEALTH DEMOGRAPHIC PROFILE SYSTEM

Dianne Stiles, University of Maryland Agricultural Experiment Station
Harold F. Goldsmith, National Institute of Mental Health
David J. Jackson, National Institute of Mental Health
James Longest, University of Maryland Agricultural Experiment Station
Peter Hurley, National Center for Health Statistics
Joseph P. Barbano, National Center for Health Statistics
Susan Doenhoefer, National Institute of Mental Health
Wayne Johnson, National Institute of Mental Health

THE 1980 HEALTH DEMOGRAPHIC PROFILE SYSTEM

This paper is an introduction to the 1980 Health Demographic Profile System (HDPS). It presents a description of the context within which HDPS was created, its present form, and some suggested uses. Much of this discussion is extracted from two larger publications: National Institute of Mental Health, Series BN No. 3, Small Area Social Indicators, by Goldsmith, H.F.; Lee, A.S.; and Rosen, B.M. DHHS Pub. No. (ADM)82-1189. Washington, D.C.: Superintendent of Documents, U. S. Government Printing Office, 1982; and National Institute of Mental Health, Series BN No. 4, The Health Demographic Profile System's Inventory of Small Area Social Indictors, by Goldsmith, H.F., Jackson, D.J.; Doenhoefer, S.; Johnson, W.; Tweed, D.L.; Stiles, D.; Barbano, J.P.; and Warheit, G. DHHS Pub. No. (ADM)84-1354. Washington, D.C.: Superintendent of Documents, U. S. Government Printing Office, 1984.

THE DECENNIAL CENSUS

The decennial census, which began in 1790 to produce population counts for Congressional apportionment, was soon recognized as an important resource for information about the American population. The users of and uses for the information increased over time; so did the magnitude and variety of the census data collected, and the number and variety of forms of the data published by the Census Bureau.

This information is made available in the form of maps, graphs, and tables of statistics published in books, microfiche, or computer tapes and summarized for a variety of geographic units (see Bureau of the Census, 1982, 1983). A richly detailed portrait of the nation at a single point in time created by the decennial census data can be seen from the national level through state and county levels to units as small as blocks. Since there have been nineteen censuses since 1790, a picture of a population changing across time can be created as well.

The vastness, variety, and complexity of the decennial census which permits its flexibility to meet the needs of its diverse audience can pose substantial barriers to the occasional user who has limited expertise, time, money or computer and statistical resources.

The first Demographic Profile System--the Mental Health Demographic Profile System (MHDPS)--was developed by demographers and statisticians at the National Institute of Mental Health in the early 1970's for one such group of occasional users--persons responsible for doing legislatively mandated mental health need assessment. It was assumed that carefully selected indicators from the census could be used for indirect need assessment if they were related to mental health service delivery and could be obtained for relevant levels of geography in an easily used form at a reasonable cost.

Health planners, epidemiologists, ecologists, and other scientists found the 1970 Demographic Profile System data to be valuable not only for need assessment and program planning but for basic ecological and epidemiological research as well. The extensive use of the 1970 system by both researchers and administrators resulted in the decision to develop the 1980 HDPS system.

The 1980 Health Demographic Profile System was designed, compared to the 1970 system, to have an increased capacity to meet the specific needs of researchers and administrators. This is because scientists concerned with social ecology and epidemiology need a large, complex, and flexible data base, whereas administrators often need only the minimum number of social indicators necessary to do program planning evaluation and need assessment.

It is the objective of the remainder of this paper to describe the social indicator system we have called the 1980 Health Demographic Profile System.

THE ROLES OF SOCIAL INDICATOR

Social indicators from the decennial census can serve a variety of uses. The HDPS was designed to assist in the process of population assessment by providing an understanding of empirical conditions under which a population lives, particularly insofar as conditions affect the levels of risk or needs of the population and the extent of health and mental health service demand and utilization. In particular, the system was designed to assist users of census data as they seek to accomplish the following tasks:

* to locate and order by relative need subpopulations within small areas that are targeted for special services or with special needs or behaviors and thereby provide a distribution of the relative needs of populations and subpopulations within and among service areas;

* to describe the demographic structure of small areas (including both central tendencies and heterogeneity) so as to provide a basis for inferences about sociocultural contexts of residents and thereby estimate not only the specific disabilities and needs of typical and atypical areas residents but their help seeking and utilization patterns as well.

Three sets of questions or decisions faced the designers of the original 1970 and subsequent 1980 HDPS, the same ones which face each user of census data:

* what is the relevant level of geography?

* what are the meaningful data or indicators?

* what is the best and most useable form for presenting the data?

GEOGRAPHIC UNITS

The availability of identical data for multiple levels of geography is a fundamental characteristic of HDPS, as it is for the census. Four levels of geography from among the many census geographical units were selected: state, county, minor civil division or census county division, and census tract. While administrative or political units such as states and counties are familiar to most people, minor civil divisions or census county divisions and census tracts most likely are not. A minor civil division is a subdivision of a county such as a township or election district. In 21 states which had no designated minor civil divisions the Census Bureau designated equivalent subcounty divisions, for reporting purposes, called census county divisions. A census tract is a subdivision of a metropolitan county. The Bureau of the Census requests that local "census tract committees" within metropolitan areas design their tracts so that they are homogeneous with respect to social and economic characteristics and have populations of approximately 4,000 persons. Besides MCDs (or CCDs) and census tracts, the 1980 decennial census contains data for several additional subcounty geographical units--enumerations districts, blocks and block groups. These were not included in the data base because they usually have less than 1,000 residents and consequently the Bureau often suppresses information to maintain confidentiality of individual households.

While the system contains both small (subcounty) and large (state and county) geographical units, the subcounty units are less well known and the rationale for their use should be clarified. In the recent past, epidemiologists and health service systems researchers have typically focused research and planning efforts on populations associated with large geographic units such as national, regional, or State populations. The Bureau of the Census's Social Indicators III (Klutznick et al, 1980) is an excellent example of the use of social indicators at the national level. The Health Resources Administration's The Area Resource File (U.S. Bureau of Health Professionals, 1980) is a well known planning and research data base with data limited to the Nation, States and counties. Although there are a number of classic exceptions to this rule of research on geographic units at the county level or higher--for example, Faris and Dunham's Mental Disorders in Urban Areas (1939)--the logic for conducting ecological analysis for either epidemiological or planning purposes in principle applies for "small areas" also. Reasons for the failure to pursue social indicator research at the subcounty level are, for the most part, linked to practical problems related to methodological issues that are difficult to resolve, the timely availability of data, and the cost of data processing. While health service and epidemiological researchers have been slow to realize the full importance of small area social indicators, the business community has been enthusiastic in its use of small area social indicators for purposes such as site selection and targeting of advertising (see any recent issue of American Demographics). In addition to these traditional larger units the smaller subcounty units of census tracts and minor civil divisions of Census County Divisions are included in HDPS as well. These smaller units have been included because they permit planners a closer look at the mosaic of places where people live and work. For larger administrative units such as counties, aggregate statistics have been found to hide localized subpopulations with special needs or problems. An additional benefit to the planner is the flexibility of being able to aggregate the smaller census building blocks into unique service delivery or planning areas such as Health Service Areas or Community Mental Health Catchment areas.

Theoretical as well as practical issues influenced the selection of the geographic unit contained in HDPS. From a social ecological perspective it is generally accepted that much of the important behavior related to residence can be understood and accounted for by considering three, and perhaps four, types of residential areas--the household, the neighborhood, and the local residential area, with some social ecologists adding the municipality (city, town, or county) (Greer 1962). The rationale for identifying these types of areas is the probability of the observation of different kinds of behavior and social action. These areas can be defined in

terms of census units. The most important residential area may be the neighborhood. A neighborhood can be viewed as a set of contiguous households and represents a person's (or household's) immediate residential environment. As such, the neighborhood tends to be the site of informal communication, interhousehold visiting patterns, mutual aid, and friendship. Census blocks or enumeration districts are the small areas that approximate the neighborhood. Unfortunately, as previously noted, these may have less than 1,000 residents, so the amount of data published, or available, is severely restricted in order to ensure confidentiality. Consequently, they are not in the HDPS data base.

The next larger area unit, the local residential area,' is made up of a number of neighborhoods. In metropolitan areas, the census tract, a unit with a population of about 4,000 persons, may approximate a local residential area. Outside of a metropolitan area, the MCD or CCD has many of the characteristics of the census tract.

ITEM SELECTION

Having decided the geographical units to be included in the data base, the next step in constructing HDPS consisted of identifying, in the scientific literature, the recurrent demographic indicators of individual (household) and small area characteristics considered essential for ecological and epidemiological research and needs assessment activities.

In selecting the indicators, preference was given to items that had stable statistical definitions and substantive meaning over time. Also, an emphasis was placed upon selecting time-proven indicators, such as median family income or median years of education.

Item Selection for MHDPS (1970)

The item selection process for 1980 HDPS was an extension of the 1970 process, which is described in Redick, Goldsmith, and Unger (1971, pp. 1-2) as follows:

> An important first step in the initial stage of the NIMH project was to decide what data items should be abstracted from the 1970 Census source which would provide the Community Mental Health Centers program with meaningful information about catchment areas [as a whole as well as their subpopulations]... and which might also have relevance for research in other areas. Studies in the areas of sociology, human ecology, and the epidemiology of mental disorders served as guidelines for the selection of the census items or variables. We have selected demographic and ecological dimensions that are useful in differentiating among residential subareas of American cities and that can be measured using available census data. Particular emphasis has been placed on attempting to identify areas with high risk populations, that is, populations with those characteristics which past studies have shown to be associated with high rates of mental illness and/or high incidence of social disorganization or disruption... [Social area dimensions] served as a primary baseline for selection of many of the 1970 census items or variables. The position of social area analysis 'sensu stricto' [was] used as a starting point because it has been widely recognized, empirically investigated, and critically evaluated (Berry and Rees 1969; Abu-Lughod 1969).

Social area analysis was initially based on the theoretical contention that many residence-related behaviors can be understood and accounted for in terms of three types of society-wide population characteristics or dimensions: social rank, life style, or urbanization and ethnicity. Greer, a proponent of this general theoretical position, points out that these dimensions can be used to "order and compare different neighborhoods of the metropolis and also differing cities--or the same city at different points in time" (Greer 1962, p. 31).

An examination of the literature that existed prior to 1970 suggested that the three standard social area dimensions used to differentiate areas should be increased to include separate consideration for family status, family life cycle, residential life style, familialism, residential instability, area homogeneity, and the subcomponents of social rank--economic status, educational status and social status (see Goldsmith and Unger 1970; Redick, Goldsmith, and Unger 1971).

For the 1970 system, a parallel set of the indicators was selected to represent each of the key dimensions for black, white, and Spanish heritage populations. Additional indicators were added to ensure that populations with high risk of mental or physical disability or those targeted for special services could be identified (See Goldsmith et al. 1975).

Item Selection for HDPS (1980)

Three guides were used to develop the expanded small area inventory of the 1980 HDPS. These were to select indicators that indexed the existing relationships between small area social indicators and behavior; that insured, insofar as possible, that the specific purpose or problem that brings users to the HDPS Inventory could be addressed efficiently and effectively and that had clear meaning to the user. Where possible, traditional measures were to be provided. Another key consideration was the expectation that typologies rather than single indicators were needed to characterize small areas with significantly different behavior.

As the 1980 Inventory and its accompanying distributions and tabulations were being selected from the literature, it became more recognized that behavior attributed to or associated with small residential areas often reflects not only the consequences of aggregates of similar households but distinct area (contextual) effects as well (see Goldsmith and Jackson 1981; Gray, Goldsmith, and Dupuy 1982). This meant that the social stratification, social structure, human ecology, and social indicator literature as well as the health and psychiatric epidemiology and community psychology literature, had to be explored for indicators. Further, as the system was developed, its developers recognized the limitations of indicators of average background characteristics generally used in social area analysis. The use of such indicators to categorize areas fails to recognize that atypical residents of areas may respond differently to their local residential environments than do the typical residents indexed by the usual social area indicators. This recognition led to an emphasis on selecting indicators to index the heterogeneity of small area environments (see Rosen, Goldsmith, and Redick 1979). The need to provide meaningful indicators for rural and nonmetropolitan areas as well as urban and metropolitan areas was also recognized.

An initial set of indicators available from the 1980 census was accordingly derived. Simultaneously, a panel of administrators and scientists was recruited and instructed to select what they considered the 'best' set of indicators for health and mental health need assessment. The indicators chosen by the panel and the indicators derived from the literature review were merged into a single comprehensive set of items. This set was then submitted to several panels for evaluation. These panels had expertise in the fields of psychiatric epidemiology, applied demography, social area analysis, and health and mental health service delivery. The final product of these deliberations is a set of 175 indicators usually reported as percentages, medians, and quartiles or tabulations taken directly from the census.

THE STRUCTURE OF HDPS

The structure of the original source census data and various perceived uses of HDPS produced the final structure of the HDPS system portrayed in Exhibit 1. As one moves through the various stages of the chart from top to bottom the available dataset becomes smaller, more highly processed, and easier to use.

HDPS is a reorganized abstract of two 1980 Census Summary Tape File (STF) for each of the 50 states and the District of Columbia. STF2 was the source of data for the complete U.S. population (such as age, sex, race, and marital status) and STF4 was the source of data collected on a sample basis (such as income, education, and occupational status). This split between complete count and sample data was preserved in HDPS and is reflected in the parallel structure of the left and right sides of the exhibit. A copy of these 1980 census tapes is stored in the Parklawn Computer Center's Tape Library in Rockville, Maryland; in the DCRT of the National Institutes of Health in Bethesda, Maryland; and in the Research Triangle Computer Center of the National Center for Health Statistics in Raleigh, North Carolina; however, they are not considered an active part of HDPS80.

The Master Summary Tape Files are the most complete form of the HDPS containing all of the items, distributions, and census tabulations. Data are available for total, white, black, American Indian, Asian, other (residual group) and Spanish origin populations. All of the geographic units (state, counties, MCD's/CCD's and tracts) within the state are listed sequentially by type, making processing potentially expensive, and the data documentation of the file is complex, extensive and difficult to use. Two Master Summary Tape Files, complete count and sample for each state, are stored, like the census tapes, at the Parklawn Computing Center. For potential users who do not have access to the Parklawn Computer System, in the near future access to the Master Summary Tapes will be possible through the National Technical Information Service (NTIS) of the Department of Commerce.

The most flexible and accessible parts of HDPS are the Statistical Analysis System (SAS) libraries created from the Master Summary Files. A reduced number of variables for a reduced number of race/ethnic groups (total, white, black, Spanish origin) are stored in four libraries. As before, the libraries are separated into complete count and sample data, but additionally, the indicators and their denominators have been separated from the distributions. Each library contains four parallel datasets, one for each of the levels of geography: state, county, MCD/CCD, and tract.

SAS programs have been developed to extract selected indicators from the SAS libraries and create 3 reports: a general profile, a high risk profile, and population pyramids. The profile reports are available in 2 formats: one to facilitate practical comparisons within a selected geographic area and one to facilitate comparisons across geographic units. Like the Master Summary Tape Files, the HDPS SAS Library tapes are stored at the Parklawn Computer Center. These libraries and accompanying report writing programs are also being made available to each State Mental Health Authority. Detailed descriptions of how to use the SAS Libraries and report writing programs can be found in the Health Demographic Profile System: 1980 SAS Documentation (Stiles et al, 1985).

Work has not begun yet on the INFORMS aggregation database proposed as a means of permitting aggregations of the smaller geographic units into larger planning units.

CONCLUSION

This report has described the 1980 Health Demographic Profile System. Specifically it provides information about the types of social indicators in the system, the relative accessibility of these indicators to potential system users, the rationale for the inclusion indicators, and the levels of geography for which HDPS data are available. The system is an inexpensive way for the researcher and/or planner to access information from the decennial census of population and housing. It was designed to meet the needs of a wide range of users who have different levels of sophistication. HDPS has evolved into a general purpose data base information system containing census data for all tracts, minor civil divisions (or census county divisions), counties, States, and special areas in the Nation that can be used for both applied and basic research in social ecology and social epidemiology.

REFERENCES

"Census of Population and Housing, 1980: Summary Tape File 2 Technical Documentation." Prepared by the Data User Services Division, Bureau of the Census. Washington, D.C.: The Bureau, 1982.

"Census of Population and Housing, 1980: Summary Tape File 4 Technical Documentation." Prepared by the Data User Services Division, Bureau of the Census. Washington, D.C.: The Bureau, 1983.

Faris, R.E.L., and Dunham, H.W. Mental Disorders in Urban Areas. Chicago: University of Chicago Press, 1939.

Goldsmith, H.F.; Jackson, D.J.; Doenhoefer, S.; Johnson, W.; Tweed, D.L.; Stiles, D.; Barbano, J.P.; and Warheit, G. National Institute of Mental Health. Series BN No. 4, The Health Demographic Profile System's Inventory of Small Area Social Indicators. DHHS Pub. No. (ADM)84-1354. Washington, D.C.: Supt. of Docs., U.S. Govt. Print Off., 1984.

Goldsmith, H.F., and Jackson, D.J. "Community Dissatisfaction and Depressed Mood in Suburbia." Evaluation and Program Planning 4:95-105, 1981.

Goldsmith, H.F.; Lee, A.S.; and Rosen, B.M. National Institute of Mental Health, Series BN No. 3, Small Area Social Indicators. DHEW Pub. No. (ADM) 82-1189. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1982.

Goldsmith, H.F., Rosen, B.M., Shambaugh, J.P., Stockwell, E.G., and Windle, C.D. National Institute of Mental Health, MHDPS Working Paper No. 24, "Demographic Norms for Metropolitan, Nonmetropolitan, and Rural Counties." Adelphi, Md.: Mental Health Study Center, July 1975.

Goldsmith, H.F., and Unger, E.L. National Institute of Mental Health, Laboratory Paper No. 35, "Differentiation of Urban Subareas: A Re-examination of Social Area Dimensions." Adelphi, Md.: Mental Health Study Center, November 1970. 49 pp.

Gray, L.C.; Goldsmith, H.F.; Livieratos, B.B.; and Dupuy, H.J. Individual and contextual social status contributions to psychological well-being. Sociology and Social Research 68(Oct.):78-95, 1983.

Greer, S.A. The Emerging City: Myth and Reality. New York: Free Press of Glencoe, 1962.

Klutznick, P.M., Slater, C.M., and Barabba, V.P. Social Indicators III: Selected Data on Social Conditions and Trends in the United States. A publication of the Federal Statistical System, Washington,D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1980.

National Academy of Sciences, Toward an Understanding of Metropolitan America. Report of the Social Science Panel. San Francisco: Canfield Press, 1974.

Redick, R.W.; Goldsmith, H.F.; and Unger, E. National Institute of Mental Health, Series C No. 3, 1970 Census Data Used to Indicate Areas With Different Potentials for Mental Health and Related Problems. DHEW Pub. No. (HSM)73-9058. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1971.

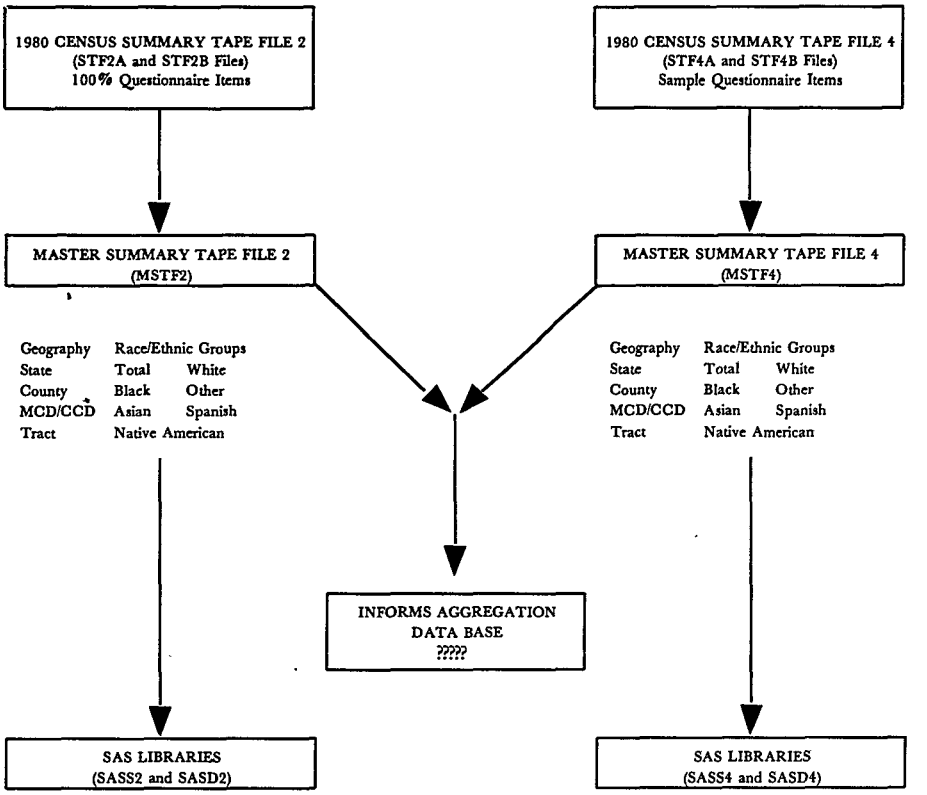
Rosen, B.M.; Goldsmith, H.F.; and Redick, R.W. Demographic and social indicators from the U.S. Census of Population and Housing: Uses for mental health planning in small areas. World Health Statistics Quarterly 32(1), 1979.

Stiles, D.; Jackson, D.J.; Goldsmith, H.F.; and Longest, J.W. "Health Demographic Profile System: 1980 SAS Documentation." College Park, Md.: University of Maryland Experiment Station. Report prepared in cooperation with the U.S. Department of Agriculture, July 1985.

U.S. Bureau of Health Professions. The Area Resoursce File. Washington, D.C.: Sup. of Docs., U.S. Govt. Print. Off., Nov. 1980.

Figure 1

Components of the Health Demographic Profile System[1]



(SASS2: Selected Statistics and Denominators)
(SASD2: Selected Distributions)
("Race" Groups: Total, White, Black, Spanish)

SASS2.STAT04	SASD2.DIST04
SASS2.STAT11	SASD2.DIST11
SASS2.STAT12	SASD2.DIST12
SASS2.STAT32	SASD2.DIST32

State Files
County Files
MCD/CCD Files
Census Tract Files

(SASS4: Selected Statistics and Denominators)
(SASD4: Selected Distributions)
("Race" Groups: Total, White, Black, Spanish)

SASS4.STAT04	SASD4.DIST04
SASS4.STAT11	SASD4.DIST11
SASS4.STAT12	SASD4.DIST12
SASS4.STAT32	SASD4.DIST32

[1] Developed by the Division of Biometry and Epidemiology, National Institute of Mental Health, in cooperation with the National Center for Health Statistics and the U.S. Department of Agriculture.

QUALITY OF CARE IN CALIFORNIA LONG-TERM CARE FACILITIES

William E. Wright, California Department of Health Services

This paper focuses on the utility of several existing administrative data systems for developing surrogate measures of quality of care in skilled nursing homes. The study orginated from an NCHS grant to develop a Cooperative Health Statistics System in California. The opinions and conclusions expressed in this paper are the author's and do not necessarily represent the policies of the California Department of Health Services.

The issues of cost, accessibility, and "quality of care" in nursing homes have prompted many governmental inquiries. In 1980, in response to a mandate from the state legislature, the California Health Facilities Commission (CHFC), the state agency that collects and disseminates fiscal information on licensed hospitals and skilled nursing facilities, created a Long-Term Care Standards Task Force. This Task Force was charged with developing "effectiveness" standards for nursing homes in California. After several months of investigation, the Task Force found that direct measures of quality of care which could provide a basis for developing standards or for monitoring changes over time were not available.

In retrospect, it should not be surprising that the Task Force found so little empirical information about nursing home care. In his review, Rango (1) reported that most of our information on quality of care comes from public testimony at legislative hearings and from anecdotal reports in the mass media. There are few empirically sound studies based on representative samples of nursing homes. Furthermore, quality of care has many subjective aspects and there is no consensus in the literature on how to measure it. Ultimately it must focus on some set of outcome measures such as degree of rehabilitation, length of life, or physical and emotional comfort provided. These outcome measures require observation of patients which can then be aggregated for each nursing home to measure the overall degree of "quality care" that the nursing home delivers to its patients. Unfortunately, patient specific information is extremely expensive to collect; therefore it is seldom reported.

Data Sources

Administrative data sets on long-term care facilities have existed in California for several years. Facility expenditure reports, which also contain employee hours and staff turnover, have been reported to the CHFC since 1977. The Office of Statewide Health Planning and Development (OSHPD) has collected staffing and utilization statistics since 1975. Information on compliance with Medicare/Medicaid regulations has been collected since the late 1960's by the Department of Health Services' (DHS) Licensing and Certification Division. However, the inspection survey reports have not been stored in a central location and they are not computerized; they exist only as paper files in DHS field offices. The DHS also issues citations when serious violations of regulations are found either as a result of the annual inspection or as a result of an inspection following a complaint. Citations are issued for violations which present an imminent danger to patients or guests (Class A) or which have a direct or immediate relationship to the health, safety, orsecurity of patients (Class B). Citation information exists as a manually operated paper file in the DHS central office.

In December, 1981 the DHS in collaboration with the CHFC and the OSHPD began to study the feasibility of combining existing data on long-term care facilities from the three departments into one data-base. The purpose was to examine the relationships among facility characteristics, patient characteristics, and quality of care. A specific objective was to collect facility survey reports from DHS field offices for all facilities in the state for 1980 which was the last year in which there was a common survey form for all facilities. To our knowledge, these reports had not been previously collected in a central location or aggregated statewide to study quality of care. Also, data on long-term care facilities from all three departments had never been merged into a single data-base. The Center for Health Statistics in the DHS was directed to implement this study and to develop a computerized long- term care facility data-base.

Sample

In 1980, there were approximately 1,200 facilities licensed by the DHS to provide long-term health care. This number excludes hospital-based beds that are devoted to LTC patients and non-health care facilities such as residential group homes, bed and board homes, and other "oversight" facilities. We then excluded the five facilities that were owned by state and local governments. The sample was further reduced to facilities that provided over 50% skilled-nursing (SNF) care. Facilities that provided intermediate nursing or primarily care for the mentally disordered or developmentally disabled were also excluded. This left a sample of 1,058 facilities that provided SNF care during calendar year 1980.

Method

Data files from the three departments were collected and merged into a single file and various measures from each of the data sources were selected for further analysis. The primary focus of this paper is on the inspection survey data. The 1980 Skilled Nursing Survey Report (HCFA-1569) contained over 300 elements that were evaluated during the inspection for compliance with federal regulations. These elements focused primarily on physical and process aspects of facility operations such as record-keeping, existence of oversight committees, and qualitifications of staff. The survey also contained a patient census of 13 items about patient conditions and dependency as the number of incontinent patients and the number of bedfast patients.

The initial set of inspection deficiencies was reduced to 131 items that were grouped into 12 categories based on similarity of content. These 12 categories were correlated with other measures in the file and then were reduced to four categories: nursing services (11 items), patient activities (6 items), physical environment (15 items) and infection control and housekeeping (10 items). The items were chosen for the following reasons: (1) the content was related more to patient care than to facility operations; (2) frequency of occurrence, i.e., in general, if less than five percent of the facilities were deficient on an item it was excluded; and (3) among the twelve categories of variables, the four chosen had the highest correlations with the other study variables.

The results which follow are primarily descriptive and the analyses were not guided by theories of institutional or economic behavior. The primary intent is to describe the inspection deficiency variable and its relationship with other facility characteristics.

Results

Table 1 describes the variables used in this study and presents several measures of central value. The average facility had 90 beds and 95% occupancy, was 70% funded by Medi-Cal (California's Medicaid program), and 87% of the facilities were proprietary. The small percentage of non-profit facilities may make policy arguments over restricting the proprietary facilities' share of the market moot. In 1980 California skilled nursing facilities provided an average of 2.6 nursing hours per patient day on an average expenditure of $34.38 and had 130% staff turnover. In an effort to include some rough approximation of "patient mix", we included a few patient descriptor variables. Aggregated across facility, the average patient age was 80 years and the average length of stay for patients in residence on the day of the census was 18 months. (Note: Length of stay based on a sample of discharges is considerably less than length of stay based on a census of patients (2)). Nineteen percent of the patients required full assistance in eating and nearly six percent had decubitus ulcers. The patient turnover for the year was 126% which, interestingly, was less than the staff turnover. During the course of the licensing inspection the average facility was deficient on six of the 42 inspection items. Some of the deficiencies were serious enough so that 40% of the facilities were cited for one or more violations and almost 25% were cited for two or more violations.

Table 2 presents variable means grouped by facility ownership type and by geographic area. In general, non-profit facilities were smaller and had lower Medi-Cal utilization rates which meant their income was higher than proprietary facilities. They also cared for patients who were considerably older than patients in proprietary facilities and the average length of stay was longer. The non-profit facilities provided more nursing hours and expended more dollars per patient day. They had somewhat lower staff turnover rates and received fewer inspection deficiencies and citations. By geographical area, facilities in metropolitan areas (Los Angeles and San Francisco) were larger and received more deficiencies than facilities in other areas. Facilities in metropolitan and rural areas had higher Medi-Cal utilization rates and rural facilities had lower patient and staff turnover and they spent less per patient day than facilities in other areas.

The intercorrelations among all of the study variables are listed in Table 3. Most of the correlations are small which implies that if the measures are reliable then they are describing relatively independent facility attributes. The largest correlations were between expenditures and nursing hours (.62), occupancy rate (-.52), and Medi-Cal utilization (-.48). The variable inspection deficiencies had small but statistically significant correlations with all of the other variables in this set.

The next step was to examine the relationship between inspection deficiencies and all of the other variables through multiple regression analyses. For these analyses, geographical area and ownership type were included as dummy variables. These analyses were based on an hierarchical decomposition method. The variables were introduced in a specific order and any shared variance among variables was assigned to the first introduced variable. In this way the relative contribution of each variable to the variance in inspection deficiencies could be judged independent of the contribution of all variables preceding it in the hierarchy and inclusive of all variables following it.

The results of the regression analyses are shown in Table 4. Size was the first variable in the hierarchy and it explained a little more than two percent of the variance in inspection deficiencies (larger facilities received more deficiencies). Controlling for size, the number of on-site health services (more services, fewer deficiencies) and the occupancy rate (higher occupancy, fewer deficiencies) contributed an additional one percent each to explaining or predicting the variance in inspection deficiencies. Medi-Cal utilization was the largest explanatory variable in the equation (higher utilization, more deficiencies), accounting for nearly five percent of the variance. Controlling for the preceding four variables, geographical area added nearly two percent to the variance (facilities in metropolitan and urban counties received more deficiency ratings than facilities in suburban and rural counties). Ownership type added another one percent (other, non-profit facilities received fewer deficiencies than the other three types).

The five patient characteristic variables were added next to the hierarchy. Controlling for the preceding variables, length of stay contributed two percent to inspection deficiencies (longer length of stay, fewer deficiencies) and decubiti contributed one percent (more patients with decuitus ulcers, more deficiencies). As an aside, since the average length of stay was 18 months, it is hard to find support in these data for the argument heard in some quarters that decubitus ulcers are not related to care received in the nursing home but are the result of patients' conditions

prior to arrival in the facility. Two other variables, patient age and percent patients requiring full assistance with eating, contributed statistically significant but relatively insignificant explanatory power.

Of the remaining four variables that were included in the regression analysis, only staff turnover added any additional explanation (higher staff turnover, more deficiencies). Overall, the 14 predictor variables in the multiple regression explained 17.7% of the variance in the dependent variable.

In a further effort to explore the interrelations between the predictor variables and inspection deficiencies, a pattern analysis was performed. Each of the eleven statistically significant predictor variables from Table 4 was dichotomized at the median (geographical area was dichotomized as metropolitan or urban county versus others and ownership was dichotomized as proprietary versus others) to create a high and low deficiency prone group for each variable. For example, facilities that had more than 87 beds, the median bed size, were more prone to receiving deficiencies than facilities with less than 87 beds and facilities with more than five on-site health services were less prone to deficiencies than facilities with less than five services. Patterns were constituted of facilities which were in the deficiency prone group for any number of the 11 predictor variables. Table 5 shows the mean score on the inspection deficiency variable for each pattern. The optimal pattern contained facilities that were not in the deficiency prone group on any of the 11 variables, the one departure pattern contained facilities that were in the deficiency prone group on any one of the 11 variables, and so forth. With two exceptions an increase in the number of departures was associated with an increase in the mean number of deficiencies. The exceptions were the optimal pattern which only contained two facilities and cannot be considered a stable pattern and the nine departures pattern which fell between the seven and eight departures patterns. While pattern analysis does not explain a greater percent of the variance in inspection deficiencies ($\bar{R}^2$=.144) than multiple regression, it does give a better sense of the relative independence of the eleven predictor variables.

The regression analysis with inspection deficiencies contained three variables, nursing hours, staff turnover, and expenditures, that have been used by other researchers as surrogate measures of quality of care. In order to make some comparison among our five different measures of quality of care, a separate multiple regression analysis was performed for each using the 11 facility and aggregated patient characteristic variables as predictors. Table 6 presents the results of those analyses. (Note: Table 6 shows the influence of each variable controlling for all other variables. Table 4 showed the influence of each variable controlling only for those varaibles above it in the hierarchy.) The explanatory power of the 11 predictors ranged from 10.7% of the variance in citations to 54.4% of the variance in expenditures and each of the five measures of quality of care had a somewhat different pattern of association with the predictor variables.

Discussion

How good are these results and should inspection deficiency data be pursued further? For the expenditure variable our results compare favorable with previous research. In her review of nursing home cost studies, Bishop (3) reported that other multiple regression studies have explained from 47% to 77% of the variation in average costs. So our predictor variables seem adequate, at least for the expenditure variable.

The low multiple correlation with inspection deficiencies could be due to two factors. First, the inspection deficiency variable may not be reliable. The best way to assess that would be a test-retest measure over a short time period. Unfortunately such measurements are not available from administrative data systems and will require special studies. Second, inspection deficiencies may have a small association with these predictors and our other quality of care variables because it taps a different dimension of quality of care. Based on the content of the items, the inspection deficiencies variable is more closely related to process measures of facility behavior toward patients than the other variables in this study. It certainly seems to reflect qualities that many of us would look for in a nursing home: twenty-four hour nursing services, an active program of rehabilitative care which is performed daily, nursing staff that are aware of patients' nutritional needs, proper drug administration, adequate and meaningful patient activities, a clean and safe physical plant, and adequate procedures for infection control. I believe that, if it is reliable, inspection deficiencies is a better measure of nursing home behavior toward patients than are the other measures of facility characteristics.

There are at least three reasons for pursuing and refining information contained in the HCFA annual inspection reports. First, these data are becoming increasingly available. Another paper being presented at this conference (4) describes efforts to extricate long-term care facility data from the federal government's Medicare/Medicaid Automated Certification System. Second, these data are potentially available in many states over several years both in the past and potentially in the future and could be used to evaluate the impact of policy decisions. Few other process-focused quality of care measures are likely to have this continuity. Finally, inspection deficiency data have an important heuristic value for researchers. We know that existing administrative data systems do not solve many measurement problems or provide data for legitimate concerns such as the impact of fiscal policy on quality of care. An objective for interested researchers is to keep making that point. Using limited administrative data to outline areas of genuine public health concern and bringing it to the attention of policy-makers and an interested public gives us a vehicle for keeping the measurement issue alive. Inspection deficiency data can serve that purpose.

TABLE 1
VARIABLE DESCRIPTIONS

VARIABLE NAME	DESCRIPTION	n	MEDIAN	MEAN	STANDARD DEVIATION
Size	Average number (monthly) of licensed beds.	1058	87	89.8	47.82
On-site Health Services	Number of health services maintained in the facility with either facility or contract personnel.	1058	4.5	6.0	4.49
Occupancy (%)	Total number of patient days ÷ (size * number of days in the year).	1050	96.9	94.8	6.05
Medi-Cal Utilization (%)	Percentage patient days reimbursed by Medi-Cal.	1001	77.0	70.2	21.4
Geographical Area	Four categories based on county population and density.				
Ownership	Four categories based on ownership type.				
Patient Age	Average age of patients in the facility on the census date.	1036	81.0	80.0	5.08
Length of Stay (mos.)	Average length of stay (months) of patients in the facility on the census date.	1045	18.2	18.2	4.40
Eating Assistance (%)	Percentage of patients requiring full assistance in eating on the inspection date.	932	19.0	19.5	8.61
Decubiti (%)	Percentage of patients with decubitus ulcers on the inspection date.	933	5.0	5.8	4.43
Patient Turnover (%)	(Number of discharges during the year ÷ the monthly average number of patients) * 100.	1054	113.0	126.5	74.01
Nursing Hours (per day)	Total hours worked by nurses, aides and orderlies ÷ number of patient days.	1054	2.5	2.6	0.52
Staff Turnover (%)	((Total number of persons employed during the year, including part-time and temporary ÷ average, per pay period, number of employees) * 100) - 100.	1054	129.8	138.6	74.58
Expenditures (per day)	Total facility expenses ÷ number of patient days.	1052	32.52	34.38	7.00
Inspection Deficiencies	Sum of 42 (met, not met) items from the Medicare/Medicaid Skilled Nursing Facility Survey Report (HCFA-1569).	968	5.0	6.1	4.03
Citations	Number of serious violations of state and federal regulations.	1058	0	1.4	3.21

Category	n	%
Metropolitan	390	36.9
Urban	320	30.2
Suburban	217	20.5
Rural	131	12.4

Category	n	%
Proprietary-Chain	528	49.9
Proprietary-Individual	388	36.7
Church Affiliated	62	5.9
Other-Nonprofit	80	7.6

Sources: California Health Facilities Commission, Long-Term Care Disclosure Report, 1980 Fiscal Year. Office of Statewide Health Planning and Development, Skilled Nursing Facility Annual Report, 1980. Department of Health Services, Licensing and Certification Division, Long-Term Care Facility Files, 1980, and Citation Reports, 1980.

TABLE 2
VARIABLE/ MEANS BY FACILITY OWNERSHIP AND GEOGRAPHICAL AREA

VARIABLES	OWNERSHIP				GEOGRAPHICAL AREA			
	P-C (n=528)	P-I (n=388)	C-A (n=62)	ON (n=80)	M (n=390)	U (n=320)	S (n=217)	R (n=131)
V1. Size	99.0	84.3	61.8	77.8	94.8	88.8	84.1	87.1
V2. On-site Health Services	5.8	6.3	5.3	5.9	6.7	5.8	5.7	4.8
V3. Occupancy (%)	94.9	95.6	94.7	90.5	93.5	94.9	96.0	96.5
V4. Medi-Cal Utilization (%)	74.0	69.8	47.9	59.6	73.8	64.3	69.1	75.5
V5. Patient Age	79.1	79.8	84.0	84.1	79.2	80.9	80.5	79.6
V6. Length of Stay (mos.)	17.7	18.4	19.5	19.2	17.5	18.3	18.6	19.1
V7. Eating Assistance (%)	19.0	19.8	22.3	19.0	17.8	20.0	20.3	21.4
V8. Decubiti (%)	6.5	5.3	4.6	4.3	5.9	6.3	5.5	5.0
V9. Patient Turnover (%)	131.3	121.5	115.2	127.2	133.6	122.3	129.8	109.9
V10. Nursing Hours (p.d.)	2.5	2.5	3.1	3.2	2.6	2.6	2.6	2.6
V11. Staff Turnover (%)	151.7	134.0	105.5	99.4	141.4	137.3	144.3	124.0
V12. Expenditures ($p.d.)	32.93	33.32	41.76	43.68	34.73	35.27	33.98	31.82
V13. Inspection Deficiencies	6.7	5.8	4.3	4.3	6.8	5.8	5.2	5.7
V14. Citations	1.9	1.0	0.6	0.7	1.4	1.6	1.0	1.7

Legend: P-C = Proprietary-Chain C-A = Church Affiliated M = Metropolitan S = Suburban
P-I = Proprietary-Individual ON = Other Nonprofit U = Urban R = Rural

TABLE 3
INTERCORRELATIONS AMONG INSTITUTIONAL, PATIENT, AND QUALITY OF CARE VARIABLES

VARIABLES	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V4
V1	1.00													
V2	.10	1.00												
V3	-.10	-.04	1.00											
V4	.12	-.12	.12	1.00										
V5	-.25	.06	-.01	-.41	1.00									
V6	-.10	-.01	.30	.06	.13	1.00								
V7	-.07	-.03	.07	-.17	.15	.14	1.00							
V8	.12	.05	-.13	-.01	-.06	-.26	.14	1.00						
V9	.16	.09	-.41	-.18	-.01	-.46	-.07	.17	1.00					
V10	-.24	.03	-.25	-.38	.30	-.04	.14	-.01	.08	1.00				
V11	.08	-.12	-.06	.18	-.13	-.16	-.06	.16	.09	-.19	1.00			
V12	-.16	.12	-.52	-.48	.29	-.12	.01	.07	.37	.62	-.17	1.00		
V13	.17	-.08	-.08	.23	-.22	-.18	-.16	.17	.07	-.16	.22	-.15	1.00	
V14	.13	-.01	-.11	.13	-.15	-.13	-.07	.21	.02	-.10	.21	-.05	.34	1.00

Note: Sample sizes ranged from 932 to 1,058. With n = 1,000, a correlation coefficient of .062 is statistically significant at $p \leq .05$. Variable labels are in Table 2.

TABLE 4
SUMMARY OF MULTIPLE REGRESSION ANALYSES
ON INSPECTION DEFICIENCIES

VARIABLE (Direction)	R^2	R^2 INCREASE	F
Size (+)	.022	.022	22.82**
On-site Health Services (-)	.032	.010	10.37**
Occupancy (-)	.046	.014	14.52**
Medi-Cal Utilization (+)	.093	.047	48.76**
Geographical Area	.112	.019	19.71**
Ownership	.123	.011	11.41**
Patient Age (-)	.129	.006	6.22*
Length of Stay (-)	.149	.020	20.75**
Eating Assistance (-)	.156	.007	7.26**
Decubiti (+)	.166	.010	10.37**
Patient Turnover (0)	.166	.000	0.00
Nursing Hours (0)	.166	.000	0.00
Staff Turnover (+)	.176	.010	10.37**
Expenditures (0)	.177	.001	1.04

* $p \leq .05$

** $p \leq .01$

TABLE 5
PATTERN ANALYSIS OF INSPECTION DEFICIENCIES

PATTERN	MEAN	SAMPLE SIZE	STANDARD DEVIATION	AGGREGATED PATTERNS	MEAN	SAMPLE SIZE	STANDARD DEVIATION
Optimal	6.0	2	1.41				
1 Departure	3.1	8	2.03				
2 Departures	3.8	37	2.94	0-4	4.1	207	3.07
3 Departures	4.1	65	3.01				
4 Departures	4.3	95	3.26				
5 Departures	5.2	139	3.29				
6 Departures	6.3	145	3.53	5-7	6.0	434	3.78
7 Departures	6.5	150	4.30				
8 Departures	8.3	107	4.16				
9 Departures	7.4	72	4.03	8-11	8.3	236	4.13
10 Departures	9.4	45	4.18				
11 Departures	9.5	12	3.37				

TABLE 6
STANDARDIZED REGRESSION COEFFICIENTS FOR
SURROGATE MEASURES OF QUALITY OF CARE

PREDICATOR VARIABLES	QUALITY OF CARE MEASURES				
	Nursing Hours	Staff Turnover	Expenditures	Inspection Deficiencies	Citations
Size	-.114		-.132	.065	
On-site Health Services		-.123	.083	-.081	
Occupancy	-.151		-.353		-.135
Medi-Cal Utilization	-.264	.175	-.316	.163	.100
Geographical Area:[a]					
Metropolitan	-.174	.113		.109	
Urban	-.144	.125	.114		
Suburban	-.114	.128			
Ownership:[a]					
Proprietary-Individual		-.076			
Church Affiliated	.142		.181		
Other-Nonprofit	.242	-.102	.262		
Patient Age	.068				
Length of Stay		-.116	.056	-.120	
Eating Assistance	.074			-.103	
Decubiti		.105	.075	.110	.176
Patient Turnover		.079	.162		
Summary Statistics:					
R	.531	.366	.738	.407	.327
R^2	.282	.134	.544	.166	.107
F	22.42**	8.88**	68.08**	11.41**	6.84**

Note: Only those coefficients significant at $p \leq .05$ are shown; n = 873 facilities for each regression analysis.

[a] Each variable is a dichotomous measure. The reference group for geographical area is rural and the reference group for ownership is proprietary-chain.

** $p \leq .01$.

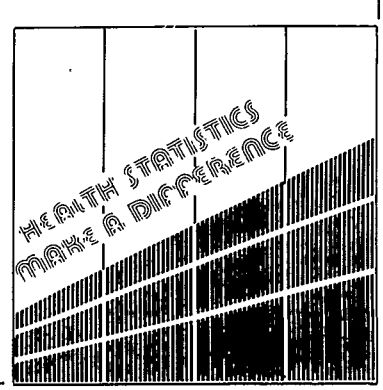
Acknowledgements
The efforts of Jeanne Kochanski in preparing this manuscript are gratefully acknowledged.

References

1. Rango, N. Nursing-home care in the United States: Prevailing conditions and policy implications. The New England Journal of Medicine, 1982, 307(14), 883-889.
2. National Center for Health Statistics, Discharges from Nursing Homes: 1977 National Nursing Home Survey (Vital and Health Statistics: Series 13, No. 54). Hyattsville, MD: U.S. Dept. of Health and Human Services, August, 1981.
Green, M. & Wright, W. E. Data Matters: California Nursing Homes 1980-Inquiries and Issues. Sacramento, CA: California Dept. of Health Services, Center for Health Statistics, May, 1984.
Chung, A. Data Matters: Discharges from Long-Term Care Facilities and Other Related Issues--California, 1983. Sacramento, CA: California Dept. of Health Services, Center for Health Statistics, March 1985.
3. Bishop, C. E. Nursing home cost studies and reimbursement issues. Health Care Financing Review, 1980 1(4), 47-64.
4. Friedlob, A. S., Cornelius, E. S., and Dickerson, R. Long-term care data from the Medicare/Medicaid Automated Certification System: Applications to health services research and policy analysis. Paper presented at the 1985 Public Health Conference on Records and Statistics, Washington, D.C., August 13-15, 1985.

Session K

Perinatal and Infant Mortality



PREVENTABILITY OF FETAL DEATH DURING LABOR: EPIDEMIOLOGIC STUDIES AND ONGOING SURVEILLANCE USING NEW YORK CITY VITAL RECORDS

John L. Kiely, Nigel Paneth, and Mervyn Susser

Columbia University

For the last few years, with the help of the Bureau of Vital Statistics of the New York City Health Department, we have been studying the epidemiology of fetal death during labor. The two major objectives of this research are: First, to discover whether there are ways in which fetal death during labor can be prevented; and second, to develop a method of using the rate of fetal death during labor as an epidemiological index of level of obstetric care. Ultimately we would like to use intrapartum fetal death rates as a surveillance tool for monitoring the quality of obstetric care.

When we began this research we had little understanding of fetal death during labor as a distinct epidemiological entity. This lack of knowledge made it difficult to immediately begin to use fetal death in labor as an outcome in assessments of the effects of obstetric interventions. We therefore first carried out a detailed study of the epidemiologic features of intrapartum fetal death, and compared those features to those of other components of perinatal mortality.

Intrapartum fetal deaths were defined as fetal deaths born at 24 completed weeks of gestational age or later in which it was recorded on the fetal death certificate that the fetus died during labor. Information on death during labor comes from a question on the fetal death certifcates which says: Did fetal death occur before labor? or during labor? This question is answered on 85% of fetal deaths. In 15% of fetal deaths, therefore, we do not know whether they were antepartum or intrapartum.

Antepartum fetal deaths were defined as stillbirths delivered at 24 completed weeks of gestational age or later in which fetal demise was recorded as having occurred before the onset of labor.

Neonatal mortality we divided into two components, deaths in the first four hours and deaths from the fifth hour to the 28th day, as we hypothesized that fetal death during labor would have epidemiologic features similar to neonatal deaths in the first few hours.

All late fetal deaths and neonatal deaths in which the cause of death was recorded as a congenital anomaly were excluded from the four components that I have listed so far, and were analyzed as a separate entity.

This study was based on 320,726 single births (>500 grams) that occurred in New York City in 1976 through 1978. The information on which the analyses were based is routinely collected by the New York City Department of Health. There were 6,092 fetal and neonatal deaths during the three-year-study period. 360--or 6 percent--were recorded as having occurred during labor. Thus the rate of intrapartum fetal death in singletons >500 grams was 1.1 per 1,000 births. Of the entire group of deaths, 33.5 percent were fetal deaths before labor, and 42.7 percent were neonatal deaths. 11.1 percent were attributed to congenital anomalies.

RESULTS

Table 1

Adjusted* Relative Risk of Blacks for Five Components of Perinatal Mortality, with Nonblacks as Reference Group, New York City Single Births 1976-1978

Outcome:	Adjusted* Relative Risks (95% Confidence Limits)
Late Antepartum Fetal Deaths	1.07 (0.97-1.18)
Intrapartum Fetal Deaths	1.11 (0.89-1.38)
Neonatal Deaths in the First Four Hours	1.63 (1.42-1.87)
Neonatal Deaths from the Fifth Hour to the 28th Day	1.31 (1.19-1.44)

*Adjusted for type of service (public vs. private), mother's education, and marital status using maximum likelihood logistic regression.

In Table 1 we show the relationship of race to four components of perinatal mortality. The numbers are relative risks in which blacks have been compared to nonblacks. The relative risks have been adjusted by logistic regression for mother's age, parity, marital status, whether the mother was delivered on a public ward or by a private physician, and prior fetal loss.

The conspicuous result from our analyses of race is that it has almost no association with fetal death. The relative risk for fetal death before labor was 1.07 and the relative risk for fetal death during labor was 1.11. However, black infants had a definitely raised risk of neonatal death as compared to nonblacks. The relative risk of 1.63 indicates that this was especially true in the first four hours of life, but there was also an effect on neonatal deaths from four hours to 28 days.

We have also examined the association of high parity with the four components of perinatal mortality (1). Mothers of high parity have been compared to women with 1 to 3 previous livebirths, controlling for maternal age, prior fetal loss, private vs. public service, marital status, and race. Intrapartum fetal death is unique in this analysis, as it is the only outcome that was strongly affected by high parity. In mothers with 4 previous livebirths, the rate of intrapartum death was 1.66 times the rate in women with 1 to 3 previous livebirths. Furthermore, in women with 5 or more previous livebirths the rate was twice as high as in women with 1 to 3 previous livebirths. Neither antepartum stillbirths, nor neonatal deaths, nor congenital anomaly

deaths were affected by high parity.

We have also explored the relationship of older maternal age to the four components of perinatal mortality (1). We chose 25 to 29 years of age as the reference group for our analyses of the effects of increasing maternal age. Relative risks were computed after adjusting for parity, prior fetal loss, type of service, legitimacy status, and race. Antepartum fetal death stands out as the outcome that is most strongly affected by older maternal age, with a relative risk of 1.45 in 30 to 34-year old women and a relative risk of 2.3 in women over 34.

Older maternal age had no independent association with fetal death during labor--the relative risk for women over 34 was 1.04. Older maternal age did, however, have a positive association with congenital anomaly deaths. There was no excess risk of this outcome in the early 30's, but women aged 35 or over had almost a 50 percent excess risk of congenital anomaly deaths compared to 25 to 29 year-old women. The excess of congenital anomaly deaths in women over 34 was due entirely to three groups of malformations from I.C.D. 8: congenital syndromes affecting multiple systems, anomalies of the heart and circulatory system, and neural tube defects.

We also have analyzed the relationship of prior fetal loss to four components of perinatal mortality. By prior fetal loss, we mean previous spontaneous abortions or late fetal deaths. Prior fetal loss had a strong association with antepartum and intrapartum stillbirths and with neonatal deaths. For 2 or more prior fetal losses, the relative risks ranged from 2.8 to 3.3. However, there was no association between prior fetal loss and congenital anomaly deaths.

TABLE 2

Table 2 shows the association of two biologic factors to the components of perinatal mortality. Of all the variables we analyzed, premature separation of the placenta and prolapsed umbilical cord were the two that were most strongly related to fetal death during labor. Compared to births without the complication, births with premature separation of the placenta had a relative risk of 33.8 for fetal death in labor. Although it is associated with other perinatal outcomes, premature separation of the placenta did not relate nearly as strongly to antepartum fetal deaths or to neonatal deaths. Prolapsed cord related very strongly to fetal death in labor, with a relative risk of 36 for those with the complication compared to those without it. Again, prolapsed cord did relate to other outcomes, but its association with fetal death in labor is most striking.

TABLE 3

In Table 3, we show the relationships of some other complications of pregnancy, labor, and delivery to fetal death in labor. The first column of numbers are relative risks. They represent the rate of fetal death in labor in births that experienced each biologic risk factor relative to the rate in births that did not experience that risk factor. All of the complications listed in the slide have significant associations with fetal death in labor.

When analyzing the effect of any factor on perinatal outcome, it is instructive to consider the possibility that birthweight is acting as an intervening variable. The factor may cause low birthweight, or it may merely be more common among low birthweight births, but much of the raised risk of mortality associated with a certain characteristic may be due to a tendency for births with that characteristic to have lower birthweights. You can see this phenomenon in the second column of numbers, where we have controlled for birthweight and estimated adjusted relative risks.

TABLE 4

In Table 4 we show the relationship of the three levels of maternity services in New York City to fetal death. The relative risks for fetal death during labor in intermediate-level hospitals and community hospitals compared to perinatal intensive care units indicate that the rate of fetal death during labor increases as intensiveness of perinatal care decreases (2).

However, I'd like to call your attention to the second and third columns of relative risks. These provide partial validation of fetal death in labor as an epidemiological entity.

We hypothesized a priori that variation in intrapartum fetal death rates could not be accepted as valid unless antepartum fetal death rates were very similar in the three hospital levels. Indeed, we did find that this was so. As you can see in the second and third columns of relative risks, for both antepartum fetal deaths and fetal deaths in which it was unknown whether the death took place before or during labor, there was a small, nonsignificant gradient of risk, with intensive care units having the lowest rate and community hospitals having the highest.

As a result of these analyses, we concluded that separation of fetal death during labor from other components of perinatal mortality make for greater precision in the analysis of adverse pregnancy outcomes. Also, our findings for the relationship between hospital level and intrapartum fetal death suggest that it is a valid epidemiological indicator of quality of obstetric care. Our goal now is to begin using intrapartum fetal death rates as a device for monitoring individual hospitals and for monitoring changes over time. One of the things we are now in the process of doing is looking at time trends in intrapartum fetal death rates in New York City and attempting to relate these to time trends in rates of cesarean section and other obstetric interventions.

In 1979, the World Health Organization and the International Federation of Gynecology and Obstetrics recommended that more effort be put into collecting good vital data on late fetal deaths, especially data on fetal death during labor (3). We agree with this recommendation and we encourage State Departments of Health to try to collect as complete and reliable data as possible on fetal deaths, since intrapartum fetal death rates have a lot of potential as a means of monitoring the outcome of obstetric interventions in the population.

References

1. Kiely JL, Paneth N, Susser M. An assessment of maternal age and parity in different components of perinatal mortality. American Journal of Epidemiology, (in press)

2. Kiely JL, Paneth N, Susser. Fetal death during labor--An epidemiological indicator of level of obstetric care. American Journal of Obstetrics and Gynecology, (in press)

Table 2

Relationships of Two Biologic Risk Factors to Components of Perinatal Mortality

	Relative Risks			
			Neonatal Deaths	
Risk Factors	Fetal Death Before Labor	Fetal Death During Labor	First 4 Hours	4 Hours–28 Days
Premature Separation of the Placenta	18.0	33.8	12.7	10.7
Prolapsed Cord	8.5	36.0	6.7	5.6

Table 3

Associations Between Selected Biologic Risk Factors and Fetal Death During Labor

	Relative Risk	Birthweight-Adjusted Relative Risk
Abnormal Uterine Bleeding	9.1	2.1
Pre-Eclampsia, Eclampsia, or Hypertensive Disease	2.5	1.5
Breech Position During Labor	12.2	2.9
Duration of Labor 20 Hours	3.0	2.5
Placenta Previa	7.2	1.7 (NS)
Prolonged Rupture of Membranes	2.5	0.7 (NS)

Table 4

Adjusted* Relative Risks for Fetal Death by Hospital Level

(Single Births >1000 grams)

	Fetal Death During Labor	Fetal Death Before Labor	Unknown Time of Death
Perinatal Intensive Care Units (Level 3)	1.00	1.00	1.00
Intermediate Level (Level 2)	1.35**	1.06	1.05
Community Hospitals (Level 1)	1.61***	1.13	1.18

**p 0.10
***p 0.001

*Adjusted for Birthweight, Gestational Age, Prior Fetal Loss, Parity, Type of Service, and Marital Status

CREATING A COMPLETE MATERNAL PREGNANCY HISTORY RECORD USING NORTH CAROLINA DATA

Paul W. C. Johnson, State of North Carolina

The state of North Carolina has had available since 1959, in one form or another, a consolidated birth-infant death file. This data has been useful in various areas of research. I will discuss the problems encountered in developing a somewhat similar system. Instead of linking together a birth and death record dealing with events which are within one year of each other, we want to link together birth and fetal records for the same woman over a period of years.

The linkage makes use of five variables:

1. Mother's name
2. Mother's race
3. Birth order
4. Type of previous delivery
5. Month and year of previous delivery

A snapshot of the process involves three phases:

1. Accommodate multiple births, primarily twins and triplets, by producing one record for each pregnancy,
2. Link together multiple pregnancies for the same woman within the same year,
3. Link these women to previous years developing a maternal or pregnancy record for each woman.

I selected the two years 1975 and 1976 for a pilot study for three reasons. First they are the earliest years having all the information needed for linking. Second, the quality of the data has improved over time, so whatever I find here would serve as a base year and could only get better as I near 1984 in my work. The third reason is the fact that it is easier to conceptualize what is to be done by moving forward rather than backward in time.

For 1975 I have 83,055 birth and fetal records which reduces to 82,250 pregnancies among 82,157 women. Similarily for 1976 the 82,829 births and fetals reduce to 81,888 women. Of this number 8438 (10.3%) indicated their previous pregnancy terminated in 1975.

Using the five linkage variables I was able to computer match 4201 (49.8%) women. These were exact matches on all 5 variables. Then manually, so I could take into account misspellings, keying errors, and poor recall on the mother's part, I located another 522 (6.2%). This gave me a final total of 4723 or 56%.

In the process I found and corrected errors at all 3 phases of linking:

1. 1975 — 1.1% (18 in 1604)
   1976 — 0.2% (4 in 1708)
   0.7% (22 in 3312)
2. 1975 — 4.3% (10 in 235)
   1976 — 4.6% (11 in 238)
   4.4% (21 in 473)
3. 6.2% (522 in 8438)

Since errors in phases one and two should not have happened, the following discussion is limited to errors in the final phase — explicitly errors in trying to link across years. Of the 522 errors in phase three:

447	had one descrepency	447
34	had two	+ 68
41	had three	+ 123
522		638

These descrepencies were of the following types:

416	disagreed on pregnancy history (birth order). In fact, 35 of these were unrecorded miscarriages, i.e. fetal deaths under 20 weeks gestation and not required to be reported
102	disagreed on month previous event occurred
73	disagreed as to the last event being a birth or fetal death
29	disagreed on race of mother
18	disagreed on spelling mother's name
638	

This is on the 522 linkages I made manually, what about the 3715 women who indicated on their 1976 record that the prior event happened in 1975 and I am unable to find? What explanation is there for not finding these prior events? On basis of the 522 events I manually matched, there exists a certain amount of incorrect information being supplied by the mother. I cannot explain why this happens nor can I get an accurate measure of it now. But other sources of non-linkage can be identified and quantified. Of the 3715 women I could not link:

2831	said the 1975 event was a fetal death, so how many of these are miscarriages? If they are, I will have to search earlier years for another event for this woman.
93	reported being non-residents,
33	records came from events reported in another state,
1481	were non-native to N.C. and possibly not in N.C. in 1975.

Taking all this into consideration, the miscarriages and the previous record not being available in my state for one of several reasons, the 4723 linked records between 1975 and 1976 may be all one could expect to get. A system such as this lends itself to many potential areas of research.

Questions that could be addressed are:

1. What is the annual rate for miscarriages?
2. How accurate is the pregnancy history information?
3. What is the impact of in-migration on such a system?
4. Will the fraction of linked records improve when I use 1983 and 1984 data?
5. In deciding when to seek prenatal care and the frequency of visits, what was the influence of the prior event?
6. Is the woman's birthing history predictive of current outcomes:
   a. birthweight?
   b. apgar scores?
   c. number and types of malformations?
   d. infant mortality?

So, in conclusion,

1. Yes, creating a maternal history is feasible using a computerized system. The time and cost will be lower than a manual system in this area, but some manual verification will always be necessary.
2. Yes, there are problems: there may be only 60% or so of the records linkable in two adjacent years, so a 30 year history may be very rare.
3. Yes, we will continue with the pilot study using at least the 1983 and 1984 data.
4. Yes, we are interested in pursuing various areas of research, especially the impact miscarriages and in-migration have on such a system.

THE APGAR SCORE AS A PREDICTOR OF NEONATAL MORTALITY

Winslow J. Bashe, Jr., Wright State University SOM

After its introduction in 1953, the Apgar score was quickly accepted as a quantitative index of newborn status and usually interpreted as a measure of severity of birth asphyxia. As originally described [1-3], the score was derived from values assigned to each of five clinical manifestations-heart rate, respiratory effort, reflex irritability, muscle tone and color observed one minute after delivery. The system was designed so that the lower the combined score, the more profound the problem and the poorer the prognosis, including mortality. In 1964, it was shown that the same score at an assessment made five minutes after delivery predicted higher mortality rates[4-5].

Since that time, the Apgar score has been widely used as a measure of the quality of care and as one of the determinants of neonatal outcome. During this period, major changes have taken place in perinatal care accounting for much of the marked reduction in neonatal mortality which currently is less than half of that in force when the original observations were reported[6]. Despite these changes, only one recent study has reassessed the ability of the five minute Apgar score to predict mortality in a large population of newborns[7].

This report details the differences in mortality rates associated with one and five minute Apgar scores in birth weight groups and race, compares these rates with those originally reported and identifies differences in mortality of those with selected clinical diagnoses in relation to their Apgar score ratings.

Methods and Materials

The data are based on information reported to the Ohio Department of Health on newborns with one or more clinical problems delivered in Ohio maternity units from January 1, 1978 through December 1, 1981. The 127,466 records were compiled from three sources: (1) reports from the maternity units; (2) reports from regional reference neonatal intensive care units (NICU's) on transferred newborns and; (3) deaths unreported by these units but identified in the birth-deaths certificate matched file. Records on transferred newborns were merged with those of the corresponding neonates reported by the maternity units so that final diagnoses were those of the NICU. Hospital of birth scores and birth weights were retained in the merged records. Surveillance, was continued until the child left hospital care so that mortality includes some post neonatal deaths.

The criteria governing reporting were very broad but required information on all newborns with scores of six or less and birth weights of 2500 grams or less. Those exceeding these limits had other problems requiring reporting. Many of these were minor but because those under study excluded 80% without a reportable problem, the mortality rates calculated for larger newborns with scores greater than six are artificially higher. Estimates of these mortality rates, where applied, assumed that all unreported newborns had scores greater than six with the rate based on the actual number of live births in the weight class.

Grouped Apgar scores (0-3) and (4-6) follow the designations applied in the International Classification of Disease Ninth Revision (ICD-9)[8].

OBSERVATIONS

Table 1 shows the mortality rates associated with individual one-minute Apgar scores. At each score, the mortality rate decreased as the birth weight increased. For example, at score 0-1, it varied from 94.8 percent for those less than 750 grams to 13.3 percent at birth weights greater than 4000 grams. At Apgar scores 9-10, it ranged from 36.4 percent at less than 750 grams to 0.2 percent for those greater than 4000 grams. These relationships also existed for those without Apgar scores and for all scores.

A similar pattern was seen for mortality rates associated with Apgar scores within birth weight groups, for those with unstated birth weights and for all weights. Increases in Apgar scores resulted in decreased mortality except for those less than 1500 grams with scores four to six where there was little variance or a higher score resulted in a slight increase in mortality rate. There was a major increase in survivorship of those with higher scores as compared to those in the low ranges. For example, at 1000-1500 gram birth weight, the mortality rate with Apgar score 0-1 was associated with a 45 percent mortality rate versus a 6 percent mortality in those with scores of 9-10.

Table 2 gives the mortality rate associated with individual 5-minute Apgar scores. The pattern was the same as that observed for one-minute scores, except that in almost every Apgar score-birth weight group the mortality at the five-minute score was much higher than that observed for an equal one minute score. For example, in 1000-2500 gram birth weight group, a five-minute score of 0-1 was associated with a 65.3 percent mortality as opposed to a 27.8 percent at a one-minute reading. At the same five-minute Apgar score, 80% of all birth weights died.

For both one and five minute readings, the reductions in mortality associated with increases in Apgar scores and/or birth weight resulted in relatively uniform variances between contiguous birth weight-Apgar score cells. There was a considerable difference in mortality rate between those with a one or less and a three score, but were relatively smaller between those with scores of three and four, four and six, and six and seven.

Because both Apgar scores and birth weight are closely correlated in their ability to predict mortality, the frequencies are heavily skewed to represent these relationships, with relatively few members in the cells that diverge from this association.

The differences in mortality rates associated with one and five minute Apgar scores are more clearly seen in Table 3 with the scores placed in the conventional groups of less than four, four to six and greater than six. In every birth weight group in which the stated score was six or less, the mortality rate associated with the five minute score was higher. For those with scores greater than six, the differences in mortality when present, were small and at higher birth weights about equal. This held true for observations restricted to the study group and on estimates based on total live births. No practical differences were noted in those with scores where not stated.

Table 4 compares the mortality rates associated with five minute scores in white and non-white newborns. At scores of less than four, non-white newborns had lower mortality rates in all birth weight groups-and notably for those 1000-2500 gram birth weight. At scores four or greater and unstated, the non-white superiority held for those 750-2500 grams but not for lower or higher birth weights for observations restricted to the study group or based on estimates. For all birth weights, mortality rates of those with scores of six or less were equal, for scores greater than six, whites had lower mortality rates.

The data on one-minute scores reported by Apgar and James[3] excluded the live births less than 500 grams. In Table 5 the data were adapted to match this birth weight distribution. A significant reduction in mortality was noted in each Apgar score-birth weight group except for those whose scores were less than four and birth weights greater than 2500 grams. For all stated weights, however, the mortality rates were greater in 1978-81 at scores less than six.

In Table 6, the 1978-81 mortality rates associated with five minute scores are compared to those reported by Drage et al[4] in 1964 on live births prior to that date. At scores less than four, the 1978-81 rates were equal or higher in those less than 2501 grams, but higher in those greater than 2500 grams and all stated weights. For scores greater than six and all stated scores, the 1978-81 rates were lower. Reductions, when noted, were much smaller than those based on one minute scores.

Birth-death certificate matched files are limited in defining the role of Apgar scores as predictors of mortality because they contain little diagnostic information on the deaths. Because this file also contained morbidity data, the relationship of Apgar scores to mortality of those with clinical diagnoses was examined. In Table 7, the newborns have been partitioned into groups determined by the lowest one and/or five minute readings of the paired one and five minute scores and the condition specific mortality rates compared in terms of their Apgar score designations. It was not practical to separate these conditions into exclusive groups because many had multiple overlapping diagnoses, especially those in NICU's. In every diagnostic category, those with one or both scores of less than four or no score had much higher mortality rates than could have been predicted if Apgar scores had been ignored. Except for congenital defects, scores of four to six had little discriminative value and scores of greater than six enhanced the probability of survival.

DISCUSSION

Sharp reductions of mortality rates associated with one-minute scores compared to those originally reported were noted in most birth weight groups. Improvements in mortality at five-minute readings were much smaller and restricted to scores of four or greater. As birth weight increased and/or Apgar scores decreased, higher mortality rates prevailed in 1978-81, resulting in higher mean-mortality rates for those with scores of six or less. These anomalous observations led to an examination of potential contributing factors.

Those with lower scores in the original studies had a greater proportion of higher birth weights. When the data were standardized to match Ohio's 1978-81 birth weight distributions, the adjusted mortality rates of the earlier studies were higher for those with low one-minute scores and only slightly lower for those with five minute scores of less than four and higher for those with scores of four to six.

While this accounted for gross effects related to birth weight, it did not explain the higher 1978-81 mortality rates in the Apgar score-birth-weight groups. The differential in mortality rates be-

tween those with scores of one or less and three suggested that the same factor might be operative within the group itself, viz, a greater proportion of those with higher scores and attendant lower mortality. The discrete frequencies of individual scores were not given in the early five minute study but comparisons of the Apgar one-minute data with those of the 1978-81 for birth weights greater than 2500 grams revealed no significant difference in the proportionate distribution of individual scores less than four and a lower mortality rate in the earlier study for each score.

A second potential factor lay in the observation that in the 1978-81 data those with unstated one-minute scores at higher birth weights had mortality rates greater than those associated with Apgar scores of two. This suggested that a large share of these would have been less than four if reported. When all were converted to 0-3 rating, it failed to increase the proportion of those with high birth weights and increased the mortality rate.

Mortality rates associated with low five minute scores were higher than those reported on North Carolina newborns during 1978-80[7], mostly due to higher rates at birth weights greater than 2500 grams. When standardized for birth weight, the adjusted rates were only slightly higher for scores of less than four and equal for scores of 4-6. Other analyses of this data base had indicated that some of the variances in mortality rates may be due to underreporting of incidence[8]. Although the incidence rate of scores less than four at five minutes was only 1.4 per 1000 live births lower than those recorded on 1978 national live birth certificates[9], if translated into frequencies would have lowered the mortality rates below that reported for North Carolina. This suggested that underreporting high birth weight survivors with low Apgar scores may have contributed to higher mortality, but did not account for all of it. Differences in birth weight distribution precluded a direct comparison of North Carolina's low score-high birth weight mortality rate with that of the original report, but that available suggests that it also would have been higher, or at least not lower.

Both original reports were based on somewhat small selected populations. That for one-minute score came from a single hospital and excluded births less than 500 grams. The five-minute score data were derived from reports of newborns under study for the late effects of brain damage contributed by 13 university related hospitals. Neither accounted for race. Although the Ohio data is statewide, it too is selected in that it is mostly morbidity or risk based and contains some postneonatal deaths. Differences in demography, data collection and interpretations of clinical observations needed to apply a score may have affected the occurrence and distribution of scores, birth-weights and their relationship to mortality.

Higher mortality rates associated with low five-minute scores as compared to one-minute scores is accomplished at some expense in sensitivity. A five-minute score of less than four predicted only 35 percent of all deaths, and only four percent were not accompanied by an equal or higher one-minute score. Five minute scores of 4-6 predicted an additional 19 percent of deaths but only 2 percent did not have an equal or lower one minute score. Since low one minute scores predict almost all deaths as an equal five-minute score and many others not discerned by the latter, attempts have been made to use paired scores as a predictor.

Jennet and his associates[10] formulated an Apgar index which is based on both scores, differences and the direction of differences between the paired scores. When these indices were applied to the scores in this data base, there was no consistent relationship between the index and the associated mortality rate. This probably occurred because the index does not account for factors such as birth weight. Up to now we have been unsuccessful in defining the independent effects of these three variables upon mortality.

When paired one and five minute scores were applied to specific diagnostic conditions, a high proportion of deaths and distinctly higher mortality rates were associated with scores of less than four but scores of four to six appeared to have little discriminative value. Restricting the associations to five minute scores greatly decreased the number of deaths of those with scores of less than four and increased the number with scores of four to six. Mortality rates in each Apgar score group were significantly increased. For those with RDS, they rose from 35 to 55 percent at scores of less than four and from 10.9 to almost 23 percent at scores of four to six. This again shows how well Apgar scores discern the potential for mortality and the greater specificity of the five-minute score especially at ratings of four to six.

The effect of immaturity was also noted when mortality rates were calculated using gestational age instead of birth weight as a criterion. We chose birth weight not only because it appeared more reliable but because the literature available for comparisons was based on this factor. Lower female mortality rates as compared to those for males were mirrored by lower rates at higher Apgar scores and lower birth weights.

The primary purpose of this report was to provide the user an updated version of the factors which affect mortality rates predicted by Apgar scores. Stratification of the data restated earlier observations that the one and five minute scores and birth weight (as a surrogate for immaturity) are highly correlated in this role. This, in combination with differences in related demographic characteristics such as race, suggest that all these factors be weighed when employing this type of data for program development assessing the quality of care, or in making decisions in a clinical setting.

REFERENCES

1. Apgar,V. Proposal for a New Method of Evaluation of the Newborn Infant. Anesth. and Analg. 32:250-267. 1953.
2. Apgar,V., Holaday,D.A., James,L.S. and Weisbrot,I.M. Evaluation of the Newborn Infant Second Report. JAMA 168:1985-88. 1958.
3. Apgar,V. and James,L.S. Further Observations of the Newborn Scoring System. Am.J.Dis.Child. 104:419-427. 1962.
4. Drage,J.S., Kennedy,C. and Schwarz, B.K. The Apgar Score as an Index of Neonatal mortality. A Report from the Collaborative Study of Cerebral Palsy. Obstet. and Gynecol. 24:222-230. 1964.
5. Drage,J.S. and Berendes,H. Apgar Scores and Outcome of the Newborn. Pediatr.Clin. North Am. 13:635-643. 1965.
6. Vital Statistics. 1982 Annual Report. Ohio Department of Health, Columbus, OH 1984.
7. Atkinson, D. An Evaluation of Apgar Scores as Predictors of Infant Mortality. N.Carolina Med. J. 44:45-54. 1983.
8. Bashe,W.J.,Jr., Platt,L.J. and Quilty, J.C.,Jr. Regionalization of Neonatal Care in Ohio, II. Regional and Subregional Incidence and Outcomes of Selected Neonatal Problems. Unpublished data.
9. Querec,L.J. Apgar Score in the United States, 1978. Monthly Vital Statistics Report, National Center for Health Statistics 30: No. 1. Supplement 1981
10. Jennet,R.J., Warford,H.S., Kreinick, C. and Waterkotte,C.W. Apgar Index: A Statistical Tool. Am.J.Obstet. & Gynecol. 140:206-212, 1981.

ACKNOWLEDGEMENTS

The data used in this report are a by-product of a cooperative project primarily aimed at assessing the status of the regionalization of neonatal care in Ohio. Data entry, management and analysis for this project were largely supported by the Ohio Department of Health (ODH) Grant #750 to Wright State University, with the collection from hospitals a responsibility of the perinatal unit of the MCH Division of the ODH. Space does not permit individual acknowledgement of all who worked on the project over seven years, but several at the Department of Health require recognition. They are Drs. Antoinnette, P. Eaton and later James F. Quilty, Jr., Chiefs of the MCH Division for their continued support of the program and Mr. James Cox, Chief of the Division of Data Services and Mrs. Mary F. Smith, Head of its Statistical Unit for providing the vital statistics data needed for the analyses.

TABLE 1

Relationship of One Minute Apgar Scores and Mortality Rates of Newborns by Birth Weight Class, Ohio 1978-81

1" Apgar Score	Birth Weight (Grams) <750 No.*	<750 %**	750-999 No.	750-999 %	1000-1500 No.	1000-1500 %	1501-2500 No.	1501-2500 %	2501-4000 No.	2501-4000 %	>4000 No.	>4000 %	Unstated No.	Unstated %	All Weights No.	All Weights %
0-1	752	94.8	366	69.1	413	45.0	619	27.8	1092	14.7	173	13.3	89	41.6	3504	44.4
2	414	90.6	273	58.4	399	36.6	705	16.7	1626	7.8	250	5.6	67	32.9	3734	25.6
3	201	84.6	212	52.4	382	25.9	875	11.0	2130	3.7	353	0.9	50	32.0	4203	13.7
4	114	67.5	167	44.3	443	15.8	1199	6.6	3159	2.4	538	1.1	37	10.8	5657	6.7
5	56	71.4	150	38.6	519	16.8	1808	3.5	4992	1.4	841	0.8	49	14.3	8415	4.0
6	52	46.2	137	43.8	535	12.5	2718	1.9	8514	0.9	1447	0.6	50	14.0	13453	2.3
7	23	56.5	73	24.7	547	11.5	4782	1.5	9551	1.3	2172	0.5	68	7.4	17216	1.8
8	23	39.1	54	24.1	432	8.6	7400	1.1	21571	0.8	4829	0.3	92	5.3	34401	0.9
9-10	11	36.4	28	28.6	184	6.0	5497	0.5	20514	0.7	5103	0.2	114	4.4	31451	0.7
Unstated	202	92.6	338	84.6	335	40.9	902	14.7	2357	9.3	345	8.7	953	10.7	5432	20.1
All Scores	1848	87.2	1798	57.8	4189	21.4	26505	3.4	75506	1.7	16051	0.8	1569	13.4	127466	4.7

*Number of Newborns in Birth Weight-Apgar Score Class
**Percent Mortality in Birth Weight-Apgar Score Class

TABLE 2

Relationship of Five Minute Apgar Scores and Mortality Rates of Newborns by Birth Weight Class, Ohio 1978-81

5" Apgar Score	Birth Weight (Grams) <750 No.*	<750 %**	750-999 No.	750-999 %	1000-1500 No.	1000-1500 %	1501-2500 No.	1501-2500 %	2501-4000 No.	2501-4000 %	>4000 No.	>4000 %	Unstated No.	Unstated %	All Weights No.	All Weights %
0-1	660	97.7	172	87.2	142	69.7	173	65.3	188	48.9	34	38.2	40	55.0	1409	80.4
2	311	93.2	119	67.2	123	56.9	164	41.5	229	29.3	36	33.3	23	34.7	1005	59.2
3	133	91.7	133	63.2	126	44.4	212	28.3	307	21.5	40	7.5	37	51.4	988	41.5
4	108	69.4	143	65.0	201	32.3	275	15.3	499	13.2	75	10.6	38	39.5	1339	27.2
5	90	70.0	184	47.8	346	28.0	569	12.5	941	7.0	127	4.7	28	21.4	2285	17.4
6	108	68.5	214	38.3	552	18.1	1201	7.0	2147	3.6	306	1.6	45	22.2	4573	9.4
7	57	57.9	210	40.0	711	15.3	2424	3.2	4570	2.0	630	1.3	59	18.6	8661	4.8
8	46	52.2	132	32.6	832	10.9	5492	1.9	12591	1.2	2254	0.4	103	4.9	21450	2.0
9-10	46	39.1	116	24.1	792	8.1	15012	0.9	51670	0.7	12206	0.2	242	5.0	80084	0.8
Unstated	289	92.7	375	82.1	364	40.1	983	16.1	2364	9.2	343	10.0	954	10.7	5672	21.5
All Scores	1848	87.2	1798	57.8	4189	21.4	26505	3.4	75506	1.7	16051	0.8	1569	13.4	127466	4.7

*Number of Newborns in Birth Weight-Apgar Score Class
**Percent Mortality in Birth Weight-Apgar Score Class

TABLE 3

Comparison of Mortality Rates of Newborns with One and Five Minute Apgar Scores by Birth Weight Classes, Ohio 1978-81

Apgar Score	Time	Mortality Rate* in Birth Weight Class <750 gm.	750-999 gm.	1000-1500 gm.	1501-2500 gm.	2501-4000 gm.	>4000 gm.	Unstated	All Weights
<4	1"	91.9*	61.5	35.6	17.6	7.5	5.1	35.4	26.9
	5"	95.7	74.1	57.5	43.9	31.1	25.5	49.0	62.9
4-6	1"	63.5	42.3	14.7	3.4	1.4	0.7	13.2	3.7
	5"	69.3	48.6	23.8	9.6	5.8	3.7	27.9	14.6
>6	1"	45.6	25.2	9.5	1.0(0.7)**	0.9(0.1)	0.3(0.5)	5.5	1.0(0.1)
	5"	50.3	33.8	11.3	1.4(0.9)	0.9(0.1)	0.3(0.1)	6.9	1.3(0.2)
Unstated	1"	92.6	84.6	40.9	14.7	9.3	8.7	10.7	20.1
	5"	92.7	82.1	40.1	16.1	9.2	10.0	10.7	21.5
All Scores		87.2	57.8	21.4	3.4(2.5)	1.7(0.2)	0.8(0.2)	13.4	4.7(0.9)

*Rate Per 100 Newborns in Birth-Weight Apgar Score Class
**Rates in Parentheses-Estimates based on all Live Births in Birth-Weight-Apgar Score Class

TABLE 4

Relationship of Five Minute Apgar Scores & Mortality Rates of White & Non-White Newborns by Birth Weight Class, Ohio 1978-81

Apgar Score	Race	Mortality Rate* in Birth Weight Class <750 gm.	750-999 gm.	1000-1500 gm.	1501-2500 gm.	2501-4000 gm.	>4000 gm.	Unstated	All Weights
<4	NW	94.6	69.9	43.1	29.0	22.5	20.0	42.1	62.1
	W	96.4	76.2	64.2	48.9	33.5	26.3	50.6	63.2
4-6	NW	70.5	40.4	18.1	7.4	5.0	7.0	25.7	14.3
	W	68.7	52.1	25.8	10.4	6.1	3.1	22.9	14.6
>6	NW	65.5	29.1	7.8	0.9(0.1)**	1.0(0.1)	0.3(0.2)	8.7	1.5(0.3)
	W	41.5	36.5	12.7	1.5(1.1)	0.9(0.1)	0.3(0.1)	6.4	1.3(0.2)
Unstated	NW	93.4	75.9	29.7	9.8	9.0	11.3	17.9	25.6
	W	92.3	84.9	43.6	16.3	9.0	9.7	10.1	20.3
All Scores	NW	88.3	51.1	15.7	2.3(0.2)	1.8(0.3)	1.9(0.6)	17.6	6.2(0.2)
	W	86.6	61.1	23.6	3.8(0.3)	1.6(0.2)	0.8(0.2)	12.2	4.4(0.1)

*Per 100 Newborns in Birth Weight-Apgar Score Class
**Rates in Parentheses-Estimates based on all Live Births in Birth Weight-Apgar Score Class

TABLE 5

Mortality Rates Associated with One Minute Apgar Scores by Birth Weight Classes, 1952-60* and Ohio 1978-81

Apgar Score	Year	Mortality Rate** in Birth Weight Class 500-999 grams	1000-1500 grams	1501-2500 grams	>2500 grams	All Stated Weights
<4	1952-60	87.7	68.5	26.1	3.7	12.1 (28.7)***
	1978-81	74.9	35.6	20.5	7.2	23.1
4-6	1952-60	61.1	32.7	7.0	5.5	2.3 (9.5)***
	1978-81	47.2	15.0	3.4	1.3	3.6
>6	1952-60	83.3	24.1	2.7	0.2	0.4 (0.5)***
	1978-81	31.0	9.5	0.6	0.1	0.1
All Stated Scores	1952-60	82.5	46.9	6.6	0.5	1.5
	1978-81	64.4	19.1	2.1	0.2	0.6

*Adapted from Apgar and James[2]
**Per 100 Live Births in Birth Weight Class-Apgar Score
***Mortality Rates in Parenthesis Adjusted on 1978-81 Ohio Birth Weights

TABLE 6

Mortality Rates Associated with Five Minute Apgar Scores by Birth Weight Class Pre-1964* and Ohio 1978-81

Apgar Score	Year	Mortality Rate** in Birth Weight Class <2000 Grams	2001-2500 grams	>2500 grams	All Stated Weights
<4	Pre 1964	77.9	29.6	15.4	35.4 (58.2)***
	1978-81	78.6	42.1	30.6	63.5
4-6	Pre 1964	39.1	11.1	3.4	9.5 (17.4)***
	1978-81	29.4	7.1	5.6	14.5
>6	Pre 1964	10.9	0.9	0.3	0.5 (.5)***
	1978-81	6.9	0.6	0.1	0.2
All Stated Scores	Pre 1964	30.1	2.1	0.6	1.4
	1978-81	22.5	1.2	0.2	0.7

*Adapted from Drage, et al
**Per 100 Live Births in Birth Weight Class-Apgar Score
***Mortality Rates in Parenthesis Adjusted on Ohio 1978-81 Birth Weights

TABLE 7

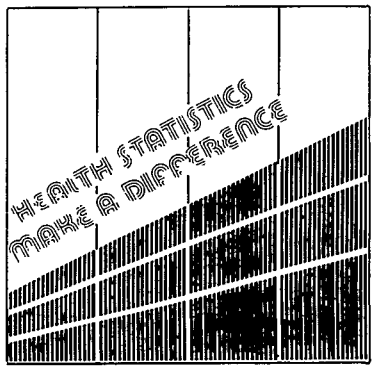
Deaths and Mortality Rates of Newborns with Selected Clinical Diagnosis and Associated One and Five Minute Apgar Score Pairs - Ohio, 1978-81

Clinical Diagnosis	Deaths and Mortality Rates* at Lowest 1" and/or 5" Score: Score<4 Number	Rate	Score 4-6 Number	Rate	Score >6 Number	Rate	Score Unstated Number	Rate	All Scores Number	Rate
Respiratory Distress Syndrome (RDS)	988	35.0	463	10.9	266	4.6	203	34.0	1920	14.3
Non-RDS Respiratory Disease	347	15.4	107	2.3	110	1.1	60	11.4	624	3.6
Congenital Defect	716	52.3	437	23.8	176	2.9	229	32.0	1557	15.6
Other Diagnosis Excluding Immaturity	290	16.5	108	2.0	176	0.5	180	7.7	754	1.6
No Diagnosis Except Immaturity	1122	27.0	109	0.9	53	0.2	361	35.0	1645	3.9

*per 100 Newborns with Diagnosis and Apgar Score Pair

Session L

Surveillance Systems for Health Events



POPULATION SURVEILLANCE FOR RARE HEALTH EVENTS

Tim E. Aldrich, University of Texas School of Public Health - Assigned to Oak Ridge National Laboratory

Cindy C. Wilson, Arizona State University

Clay E. Easterly, Oak Ridge National Laboratory

I. Need For Public Health Policy

The purpose of this paper is to describe a role for health statistics in the process of environmental policy formulation and decision making. The specific concern is with the use of population surveillance for identifying sources of environmental hazards, and for monitoring health patterns around sites of public health concern (e.g., toxic waste disposal sites, or oint source industrial locations). There is an increasing need for reasoned public health actions in the management of environmentally related health risks. With increasing frequency, modern society is faced with trade-offs when deciding about accepting risks. The process of public health policy formulation should not be based on dichotomous choices (e.g., all or nothing). Rather, policy decisions should be based on a continuous perspective of risk; for risk is truly probabilistic, as it exists in nature.

Historically, the available options for addressing unacceptable health standards or practices have been: 1) for occupational exposures - strikes, workers compensation, negotiation for improvements, or development of a preventive plan, and 2) for residential exposures - relocation, homeowner compensation, negotiation of acceptable improvements, or development of a prevention plan (Hanlon, 1984). Each of these actions implies a certain level of agreement between the contending parties on the existence of an unhealthful or undesirable situation, and of the definition of what constitutes acceptable health conditions. In modern society, it is more common for there to be some level of recognized uncertainty over both of these criteria.

There is often great uncertainty over the presence of a meaningful risk(s) (e.g., measurable) from many so called "environmental hazards". This doubt is fueled by a lack of understanding of the biologic mechanisms for many controversial health events and the absence of exposure metrics or the assessment of total dose. Often animal and human data upon which decisions may be based are in conflict with one another. Straightforward, biological and environmental reasoning is often complicated with questions about "worst case senarios" and considerations for high-risk individuals. There is also the likelihood that individuals will be exposed to other unhealthful factors (e.g., cigarette smoke) that will confound judgement regarding the risk(s) posed by the environmental factor of interest.

Finally, a community's response to a possible environmental "risk(s)" is another critical consideration. This response is not always one of "lets get rid of it". Increasingly, there are examples in which communities wish to accept what may be considered a low level health risk(s) in favor of retaining the economic benefits associated with the emission source (Ruckelshaus, 1984).

II-Health Belief Model

The Health Belief Model provides a conceptual basis for identifying and studying factors that influence personal decision making on health related issues. The elements of the Health Belief Model are useful for studying the environmental health circumstance of interest in this paper. That is when, what is commonly referred to as the "public domain" becomes a force field with positive and negative forces acting upon it (Becker, 1974). People must first perceive that they are susceptible to health risks as a result of an environmental exposure. This implies some appreciation for the severity of the exposure, the level of risk associated with it, and the relative incidence of the health event in question.

Decisions about environmental health issues frequently become highly politicized and publicized. Media representations of health issues often modify the public's

perception of environmental matters. Political priorities will often modify public health actions as well. Communities with different demographic and social configurations may respond differently to similar environmental questions; some may be apathetic, others zealous. Not infrequently, the visibility of one personality or the action of a special interest group will prompt or delay environmental action.

There are economic considerations related to the making of an environmental health decision: the cost of alternative commodities, the engineering feasibility of the proposed changes or clean-up, and the potential loss of jobs and revenues. Currently, there are communities in Washington state, Texas, Tennessee and Arizona where individuals say "we accept both" [the risk and the benefit] (Ruckelshaus, 1984). They want as safe an environment as possible, but they also want the jobs and economic benefits from the industry implicated as posing an environmental risk(s). In some settings, there may be an impasse in forming environmental health policy or decision making simply because "proof" is not available with regard to the presence of an elevated disease risk(s).

III-Risk Assessment versus Risk Management

Epidemiology is the study of disease risk(s), which is by definition, the likelihood of morbidity or mortality (Lilienfeld, 1980). It is the specific public health discipline directed toward the assimilation of factual information for the purpose of measuring the probability of disease occurrence. Epidemiology is also concerned with the causal inferences that are a part of studying disease processes in human populations. Disease prevention is the express objective in epidemiology, which uses both biological and statistical reasoning. "Risk assessment" has emerged as a separate, theoretical entity from classical epidemiology. This paper is directed in part, to suggesting creative uses of population-based disease surveillance as an adjunct source of epidemiologic reasoning to risk assessment, in lieu before epidemiologic study results are avialable.

When no human data is available, risk assessment is a highly theoretical, often mathematical endeavor to estimate the increased risk of a certain disease, as a result of exposure to some agent. Risk assessment is performed for a specific agent(s), using dose-response inferences, based upon models extrapolated from animal or cellular experiments. Risk management is the policy and implementation component of an environmental health decision, it includes weighing the engineering feasibility, economic impact and political efficacy of action. Implicit in these two risk oriented activities are definitions for what constitutes an acceptable risk(s), what portion of the disease experience of a community is attributable to a specific environmental exposure, and the existence of a decision rule indicating when official public health action is appropriate (Ruckelshaus, 1984).

Public health policy decisions are influenced by the relative frequencies of the exposure and the health events of interest. Primary-level disease prevention is going to become increasingly complex as low-level environmental exposures are linked to health effects (Nasseri, 1979). This is especially true when the low-level exposure is ubiquitous or nearly so, as with drinking water contaminants and energy related technologies. The absence of appropriate exposure metrics or the ability to assess total dose will add to the uncertainty of public health decision making.

Public health policy is difficult to formulate in a setting where the disease of interest is relatively common, e.g., lung or colon cancer. This difficulty is due in part to the lifestyle risk factors that have identified for these processes (e.g., smoking, diet). Less common cancers or other infrequent health events may serve more productively than "common" health effects to signal biological activity from a circumscribed, ambient environmental exposure (Aldrich et al, 1983). In fact, health events that are quite rare, are frequently in question with studies of environmental factors (e.g., liver cancer, brain cancer, certain birth defects).

This rarity of disease cases is an obstacle to research and an additional source of uncertainty for decision making. A path of policy-making restraint is recommended for a setting that involves rare health events. This course would include a search of evidence for the existence of an environmental risk(s). Vital health statistics have been used for identifying

possible sources of environmental hazards, and for surveillance of health patterns in populations living around sites of public health concern. One goal for the use of population surveillance data is to assist risk assessment, and risk management (e.g., public health policy formulation), before the presence of a health effect is manifest. The recognition of unusual aggregates of rare disease events, also called "case clusters", is one means of early identification for possible areas of environmental health concern; and "case cluster" recognition is a focus of this paper.

IV-Space-Time Disease Clusters

This method utilizes causal reasoning applied to case data to detect a change in the pattern of disease experience in a defined population at risk (Caldwell and Heath, 1976). A phrase that is often applied to case aggregates is "space-time clusters". Clusters of rare health events are encouraged for use as sentinel phenomena for early inferences regarding evidence for biological activity or the presence of some detrimental environmental factor. Common health events (e.g., breast and colon cancer) are less useful with these techniques because of the confounding, lifestyle risk factors that are already identified for these diseases (Aldrich, et al., 1983).

Spatial clusters represent the concept of geographical pathology, that is the simple proximity or dispersion of cases. This spatial reasoning may be likened to a scatter plot on a conventional x,y axis. There is of course, the need to adjust the spatial pattern for the underlying distribution of the population of interest. Temporal clusters are characterized by the classic "epidemic peak", where, for an infectious disease, there is a time period of high disease frequency, preceded and followed by relatively constant (and lower) rates. Space-time clusters are characterized by both spatial proximity, and a "peak" in time. Use of both space and time attributes may be helpful for recognizing an aggregate of disease events. With extremely rare events (e.g., where the disease incidence is uncertain), space-time patterns may still go undetected (Aldrich, 1984).

Refined statistical techniques have been developed for detecting disease clusters (Langmuir, 1965). One technique to analyze spatial data uses a modified Chi-square approach to test the frequency of disease events in user defined "cells" (Pinkel and Nefzger, 1959). Techniques for detecting temporal clusters use the strategy of a fixed time interval to "scan" along a time line to identify a period of increased occurrence (Ederer, et al., 1966; Wallenstein, S., 1980; Weinstock, M.A., 1981; Naus, J.I., 1982). Most methods however, have been developed for detecting space-time clustering (Knox, 1964; Barton et al., 1965; Mantel, 1967; Chen et al., 1982). These techniques vary widely in their approach, are quite complex and can be difficult to apply (Smith, 1982; Aldrich et al, 1983).

With the methods referred to above, there is often a dilemma of small numbers of cases (Aldrich, 1984). A prototypic disease experience of this type can be described using data from a Florida cancer cluster report involving a rare embryonal, pediatric tumor (Aldrich et al., 1984). In this instance, there were 11 cases of Endodermal Sinus tumor in the entire state of Florida over a ten-year period. Five of the eleven cases were aggregated in one small residential area, in the northeast corner of the state; four of these cases occurred in a two year period. This provocative cluster could have been detected by several of the clustering techniques mentioned above, yet detection would have been "after the fact". There is a means for recognizing a shift in the pattern of disease occurrence at an earlier time.

This problem may be likened to viewing cases within a multi-dimensional matrix (see Figure 1) (Aldrich, 1984). The matrix is defined by descriptive characteristics available for both the cases and the population from which they are taken. For example, consider age, race, sex, county of residence, year of diagnosis, etc. Convenient census data are available for the underlying population, e.g., age, race, sex characteristics, by county, by year. The cases that have accumulated by a certain date may be viewed as an ordered sample of size N, drawn with replacement from an available population of M "cells". These "cells" are defined by the descriptive characteristics mentioned before, e.g., the "cell" of individuals represented by non-white females, under the age of 20, in one Florida county. The probability (Pr) of K of the N cases being from this one cell is given in

Equation 1 (Parzen, 1960).

Equation 1:

$$Pr = \frac{\binom{N}{K}(M-1)^{N-K}}{M^{N}}$$

The utility of this "model" technique is demonstrated by comparing it with the other tests for space-time clustering (See Table I). This "model" approach has the advantage of sequential reasoning (see Table II), to serve as a "trip-wire" for indicating when there is a possible shift in the pattern of a rare event(s) occurrence. In Table II, two cases indicate that some aggregation may be occurring, even at the first level that test is recommended (e.g., with 3 cases). However, with the occurrence of the next case, the pattern dissipates. This absence of a pattern continues with the occurrence of the 5th and 6th cases. After the 7th case, there again is some basis for public health concern (e.g., a "watch" status is suggested when the probability falls below 1/M). This concern is substantiated by the continuing pattern of the disease occurrence during successive time intervals. Decision rules for conducting an investigation may vary with the relative disease frequency and severity (Aldrich, 1984).

With the many population based cancer and birth defect registries that are operating today, this "model" method may prove quite useful for the indentification of new sources of environmental hazards, and for surveillance of health patterns around sites of public health concern. This "model" may be used with some of the more elaborate statistical methods for detecting space, time or space-time clusters, or with others discussed at this meeting. This approach may also be applied in those situations where the underlying incidence rate of a health event is unknown or the population at risk is uncertain.

V. Summation

The use of health statistics data should be as one element in policy formation and decision making related to environmental health risks. Further, the practice of population surveillance is recommended when a choice is made to accept a low-level environmental risk(s), for which there are uncertain health effects and strong advantages. Especially encouraged is attention directed to the patterns of occurrence among rare health events, with environmental implications. Further work is needed on the use of health statistics with public health policy decisions.

References:

Aldrich, T. E. (1984). Detecting Space-Time Aggregation of Rare Events, an Abstract, Am. J. Epidem. 120(3), p. 464.

Aldrich, T. E., Glorieux, A., and Castro, S. (1984). Florida Cluster of Five Children With Endodermal Sinus Tumors, Oncology, 41(4), p. 233-38.

Aldrich, T.E., Garcia, N., Zeichner, S., and Berger, S. (1983), Cancer Clusters: A Myth or A Method ?, Med. Hypoth., 12(1), p. 41-52.

Becker, M.H. (1974). The Health Belief Model and Personal Health Behavior, Pub. Charles B. Slack, Inc., Thorofare, N.J.

Caldwell, G. G. and Heath, C.W. (1976). Case Clusters in Cancer, So. Med. J., 69, p. 1598-1602.

Chen, R., Mantel, N. and Isaccson, C. P. (1982). A Monitoring System For Chronic Diseases, Meth. Inform. Med., 21, p. 86-90.

Ederer, F., Myers, M.H., and Mantel, N. (1966). A Statistical Problem in Space and Time: Do Leukemia Cases Come in Clusters", Biometrics, 20, p. 620-38.

Hanlon, J. J. (1984). Public Health: Administration and Practice 7th Ed., Pub. C. V. Mosby Co., St. Louis

Knox, E.G. (1964). The Detection of Space-Time Interactions, Appl. Stat. 13, p. 25-29.

Langmuir, A. D. (1965). Formal Discussion of the Epidemiology of Cancer: Space-Temporal Aggregation, Cancer Res., 25, p. 1384-86.

Lilienfeld, A. M. (1980). Foundations of Epidemiology, 2nd Ed., Pub. Oxford University Press, New York.

Mantel, N. (1967). The Detection of Disease Clustering: A Generalized Regression Approach, Cancer Res., p. 209-20.

Nasseri, K. (1979). "Letter to the Editor", Int. J. Epidem., 8(4), p. 389-90.

Naus, J.I. (1982). Approximations for Distributions of Scan Statistics, J. Am Stat. Assoc., 77, p. 177-83.

Parzen, E. (1960). Modern Probability Theory and Its Application, Pub. John Wiley and Sons, Inc., New York.

Pinkel, D. and Nefzger, D. (1959). Some Epidemiological Features of Childhood Leukemia in Buffalo, N.Y. Area, Cancer, 12, p. 351-58.

Ruckelshaus, W.D. (1984). Risk In A Free Society, Risk Analysis, 4 (3), p. 157-62.

Smith, P.G. (1982). Spatial and Temporal Clustering, in Cancer Epidemiology and Prevention, Ed. Schottenfeld D. and Fraumeni, J., pub. W.B. Saunders and Co., Philadelphia, p. 391-407.

Wallenstein, S. (1980). A Test For THe Detection of Clustering Over Time, Am. J. Epidem., 111, p. 367-72.

Weinstock, M.A. (1981). A Generalized Scan Statistic for the Detection of Clusters, Int. J. Epidem., 19, p. 177-83.

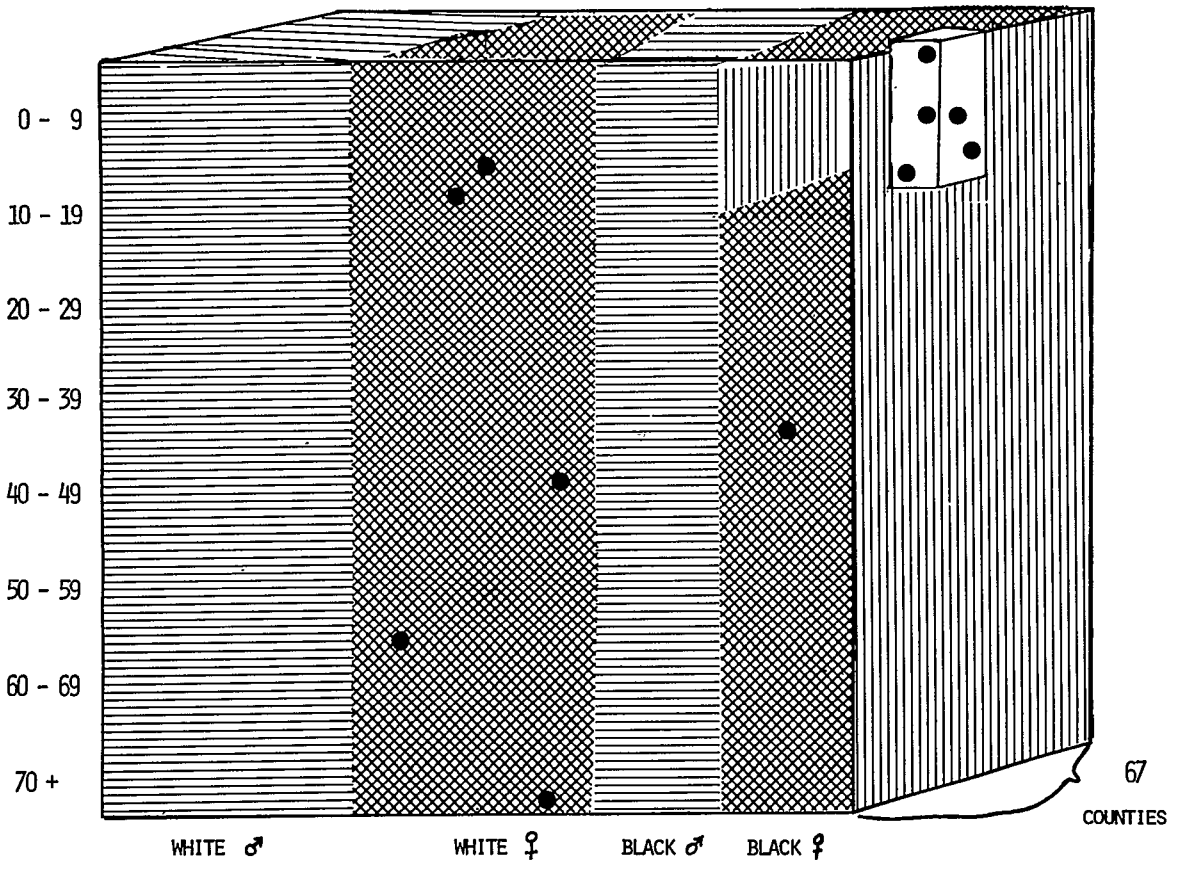


Figure 1 - Schematic Diagram of a Multi-dimensional Matrix. The sample dimensions shown are age (Vertical), Race/Sex (Horizontal), and County of Residence (Depth). More elaborate dimensionality is possible, e.g., time attributes.

Table I - Comparison of the Model to Other Methods

Test	Finding	Comment
Model (Aldrich, 1984)	p < 0.001	5 of 11 cases in 1 of > 66 cells
Pinkel and Nefzger (1959)	p = 0.141	67 county cells, 10 1 - year periods
Knox (1964)	0.10 < p < 0.05	4 observed / 2 expected
Barton and David (1965)	p < 0.05	Q = 0.733 Var. = 0.776 F = 9.33 (4,3 d.f.)
Ederer, Myers and Mantel (1965)	0.10 < p <0.05	Five 2 - year intervals chi-square = 2.875
Mantel (1967)	p < 0.50	Test of B, t = 0.4354
"Scan" Statistic (see references)	p < 0.995	11 cases in 8 time periods, 3 case maximum
Chen, Mantel and Isaccson (1982)	p < 0.05	9 of 11 cases met criteria for being "close"

Table II - Simulation of the "Model" As A Sequential Test

Number of Cases (number of cases in suspected cluster)	p-value of Model test (M = 640 cells)	
3 cases (2)	8.977 x 10-5	
4 cases (2)	0.188	
5 cases (3)	0.012	
6 cases (4)	0.164	
7 cases (2)	2.154 x 10-5	*
8 cases (3)	3.612 x 10-7	#
9 cases (4)	1.523 x 10-9	
10 cases (5)	5.667 x 10-12	

* - Criteria for suspecting a cluster exists is recommended as 1/M. In the first instance the pattern was not continued through successive time periods (evidence for consistentcy).

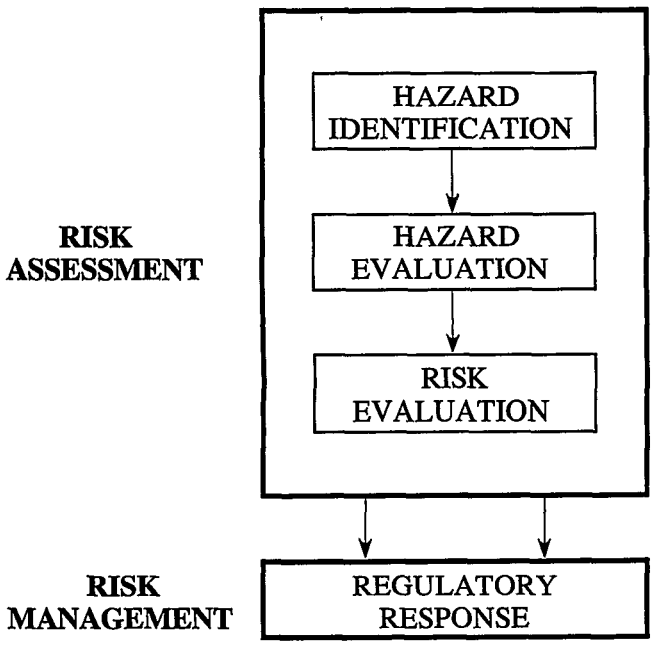
- Some level of public health response would be indicated at this point (Aldrich, et al., 1983; Aldrich, 1984)

HEALTH STATISTICS SURVEILLANCE SYSTEMS FOR HAZARDOUS SUBSTANCE DISPOSAL

Jay H. Glasser, The University of Texas Health Science Center at Houston

Hazardous waste site monitoring presents several challenges to health statistics. Firstly, monitoring requires an admixture of scientific knowledge within context of a political and regulatory environment. Figure 1 drawn from C.N. Park and R.D. Snee (1983) succinctly show the multiple steps involved in monitoring. Broadly both a risk assessment process and a risk management strategy is required. The process is bidirectional; the risk management and regulatory response must be consonant with the scientific base, conversely the identification and evaluation of hazardous substances must be translated into a regulatory process.

Figure 1. The Four Major Steps In The Process Of Risk Assessment and Risk Management.



Source: The American Statistician, November 1983, Vol. 37, No. 4

The critical factor of the causal link or association itself connotes a long trail of scientific evidence from multiple sources as presented by C.N. Park and R.D. Snee op cite (Figure 2) including genetic, chemical and toxicological, field and laboratory epidemiology and biomathematical modeling. Furthermore as illustrated in Figure 2 we are dealing with a "moving target" assessing risks as the scientific understanding and knowledge base also change.

For health statisticians, environmental engineers, and epidemiologists, a variety of quantitative studies and data sources need to be interdigitated. (see Figure 3).

The quantitative base can and often does draw open both available data as well as special studies. The special studies are often necessitated because of the site specific conditions, and the characteristics of the (small area) surrounding population and ecology at risk of the particular hazard wastes and derivatives generated in the storage or incineration process.

Figure 2. Data Used In The Risk Assessment Process.

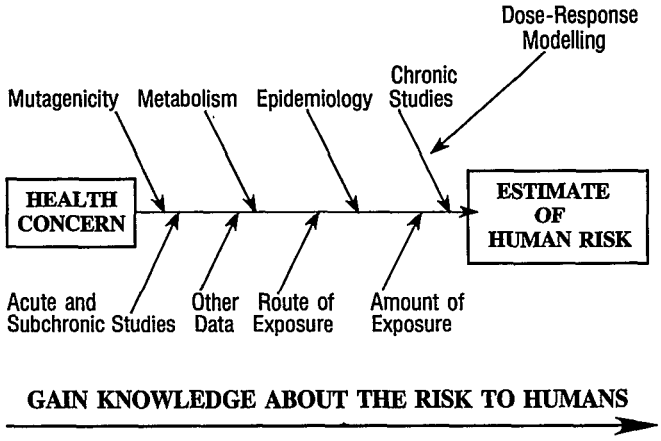
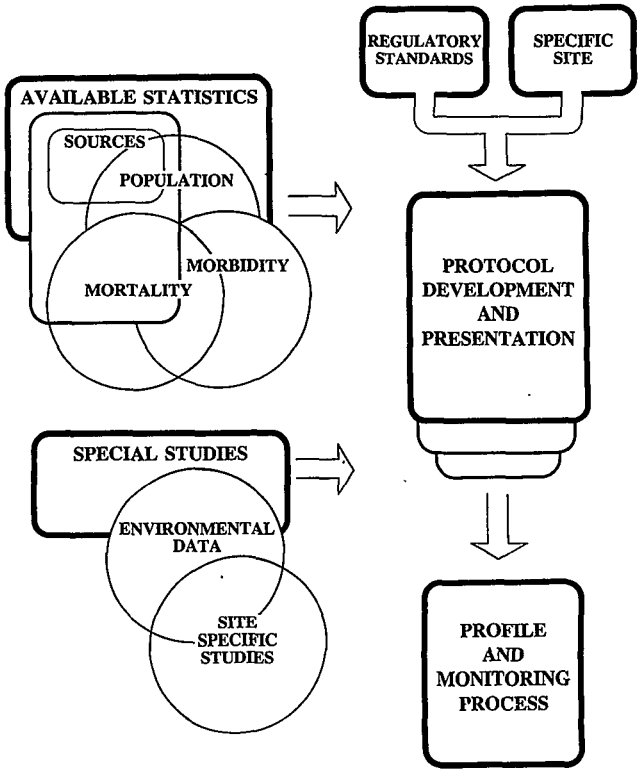


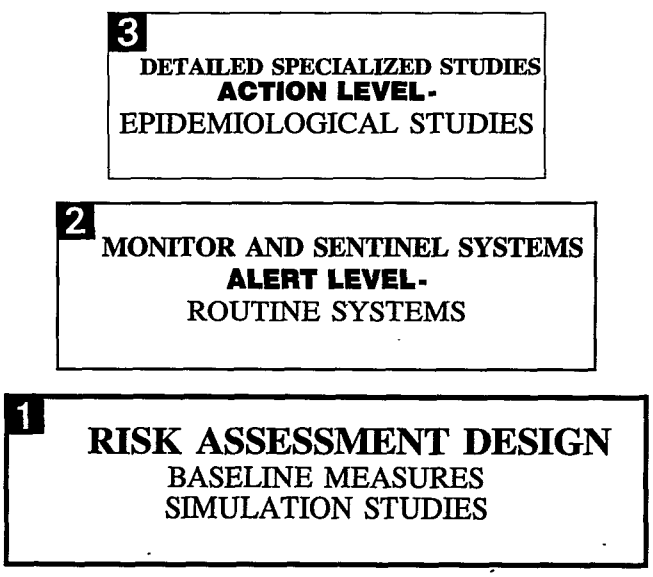
Figure 3. Development Of The Health Profile and Monitoring Process.



The challenges for monitoring multicausal and interacting factors that operate in the small area can be either potentially intensive or attenuated exposure effects and are therefore manifested in several types of health effects, short-term and long-term. To accomodate both the "usual" hazards and the potential for unknown untoward circumstances the hazardous waste monitoring profile must be responsive to measurement challenge of the multiple exposure , time effects, and the underlying at risk populations; no small task.

A surveillance system must evolve in the practical context of the entire spectrum of scientific and regulatory structure described above. A flexible response is also a necessary prerequisite. To this end various concerned parties have proposed what amounts to a three tiered structure for monitoring (see Figure 4). The base line establishes the current context both information wise and strategy wise. The monitor and surveillance component carries out the ongoing routine study of conditions. This level provides an alert level, where chosen sentinel events are either seen as within tolerable bounds or exceeding critical levels as determined during the base line periods, or revised subsequently by new knowledge or regulations. Any rise above the proscribed levels creates an alert, that triggers further investigation. The third level is an action level where detailed study is indicated by the monitoring and alert level surveillance.

Figure 4. Three Tiered Information Capability.



The hierarchical three tiered system is a sensible approach to dealing with the multifactor and time series tasks embedded in the hazardous waste surveillance. Figure 5 illustrates several of the key factors that must be untangled, and further underscores the need for both routine monitoring and special study capability. Included in the hierarchical structure must be the analytic capability to provide the appropriate sentinel measures and their interpretation. Figure 6 provides an illustration of the need to measure rare events in small area populations. The use of relative risks and odds ratios provide specific rates. The data and chart in Figure 6 are drawn from a study of Love Canal and amply demonstrate the small area methodological and base line assumptions that are often embodied in such studies. In addition Love Canal itself, underscores the social and political context within which controversy may be an inevitable partner to the quantitative results.

Figure 5. Factors In The Web Of Causation

- **UNDERLYING DIFFERENCES IN POPULATIONS AT RISK**
- **CHANCE**
- **EFFECTS OF OTHER EXPOSURES**
- **INFORMATION BIASES**
- **LEVEL OF ACTUAL POSITED EXPOSURES**
- **MULTIPLE EXPOSURES - MULTIPLE EFFECTS**

Figure 6. Maternal Age And Number Of Miscarriages (Observed And Expected*) Among Residents Of The Love Canal

Maternal Age	Number of Pregnancies	Number of Miscarriages		Relative Odds Ratio Observed/Expected
20	2	0	0.212	0.00
20 - 24	13	0	1.852	0.00
25 - 29	28	3	3.550	0.85
30 - 34	19	6	2.677	2.24
35 - 39	15	8	3.104	2.58
All Ages	**77**	**17**	**11.395**	**1.49**

*Based on Warburton and Fraser: "Spontaneous Abortion Risks in Man: Data from Reproductive Histories Collected in a Medical Genetics Unit." *Human Genetics* Vol. 16, No. 1, 1964, Page 8.

Much demographic and statistical work does provide a basis for the use of both routine data in support of monitoring in the face of admittedly difficult circumstances. The use of multiple measures can realistically present a picture that mirrors both reality and yet can convey the components of the health profile to an informed public. For example, Figure 7 displays the use of both standardized morbidity and standardized proportional morbidity ratios that compare exposed county areas to state levels (unpublished specimen analyses by the author). The standardized rates provide a comparative view, and the collateral use of the proportional ratios provides a check where accurate denominator data is potentially a problem.

Figure 7. Standardized Ratio Of Malignant Neoplasm Hospital Discharges For Aggregate Counties: 1979

Neoplasm Sites (ICD - 9)	Standardized Morbidity Ratio* [$x^2(1)$]	Standardized Proportional Morbidity** Ratio
Digestive (150-159)	.542 (7.73)+	.696 (2.65)
Respiratory (160-165)	.841 (1.29)	1.090 (0.32)
Leukemia (204-208)	1.304 (1.42)	1.362 (1.92)
Other Sites	.813 (8.08)	.999 (0.00)
All Sites (140-208)	.810 (12.08)	1.000 (--)

*computed using the 1979 Missouri age-specific morbidity rates as the standard.

**computed using the 1979 Missouri age-specific proportions of all neoplasms as standard.

+significantly different from unity at .05 level.

Much of the statistical basis for the monitoring profile performance may be established during the baseline period. This enhances the impartiality of the system and at the same time establishes the ground rules by which interpretations will be made on an a priori basis. Figure 8 (R.J. Hardy et al, 1983) provides an illustration of the manner in which the alert and action levels for the SMR ratios cited previously may be established within the context of the three tiered system. The z score level may be chosen consistent with the stringency desired, the time periods used to compare trends, and build in considerations based on the anticipated costs and benefits of declaring alert and action levels. The expected values (E) are generated by the appropriate comparison (non-exposed) risk specific population.

The alert and action levels may be expressed as critical numbers of deaths (or correspondingly morbidity cases) for given situations (R.J. Hardy, op cite) and provide a further convenient method of defining action levels in the surveillance system (Figure 9).

No system will ever meet and simultaneously satisfy optimal properties of scientific, political, and pragmatic operational standards. The problems are many and may be abundantly illustrated in the monitoring of incinerator sites. Here the variable potential exposures are complicated by atmospheric, topographical and population spatial clusters. Yet the use of either direct or modelled exposure data may be overlayed on the population and geographic small area distributions to provide a basis for quantitative surveillance. Figure 10 displays the overlay of potential exposure contours on the spatial population distribution. In this case a natural grouping of low exposure may be compared to a potentially higher exposure group with geographical proximity (unpublished specimen analysis by the author).

Figure 8. Summary Table Of The Number Of Deaths Required (Or The Magnitude Of SMR Required) For An Alert* Or Action To Be Taken For Various Values Of The Expected Number Of Deaths

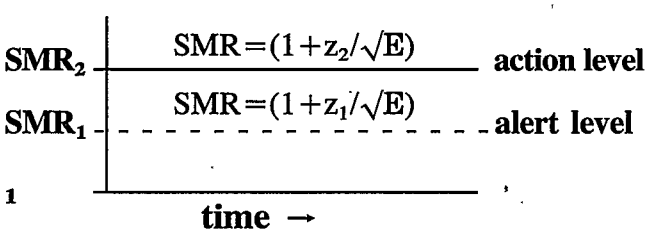
Expected Number of Deaths	Number of Deaths For an Alert	For Action
0.050	1 (20)**	3 (60)
0.100	2 (20)	3 (30)
0.200	2 (10)	4 (20)
0.400	2 (5)	4 (10)
0.500	2 (4)	5 (10)
1	3 (3)	6 (6)
2	5 (2.5)	9 (4.5)
4	8 (2)	12 (3)
5	9 (1.8)	14 (2.8)
10	15 (1.5)	22 (2.2)
15	21 (1.4)	29 (1.93)
20	26 (1.3)	36 (1.8)
25	33 (1.32)	43 (1.72)
30	38 (1.27)	49 (1.63)

*Action and alert levels correspond to $p_2 = 0.001$ and $p_1 + p_2 = 0.09$ for a two-year error level of 0.01.

**( ) corresponding SMR associated with the specified expected number of deaths and observed number of deaths.

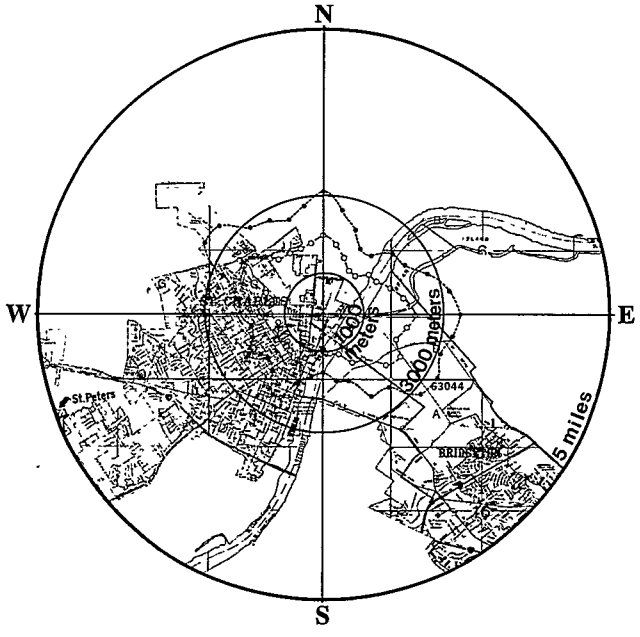
Source: Hardy, RJ, *Monitoring for Health Effects of Low-Level Radioactive Waste Disposal: A Feasibility Study,* 1983.

Figure 9. Setting Statistical Criteria For Interventions



Source: Hardy, RJ. *Monitoring for Health Effects of Low-Level Radioactive Waste Disposal: A Feasibility Study,* 1983.

Figure 10. Particulate Emissions Contours (All Sources) Empirical Example Plant Annual Concentration (Incinerator Assessment) For .2 And .4 Micrograms Per Cubic Meter And Five Mile Radius (Detail Of Surrounding Site Communities).



—— .4 micrograms per cubic meter
o—o .2 micrograms per cubic meter

The conclusions are several. The problems of hazardous waste sites and incinerators is a vexing one; socially and health wise an important one. In the context of epidemiology and environmental sciences it is one with multiple factor and risk considerations. The data and quantitative studies available are usually not equal to the task, and their limitations may be and should be clearly stated. However existing data and methods may be brought together within a regulatory context to provide a sensible basis and strategy for objective surveillance that can serve the legitimate goals of a complex industrial and pro health promotion society. Small area health statistics and analyses will play an important role in any credible surveillance system.

THE ROLE OF HEALTH INSURANCE DATA IN EVALUATING OCCUPATIONAL MORBIDITY

Jerome Wilson, National Cancer Institute

Introduction

Although epidemiologic studies play an important role in the elucidation of health hazards in industrial work environments, there is currently very little information available on morbidity events associated with industrial exposures among occupational groups. In large measure, this paucity of data on illnesses not associated with death reflects the fact that practically all occupational studies completed to date have focused on mortality. It may be, however, that morbidity is able to offer a greater degree of sensitivity than mortality in measuring the potential biological effects of occupational exposures. The purpose of this paper is to discuss the use of medical insurance data as a resource for occupational morbidity studies and to present an example of one such study.

Mortality data are limited as a measure of the amount and characteristics of ill health in any population, especially when considering nonfatal diseases. For an industrial population, the limitations of mortality are even more apparent since "good health" is usually required in order to maintain employment. Mortality yields only one event, given that death is at the end of a continuous time line from birth to death (Figure 1). Mortality is easy to count since it is a non-repeating event, but making causal inference is more complex. Morbidity allows for the examination of several outcomes over time, but is more difficult to detect and count.

Employee medical insurance claims represent an untapped resource for occupational morbidity studies. Although insurance claims are primarily used for payment of benefits, the medical diagnosis is an important part of this document. Medical insurance coverage has been a part of employee benefits for several decades in this country. The physician reported diagnosis can be abstracted and coded to the International Classification of Diseases, adapted (ICDA).

Lack of Research

Limited epidemiologic research has been conducted in the area of occupational morbidity, possibility because of the difficulties in describing and measuring broad categories of illness, and the general lack of readily retrievable data.

The amount of ill health in a population may be measured by: (1) the number of individuals who have a disease event; (2) the number of disease events; and (3) the number of individuals who experience multiple disease events (Dorn, 1957). During a fixed interval of time, one person may experience one or more disease episodes.

Most studies in occupational epidemiology have used the cohort mortality approach. Mortality studies, however, may not be the most effective tool for studying health outcomes in working populations. There are, for example, nonfatal health conditions that are not likely to appear on the death certificate, but which are potentially related to the occupational exposure and need to be evaluated. Research that focuses primarily on morbidity can potentially provide a more complete understanding of the health problems associated with industrial exposures.

Medical Insurance Claims

Medical insurance claims represent existing data that are collected from routine company operations and not from special epidemiological studies. These claims contain the diagnosis, treatment, and some demographic data. The diagnosis is certified by the attending physician. The requirement to have all claims certified by a physician is an important one. Physicians are most familiar with medical morbidity since it is the basis of clinical medicine.

Medical insurance claims represent a logical data source for morbidty studies, especially nonfatal acute and chronic disease of relatively short latency. This approach allows one to evaluate the disease experiences of workers while employed in a particular industry, thereby decreasing the interval between potential exposures and disease manifestation. The longer the interval between exposure and outcome, the poorer the probability of being able to make a casual association.

The advantages of using the insurance records for identifying morbidity (non fatal disease outcomes) are as following:

1. Diagnosis is made by a physician.
2. Complete reporting of cases among employees is insured because a claim must be filed to obtain benefits.
3. These records have potential for use in an ongoing occupational helath surveillance system.

Limitations

Some potential problems with morbidity data include record keeping, diagnosis, coding, and recurring episodes of nonindependent diseases (Tjalma, 1972). Data collection and processing is by no means a small task; therefore, data management requires considerable effort (Barrett, 1977).

The quality of the data is influenced by several factors: purpose of the recording, frequency of recording, persons recording, geographic location, physical setting, date of

recording, number of diagnoses recorded at one encounter, continuity of care, the interval between service and recording, and the recording system itself (Anderson, 1980; Kerr, 1978)).

The major disadvantage associated with the study of nonfatal disease outcomes is that data bases, are not well developed in the United States at the present time except possibly for cancer registries (i.e., SEER). Access to medical information is limited, making it difficult to validate end points. No national or regional statistics are collected and validated. One data base that comes close to this is the National Health Interview Survey Data, which is based on self-reported diagnosis. There may be increased cost and time associated with morbidity studies, especially cohort studies.

Example from Uranium Workers

To evaluate the usefulness of health insurance data for etiologic research, insurance data from a cohort of uranium workers was studied to examine the relationship between nonmalignant respiratory diseases and uranium exposure. A significant difference in the probability of developing a nonmalignant respiratory disease among three exposed groups was observed on the basis of health insurance data (Wilson, 1983), whereas no association was found when the analysis was restricted to mortality data alone.

The study cohort consists of all white males who were first hired between January 1, 1952, and December 31, 1972, and who have at least three months of continuous employment. The restriction to those working a minimum of three months is related to a 90-day employment requirement to qualify for medical insurance benefits. The cohort was enumerated through company rosters and further defined with the use of personnel records. All disease events in this study were taken from employee medical insurance claims. Thus, only physician-diagnosed disease events were included. Figure 2 illustrates how insurance claims were processed.

Summary

The primary purpose of this paper is to stimulate interest in a largely unexplored area of occupational epidemiology, which offers great potential for researchers and industrial management. The insurance industry in this country would be an excellent source of information for developing a large data base on morbidity and health surveillance.

Despite the complexities and difficulties associated with morbidity studies, there are several advantages to conducting such studies. First, morbidity studies are a natural complement to mortality studies, given that mortality is the ultimate morbid event. Second, morbidity offers an opportunity to shorten the interval between exposure and development of disease, thereby making it possible to gain new knowledge.

Health insurance data can be used for cross-sectional, case-control, or cohort studies of the relationship between the occupational environment and health (Smith, 1983). This approach to the study of occupational disease is in need of additional research (Goldberg, 1982).

Even in light of these limitations, we have demonstrated that morbidity data can be useful in the evaluation of health outcomes in an occupational setting. We are proposing that morbidity and mortality studies be conducted together. Moreover, these data can offer sound preliminary evidence in a case where the cohort is small and follow-up relatively short; a study of disease episodes is likely to be more informative than mortality events.

We conclude that health insurance data can play a role in epidemiologic studies by highlighting associations of certain occupations and exposures with recurring nonindependent disease events, and by permitting the evaluation of precancerous disease states, co-morbidity, and competing risks.

References

Anderson, J., 1980. Reliability of Morbidity Data in Family Practice. J. of Family Practice 10: 677-683.

Barrett, C.D., and Belk, H.D., 1977. A computerized Occupational Medical Surveillance Program. J. of Occupational Medicine 19: 733-736.

Dorn, H.F., 1957. A classification system for Morbidity Concepts. Public Health Reports 72: 1043-1048.

Goldberg, M., Blane, M., Chastang, J.F., Blane, C., and Sommer, M., 1982. The Health Data Base of a Nationwide Company - Its use in Epidemiological Studies. J. Occupational Medicine 24: 47-52.

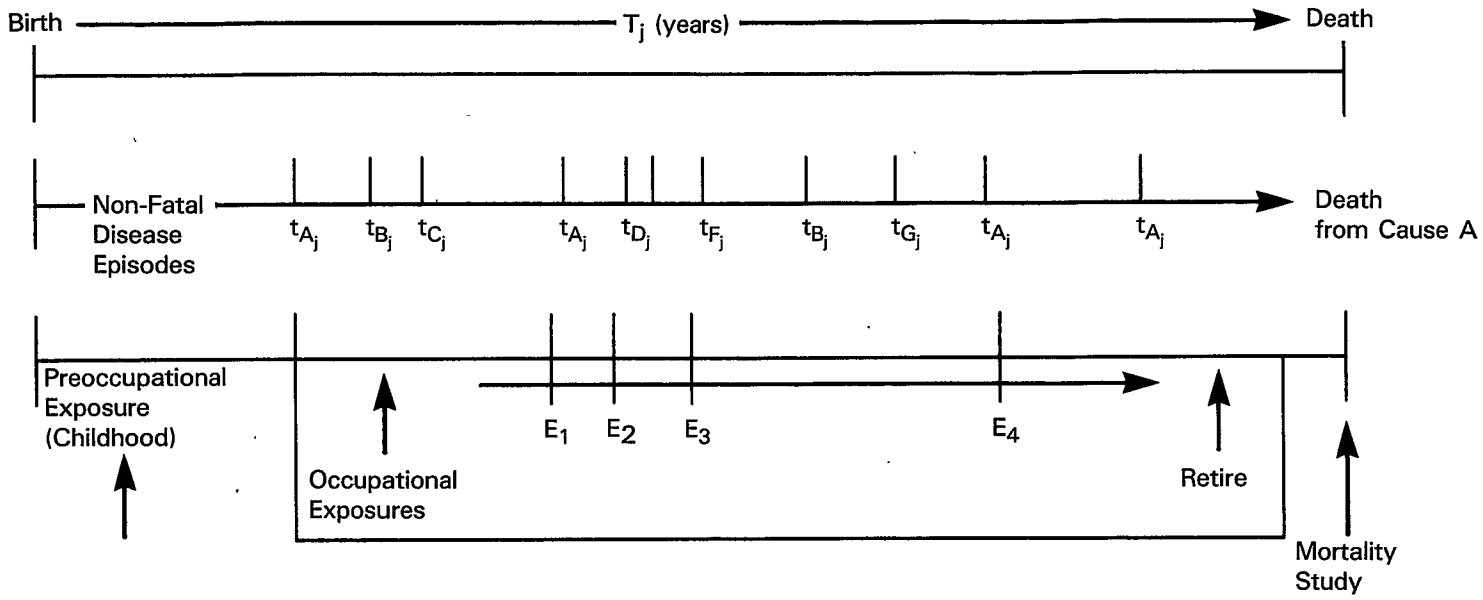
Kerr, P.S., 1978. Recording Occupational Health Data for Future Analysis. J. of Occupational Medicine 20: 197-203.

Smith, A.H., 1983. Factors in the selection of control groups. In: Chiazzen, L.,Jr., Lundin, F.E., Walkins, D., eds. Methods and Issues in Occupational and Environmental Epidemiology. Ann Arbor Science, Ann Arbor, Michigan 11: 107-115.

Tjalma, R.A., and Braun, J.L., 1972. Biomedical Data Resources: A limiting factor in Analytic Epidemiology. Environmental Health Perspectives October, pp. 67-72.

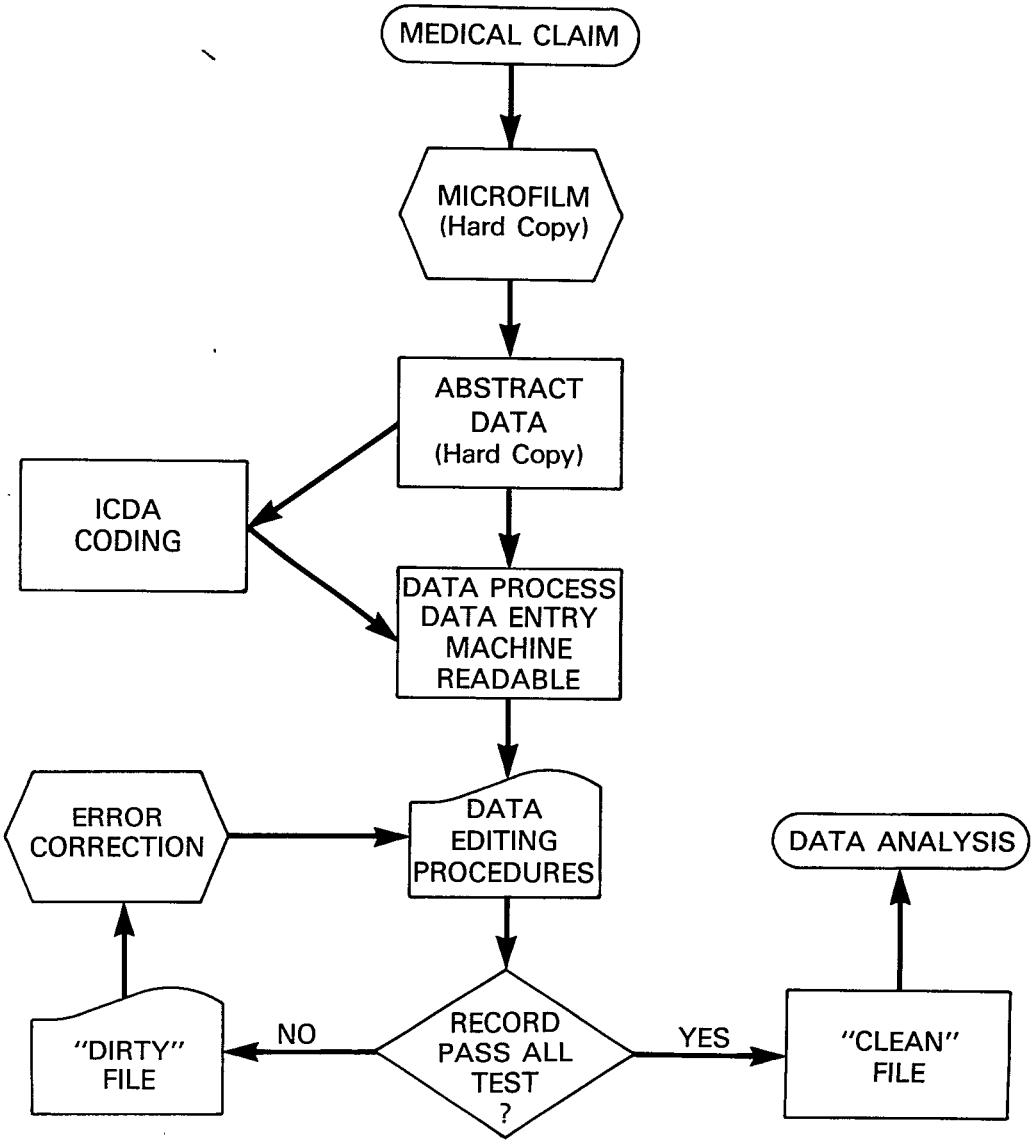
Wilson, J., 1983. An Epidemiologic Investigation of Nonmalignant Respiratory Disease Among Workers at a Uranium Mill, Ph.D., Dissertation, University of North Carolina, Department of Epidemiology, 1983.

Figure 1 Theoretical Life-Time Exposure-Disease Events Scheme



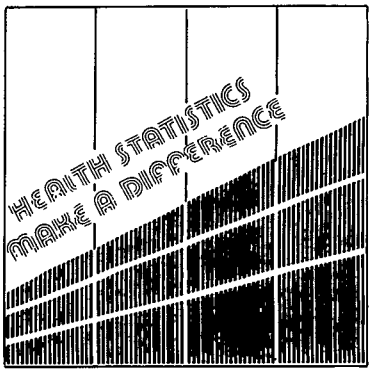
1. t_A = Diseases "A" at time "t" and event "j."
2. E_1 = Exposure to agent 1, 2, 3, etc.
3. This scheme emphasizes occupational exposures but other environmental exposures are considered, [homes, personal habits (i.e., smoking, drinking, etc.), urban, rural, sex, and race].

FIGURE 2.
FLOW CHART FOR MORBIDITY DATA DEVELOPMENT



Session M

Developments in Statistical Methodology for Complex Sample Surveys



MIXING MICRO AND MACRO DATA: STATISTICAL ISSUES AND IMPLICATIONS FOR DATA COLLECTION AND REPORTING

Mark S. Kamlet, Steven Klepper, Carnegie-Mellon University, Richard G. Frank, The Johns Hopkins University

1. Introduction

This paper discusses the statistical estimation problem that arises when data on individuals are augmented with data pertaining to groups of which the individuals are members. We begin by clarifying the contexts in which this problem arises. We then discuss the nature of the statistical estimation problem, focusing on regression analysis. Finally, we discuss a relatively simple procedure that can be employed to correct the problem.

We illustrate the statistical problem in the context in which it arose in our own research. We have been involved in a project, funded by the Environmental Protection Agency, that has involved in part an investigation of the health effects of occupational exposure to several specific pollutants. We have employed two primary data sets in this work. One is the 1980 Health Interview Survey (HIS), a large, stratified cluster sample of 25,000 U.S. households comprised of some 100,000 individuals. The HIS is collected annually by the National Center for Health Statistics (NCHS). It provides detailed information on health status and various other demographic, economic, and health status characteristics for individuals in the sample. The 1980 HIS was particularly useful for our interests because it contains supplementary information for a subset of the sample on smoking and occupational history. The other primary data set is the National Occupational Hazard Survey (NOHS) conducted in the mid-1970's by the National Institute for Occupational Safety and Health. It provides estimates of the exposure of workers in various industrial and occupational categories to different occupational pollutants.

In investigating the health effects of occupational pollution exposure, we wanted to estimate a "health-production function" (to use the terminology of economics) or "dose-response function" (as it is more commonly referred to in epidemiology and biostatistics), in which some continuous measure of health status for each individual i, y_i, is regressed on a vector of nonpollution variables that may influence health, $\underline{z}_i$ (column vectors are underlined), as well as the level of exposure of the individual to pollution, x_i^*. The variables y_i and $\underline{z}_i$ come from the HIS. The pollution exposure variable was constructed by matching each individual's occupation as recorded in the HIS survey with the pollution exposure for that occupation reported in the NOHS survey.

The problem this raised was that the constructed pollution variable did not measure the pollution exposure of each individual in the sample. Rather, it uses the expected level of exposure of individuals in similar occupations or industries to measure individual exposure. This sort of situation arises frequently in health research which utilizes large surveys such as those collected by NCHS. For instance, we presented a paper at this conference (Frank, Kamlet, and Klepper; 1985) in another session reporting the substantive results of the research described above. In that session all of the papers involved linking individual health outcomes to (among other variables) some average or estimated level of individual exposure based on factors such as expert opinions about occupational exposure, distance of residence from a source of pollution, etc. Another common context in which this problem occurs is when individual health status is related to (among other variables) a measure of air pollution based on fixed-site monitors. Once again, the actual level of exposure for the individuals in the sample is not known. All that is known is the average or expected exposure level for all individuals within a given geographical area around the monitor.

Another example of this problem is when some individual-level variables in a data set cannot be reported on an individual basis. This may be due, for instance, to confidentiality restrictions of the sort often involved in NCHS surveys, such as HIS, in which individual-level data cannot be reported and instead only averages from the primary sampling unit (e.g., census tract) or even larger aggregations (e.g., county) can be provided.

2. The Nature of the Statistical Problem

In all of these instances, one does not have individual-level data for some key explanatory variable, only information about the variable for the group that the individual is a member of -- his occupation, his census tract, his location relative to a pollution source or fixed-site monitor, etc. What is done in practice when this situation arises? Typically, the problem is ignored. The researcher uses the associated group-level information instead of the (unobserved) individual-level data. In the case of our research, for instance, lacking a measure of the degree of occupational exposure of a given individual to a pollutant, one would use the average exposure of individuals in the same occupation or industry.

However, substituting group for unobserved individual-level data is not an innocent practice. It introduces a measurement error in the pollution exposure measure whenever individual exposure differs from the average exposure of the group to which the individual belongs. It has long been recognized that measurement error in explanatory variables, even if of a "classical" or white-noise nature, can lead to inconsistency in parameter estimation. In a regression context, for example, classical measurement error in a single variable leads to an attenuated coefficient estimate for that variable. It also leads to inconsistent estimates of the effects of other explanatory variables that are correlated with the mismeasured variable, possibly even causing such estimates to be the opposite sign of the true coefficients. Note that the problem has nothing to do with sample size. Since the estimates are inconsistent, increasing the sample size merely provides more precise estimates of the wrong coefficients.

To our knowledge, no one has explored the nature of the measurement error when group-level data are used in place of individual-level data for some explanatory variables. It turns out, as we discuss below, that the nature of the resulting measurement error is decidedly nonclassical (i.e., not white-noise). The fact that the measurement error is nonclassical

does not, however, mitigate its consequences. Indeed, a number of researchers have questioned the findings of studies using group-level data in place of individual-level data. They suggest that more effort should be directed toward compiling individual-level data on variables such as pollution exposure, diet, and exercise. If in fact such measures are needed to make reliable inferences, it reduces significantly the usefulness of such large-scale NCHS surveys as HIS, the Health and Nutrition Survey and the National Medical Care Utilization and Expenditure Survey in shedding light on the health effects of pollution as well as many other important health policy questions.

In a sense, we strongly disagree with both camps in this debate. We believe that the statistical problems that arise from augmenting individual-level data with group-level data are important and should not be ignored. At the same time, we do not believe the nature of such problems necessarily implies the need to resort to individual-level observations for all variables whenever the dependent variable is individual level. There are some clear disadvantages with requiring individual-level data for all explanatory variables. One disadvantage is cost. Often the collection efforts required, such as using personal pollution monitors to measure individual pollution exposure, are orders of magnitude more expensive than group-level variable collection, particularly when group-level data are available from previous studies. Perhaps more important, the flexibility of data sets is substantially reduced when individual-level data are required for all explanatory variables. Under such a requirement, all variables that might ever be used in an analysis must be collected at the same time. The collection effort will be rendered useless if, after the fact, it is decided that individual-level data on some variable are needed but were not collected. One of the great promises of large-scale health-related data sets is their potential use in addressing research questions that were not completely anticipated at the time of their collection.

The answer to this problem from our perspective is neither to ignore the statistical problems resulting from mixing micro- and macro-level data nor to dismiss any statistical effort that augments micro- with macro-level data. A better approach is to examine the nature of the statistical problems involved and to determine what procedures might be employed to lessen or eliminate the problems. We pursue this approach in Kamlet and Klepper (1985). In the following discussion we summarize some of the results of that analysis.

3. A Model of Mixing Macro- and Micro-Level Data

To understand more clearly the nature of the statistical problems that arise when micro- and macro-level data are mixed, consider the health-production or dose-response function:

$$(1) \qquad y_i = \xi_1 x^*_{ij} + \underline{\xi}_2' \underline{z}_i + \epsilon_i ,$$

where i refers to individual i, y_i is some continuous health-status measure, $\underline{z}_i$ is a (K-1)x1 vector of explanatory variables, x^*_{ij} is some other true explanatory variable (say exposure to a given pollutant), ξ_1 is a scalar coefficient, $\underline{\xi}_2 \equiv (\xi_2, \xi_3, \ldots, \xi_K)'$ is a (K-1)x1 vector of coefficients, ϵ_i is a classical disturbance term and the j subscript is a "superfluous" subscript indicating the group (say occupation) of which individual i is a member.

Suppose interest focuses on the coefficient of the pollution variable, ξ_1. Statistical difficulties arise if instead of observing the pollution exposure for each individual i in group j, only the expected or average exposure for all individuals in group j is observed. For example, suppose, as in our research, that one knows only the average exposure of all individuals in occupation j, j=1,2,...,J, but not the specific exposure of each individual in each occupation j. Letting $x_{ij} = x_j = E(x^*_{ij} \mid j)$ represent the average or expected exposure of individuals in group j, x_{ij} can be related to x^*_{ij} as

$$(2) \qquad x^*_{ij} = x_{ij} + w_{ij},$$

where w_{ij} represents the difference between individual i's exposure and the mean exposure of individuals in group j.

As we noted earlier, one approach to dealing with this problem is to ignore it. Observations on x_{ij} are substituted for x^*_{ij} and y_i is regressed on x_{ij} and $\underline{z}_i$. To analyze the implications of this approach, we can use (2) to substitute for x^*_{ij} in (1), which yields

$$(3) \qquad y_i = \xi_1 x_{ij} + \underline{\xi}_2' \underline{z}_i + \epsilon_i + \xi_1 w_{ij}$$

The regression of y on x and $\underline{z}$ will (consistently) estimate ξ_1 and $\underline{\xi}_2$ if and only if the composite disturbance term $(\epsilon + \xi_1 w)$ is uncorrelated with x and $\underline{z}$. Since ϵ is uncorrelated with all variables in the model by construction, the regression of y on x and $\underline{z}$ will yield consistent coefficient estimates as long as w is uncorrelated with both x and $\underline{z}$.

Consider first the correlation between x and w. It can be shown that these two variables are uncorrelated. Intuitively, knowing x_{ij} (e.g., the mean pollution exposure for individuals in occupation j) does not indicate anything about whether individual i's exposure is above or below his group mean. Conversely, knowing w_{ij} is not informative about the group to which individual i belongs, hence it is not informative about x_{ij}. Note that this result is in sharp contrast to the classical measurement error case. When the measurement error in a variable is classical, the measured variable can be related to its true counterpart by

$$(4) \qquad v_i = v^*_i + \mu_i,$$

where v^*_i is the true variable, v_i is the measured variable, and μ_i is a white-noise error term. In this case, if v_i is substituted in a regression in place of v^*_i, the resulting composite error term is correlated with v_i, which in turn leads to inconsistent coefficient estimates.

While x and w are uncorrelated, it turns out that $\underline{z}$ and w are correlated, which means that the estimates from the regression of y on x and $\underline{z}$ will not be consistent. To see how this correlation arises, consider an observation i for which w_{ij} is positive--i.e., individual i's pollution exposure is above his group average. This, in turn, implies that the expected value of x^*_{ij} conditional on w_{ij} being positive is greater than the unconditional expected value of x^*_{ij}. Suppose also that x^* and z_k are

positively correlated, where z_k is one of the elements of $\underline{z}$. This implies that the expected value of z_k conditional on x^*_{ij} being positive is greater than its unconditional expectation. But this means that z_k and w will be positively correlated since positive w values imply, on average, greater than average x^* values. In general, the sign of the correlation between any variable z_k and w will be the same as the sign of the correlation between z_k and x^*.

To analyze the nature of the correlation between z and w more precisely, we assume the following condition holds for all the nonpollution variables in the regression:

$$E(z_k \mid j) - E(z_k) = \frac{\text{cov}(x^*, z_k)}{\text{var}(x^*)} [E(x^* \mid j) - E(x^*)].$$

This assumption will be fulfilled for z_k whenever x^* and z_k covary in the same way within each group j as they do across groups. Writing z_k as

$$z_k = \alpha + \beta x^* + \mu,$$

where μ is a classical disturbance, it can be shown that this assumption will hold as long as $E(\mu \mid j) = 0$ for all j. This will be satisfied whenever the conditional expectation of z_k given x^* is linear in x^*. A sufficient, but not necessary, condition for this to hold is that x^* and z_k are jointly normally distributed.

Under the above assumption, it can be shown that the correlation between w and z_k can be expressed as

$$\rho(w, z_k) = (1-v)\, \rho(x^*, z_k), \tag{5}$$

where

$$v \equiv \frac{\text{var}(x_{ij})}{\text{var}(x^*_{ij})}. \tag{6}$$

Equations (5) and (6) indicate that the correlation between w and z_k depends on the correlation between x^* and z_k and the variance of x relative to the variance of x^*. If in fact x^* were uncorrelated with each of the z_k then w and $\underline{z}$ would be uncorrelated and the regression of y on x and $\underline{z}$ would yield consistent estimates of ξ_1 and $\underline{\xi}_2$. However, the z_k are typically included in the regression specifically because they are thought to affect health and also are likely to be correlated with x^*. In light of this, the substitution of x for x^* will in general cause w and z_k to be correlated, which means that the regression of y on x and $\underline{z}$ will yield inconsistent coefficient estimates.

Note that the correlation between w and z_k also depends on v. When v = 1, the correlation between w and z_k is zero. In this case x and x^* have the same variance, which can occur only when each group contains only one individual--i.e., when x^* is observed. As v gets smaller, the group definitions get coarser, which in turn leads to a greater correlation between x and z_k. In the extreme case in which v is zero, the only group-level data that are available are for the entire sample. Then x takes on the same value for all observations in the sample and the correlation between w and z is maximized.

Define $\underline{\beta}$ to be the vector of coefficients estimated by the regression of y on x and $\underline{z}$. As we noted above, if any of the z_k are correlated with x^* then $\underline{\beta}$ will differ from the vector of coefficients $(\xi_1, \underline{\xi}'_2)'$ defined by (1). It is possible to solve out for $\underline{\beta}$ in terms of ξ_1 and $\underline{\xi}_2$ as follows (see Kamlet and Klepper (1985)):

$$\underline{\beta} \quad \begin{pmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_K \end{pmatrix} = \begin{pmatrix} [(1-\phi)/v]\xi_1 \\ \xi_2 + \phi\theta_2\xi_1 \\ \xi_3 + \phi\theta_3\xi_1 \\ \cdot \\ \cdot \\ \xi_K + \phi\theta_K\xi_1 \end{pmatrix} \tag{7}$$

where

$$\phi = \frac{(1-v)\text{var}(x^*)}{(1-v)\text{var}(x^*) + v\ \text{var}(\pi)}$$

and π and the coefficients θ_i, i=2,3,...,K, are defined by the auxiliary regression of x^* on $\underline{z}$:

$$x^* = \theta_2 z_2 + \theta_3 z_3 + \cdots + \theta_K z_K + \pi.$$

It can be shown that $0<(1-\phi)/v<1$, which implies that the coefficient of x defined by the (population) regression of y on x and $\underline{z}$ will be the same sign but smaller in absolute value than ξ_1. Thus, the regression of y on x and $\underline{z}$ will tend to underestimate the true effect of the pollution variable on health. The coefficients of the other variables will also differ from their true counterparts, and in fact might even be the opposite sign of the true coefficients. Note these results are qualitatively similar to the classical errors-in-variables model (cf. Garber and Klepper (1980)) despite the fact that the measurement error that results from mixing individual and group-level data is decidedly nonclassical.

4. A Consistent Estimator for Mixed-level Data

Although estimates from the regression of y on x and $\underline{z}$ are inconsistent, with sufficient information it is possible to use data on y, x, and $\underline{z}$ to develop consistent estimates of the true coefficients (i.e., the coefficients defined by (1)). To see how this can be done, suppose that x^* could be observed and suppose that a random sample of N observations is drawn on y, x^*, and $\underline{z}$. Let $\underline{y}$ be the Nx1 vector of observations on y, $\underline{x}^*$ be the Nx1 vector of observations on x^*, and Z be the Nx(K-1) matrix of observations on $\underline{z}$. Using $\underline{y}$, $\underline{x}^*$, and Z, the ordinary least-squares estimator of $\underline{\xi} \equiv (\xi_1, \underline{\xi}_2')'$ is defined as

$$\hat{\underline{\xi}} = ([\underline{x}^*, Z]'[\underline{x}^*, Z])^{-1}[\underline{x}^*, Z]'\underline{y}.$$

As the sample size grows large, it is easy to demonstrate that $\hat{\underline{\xi}}$ converges to $\underline{\xi}$:

$$\text{plim}\hat{\underline{\xi}} = V([x^*,\underline{z}'])^{-1}\text{Cov}([x^*,\underline{z}'],y) = \underline{\xi},$$

where $V([x^*,\underline{z}'])$ is the population covariance matrix of the Kx1 vector of variables $(x^*,\underline{z}')'$ and $\text{Cov}([x^*,\underline{z}'],y)$ is the population covariance vector between y and $(x^*,\underline{z}')'$.

Unfortunately, since x^* is not observed, $\hat{\underline{\xi}}$ cannot be computed. Letting $\underline{x}$ be the Nx1 vector of observations on x, the analogous estimator to $\hat{\underline{\xi}}$ that is computable is

$$\hat{\underline{\beta}} = ([\underline{x},Z]'[\underline{x},Z])^{-1}[\underline{x},Z]'\underline{y}.$$

The probability limit of $\hat{\underline{\beta}}$ is

$$\text{plim}\hat{\underline{\beta}} = V([x,\underline{z}'])^{-1}\text{Cov}([x,\underline{z}'],y).$$

However, since $V([x,\underline{z}']) \neq V([x^*,\underline{z}'])$ and $\text{Cov}([x,\underline{z}'],y) \neq \text{Cov}([x^*,\underline{z}'],y)$, $\hat{\underline{\beta}}$ will not (consistently) estimate $\underline{\xi}$.

This problem can be remedied, however, if a consistent estimate of v is available. To see this, consider how $V([x,\underline{z}'])$ and $\text{Cov}([x,\underline{z}'],y)$ are related respectively to $V([x^*,\underline{z}'])$ and $\text{Cov}([x^*,\underline{z}'],y)$. It is easy to demonstrate that

$$(9)\quad V([x^*,\underline{z}']) = \begin{pmatrix} V(x^*) & \text{Cov}(x^*,\underline{z})' \\ \text{Cov}(x^*,\underline{z}) & V(\underline{z}) \end{pmatrix} = \begin{pmatrix} V(x)/v & \text{Cov}(x,\underline{z})'/v \\ \text{Cov}(x,\underline{z})/v & V(\underline{z}) \end{pmatrix}$$

$$(10)\quad \text{Cov}([x^*,\underline{z}'],y) = \begin{pmatrix} \text{Cov}(x^*,y) \\ \text{Cov}(\underline{z},y) \end{pmatrix} = \begin{pmatrix} \text{Cov}(x,y)/v \\ \text{Cov}(\underline{z},y) \end{pmatrix}$$

This suggests using the following estimator for $\underline{\xi}$:

$$\hat{\underline{\xi}}^* = \begin{pmatrix} \underline{x}'\underline{x}/v & \underline{x}'Z/v \\ \underline{x}'Z/v & Z'Z \end{pmatrix}^{-1} \begin{pmatrix} \underline{x}'\underline{y}/v \\ Z'\underline{y} \end{pmatrix}$$

Since $\text{plim}(\underline{x}'\underline{x}/N)=V(x)$, $\text{plim}(\underline{x}'Z/N)=\text{Cov}(x,\underline{z})$, $\text{plim}(\underline{x}'\underline{y}/N)=\text{cov}(x,y)$, $\text{plim}(Z'Z/N)=V(\underline{z})$, and $\text{plim}(Z'\underline{y}/N)=\text{Cov}(\underline{z},y)$, it follows from (9) and (10) that

$$\text{plim}\hat{\underline{\xi}}^* = V([x^*,\underline{z}'])^{-1}\text{Cov}([x^*,\underline{z}'],y)=\underline{\xi}.$$

Thus, given v it is possible to construct a (consistent) estimator of $\underline{\xi}$ using the data on y, x, and $\underline{z}$. Note that all of the above results hold if a consistent estimator is used in place of the population value of v. Thus, all that is needed to construct a (consistent) estimator of $\underline{\xi}$ is a (consistent) estimate of v.

5. The Need to Estimate v

The only practical problem in constructing $\hat{\underline{\xi}}^*$ is in obtaining a (consistent) estimate of v. Surprisingly, for some problems it turns out that v can be estimated from just observations on x. This is the case, for example, whenever x^* is a dichotomous variable. If x^* is coded as 0 or 1, then for each group j, x_j measures the fraction of individuals in the group for which $x^*=1$. Equivalently, $x_j=\text{prob}(x^*_{ij}=1 \mid j)$, where $\text{prob}(x^*_{ij}=1 \mid j)$ denotes the probability that x^*_{ij} assumes the value 1 given that individual i belongs to group j. Since x^*_{ij} is a binomial variable (both within each group j and across all groups), it follows that the variance of x^*_{ij} within group j must equal $\text{prob}(x^*_{ij}=1 \mid j)(1-\text{prob}(x^*_{ij}=1 \mid j))=x_j(1-x_j)$. Consequently, it is possible to estimate the variance of x^*_{ij} within each group j given only observations on x_{ij}. This is all that is needed to estimate v, as v can be expressed as

$$v = \frac{\text{var}(x)}{\text{var}(x^*)} = \frac{\text{var}(x)}{\text{var}(x)+\sum_j x_j(1-x_j)}$$

Even if v cannot be estimated from data on x, it may still be possible to exploit the proposed approach. For example, suppose observations on x^* are not reported because of confidentiality restrictions, but instead the mean value of x^* within various groups is reported. In general, it will not violate confidentiality restrictions also to report the variance of x^* within the same groups. As we saw above, this is all the additional information that is needed to estimate v. Indeed, this suggests that in reporting health statistics it will often be valuable to report within-group variances as well as within-group means.

In other cases estimates of within-group variances may be obtainable from the data set that yields the within-group means. One possibility is to develop such estimates through selective sampling of individuals within each group. While this approach can be expensive, it is obviously less expensive than collecting individual data for the entire sample.

Even when estimates of v can be obtained only through additional data collection, the proposed approach can be used first to determine the value of such an effort. A sensitivity analysis can be performed to see how sensitive the coefficient estimates are to different choices for v. Klepper (1985) demonstrates that such an approach can be used to develop bounds on the coefficient estimates.

6. Conclusion

Given space limitations, we have not discussed a number of important issues regarding the mixing of macro- and micro-level data. These issues are developed further in Kamlet and Klepper (1985). Here we briefly review a few of the more important remaining issues.

First, it is important to realize that mixing micro- and macro-data occurs not just in regression analysis, but in many other contexts as well. For example, the same kinds of problems can surface in logit, probit, and tobit analysis as well as other limited dependent variable models. As for regression models, given a consistent estimate of v, it is possible to develop consistent coefficient estimators for these

models. While such estimators are derived directly from the likelihood function of the data, they are based on the same approach described above for regression models.

Second, while our entire discussion was cast in the context of a model in which only one variable was observable at a group rather than individual level, our approach can be generalized to the case of multiple group variables.

Lastly, we discussed only how a consistent estimator of the true coefficients could be computed. However, in order to perform hypothesis tests, standard errors for the estimators are required. Kamlet and Klepper (1985) demonstrate how asymptotic standard errors can be computed for the proposed estimator and analyze the asymptotic efficiency of the estimator.

References

Frank, R., and Kamlet, M., and Klepper, S. (1985). "The Impact of Occupational Exposure to Toxic Material on Prevalence of Chronic Illness," presented at the 1985 Public Health Conference on Records and Statistics.

Garber, S., and Klepper, S. (1980). "Extending the Classical Normal Errors-in-Variables Model," *Econometrica*, 48 (September): 1541-1546.

Kamlet, M. and Klepper, S. (1985). "Mixing Individual and Group-Level Data," Department of Social Sciences Working Paper, Carnegie-Mellon University.

Klepper, S. (1985). "Applying the Classical Errors-in-Variables Model to Dichotomous and Grouped Variables," Department of Social Sciences Working Paper, Carnegie-Mellon University.

CUMULATIVE LOGIT ANALYSIS FOR SURVEY DATA

James M. Lepkowski and J. Richard Landis, The University of Michigan

1. Introduction

In the analysis of a single sample survey data set, data may include measures reported on nominal, ordinal, discrete numerical, and quantitative measurement scales. Models developed to characterize the relationships among the survey measures must depend, in part, on the type of scale used to record the measurements. In addition, the nature of the distribution of the measure to be predicted determines the appropriate form of the model and the estimation strategy to be used to estimate the parameters in the model.

Consider, for example, a survey in which data are collected on medical care expenditures during a fixed time period, and suppose that a model is to be developed to characterize the relationship between medical care expenditures and several predictors such as age, race, and sex. There are several features of medical care expenditures which pose problems for almost any analytic method to be applied to the data.

For one, a significant proportion of the population will be without expenditures during the period, not because they received services without charge, but because they did not receive any services. In the data set, the charge item for individuals with no use of medical care services will appear as a zero value (or an inapplicable code). A model which characterizes the relationship between charges and other measures collected in the survey will need to account for the limited dependent variable (*i.e.*, limited to those with medical care use).

A second problem with an item such as charges is the skewed nature of the distribution. The distribution of charges will be highly skewed to the right, with some individuals experiencing catastrophic charges relative to the other individuals. Some charges may be so extreme that a logarithmic transformation may not eliminate the skewness.

One approach to model development and estimation for data with a limited dependent variable and a skewed distribution would be to consider multi-part model with various transformations of the data (see, for example, Duan *et al.*, 1982). One part provides a model characterizing the relationships among important predictors and the proportion of persons without charges; a second part provides a model for a transformation of charges among users of medical care. The multi-part model is less satisfying than a single model, though, since it may result in different predictors for each part and will provide a mixed model of proportions and transformed observations.

The limited dependent variable methods proposed by Tobin (1958) and Heckman (1976), prominent in the econometric modelling literature, may also be used for the development of suitable models. The dependent variable (*i.e.*, charges) can be transformed to eliminate the skewness. These limited dependent variable methods are applied to survey data, but the application does not take the complexity of the survey sample design into account.

An alternative approach to these methods is to consider the charge item in terms of an ordinal measure with meaningfully selected cutpoints, one of which can be assigned to the "zero" category. Modelling methods are available for ordinally scaled measures, levels of which can be regarded as coming from some underlying continuous scale which may not satisfy any standard distribution.

For example, McCullagh (1982) discusses the application of proportional odds and proportional hazard models to the analysis of ordinally scaled response variables. Let $\mathbf{X}$ denote a vector of factor or predictor variables, and let $c_j(\mathbf{X})$ denote the ***cumulative probability*** of the first through the *jth* category of the response at factor level $\mathbf{X}$, for $j = 1, \ldots, t$. Let

$$\kappa_j(\mathbf{X}) = c_j(\mathbf{X}) / \left(1 - c_j(\mathbf{X}) \right)$$

denote the odds that the response is less than or equal to the *jth* response category. A proportional odds model for the odds ratio κ_j is

$$\kappa_j(\mathbf{X}) = a_j \exp\left(-\beta'\mathbf{X}\right),$$

where a_j is a constant and β is a vector of unknown parameters.

The proportional odds model is equivalent to a linear logistic model since it specifies a constant difference between the logarithm of the odds ratios $\kappa_j(\mathbf{X})$ for two different levels of the factor variable represented by $\mathbf{X}$. A variety of strategies may be used to obtain estimates for the parameters a_j and β, including maximum likelihood and weighted least squares.

The regression parameter describes how the logarithm of the odds ratios are related to the factor variables in $\mathbf{X}$. The model may be further extended to include factorial arrangements among the factors and location and scale parameters for each row of the table. Thus, a general multivariate regression model approach is available for the analysis of ordinally scaled response variables. These models may be also adapted to the analysis of data with limited dependent variables and skewed distributions.

In the next section, a methodology for developing models to explain variation in cumulative proportions similar to $\kappa_j(\mathbf{X})$, referred to as ***cumulative logit analysis***, is described. The method is extended in section 3 to the analysis of data from a complex sample survey. An application of the methodology to data from the National Medical Care Utilization and Expenditure Survey is given in section 4. Section 5 concludes with a discussion of the limitations of the methodology for survey data and extensions of the methodology to problems. other than those discussed in the paper.

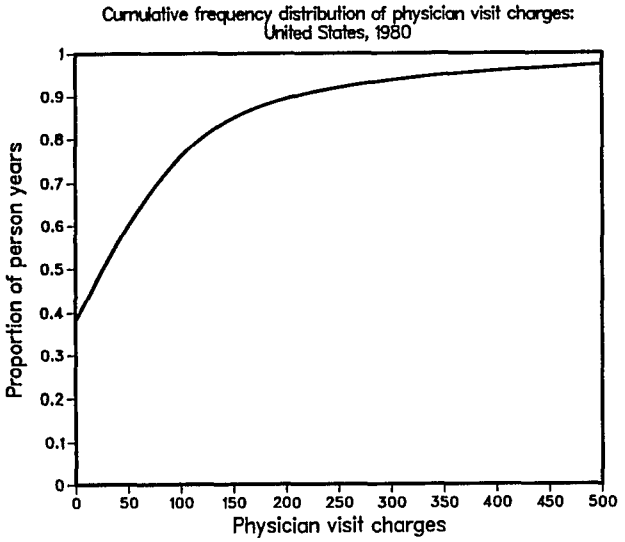
2. Cumulative Logit Analysis

Consider a cumulative frequency distribution for a characteristic such as charges for physician visits during a one year period as illustrated in Figure 1. A model is to be developed which summarizes the relationship between physician visit charges and various predictor variables or factors. As observed previously, the distribution for charges for this type of medical care has a large proportion of the population with no charges (approximately 40 percent) and a skew distribution characterized by the slow approach of the cumulative distribution to the limiting cumulative frequency of 1.0.

For ordinally scaled types of response data, the cumulative logit methodology uses the categories of the response variable to create a series of cutpoints or "thresholds". The term threshold is borrowed from the dose-response framework for analyzing biological data and used in this discussion to denote values of a response variable for which there is a substantive reason to expect a critical change in the reponse. For discrete numerical or quantitative measurements, thresholds must be created as clearly specified intervals on the measurement scale.

Suppose that t threshold values are to be selected which divide the distribution of physician visit charges into $t + 1$ groups. The number of thresholds t can be chosen to characterize the distribution of the measure sufficiently, but not so many values that the relative frequency of the

Figure 1



distribution within an interval defined by the thresholds becomes small. A reasonable number of thresholds is typically from three to six.

Preferably, the threshold values should be chosen according to substantive types of criteria, rather than statistical considerations alone. For example, for physician visits charges, the modal typical charge for a single visit may be known, and threshold values selected according to the number of visits at the typical value. One visit may represent a physician visit by a healthy individual for a physical examination or for a minor, but temporarily debilitating, acute illness episode. Two visits may represent an initial visit for an illness plus a follow-up visit. Three to four visits may represent more serious types of illness, while five or more visits represent persons with poorer health. The charges corresponding to a fixed number of visits which distinguish differences in the health status of the population can thus be chosen as threshold values.

Suppose that, in addition to being classified with respect to selected thresholds, each element of the population is further classified into s subpopulations based on categories of a predictor or set of predictor variables. Categories of the predictor variables may be defined for nominally, ordinally, or quantitatively scaled measures, and the subpopulations may be formed from a single predictor or cross-classification of a set of predictor categories.

Suppose that a simple random sample of size n is selected from the population of interest. Let n_{ij} denote the number of sample elements classified into the i*th* subpopulation and have values of the dependent variable which is less than or equal to the j*th* threshold value. Let the number of observations in the i*th* subpopulation be denoted as $n_{i\cdot}$, and let $c_{ij} = n_{ij} / n_{i\cdot}$ denote the cumulative proportion of sample elements in the i*th* subpopulation that have a value of the dependent variable that is less than or equal to the j*th* threshold. The logit of the cumulative probability c_{ij} is

$$\lambda_{ij} = \log_e \left[\frac{c_{ij}}{1 - c_{ij}} \right].$$

Consider a linear model for the logits of the cumulative probabilities which summarizes variation in the logits in terms of a series of line segments from the first to the last logit for each subpopulation. In particular,

$$\lambda_{ij} = \mu + \alpha_i + \sigma_j + \tau_{ij} + \epsilon$$

where μ denotes an overall mean cumulative logit, α_i denotes an effect of the i*th* subpopulation, σ_j denotes a common or mean slope as an increment from the $(j - 1)$*th* to the j*th* logit, and τ_{ij} denotes an effect of the i*th* subpopulation on the j*th* slope.

For this model, there are a variety of hypotheses that correspond to those of interest in the analysis of covariance or profile analysis: (a) Are there differences in "intercepts" across the subpopulations (*i.e.*, are the α_i nonzero)? (b) Are the line segments parallel across subpopulations (*i.e.*, are the τ_{ij} equal to zero)? (c) Are the line segments coincident across subpopulations (*i.e.*, are the α_i and the τ_{ij} all simultaneously equal to zero)?

These hypotheses may also be interpreted in terms of the subpopulation cumulative frequency distributions themselves. If there is no difference among the intercepts across subpopulations, the cumulative distributions have the same proportion with a "zero" value for the response variable. If the cumulative logit line segments or curves are parallel, the relative frequency distributions are similar in shape across the subpopulations (*i.e.*, they have similar amounts of dispersion and skewness once differences in the "zero" category are accounted for).

To test hypotheses about the cumulative logit model, a weighted least squares methodology can be applied. Consider a vector of the $s \times t$ cumulative logits $\mathbf{F}' = [\ \lambda_{11}, \ldots, \lambda_{1t}, \lambda_{21}, \ldots, \lambda_{2t}, \ldots, \lambda_{s1}, \ldots, \lambda_{st}\ ]$. Following Grizzle, Starmer, and Koch (1969), the linear model $E_A\{\mathbf{F}\} = \mathbf{X}\,\mathbf{B}$ is to be fit to the vector of cumulative logits $\mathbf{F}$, where $E_A\{\cdot\}$ denotes asymptotic expected value of the argument $\{\cdot\}$, $\mathbf{X}$ denotes an $st \times g$ matrix of constants, and $\mathbf{B}$ denotes a $g \times 1$ vector of parameters.

The parameter vector $\mathbf{B}$ is estimated using the weighted least squares estimator $\mathbf{b} = (\mathbf{X}'\,\mathbf{V}_\mathbf{F}^{-1}\,\mathbf{X})^{-1}\,\mathbf{X}'\,\mathbf{V}_\mathbf{F}^{-1}\,\mathbf{F}$ where $\mathbf{V}_\mathbf{F}$ denotes the estimated variance-covariance matrix for $\mathbf{F}$. $\mathbf{F}$ is computed as a series of linear and logarithmic transformations to a vector of cumulative proportions c_{ij}. Similarly, an estimate of $\mathbf{V}_\mathbf{F}$ is obtained by the application of a series of transformations to the variance-covariance matrix of the c_{ij}. The estimation of $\mathbf{V}_\mathbf{F}$ requires the use of Taylor series approximation methods which can be conveniently summarized in matrix operations (see, for example, Landis *et al.*, 1976).

The goodness of fit of the model $\mathbf{XB}$ to the observed vector of cumulative logits can be tested using the Wald statistic $Q = (\mathbf{F} - \mathbf{X}\,\mathbf{b})'\,\mathbf{V}_\mathbf{F}^{-1}\,(\mathbf{F} - \mathbf{X}\,\mathbf{b})$, which, under the hypothesis that the model fits the data adequately, has an asymptotic chi-square distribution with $(st - g)$ degrees of freedom (Wald, 1943).

In order to test hypotheses about the parameters in the model, hypotheses of the form $H_0 : \mathbf{C}\,\mathbf{B} = 0$, where contrast matrix $\mathbf{C}$ is a $(d \times g)$ matrix of constants, may be tested using the test statistic $Q_C = (\mathbf{C}\,\mathbf{b})'[\mathbf{C}\,(\mathbf{X}'\,\mathbf{V}_\mathbf{F}^{-1}\,\mathbf{X})^{-1}\,\mathbf{C}']^{-1}\,\mathbf{C}\,\mathbf{b}$. Under the null hypothesis $\mathbf{CB} = 0$, the test statistic Q_C is asymptotically distributed as a chi-square random variable with d degrees of freedom.

The hypothesis test may indicate that a reduced model with fewer or alternative sets of parameters will explain the variation in the observed cumulative logits adequately. The vector of reduced parameters may be estimated using the

weighted least squares estimation procedure and predicted values for the cumulative logits obtained as $\hat{F} = X_R b_R$ where X_R is the reduced model matrix, and b_R is the estimated reduced parameter vector.

The methods outlined here can be implemented using weighted least squares estimation and hypothesis testing software such as GENCAT (Landis *et al.*, 1976) or the CATMOD procedure within the Statistical Analysis System (SAS Institute, 1985). The nature of the matrix X and parameter vector **B**, as well as the contrast matrix C for cumulative logit analysis, are illustrated subsequently for an analysis from a complex sample survey.

3. Cumulative Logit Analysis for Sample Survey Data

The development in the preceding section was based on the assumption that the sample selection was a simple random one. For sample surveys, the selection procedure is considerably more complex involving stratification of sampling units, multiple stages of selection, weighting, and other design features. The assumptions of simple random selection may be far from appropriate when inferences about the finite population from which the original sample were selected (or a population similar to that finite population) are of interest.

In order to incorporate the complexity of a stratified multistage probability sample design with weighted observations into the analytic methods, the estimation procedures used to obtain the original set of cumulative proportions c_{ij} and their variances and covariances must account for the sample survey design. Suppose that a sample of a_h primary sampling units is selected from the h*th* stratum, where $h = 1, \ldots, H$. Let n_{ha} denote the number of sample elements selected within the (ha)*th* selected primary sampling unit, and let w_{hak} be a weight assigned to the k*th* sample element within the (ha)*th* selected primary sampling unit to account for unequal probabilities of selection and for nonresponse and other nonsampling errors.

Consider the indicator variable

$$y_{ijhak} = \begin{cases} w_{hak}, & \text{if the } (hak)th \text{ sample element} \\ & \text{is in the } ith \text{ subpopulation and } jth \\ & \text{threshold group,} \\ 0, & \text{otherwise.} \end{cases}$$

Weighted estimates of the sample proportions c_{ij} can be obtained as

$$c_{ij} = \sum_h \sum_a \sum_k y_{ijhak} / n_{i\cdot}$$
$$= n_{ij} / n_{i\cdot},$$

where $n_{i\cdot}$ denotes the sum of the indicator variable for all sample persons in the i*th* subpopulation.

Each of these estimated cumulative proportions is a ratio of two random variables, since the denominator, $n_{i\cdot}$, is not a fixed quantity in the design, but rather a random variable. Estimation of the variance of c_{ij} and covariances among these estimated sample proportions is typically accomplished through the use of Taylor series approximations. For example, the variance of c_{ij} is estimated as

$$\mathrm{var}(c_{ij}) = (n_{i\cdot})^{-2} \left[\mathrm{var}(n_{ij}) + (c_{ij})^2 \mathrm{var}(n_{i\cdot}) - 2\, c_{ij}\, \mathrm{cov}(n_{ij}, n_{i\cdot}) \right]$$

where $\mathrm{var}(n_{ij})$, $\mathrm{var}(n_{i\cdot})$, and $\mathrm{cov}(n_{ij}, n_{i\cdot})$ are the respective estimated variances and covariance.

These latter variances and covariances are estimated by taking the stratified multistage sample design into account. For example, suppose that the first stage sample selection was made with replacement of primary sampling units (or, if without replacement, the number of primary sampling units within each stratum is large enough that the distinction between with and without replacement selection is small enough to be safely ignored) and that $a_h = 2$ for all H strata (*i.e.*, a paired selection of primary sampling units was employed). Then the variances may be estimated as

$$\mathrm{var}(n_{ij}) = \sum_h (n_{ijh1} - n_{ijh2})^2$$
$$\mathrm{var}(n_{i\cdot}) = \sum_h (n_{i\cdot h1} - n_{i\cdot h2})^2,$$

where n_{ijh1} and n_{ijh2} denote the sum of the indicator variable y_{ijhak} within the $(h1)$*th* and $(h2)$*th* selected primary sampling units, respectively, and $n_{i\cdot h1}$, and $n_{i\cdot h2}$ denote estimated subpopulation sizes in the $(h1)$*th* and $(h2)$*th* primary sampling units, respectively.

This estimation procedure for stratified multistage sample survey data is implemented in statistical software such as the PSALMS program within the OSIRIS IV Statistical Software System (Computer Support Group, 1981) or the SESUDAAN package of programs which operate under the SAS system (Shah, 1984). A more detailed presentation of these estimation procedures in the weighted least squares analysis framework outlined in the previous section is given in Landis *et al.* (1982).

4. An Illustration from a Complex Sample Survey

The National Medical Care Utilization and Expenditure Survey (NMCUES) was designed to provide data about use of and charges for medical care by the U. S. civilian noninstitutionalized population during 1980. From interviews with a panel of 17,123 persons, information was collected on health, access to and use of medical services, associated charges and sources of payment, and health insurance coverage. A complete description of the purposes and procedures of the NMCUES is available in Bonham (1983).

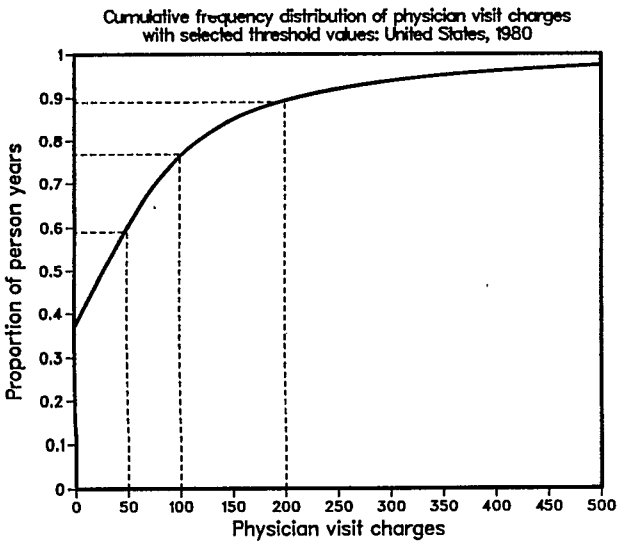
The NMCUES sample design employed a stratified multistage probability selection procedure and a weighting procedure designed to adjust estimates for unequal probabilities of selection, nonresponse, and noncoverage of the population. The complexity of the survey design requires that an analyst be familiar with a range of design features to determine an appropriate analytic methodology for estimation and inference from the survey data. A methodology similar to that described in section 3 may be used to estimate the cumulative proportions and their variances and covariances needed in the cumulative logit methodology.

Consider the problem of developing a suitable model to explain the variation of observed cumulative distributions for physician visit charges as they vary across subgroups of the population defined by age groups. It is of interest to know whether the proportion of persons with no physician visits differ across age groups as well as whether the distribution of charges differ across age groups.

A set of four thresholds were chosen (see Figure 2), yielding an ordinal response variable with five levels. One threshold, by default, corresponded to the group of persons with no physician visits in 1980. The remaining thresholds were chosen by substantive considerations of the typical charge for a physician visit in 1980 and the number of visits that might indicate differences, on average, in the need for physician care and health status. The threshold choices were also applied to the survey data to determine whether the size of the five groups defined by the four thresholds was sufficient

to provide adequate levels of precision for individual cumulative logits correponding to each threshold group in a subpopulation.

Figure 2



The cumulative proportions corresponding to the four selected thresholds in each of four age groups are shown in Table 1. The proportion of persons in each age group with no visits (*i.e.*, no physician visit charges) declines with increasing age as might be expected. (The "None" threshold group actually contains a small number of persons with physician visits in 1980 but with no reported charges for those visits.) At the same time, the cumulative proportion of persons with charges less than or equal to $200 (*i.e.*, the last threshold value) also declines with age indicating that a larger proportion of persons at older ages will have charges exceeding $200 than persons at younger ages. The cumulative frequency distribution corresponding to Table 1 and shown in Figure 3 confirm these observations about the cumulative proportions.

Table 1

Cumulative weighted proportion of person years for physician visit charge thresholds by age group: United States, 1980.

Age group	Threshold (Physician visit charges)				Sample size
	None	$50	$100	$200	
Under 18	0.372	0.724	0.865	0.956	5389
18–44	0.377	0.644	0.782	0.895	6486
45–64	0.324	0.543	0.693	0.836	3376
65 and older	0.244	0.427	0.577	0.772	1872

Figure 4 presents the cumulative logits for physician visit charge thresholds across the four age groups. The intercepts of these cumulative logit curves correspond to the proportions of the population that are without physician visit charges in 1980; a smaller logit corresponds to a smaller proportion. Persons under 18 and from 18 to 44 years of age have similar logits at the initial threshold; whereas those from 45 to 64 years and 65 or older have smaller logits.

It is difficult to determine by inspection whether the four curves are parallel, a feature that would indicate identical

Figure 3

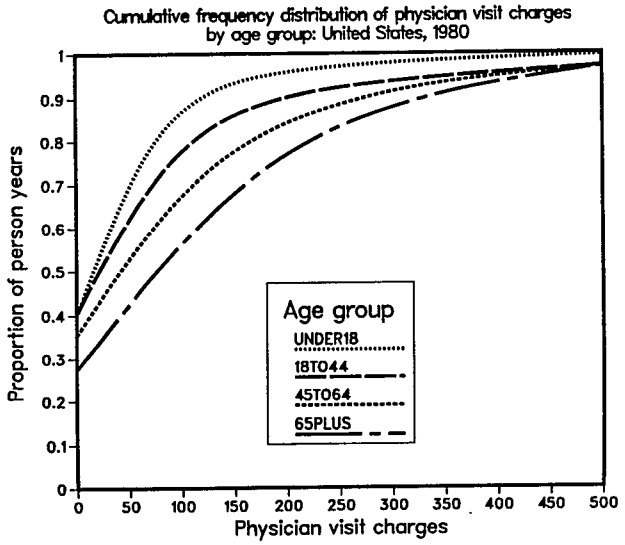
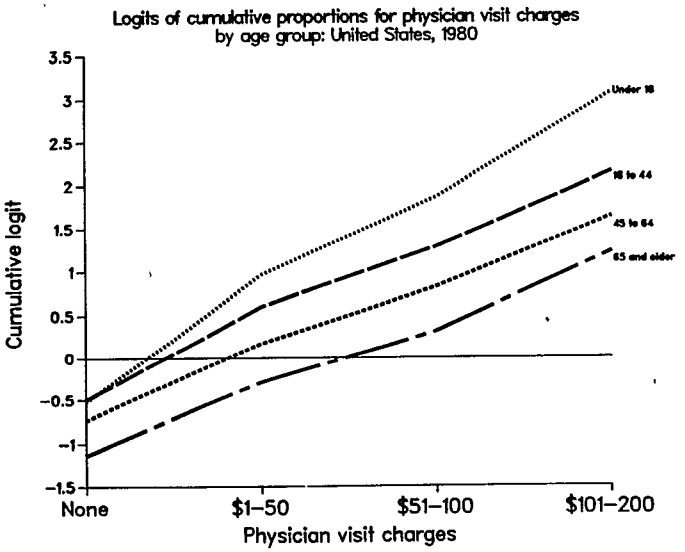


Figure 4



distributions once the initial proportion without charges are accounted for. For each of the three sets of line segments (*i.e.*, from the first to the second, the second to the third, and the third to the fourth) it appears that cumulative proportions for persons under 18 years of age increase more rapidly than the other three age groups (*i.e.*, indicating their charges are lower than the other three age groups). At the same time, the remaining three age groups appear essentially to be parallel across the various line segments. The curves might thus be summarized in terms of a model with equal intercepts for two age groups, equal slopes or increments for each subsequent threshold for three of the age groups, and a larger slope for the under 18 year age group for each segment.

Such a strategy of visual inspection, followed by model fitting, may lead to "overfitting" of models and spurious results. In addition, with the large sample sizes available for each group, seemingly small differences in slopes or intercepts can be statistically significant. The strategy used examined the parameters of a saturated model (*i.e.*, a model with as many parameters as observations) by testing various hypotheses about the intercepts and slopes shown in Figure 4.

Figure 5 presents the saturated model matrix X and parameter vector **B** examined for these cumulative logits. The blocks of four rows of X correspond to the cumulative logits for each subpopulation. The first column represents an overall

Figure 5

Saturated model design matrix X and parameter vector B

$$X = \begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & -1 & -1 & -1 & 1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \\
1 & -1 & -1 & -1 & 1 & 1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\
1 & -1 & -1 & -1 & 1 & 1 & 1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1
\end{bmatrix}$$

$$B = \begin{bmatrix}
\text{Overall mean} \\
\\
\text{Differential effect: Age } <18 \text{ vs. } 65+ \\
\text{Differential effect: Age } 18-44 \text{ vs. } 65+ \\
\text{Differential effect: Age } 45-64 \text{ vs. } 65+ \\
\\
\text{Increment for slope 1} \\
\text{Increment for slope 2} \\
\text{Increment for slope 3} \\
\\
\text{Increment effect for slope 1: Age } <18 \text{ vs. } 65+ \\
\text{Increment effect for slope 1: Age } 18-44 \text{ vs. } 65+ \\
\text{Increment effect for slope 1: Age } 45-64 \text{ vs. } 65+ \\
\\
\text{Increment effect for slope 2: Age } <18 \text{ vs. } 65+ \\
\text{Increment effect for slope 2: Age } 18-44 \text{ vs. } 65+ \\
\text{Increment effect for slope 2: Age } 45-64 \text{ vs. } 65+ \\
\\
\text{Increment effect for slope 3: Age } <18 \text{ vs. } 65+ \\
\text{Increment effect for slope 3: Age } 18-44 \text{ vs. } 65+ \\
\text{Increment effect for slope 3: Age } 45-64 \text{ vs. } 65+
\end{bmatrix}$$

mean cumulative logit, and the next three columns represent departures from that overall mean for the first three age groups.

The next three columns of X correspond to the mean slope for each of the three line segments between successive thresholds in Figure 4. The three remaining blocks of three columns each represent departure by each of the first three age groups from that overall mean slope for a given line segment. Thus, the first four columns of the matrix X (and the first four parameters of B) correspond to the intercepts of the cumulative logit curves, while the remaining columns of X (and remaining parameters in B) concern the successive slopes of the cumulative logit curves, represented as "increments" to the preceding threshold cumulative logits, and departures in the slope from the mean for each age group.

The saturated model in Figure 5 will, of course, fit the observed set of values perfectly. Nonetheless, hypotheses concerning the parameters in the model can be tested to determine whether any model reduction is possible by using the quadratic form Q_C. For example, the contrast matrix

$$C_A = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

can be used to test the hypothesis that there is no difference in intercepts across the four age groups. The results of several hypothesis test applied to the saturated model parameters is summarized in Table 2 for two sets of assumptions about the sample selection procedure.

The two design options correspond to simple random selection with weighted data (option 2) and the stratified multistage selection appropriate for the NMCUES data (option 3). Option 1 which is not shown corresponds to simple random selection with unweighted data (see Landis *et al.*, 1982, for details). The comparison of options 2 and 3 allows an assessment of the effect of the sample design on inferences about the cumulative distributions. The model development is more properly conducted using option 3.

The test for age corresponds to the hypothesis of equal intercepts for the four curves. Since Q_C is highly significant under option 3, the conclusion is that the intercepts do differ. The hypothesis concerning a "flat response" corresponds to a test that all 12 slope parameters are equal to zero (*i.e.*, the curves are horizontal); Q_C is highly significant again, as might be expected.

The parallelism hypothesis tests whether the four curves can be explained by unequal intercepts and nonzero slopes but without differing slopes for each age group. The test statistic Q_C is again quite significant. The last three hypotheses are the components of the parallelism hypothesis for each of the three successive line segments between thresholds. Each is significant indicating that the lack of parallelism is receiving a contribution from each of the three line segments.

Given that the results obtained under option 3 are so highly significant, it is perhaps not surprising that similar results are obtained under the option 2 assumptions. The ratio of Q_C values for each option shown in the last column of Table 2 indicate that the option 2 inference is somewhat "anti-conservative" compared to the more appropriate option 3 inference. These results are consistent with results for design effects in the sample survey literature and have been investigated for simpler single degree of freedom hypothesis tests for the weighted least squares methodology applied to survey data (Lepkowski and Landis, 1985).

Further hypothesis testing may be conducted to determine whether a more parsimonious model may fit the data adequately. Given the large sample sizes in this example, most of the parameters in the saturated model are highly significant and model reduction is limited. However, because the sums of squares in Table 2 are large, an alternative model development strategy similar to those applied in a standard analysis of variance can be considered.

The total sum of squares corrected for the mean under option 3 is shown in the last line of Table 2. Under the hypothesis of parallelism, approximately (270.407 / 9,289.832) × 100 = 2.9 percent of the total variation of cumulative logits is unexplained. In other words, 100 − 2.9 = 97.1 percent of the variation is explained by the parallelism model. Thus, from the perspective of explained variation alone, one would be satisfied with a model that allowed differing intercepts but parallel curves for the cumulative logits. And hence, one could conclude that the four age-specific distributions have different proportions of persons with zero charges but nearly identical relative and cumulative frequency distributions for persons with charges.

5. Discussion

The cumulative logit methodology provides a means to compare ordinal frequency distributions across subgroups of a population. The comparisons are not limited to contrasting a single measure of central tendency (*e.g.*, means, medians) and provide an opportunity to examine variation in the entire distribution of the response variable. The methodology provides a single model for the limited dependent variable

Table 2

Analysis of variance for cumulative logit analysis under two sample design analysis options for physician visit charges by age group: United States, 1980.

Source	df	Option 2 (Simple random sampling, weighted)		Option 3 (Complex sampling)		$\frac{Q_{C(2)}}{Q_{C(3)}}$
		$Q_{C(2)}$	Significance level	$Q_{C(3)}$	Significance level	
Age	3	133.869	<0.01	107.165	<0.01	1.249
Flat response	12	12388.172	<0.01	7672.638	<0.01	1.615
Parallelism	9	366.475	<0.01	270.407	<0.01	1.355
Slope 1	3	241.277	<0.01	182.998	<0.01	1.320
Slope 2	3	288.232	<0.01	220.642	<0.01	1.306
Slope 3	3	248.804	<0.01	175.231	<0.01	1.420
Total	15	12421.129	<0.01	9289.832	<0.01	1.337

problem and can be applied to measures that have nonstandard distributions and require transformations in other analytic methodologies. Reduced models with fewer parameters can be developed within the cumulative logit methodology, and predicted values for the cumulative logits (and the original cumulative probabilities) can be estimated under the reduced models. Since the methodology relies on the weighted least squares estimation and hypothesis testing approach for categorical data, it can be applied to data from complex sample survey designs taking into account the design features into the analysis.

The estimation of the variance-covariance matrix for the cumulative logits is considerably more complicated and expensive than that required under simple random sampling assumptions. In addition, the degrees of freedom utilized in the estimation process limits the number of parameters that can be specified in the linear model. For models in which the number of parameters approaches the number of degrees of freedom in the variance estimation procedure, users of the method may want to consider using an F distribution rather than a chi-square distribution for determining the significance of the Wald statistic Q and the quadratic form Q_C (see Koch and Lemeshow, 1972).

The methodology was illustrated in section 4 with only a single predictor. Several predictors can be considered by cross-classifying the categories of the predictors to form subpopulations. The linear models can then examine the nature of main and interaction effects among the predictors for intercepts as well as for slopes across the subpopulations formed by the cross-classification. Elements of the model matrix X can be chosen for ordinally scaled predictors with more than two levels to allow the investigation of linear, quadratic, and higher order effects as well.

Finally, selection of suitable threshold values was described as primarily a substantive rather than a statistical process. The effect of alternative threshold values on the development of a final model are not investigated here, but given the cumulative frequency distributions examined, it is unlikely that other choices for threshold values would have much effect on the model development. For other situations in which the cumulative frequency distributions intersect, the choice of thresholds will have an effect on the subsequent model; substantive considerations will be important in such situations to determine appropriate values for the thresholds.

References

Duan, N., Manning, W. G., Jr., Morris, C. N., and Newhouse, J. P. (1982). *A Comparison of Alternative Models for the Demand for Care.* Santa Monica, California: The Rand Corporation.

Tobin, James (1958). Estimation of relationships for limited dependent variables. *Econometrica*, **26**, 24–36.

Heckman, James J. (1976) The common structure of statistical models of truncation, sample selection and limited dependent variables, and a simple estimator for such models. *Annals of Economic and Social Measurement*, **5/4**, 475–492.

McCullagh, Peter (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society*, Series B, **42**, No. 2, 109–142.

Grizzle, James E., Starmer, C. Frank, and Koch, Gary G. (1969). Analysis of categorical data by linear models. *Biometrics*, **25**, 489–503.

Landis, J. Richard, Stanish, William M., Freeman, Jean L., and Koch, Gary G. (1976). A computer program for the generalized chi-square analysis of categorical data using weighted least squares (GENCAT). *Computer Programs in Biomedicine*, **6**, 196–231.

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, **54**, 426–482.

SAS Institute Inc. (1985). *SAS User's Guide: Statistics*, 1985 edition. Cary, North Carolina: SAS Institute Inc.

Computer Support Group, Institute for Social Research (1981). *OSIRIS IV User's Guide.* Ann Arbor, Michigan: Institute for Social Research, The University of Michigan.

Shah, B. V. (1984). Software for survey data analysis. *The American Statistician*, **38**, 68.

Landis, J. Richard, Lepkowski, James M., Eklund, Stephen A., and Stehouwer, Sharon A. (1982). A statistical methodology for analysing data from a complex survey: The first National Health and Nutrition Examination Survey. *Vital and Health Statistics*, Series 2, No. 92. DHHS Publication No. 82–1366. Public Health Service. Washington: U. S. Government Printing Office.

Bonham, Gordon S. (1983). Procedures and questionnaires of the National Medical Care Utilization and Expenditure Survey. *National Medical Care Utilization and Expenditure Survey*, Series A, Methdological Report No. 1. DHHS Publication No. 83-20001. Public Health Service. Washington: U. S. Government Printing Office.

Lepkowski, James M. and J. Richard Landis (1985). Design effects for contrasts of subclass proportions. (Submitted for publication.)

Koch, Gary G. and Lemeshow, S. (1972). An application of multivariate analysis to complex sample survey data. *Journal of the American Statistical Association*, **67**, 780–782.

THE EFFECTS OF SKEWNESS, KURTOSIS AND INEQUALITY OF VARIANCE ON PEARSON'S r AND ON MODELS BASED ON PEARSON'S r

John M. Mc Quiston, University of Montana

The Pearson Product-Moment correlation (r) has become the foundation of mathematical models where ratio scales are available and where association, description or prediction is a goal. Many factorial and path models are based on the knowledge of r. The use of Pearson is rather straightforward, one merely needs two or more variables which are ratio scales or which are interval and can be justified as acting as ratio scales. This is the common misconception.

If we examine the formula for Pearson, r_{xy}=Sum xy /Sqrt(Sum x^2 * y^2), we can see that the computations are based on the differences that exist among individuals as they are compared to one-another with respect to their deviations from a Mean, and with variation from the Mean universally for those variables. This formula then presupposes bivariate distributions for which the Mean and Standard Deviations are appropriate evaluative statistics, random Normal ones. The formula also presumes that the variances within different distributions are homogeneous, that homoscedasticity is present(1). These parameters are straightforward, actually rather innocuous. We often disregard all but the scaling parameter. Why is this? There is considerable folklore as to the effects of the violation of the other parameters of r on research results. Possibly the correlation will be smaller, lending conservative results. Perhaps error in prediction will be slightly greater. Regardless of rationale, we will tend to ignore the distributional and variance parameters(2) in favor of an easier analytical procedure and merely compute r.

This paper illustrates and examines some of the statistical and substantive outcomes that one might expect where the Pearson Product-Moment Correlation is computed without regard to distributional effects or homoscedasticity.

Methodology

A Normally distributed random distribution consisting of 2500 numbers was obtained(3) and tested with respect to skewness (Sk = .007) and kurtosis (Ku = -.088). As noted here, both approached zero, the ideal. The Mean was 0.00, and the Standard Deviation 1.00. As the goal was to test distributional effects on r, we then multiplied the numbers by 100.0 so as to produce a Normally distributed variable with a Mean of 100.0 which would contain only positive numbers. So as to obtain a relatively larger Standard Deviation, while preserving normalcy, we then set the Standard Deviation at about ten percent of the Mean (9.8) and again tested for Skewness and Kurtosis (Sk = .007, Ku = -.088). We then created various distributions from this random one by simply multiplying the randomly distributed variable by various formulii (e.g. X^2 or Sqrt X) to produce various distributions which we knew were not Normal in character and that had various Means and Standard Deviations which were different from the original one. These distributions were then correlated with one-another (4) and with the Normal one so as to produce a matrix of r's that would show the results of skewness and kurtosis on the Pearson Product-Moment correlation.

To show the effects of inequality of variance, the Standard Deviation of the Normal distribution (Mean = 100.0) was then set at different levels starting at five percent of the Mean, and ending at thirty-five percent(5). We stopped at thirty-five percent as a small number of cases received loadings that were less than zero at that point, which would render exponential transformation useless to this analysis (wraparound). Each of these resulting distributions was then transformed so that the same non-Normal list of variables was produced as in the first instance, only having different variances than before. Again, correlation matrices were produced so that the effects of variances of different magnitudes with respect to the original Mean could be compared with one-another. For instance, what is the difference between r's where a Normal distribution with a Standard Deviation of 10% of its Mean is correlated with itself squared, and where a Normal distribution having a Standard Deviation of 20% of its Mean is correlated with itself squared? Note here that the relative position of individual cases was unchanged in all distributions regardless of transformation, so if no effects were present, Pearson r's should be +1.000 for each correlated pair. Any differences between obtained r's and +1.0 would show the effects of the differences in distributions on r.

It should be noted here that when we refer to a Standard Deviation relative to the Mean we are describing the Standard Deviation of the Normal distribution from which transformed distributions were derived. Once the transformations were completed and the various new distributions were prepared, their Standard Deviations varied from that of the Normal distribution from which they were derived. The resultant correlations are parent distribution specific. They can only be related to other distributions having the same parent. Thus, we can delineate the amount of loss in efficiency and/or

TABLE 1
SIMPLE PEARSON CORRELATION COEFFICIENTS BETWEEN
THE NORMAL DISTRIBUTION AND VARIOUS OTHER DISTRIBUTIONS (N=2500)

DISTRIBUTION	Standard Deviation As A % of the Mean			
	5%	10%	20%	30%
Normal (X)(1)	1.000	1.000	1.000	1.000
X^{-5}	-.979	-.905	-.431	-.068(2)
$X^{-4.5}$	-.982	-.921	-.503	-.068(2)
X^{-4}	-.985	-.935	-.582	-.068(2)
$X^{-3.5}$	-.988	-.948	-.663	-.068(2)
X^{-3}	-.991	-.959	-.740	-.068(2)
$X^{-2.5}$	-.993	-.969	-.809	-.068(2)
X^{-2}	-.995	-.978	-.868	-.068(2)
$X^{-1.5}$	-.996	-.985	-.914	-.069
$X^{1.5}$	1.00	.999	.998	.994
X^{2}	.999	.998	.991	.980
$X^{2.5}$	.999	.995	.981	.960
X^{3}	.998	.991	.967	.935
$X^{3.5}$	.996	.986	.951	.906
X^{4}	.995	.980	.932	.874
$X^{4.5}$	.993	.974	.911	.840
X^{5}	.991	.966	.888	.804
$X^{1/2}$	1.00	.999	.997	.985
$X^{1/3}$	1.00	.999	.995	.981
$X^{1/4}$	1.00	.999	.994	.973
$X^{1/5}$	1.00	.998	.993	.968
LOG X	.999	.998	.989	.912

(1) "X" here refers to the normal distribution and its transformations. The random normal distribution used had a Shewness of .006, Kurtosis of -.080, and a Mean of 100.0 regardless of its standard deviation.

(2) The distribution here rendered the results a function of computer rounding, not of actual association.

information that we are likely to experience, given the Standard Deviation of the distribution had we not transformed it to a Normal one.

Statistical Findings

The results of the computation of correlation coefficients between the Normal distributions and twenty-one transformations of those distributions are presented in Table 1. Beginning with column one (5%), note that the values of the correlation coefficients are changed only slightly, varying from 1.0 to .979 in absolute magnitude. Note the negative correlations. These are a result of raising the Normal distribution to a negative power, the equivalent of $1.0/X^e$ where e is an exponent and X is the Normal distribution. This transformation effectively exactly reverses the ordering of the original data and reduces the values to very small fractions. Once transformed, perfect association with the Normal distribution would yield a perfect negative correlation (-1.0).

Turning to the analysis of the various transformations where the Standard Deviations were set at 10% (Table 1, Column 2) we can see that the associations are slightly weaker, ranging in magnitude from 1.0 to .905. Although weaker, we still explain 81% of the common variance between the variables, a loss of 19% from that which would have been expected from two variables which have perfect association. Still, when we were to expect 100.0% of the variance among the variables to have been explained through the use of r, 19% is an important loss. In this instance the greatest magnitude of loss of explained variance comes between the Normal distribution and a distribution which would certainly be uncommon among research ones. For X^{-5} the loss which occurred between a Standard Deviation of 5% and that of 10% was .074, one-half of one percent variance loss. We can see that all of the negative function correlations are beginning their rapid decay. If we continue to columns three and four, where the Standard Deviations were 20% and 30% respectively, we note a continuation of loss of variance accounted for regardless of the magnitude of skewness in the distribution. By the time we reach 30%, all of the negative exponential correlations approach zero and are constant. This is partly due to the distributions themselves but probably more importantly due to algorithm and rounding error within the computing process(6).

Note in Table 1 that the magnitude of difference between +1.0 and the correlation obtained is not symmetrical. That is, distributions having the same Skewness, positive and negative, result in correlation coefficients of different magnitudes when associated with the Normal distribution. Distributions which are positively skewed result in progressively smaller r's than their negative counterparts (e.g. X^2 vs. $X^{1/2}$), that difference increases as skewness increases(7). Again, the relative positioning of the Means results in large numbers among positively skewed variables which modify Sum xy in a non-linear way as compared with negatively skewed distributions. This is perhaps a function of the relative size of the numbers themselves as those having negative skew approach zero. Regardless, this is certainly perplexing as we might have thought intuitively that similarly skewed variables would produce correlations of similar size, where only the sign would change.

Thus far we have only compared the Normal distribution to some twenty-one others. We might presume then that intercorrelations among each of the twenty-one would also be effected by skewness and relative variance. Table 2 illustrates the intercorrelations among twenty-nine variations of skewness and kurtosis. The eight new distributions were added to show the effects among variables which have only a small amount of skewness. The Standard Deviation of the Normal distribution used in the transformations shown here was 20% of the Mean. The top half of the matrix consists of r's, the bottom half, the Index of Forecasting Efficiency. The Index illustrates the explanatory power of r as expressed as a percent, with 100.0% being the maximum.

The closer the skews, the greater the magnitude of r. For some correlations (noted with an *) register overflow occurred during

TABLE 2

SIMPLE ASSOCIATIONS BETWEEN THE NORMAL DISTRIBUTION
(σ= 20% of $\bar{X}$) AND VARIOUS TRANSFORMATIONS (N=2500)
PEARSON r

INDEX OF FORECASTING EFFICIENCY

	Normal	X^{-5}	$X^{-4.5}$	X^{-4}	$X^{-3.5}$	X^{-3}	$X^{-2.5}$	X^{-2}	$X^{-1.5}$	$X^{1.5}$	X^{2}	$X^{2.5}$	X^{3}	$X^{3.5}$	X^{4}	$X^{4.5}$	X^{5}	$\sqrt{X}$	$X^{1/3}$	$X^{1/4}$	$X^{1/5}$	LOG X
Normal	Normal	-.431	-.503	-.582	-.663	-.740	-.809	-.868	-.914	.998	.991	.981	.967	.951	.932	.911	.888	.997	.995	.994	.993	.989
X^{-5}	9.76		.994	.975	.940	.891	.831	.765	.698	-.394	-.362	-.333	-.307	-.284	-.262	-.243	-.226	-.473	-.488	-.496	-.500	-.520
$X^{-4.5}$	13.58	89.23		.993	.971	.934	.884	.826	.765	-.465	-.430	-.399	-.370	-.344	-.321	-.299	-.278	-.546	-.561	-.569	-.574	-.594
X^{-4}	18.68	77.69	88.28		.992	.969	.932	.884	.831	-.542	-.506	-.473	-.442	-.414	-.387	-.363	-.339	-.625	-.640	-.648	-.653	-.672
$X^{-3.5}$	25.11	65.94	76.14	87.64		.992	.969	.934	.891	-.623	-.586	-.552	-.519	-.489	-.460	-.433	-.407	-.705	-.719	-.727	-.731	-.749
X^{-3}	32.75	54.64	64.25	75.30	87.42		.992	.971	.940	-.702	-.666	-.631	-.597	-.565	-.535	-.506	-.478	-.780	-.793	-.800	-.804	-.821
$X^{-2.5}$	41.29	44.39	53.27	63.68	75.33	87.66		.993	.975	-.774	-.739	-.705	-.672	-.639	-.608	-.577	-.547	-.845	-.857	-.863	-.867	-.881
X^{-2}	50.30	35.59	43.64	53.28	64.28	76.16	88.27		.994	-.836	-.803	-.771	-.739	-.707	-.675	-.644	-.614	-.899	-.909	-.914	-.917	-.929
$X^{-1.5}$	59.41	28.35	35.56	44.37	54.61	65.87	77.57	89.10		-.886	-.857	-.828	-.797	-.767	-.736	-.705	-.674	-.940	-.948	-.952	-.954	-.963
$X^{1.5}$	93.12	8.09	11.45	15.99	21.80	28.78	36.67	45.09	53.68		.998	.992	.982	.969	.954	.936	*	.990	.986	.980	.983	.976
X^{2}	86.60	6.76	9.72	13.76	19.00	25.36	32.63	40.46	48.53	93.43		.998	.992	.983	.971	*	*	.979	.973	.970	.969	.960
$X^{2.5}$	80.42	5.69	8.30	11.90	16.61	22.39	29.06	36.32	43.87	87.15	93.67		.998	.993	*	*	*	.964	.957	.954	.951	.942
X^{3}	74.56	4.82	7.11	10.32	14.55	19.80	25.91	32.61	39.65	81.15	87.58	93.87		*	*	*	*	.947	.939	.934	.931	.921
$X^{3.5}$	69.00	4.10	6.12	8.97	12.77	17.52	23.10	29.27	35.80	75.41	81.72	87.93	*		*	*	*	.927	.918	.913	.910	.897
X^{4}	63.73	3.51	5.28	7.81	11.21	15.51	20.59	26.25	32.29	69.94	76.09	*	*	*		*	*	.905	.895	.890	.886	.873
$X^{4.5}$	58.73	3.00	4.56	6.81	9.85	13.73	18.34	23.52	29.08	64.71	*	*	*	*	*		*	.881	.870	.865	.861	.846
X^{5}	53.99	2.58	3.95	5.94	8.66	12.14	16.31	21.04	26.15	*	*	*	*	*	*	*		.856	.844	.838	.834	.819
$\sqrt{X}$	92.72	11.87	16.23	21.94	29.05	37.39	46.60	56.21	65.80	85.90	79.50	73.49	67.82	62.49	54.47	52.74	48.28		1.00	.999	.999	.998
$X^{1/3}$	90.21	12.70	17.25	23.18	30.53	39.12	48.54	58.35	68.10	83.42	77.08	71.13	65.55	60.30	55.37	50.74	46.39	97.47		1.00	1.00	.999
$X^{1/4}$	88.93	13.14	17.79	23.84	31.31	40.01	49.55	59.45	69.28	*	75.85	69.44	64.40	59.20	54.32	49.74	45.45	96.18	98.71		1.00	.999
$X^{1/5}$	88.16	13.42	18.13	24.24	31.79	40.56	50.17	60.13	70.00	81.41	75.11	69.20	63.82	58.54	53.69	49.14	44.88	95.40	97.93	99.22		.999
LOG X	85.03	14.59	19.54	25.94	33.74	42.86	52.74	62.92	72.97	78.35	72.13	66.34	60.94	55.89	51.17	46.75	42.62	92.23	94.75	96.04	96.82	

* = Register Overflow

TABLE 3

STANDARD DEVIATION OF THE NORMAL DISTRIBUTION WHEN FORECASTING
EFFICIENTY IS LESS THAN 50% OR 75% (N=2500)
INDEX OF FORECASTING EFFICIENTY IS LESS THAN 75% (r_c.867)

INDEX OF FORECASTING EFFICIENCY IS LESS THAN 50% (r_c.707)

	Normal	X^{5}	$X^{4.5}$	X^{4}	$X^{3.5}$	X^{3}	$X^{2.5}$	X^{2}	$X^{1.5}$	$X^{1.3}$	$X^{1.2}$	$X^{1.1}$	$X^{1.05}$	$X^{.95}$	$X^{.9}$	$X^{.8}$	$X^{.66}$	$X^{.5}$	$X^{.33}$	$X^{.25}$	$X^{.2}$	X^{-5}	$X^{-4.5}$	X^{-4}	$X^{-3.5}$	X^{-3}	$X^{-2.5}$	X^{-2}	$X^{-1.5}$	LOG X
Normal		7.5	10	12.5	15.0	20.0	25.0	35.0													35.0	7.5	7.5	7.5	10.0	10.0	12.5	12.5	15.0	35.0
X^{5}	35.0			20.0	20.0	15.0	15.0	15.0	12.5	12.5	12.5	12.5	12.5	12.5	10.0	10.0	10.0	10.0	10.0	10.0	10.0	5.0	5.0	5.0	5.0	5.0	7.5	7.5	7.5	10.0
$X^{4.5}$	35.0				17.5	17.5	15.0	15.0	15.0	15.0	12.5	12.5	12.5	12.5	12.5	12.5	10.0	10.0	10.0	10.0	10.0	5.0	5.0	5.0	5.0	5.0	7.5	7.5	7.5	10.0
X^{4}	35.0	35.0					20.0	20.0	17.5	15.0	15.0	15.0	15.0	15.0	12.5	12.5	12.5	12.5	12.5	12.5	12.5	5.0	5.0	5.0	5.0	7.5	7.5	7.5	7.5	10.0
$X^{3.5}$		35.0						35.0	20.0	20.0	17.5	17.5	17.5	17.5	15.0	15.0	15.0	15.0	12.5	12.5	12.5	5.0	5.0	5.0	7.5	7.5	7.5	7.5	7.5	12.5
X^{3}		35.0	35.0					35.0	25.0	25.0	22.5	22.5	22.5	20.0	20.0	20.0	17.5	17.5	15.0	15.0	15.0	5.0	5.0	7.5	7.5	7.5	7.5	7.5	10.0	15.0
$X^{2.5}$		35.0	35.0									30.0	30.0	30.0	25.0	25.0	22.5	20.0	17.5	17.5	17.5	5.0	7.5	7.5	7.5	7.5	7.5	10.0	10.0	15.0
X^{2}		30.0	35.0														30.0	25.0	22.5	22.5	22.5	7.5	7.5	7.5	7.5	7.5	10.0	10.0	12.5	22.5
$X^{1.5}$		30.0	35.0	35.0													35.0	35.0	30.0	30.0	25.0	7.5	7.5	7.5	7.5	10.0	10.0	12.5	12.5	25.0
$X^{1.3}$		30.0	35.0	35.0														35.0	30.0	30.0	30.0	7.5	7.5	7.5	10.0	10.0	10.0	12.5	15.0	25.0
$X^{1.2}$		30.0	35.0	35.0															35.0	30.0	30.0	7.5	7.5	7.5	10.0	10.0	10.0	12.5	15.0	30.0
$X^{1.1}$		25.0	35.0	35.0															35.0	35.0	35.0	7.5	7.5	7.5	10.0	10.0	10.0	12.5	15.0	30.0
$X^{1.05}$		25.0	35.0	35.0	35.0														35.0	35.0	35.0	7.5	7.5	7.5	10.0	10.0	12.5	12.5	15.0	30.0
$X^{.95}$		25.0	30.0	35.0	35.0														35.0	35.0	35.0	7.5	7.5	7.5	10.0	10.0	12.5	12.5	15.0	30.0
$X^{.9}$		25.0	30.0	30.0	35.0														35.0	35.0	35.0	7.5	7.5	7.5	10.0	10.0	12.5	12.5	15.0	30.0
$X^{.8}$		25.0	25.0	30.0	35.0	35.0														35.0	35.0	7.5	7.5	7.5	10.0	10.0	12.5	12.5	15.0	30.0
$X^{.66}$		25.0	25.0	30.0	35.0	35.0														35.0	35.0	7.5	7.5	10.0	10.0	10.0	12.5	15.0	17.5	30.0
$X^{.5}$		20.0	22.5	25.0	30.0	35.0	35.0													35.0	35.0	7.5	7.5	10.0	10.0	10.0	12.5	15.0	17.5	30.0
$X^{.33}$		20.0	20.0	25.0	30.0	30.0	35.0	35.0														7.5	7.5	10.0	10.0	12.5	12.5	15.0	17.5	
$X^{.25}$		20.0	20.0	22.5	30.0	30.0	35.0	35.0	35.0	35.0												7.5	7.5	10.0	10.0	12.5	12.5	15.0	17.5	
$X^{.2}$	35.0	20.0	20.0	22.5	30.0	30.0	35.0	35.0	35.0	35.0	35.0	35.0	35.0	35.0	35.0							7.5	10.0	10.0	10.0	12.5	12.5	15.0	22.5	
X^{-5}	12.5	10.0	10.0	10.0	10.0	10.0	10.0	12.5	12.5	12.5	12.5	12.5	12.5	12.5	12.5	12.5	12.5	12.5	12.5	15.0	15.0				17.5	15.0	12.5	12.5	10.0	7.5
$X^{-4.5}$	12.5	10.0	10.0	10.0	10.0	10.0	12.5	12.5	12.5	12.5	12.5	12.5	12.5	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0				22.5	17.5	15.0	12.5	12.5	10.0
X^{-4}	15.0	10.0	10.0	10.0	12.5	12.5	12.5	12.5	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0					22.5	17.5	15.0	12.5	10.0
$X^{-3.5}$	15.0	10.0	10.0	12.5	12.5	12.5	12.5	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	17.5	17.5	17.5	17.5	17.5	17.5						22.5	17.5	15.0	10.0
X^{-3}	17.5	10.0	12.5	12.5	12.5	12.5	15.0	15.0	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	20.0	20.0	20.0							22.5	17.5	12.5
$X^{-2.5}$	20.0	12.5	12.5	12.5	15.0	15.0	15.0	17.5	17.5	17.5	17.5	20.0	20.0	20.0	20.0	20.0	20.0	20.0	20.0	20.0	22.5	25.0							22.5	15.0
X^{-2}	22.5	12.5	12.5	15.0	15.0	15.0	17.5	17.5	20.0	20.0	20.0	20.0	20.0	22.5	22.5	22.5	22.5	22.5	22.5	22.5	22.5	17.5	25.0	25.0	25.0					17.5
$X^{-1.5}$	25.0	12.5	15.0	15.0	17.5	17.5	20.0	20.0	22.5	22.5	22.5	22.5	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	30.0	17.5	17.5	25.0	25.0	25.0	25.0			20.0
LOG X	35.0	17.5	20.0	22.5	25.0	30.0	30.0	30.0	35.0	35.0	35.0	35.0	35.0	35.0	35.0	35.0						15.0	15.0	17.5	17.5	25.0	25.0	25.0	30.0	

computation and no figures are shown. Here we see that the r's in some instances are quite close to 100% in efficiency while in others they are reduced to less than 3% efficiency. If we peruse these relationships we can readily see the power of skewness and direction on the absolute magnitude of the correlations. That is, distributions having similar skewness vary together regardless of the fact that they are not Normal. Those having quite different patterns of skewness result in relatively low measures of association. Variables having negative correlations vary positively regardless of the extent of negative skew. You should be reminded as you view this illustration, that the transformations used to produce these correlations in no way effected changes in ordering of the individual cases. Only their mathematical positions on the scales were changed. Were these data changed to ranks, all Spearman Rhos would be +1.000 except for the negative functions where they would be -1.000.

Note in Table 2 that many of the correlation coefficients are below that which might be tolerated with respect to a reasonable or tolerable amount of loss in covariation. A definition for tolerable or "reasonable" does not exist. For this illustration we have set that level quite arbitrarily, as its definition is merely for illustrative purposes. We have set two levels, one at seventy-five percent of the variance accounted for in the bivariate association, and the second at fifty percent. That is, let us suppose that we might want to know when a bivariate association between non-Normal variables would explain less than 75% or 50% of the common variance when in fact that association should have been unity. Remembering that the relative value of the Standard Deviation as compared to the Mean is important, we could then review various matricies of r's after setting the Sigma of the Normal distribution at different levels. When the resultant correlations explained less than 75% (r = .867) or 50% (r = .707) of the variance we could note the relative variance where the loss was found. Table 3 presents the results of just such an experiment where the relative variance was adjusted in 2.5% steps from 5.0% to 20%, and then in 5% steps from 20% to 35% (our wraparound point). The numbers above the diagonal represent the relative Standard Deviation of the Normal distribution where the bivariate association between the transformed pair fell below .867(75%) and those numbers below the diagonal represent the point at which the r's fell below .707(50%). Blank cells are shown where over 75% or 50% of the variance was still accounted for among the bivariate pairs when we had reached a relative variance of 35% for the Normal distribution.

Table 3 summarizes that which we have presented above. As the differences in distributions increase, and as the similarity among variances decreases an associated loss in the magnitude of r occurs. For proximous distributions explanation is still high when, if the data were Normalized, the Standard Deviations of both variables would still be 35% or less as compared to their Means. Yet consider the association between X^2 and $X^{.5}$. At a relative (Normalized) Standard Deviation of 25% we experience a loss of 25% of the associative power of the Pearson Product-Moment Correlation. Or consider, say, $X^{2.5}$ and $X^{.5}$. We lose half of the associative power at the 35% level. These skews and the values of the Standard Deviations relative to the mean are not unlike many encountered in research settings, thus we should consider these relationships rather seriously before proceeding in the computation of correlation coefficients without first normalizing each scale.

Substantive Implications

The analytical portion of this paper has dealt primarily with the statistical outcomes of the computation of Pearson's r between non-Normal interval or ratio scale variables. We have noted sometimes seriously weakened correlation coefficients. We have barely touched on the theoretical and other substantive outcomes of the adistrubutional computation of r's. Yet the substantive and theoretical implications are enormous. For instance consider the misrepresentation of relationships which may occur simply because of the positioning of the mean and Standard Deviation on the scale. For instance in social research one might find a weak positive relationship between income and education as both become more skewed, one to the right, the other to the left. This finding might then prompt a generation to leave education early in the hope of affluence. During the progressive change in the relative skews models might be prepared which would tend to reinforce the behavior.

As well, whole new techniques of residual analysis have been designed so as to determine the association of residuals after correlation with other variables. Given the presence of Skewness and Kurtosis, this analytical technique might not only be unneeded but it might mislead the researcher into believing that a particular variable was important to explanation when, in fact, it merely represented the residual of skewness not associated with the non-normal dependent variable in question. Such results would be tremendously misleading.

Many models are based on stepwise or factorial solutions. For models where the dependent variable or groups of others are skewed in similar directions, of similar amounts and perhaps where all are of different relative Standard Deviations, the stepwise entries would occur simply due to distributional effects and not because of the

power of the actual relationship among group members as based on relative scale placement. This is especially important for models where the choice and sequencing of variable entry is based on the absolute value of r. Regardless of potentially normalized placement in the selection process, variables having similar skew to that of the dependent variable would be entered first, rendering the balance of the model a matter of adjusting to the residuals and fitting to specious relationships. The resultant models would then misrepresent mathematical space as defined in the parameters of the statistics used (i.e distribution and variances) but more seriously would misrepresent the importance, direction and sequencing of the variables to the model. As well, variables which had major importance theoretically and actually, might have been omitted entirely or relegated to inferior positions in the model simply because of skewness and kurtosis in the dependent variable or one or more of the independent variables. In cases of multicolinearity among variables, ordering by size of r would probably result of the deletion of the variable which had the least similarily in skew as compared with the dependent variable even when actually it was more closely associated with it than was its intercorrelated independent companion. Even relatively small amounts of negative skew would be important here as the Pearson r decays relatively quickly under situations of negative skewness.

The problem of skewness and homosscedasticity demand that we transform data prior to analysis through the use of Pearson's r and similar models(8). Moreover, given the asymmetrical properties of r where negative skewness is concerned, the relationship between the Mean and Median should also be considered. Yet transformed variables do not lend themselves to simple explanatory declarations in the descriptions of their origins. It is one thing to understand the relationship between income and education, it is another to comprehend the relationship between the square of income as related to the square root of education. This is especially true among lay audiences. However, such descriptions can be tailored to audiences where complex relationships between and among variables exist. The problem of oral or written description is certainly not enough to warrant the violation of the parameters of the statistic used.

At the same time, a critic of the philosophy of data transformation might propose that data are merely substitutes for reality and that those data might be poor representations of that reality. The notion of sampling distributions and sampling error support this contention. Data are representations though and despite their possible flaws we use them intact and as obtained. To transform them may artificially reduce or increase their differences from their theoretically true value but we must accept that as a part of statistical error itself. Without the data we cannot attempt the analysis, without transformation we cannot use the statistic properly (9).The alternative is to shun transformation and rely on other statistics which do not require Normal distributions and homoscedasticity, for instance Spearman's Rho. Obviously this is a perfectly acceptable alternative to data transformation.

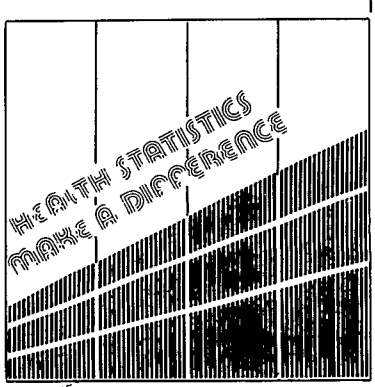
Finally, there is the issue of previous research. We have already suggested that there is relatively little in published social research to lead us to believe that transformations had been effected prior to the use of Pearson's r or like models. Similarly, these same studies do not show that transformation was unwarranted on the basis of levels of Skewness and Kurtosis. How should we treat this published information? We have shown that such results might well be misleading or erroneous if not simply inefficient. Should the models be replicated using normalized scales so as to validate them? Certainly where the theory is a major one or its use and impact is great, replication using transformed data would be wise. It is clear that the issues of Skewness, Kurtosis and Homoscedasticity are of great importance when Pearson's r is used in data analysis and subsequent description, explanation and theory construction. These important and apparantly long ignored parameters are too powerful to disregard in data analysis. Their mathematical and subsequent substantive impact is simply too great to ignore.

Notes

1.See for instance:
Guilford,J.P. Fundamental Statistics in Psychology and Education 4th Edition, Mc Graw- Hill 1965,p.108;
Blalock Jr., Hubert M. Social Statistic Revised 2nd Edition Mc Graw- Hill, 1979,pp. 389-410.
2.American Sociological Review Vols. 39 through 50.
3.This distribution was originally developed by the RAND Corporation.
4.Computations were performed using SPSS and DATRAN as a check of SPSS results.
5.Regardless of transformation, Sk=.007, Ku=-.088.
6.Different computers and software use different algorithms to reduce computing time.Under conditions of over or under-flow registers are set to a maximum or a minimum value depending on computer and algorithms independently.
7.Bohrnstedt,George and Gerald Marwell, "The Reliability of Products of Two Random Variables" in Karl F Schuessler, ed Sociological Methodology 1978, Jossey Bass 1977(sic),pp. 254-273.
8.Guilford, J.P. op. cit., p360.
9.Blalock Jr. op. cit., p343.

Session N

Issues in Diagnostic Classification



DRG REFINEMENT: A FEASIBILITY ASSESSMENT USING STAGE OF DISEASE, AGE, AND UNRELATED COMORBIDITY

Jonathan E. Conklin, SysteMetrics, Inc.

INTRODUCTION

With the ever-increasing costs of hospital care, a number of approaches to controlling costs have been explored in recent years. The most prominent is case-mix reimbursement, under which hospitals are paid on the basis of the types of patients they treat. A fundamental premise of case-mix reimbursement is that patients with similar medical problems will tend to use similar amounts of hospital resources. The most widely implemented case-mix reimbursement scheme is the hospital Prospective Payment System (PPS) developed by the Health Care Financing Administration (HCFA) for Medicare inpatient reimbursement. Under PPS, patients are classified according to 467 Diagnosis Related Groups (DRGs) and hospitals are paid according to their DRG case mix.

The DRG classification system was intended to be equitable to patients and hospitals, and it is successful in distinguishing between high and low cost diseases. The system is not without critics, however, who feel it lacks sensitivity to variations in resource use that are associated with differences in disease severity among patients. If true, this may limit the ability of PPS to provide for equitable payment of hospitals. Recognizing the potential financial risks to certain types of hospitals treating severely ill patients, Congress mandated research into the advisability and feasibility of modifications to the DRG classification system using measures of severity, as part of the 1983 statute establishing prospective payment (P.L. 98-21).

A recent pilot study (1,2) described in this paper responded to that mandate and investigated the feasibility of refining the DRG system for the Medicare population using Disease Staging, age, and a measure of unrelated comorbidity. Disease Staging (3,4,5), a computerized measure of disease severity based on disease-specific criteria, was used to assess the stage of progression of individual principal conditions (reasons for hospital admission) within each DRG. Unrelated comorbidity was defined in terms of the relationships and severities of co-occurring conditions for each patient.

This study was not based on the assumption that DRGs should be replaced by a totally different system of case mix classification; rather, it explored how the current DRG system might be improved by incorporating, where needed, further differentiation on disease severity. Specifically, the study assessed the potential impact of refinements to selected DRGs, and compared existing DRGs to alternative groupings that were based on stage of illness, comorbidity, and age.

METHODOLOGY

Analysis files were created from databases of all Medicare discharges from short-term general acute-care hospitals in Maryland (1981) and New Jersey (1979). The data elements analyzed in the study consisted of items routinely collected and computerized by hospitals. Resource consumption was defined by lengths of stay and estimates of total treatment costs. Ten DRGs were selected for analysis on the basis of their potential for further refinement. For each of these ten, the adjacent DRGs representing the same principal diagnoses or procedures but split on the basis of age and/or CCs were combined to form an "adjacent diagnosis related group" (ADRG). Table 1 presents the list of ADRGs analyzed in the study. By focusing on ADRGs as the unit of analysis, the final DRG splits (on age and/or CCs) could be easily compared to alternative groupings based on severity measures.

Analyses were conducted in two phases. In the first phase, regression models were used to assess the amount of variation in cost and LOS per ADRG that was explained by principal condition, stage of illness, unrelated comorbidity, and age. These analyses were useful in identifying which variables might be effective refinement tools for different kinds of DRGs. Comparisons were made to the proportions of variation explained by the final DRG splits within each ADRG.

TABLE 1

DESCRIPTION OF ADJACENT DRGs SELECTED FOR ANALYSIS

DESCRIPTION OF ADJACENT DRG
Major Bowel Procedures
DRG 148: Age => and/or C.C.
DRG 149: Age < 70 w/o C.C.
Pulmonary Embolism
DRG 78
Major Reconstructive Vascular Procedures
DRG 110: Age => 70 and/or C.C.
DRG 111: Age < 70 w/o C.C.
Kidney and Urinary Tract Infections
DRG 320: Age => 70 and/or C.C.
DRG 321: Age 18-69 w/o C.C.
DRG 322: Age 0-17
Cirrhosis and Alcoholic Hepatitis
DRG 202
Major Chest Procedures
DRG 75
Cerebrovascular Disorders Except TIA
DRG 14
Diabetes Mellitus
DRG 294: Age => 36
DRG 295: Age 0-36
Benign Prostatic Hypertrophy
DRG 348: Age => 70 and/or C.C.
DRG 349: Age < 70 w/o C.C.
Biliary Tract Procedures
DRG 197: Age => 70 and/or C.C. w/o C.D.E.
DRG 198: Age < 70 w/o C.C. w/o C.D.E.

The second phase of the analysis proposed specific alternative splits within each ADRG based on the significant variables in the regression analyses. Disease, stage, age, and unrelated comorbidity were used to define alternative groupings on the basis of clinical integrity and their differentiation of frequency and cost distributions. The number of alternative groups per ADRG was minimized to maintain comparability to the existing DRG splits. Regression analyses with dummy variables were used to compare the proportions of cost and LOS variance explained by these alternative splits and the existing DRG splits.

RESULTS

The results of this research demonstrate that the amount of variance explained by the severity-based regression models consistently exceed the amount of variance explained by the existing DRG splits within each ADRG. Figure 1 illustrates this finding with bar-graph representations of the cost variance explained by each classification model for the Maryland data. The proportions of variance in cost explained by the severity models are generally significant across ADRGs, with approximately 12% of variance explained on average for the cost models. These proportions are large in comparison to the proportions of variance explained by the DRG splits within ADRGs, which average about 2%.

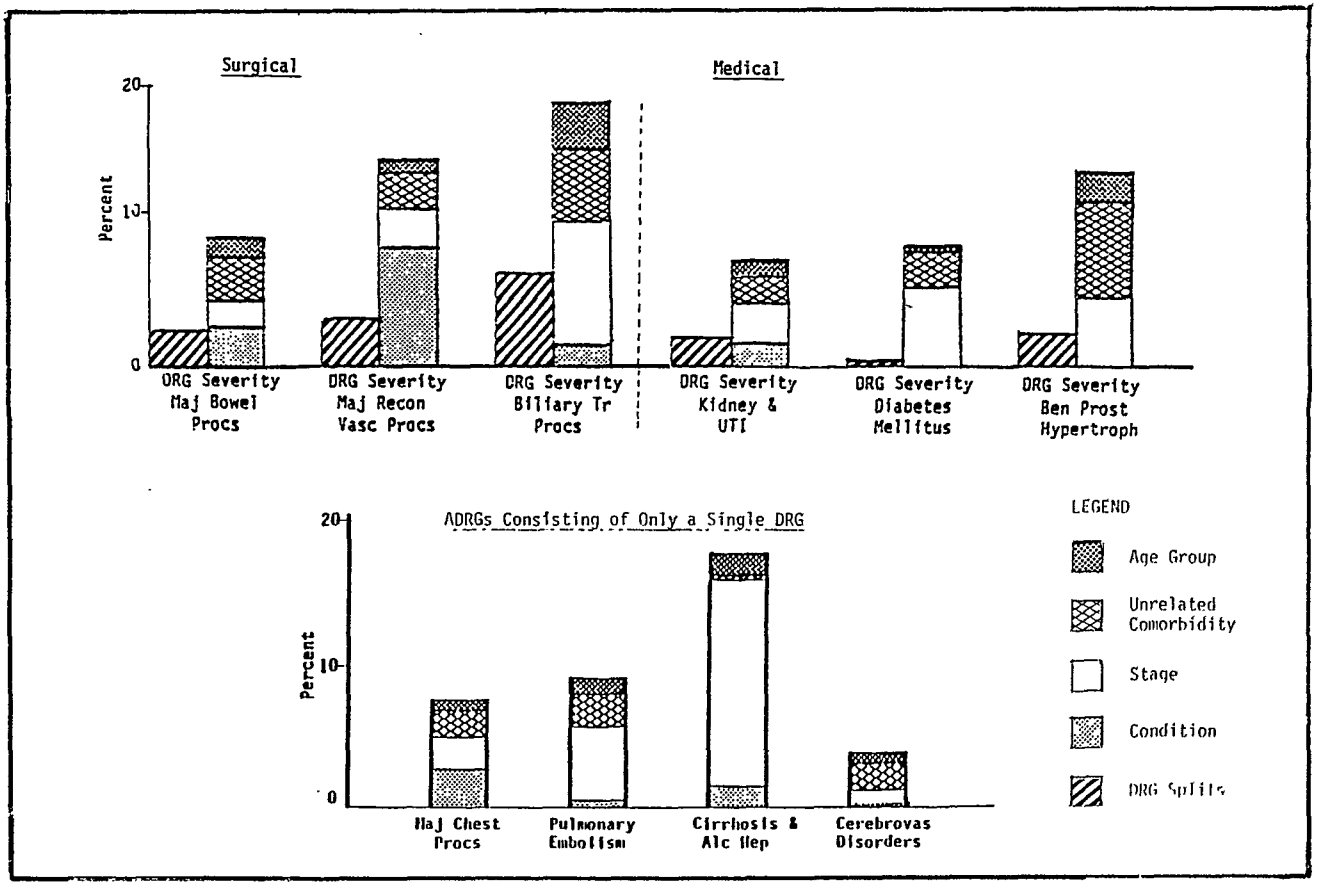
The results also indicate the incremental effects of the individual predictor variables. Principal condition accounts for significant proportions of cost and LOS variance within approximately half of the ADRGs. In general, its effects are strongest for the surgical ADRGs and weakest for the medical ADRGs. These findings indicate that, at least for some DRGs, further distinctions could be made on principal condition to obtain more homogeneous patient classifications.

Stage of illness is generally a strong significant predictor of cost and LOS within ADRGs. The proportions of variance explained by this variable alone often exceed the proportions explained by the DRG splits within ADRGs. For the selected ADRGs, then, stage of illness tends to classify patients into groups that are statistically more homogeneous than do the existing DRG splits.

Unrelated comorbidity exhibits small but significant unique effects on resource consumption within ADRGs, even after adjusting for the effects of condition and stage of illness. In fact, the variance uniquely explained by the differences in presence of unrelated comorbidity often exceeds the variance explained by the DRG splits within an ADRG. Unrelated comorbidity is consistently related to higher resource consumption, reflecting the greater complexity of treating patients with multiple unrelated problems.

FIGURE 1

PROPORTION OF COST VARIANCE EXPLAINED
BY DRG SPLITS AND SEVERITY MODEL



Age tends to add little to the explanation of variance in resource consumption for these ADRGs, within the Medicare population, once the effects of condition, stage of illness, and unrelated comorbidity are removed. Age groupings may be useful for refining patient classifications in some diseases for the general acute-care population. These analyses indicate, however, that their utility for reducing variance in resource consumption in most disease categories within the Medicare population is generally secondary to the effects of condition, stage, or comorbidity.

In the second phase of the analysis, alternative groupings were formed within each ADRG to provide comparisons to the existing DRG splits. These alternative groupings were defined by combining condition, stage, unrelated comorbidity, and age in such a way as to maximize the clinical and statistical integrity of the group definitions. Table 2 provides examples of alternative groupings for two of the ADRGs analyzed in the study. These groupings tend to cluster patients in a more clinically meaningful fashion than the existing DRG splits on age and/or CCs.

Results of the analyses reveal that the alternative groupings are consistently more effective in distributing cases and reducing variance in cost and LOS than are existing DRG splits. Figure 2 graphically compares the proportions of cost variation explained by the DRG splits, the alternative groups, and the severity regression model for each of the ten DRGs. The proportions of cost variance explained by the alternative groupings differ from one ADRG to another, but in most instances exhibit dramatic improvement over the proportions of variance explained by existing DRG splits. In addition, the small number of meaningful alternative groups within each ADRG explain nearly the same proportion of variance as that explained by the severity regression model (which includes dummy variables for all possible combinations of condition, stage, comorbidity, and age).

As a framework for interpreting these results, it is important to consider the tree-like structure of the DRG system. In this system, diseases are originally grouped by body systems (Major Diagnostic Categories or MDCs), which then branch into specific disease and procedure entities, and are finally split by age and/or presence of complications and comorbidities (CCs). This study focuses on only the final level of branching, on age and/or CCs. On average, recent studies have found that the DRG system as a whole accounts for approximately 20-30 percent of total variance in resouce consumption across patients and hospitals. The results of this study suggest that the final branching on age/or CCs accounts for only about 1-3 percent of variance in resource consumption. The alternative groupings at the same level of branching, on the other hand, account for approximately 8-10 percent of such variance without requiring significant increases in the number of groups.

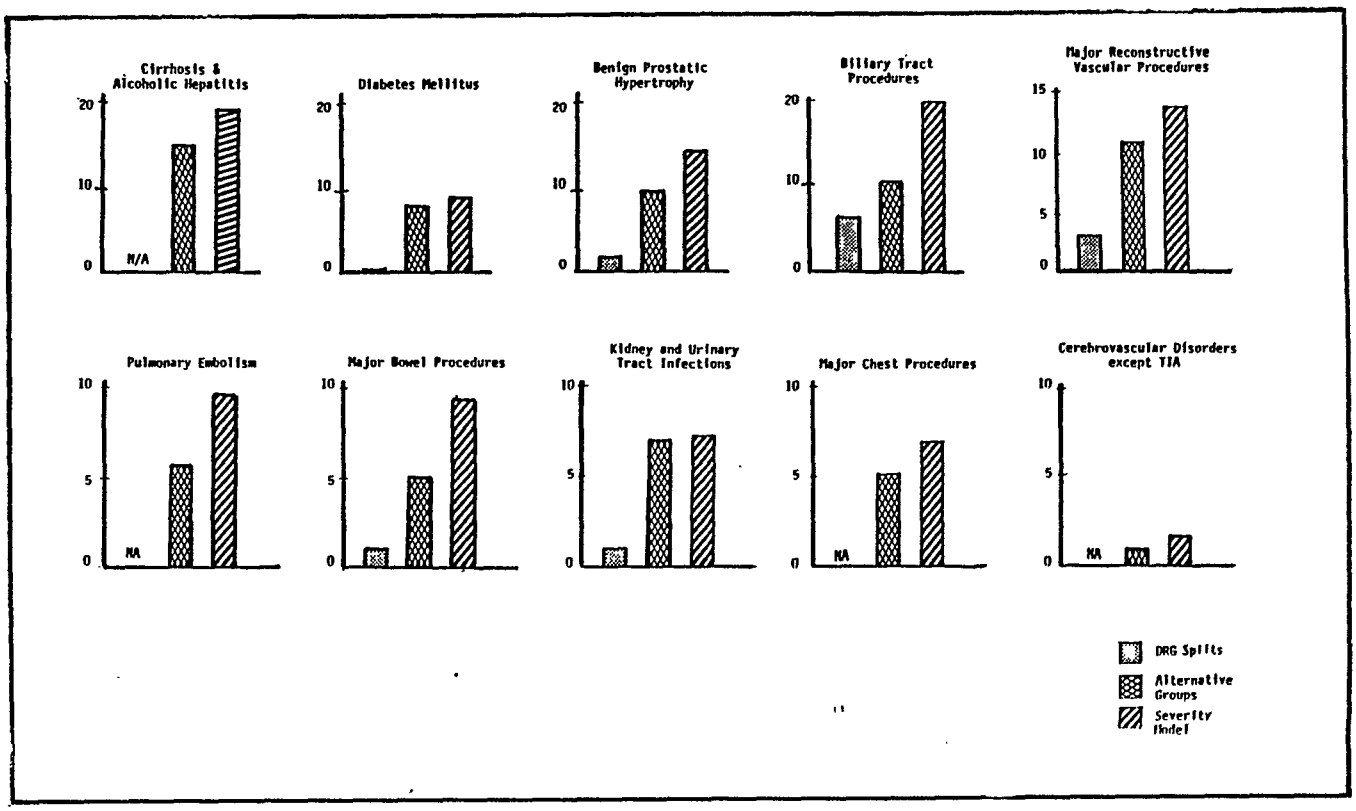
TABLE 2

EXAMPLES OF ALTERNATIVE GROUPS AND COMPARISONS TO EXISTING DRG SPLITS FOR TWO ADJACENT DRGs

		Condition	Stage	Age	Complications or Comorbidity (DRG Definition)	Unrelated Comorbidity	Frequency	COST Mean	COST SD	LOS Mean	LOS SD
ADJACENT DRG: Major Bowel Procedures											
DRG SPLITS	148	--	--	70+(or)	Yes	--	1565	$8523.61	$8053.42	23.86	18.14
	149	--	--	0-70	No	--	225	5149.21	2818.76	16.10	8.26
ALTERNATIVE GROUPS	1	--	--	--	--	No	1059	6642.03	5265.82	19.95	12.19
	2	CANCER	1	--	--	Yes	161	10584.03	9512.17	29.66	20.47
	3	CANCER	2-4	--	--	Yes	212	8200.27	6727.58	25.33	18.90
	4	OTHERS	--	--	--	Yes	358	11233.62	11268.43	27.10	24.70
						TOTAL	1790	$8099.46	$7677.73	22.89	17.40
ADJACENT DRG: Biliary Tract Procedures											
DRG GROUPS	197	--	--	70+(or)	Yes	--	917	$3945.48	$1929.43	13.44	5.45
	198	--	--	0-69	No	--	348	2884.25	1574.86	10.31	5.16
ALTERNATIVE GROUPS	1	--	1	--	--	No	852	3266.59	1585.80	11.75	5.04
	2	--	1, 2	--	--	1(yes), 2(all)	351	4247.81	1973.36	14.15	5.47
	3	--	3, 4	--	--	All	62	5606.64	2470.16	15.03	5.98
						TOTAL	1265	$3653.54	$1857.94	12.58	5.34

FIGURE 2

PROPORTION OF COST VARIANCE EXPLAINED BY
DRG SPLITS, ALTERNATIVE GROUPS, AND SEVERITY MODELS



CONCLUSIONS

These results support the conclusion that DRG refinement is feasible with existing automated data systems. Measures of principal condition, stage of illness, unrelated comorbidity, and age are easily obtained from computerized hospital discharge or claims files without further data collection. These measures can be used to form the basis of disease groupings within the existing DRG system that will enhance its sensitivity to differences in cost and LOS, thereby facilitating equitable hospital reimbursement.

REFERENCES

1. Conklin, J., Louis, D., Lieberman, J., Heinberg, J.: "Refinements to Diagnosis Related Groups Based on Severity of Illness and Age". Final Report, Contract No. HHS-100-82-0038, 1984.

2. Conklin, J.: "DRG Refinement: A Study of Alternative Groupings within Six Sets of Adjacent DRGs". Final Report, Subcontract No. 85-19 of the Cooperative Agreement No. 18-C-98489/901 between RAND Corp. and HCFA/HHS, 1985.

3. Garg, M., Louis, D., Gliebe, W., Spirka, C., Skipper, J., Parekh, R.: "Evaluating Inpatient Costs: The Staging Mechanism". Medical Care, 16, 191-201, March, 1978.

4. Gonnella, J.: Clinical Criteria for Disease Staging, Santa Barbara: SysteMetrics, 1983.

5. Gonnella, J., Hornbrook, M., Louis, D.: "Staging of Disease: A Case-mix Measurement", Journal of the American Medical Association, Vol. 251, No. 5, February 3, 1984.

PEDIATRIC DIAGNOSTIC CLASSIFICATION SYSTEM FOR REIMBURSEMENT OF CHILDREN'S HOSPITALS

Anthony J. Likovich, Jr., SysteMetrics, Inc.

Introduction

The Pediatric Diagnostic Classification System (PEDs), represents a viable alternative to DRGs for pediatric patients in both pediatric and general acute care hospitals. Disease Staging, a clinically based measure of disease severity, provides the basis of the system. The California Association of Children's Hospitals (CACH) is successfully utilizing this system of 215 groups for Medi-Cal reimbursement. Children's hospitals around the country are currently exploring the feasibility of incorporating this system into their own reimbursement strategies.

Background

CACH consists of five pediatric hospitals in California: Children's Hospital of Los Angeles, 308 beds; Children's Hospital and Health Center in San Diego, 120 beds; Children's Hospital Medical Center of Northern California in Oakland, 142 beds; Children's Hospital at Stanford, 60 beds; and Valley Children's Hospital in Fresno with 110 beds.

In the summer of 1982, the California legislature replaced "reasonable cost" reimbursement of Medi-Cal patients with selective provider contracting. All hospitals were required to negotiate a flat per diem rate of reimbursement for their Medi-Cal population. Because of their unusually heavy Medi-Cal caseloads (over 40%) and their vulnerability to changes in Medi-Cal payment policies, the CACH hospitals were exempt from the initial year contracting. In exchange for a second year exemption, CACH agreed to undertake a study to develop an alternative reimbursement system responsive to the special needs and circumstances of California's pediatric hospitals.

There were two objectives of the resulting study: Phase I - Develop a concise statement regarding resource consumption in children's hospitals compared to general acute care hospitals, and Phase II - Design a patient classification system for Medi-Cal reimbursement.

Phase I Analyses

In Phase I, resource consumption and utilization patterns were compared between CACH hospitals and California acute-care hospitals using California Hospital Financial Disclosure Reports as provided by the California Health Facilities Commission. The Disclosure Reports contained utilization and cost statistics from each non-federal hospital in California for fiscal year 1982. Eight hypotheses concerning resource consumption patterns in CACH hospitals were postulated and statistical measures were developed to test each of them. The eight areas of analysis included: the cost of treating patients, sources of payment, proportion of critical care beds and patient days, average length of stay, nursing hours per patient day, levels of nursing personnel, ancillary services, and support services.

The hypotheses were tested using statistical measures defined by ratios of input resourses consumed per unit of output. Input resources were measured by labor hours and costs. Output measures were defined by patient days, discharges, and units of service. Each measure was then applied consistently to both CACH hospitals and the California acute care hospitals' data to demonstrate where and how the CACH hospitals were different.

In addition to the analyses of statewide uniform cost data, patient-level clinical information was analyzed and interviews were conducted with operational and clinical staff from the children's hospitals to give rise to three major conclusions:

1. Children's hospitals provide a unique service to the children of California.
2. Costs in children's hospitals are necessarily and appropriately higher than costs in general acute-care hospitals.
3. Children's hospitals should be paid under a multiple rate system to protect them from dramatic savings in case mix.

An in-depth examination of the statistical analyses undertaken for Phase I can be found in "Comparison of Resource Consumption Patterns between Pediatrics and General Acute Care Hospitals - Phase I, Report on Findings and Conclusions."

Phase II Analyses

The findings of Phase I laid the philosophical groundwork for the actual design of the pediatric classification system in Phase II. The main purpose of such a system was to reduce the contracting risk to both the state of California and CACH. A flat per diem rate was not viewed by CACH as sufficiently responsive to changes in case mix intensity. Because of the incentives inherent to the per diem payment system, CACH was concerned that general acute-care hospitals may increasingly "dump" their severely ill pediatric patients onto the CACH hospitals. The influx of a large number of such patients with intensive resource requirements could place the CACH hospitals at serious financial risk. CACH felt that periodic case-mix adjustments of reimbursement would protect against such financial risk caused by dramatic increases in case-mix severity. The same kind of adjustment could simultaneously

protect the state against over-payment in times of decreases in case-mix severity.

Database Development

Patient records representing an average of twelve months from each of the five CACH hospitals were incorporated into the research database. Total cost per patient was merged with each automated discharge record. The data from the hospitals represented four different abstracting systems, necessitating conversion into a uniform file. After data exclusions of patients with missing cost data, non-pediatric patients (21 years +), patients with missing or inappropriate diagnosis codes, and extreme cost outliers, the final database consisted of 23,908 patients.

Disease Staging, as developed by Gonnella, Louis, and others (1975,1984), was applied to all records. Staging defines the progressive levels of severity for disease in terms of the events and pathophysiological observations that characterize each stage. Within a given body system, higher degrees of involvement or greater degrees of disruption are identified as more severe. Each patient's underlying staged condition and corresponding severity level (stages 1-4) were incorporated into the database.

Besides Disease Staging and the standard UHDDS data elements, additional clinical descriptors were created and added to each patient's record. To define prematurity the diagnosis codes of each patient were searched for the presence of either one of the two prematurity codes (ICD-9-CM codes 7650, 7651). A medical/surgical indicator was developed using the UHDDS list of procedure classes. Any patient having a Class 1 procedure was said to be a surgical patient; all other patients were listed as medical patients. In addition, a comorbidity indicator was developed from the Disease Staging methodology. A patient was said to have had an unrelated comorbidity if he had a secondary condition in a different body system from the underlying staged condition, and if the associated severity level of that secondary condition was higher than a pre-determined threshold level.

Methodology

Because CACH required a system that would not only meet the current needs of a per diem based system, but also a system that could be later adapted to per discharge applications for other payors, it defined discharge cost (not cost per-diem) as the dependent variable. The independent variables consisted of age (0-20) prematurity indicator (0,1), diagnosis (three digit ICD-9-CM), Medical/Surgical indicator (0,1), stage of illness (1-4), and comorbidity (0,1).

The results of all statistical analyses were subjected to intensive clinical reviews for medical meaningfulness. Physician panels from each of the five CACH hospitals were consulted for crucial input into the system. A physician also worked directly with the research staff during the analysis process.

The classification system was developed according to several basic principles in order to maximize validity and effectiveness in adjusting reimbursement rates:

- Clinical validity and parsimony
- Statistical homogeneity
- Equity
- Administrative ease, objectivity, and accuracy
- Proper incentives
- Sensitivity to future swings in actual case mix severity

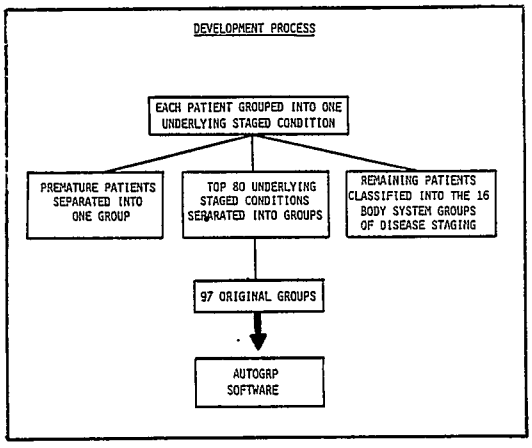
For the state and CACH, this last requirement was crucial. The system must be stable and sensitive enough to accurately reflect changes in the resource requirements of future patients.

While the computer software of the Statistical Analysis System (SAS) was used for Phase I and the preliminary analyses for Phase II, Automatic Interaction Detector (AID), as developed by Morgan and Sonquist (1963), was used in the actual development of the system. AID is a type of hierarchical clustering approach that statistically subdivides observations into disjoint exhaustive subsets to best explain the variance in a chosen dependent variable. These subdivisions proceed recursively through binary splits for one predictor at a time until a termination criterion is reached.

The interactive, computerized version of the AID method, AUTOGRP (Mills et al, 1973), was utilized for the actual construction of the system. When applying AUTOGRP to a database, the clustering parameters (variance and termination factors) must first be defined. The target group of observations, as well as the dependent and independent variables are then identified. AUTOGRP's classification analysis assesses the variance reduction caused by each independent variable. The user must then reconcile statistical and clinical validity in the construction of final groups.

The development process consisted of the following steps. Patients were grouped into 97 "a priori" groups according to their principal disease category (as defined by Staging). Patients with prematurity formed one group. Patients represented by the top 80 disease categories (N>50) were separated into 80 respective groups. The remaining patients were classified into the 16 body system groups of Disease Staging. These resulting 97 original groups were then processed through the AUTOGRP software to identify the additional splits that were statistically necessary.

Upon initial AUTOGRP processing, the subdivision and termination rules were established. An additional group split was warranted if the variance in cost could be reduced by 5% or more. Because of the requirement of constructing a system applicable



to both per diem and per discharge reimbursement, an additional group split was not allowed if the resulting per diem cost distribution was skewed. The end groups were required to have cell sizes of at least 10. Finally, the resulting statistical groups were extensively reviewed for medical meaningfulness. This development process yielded a completely automated system of 215 mutually exclusive, exhaustive, and medically meaningful groups.

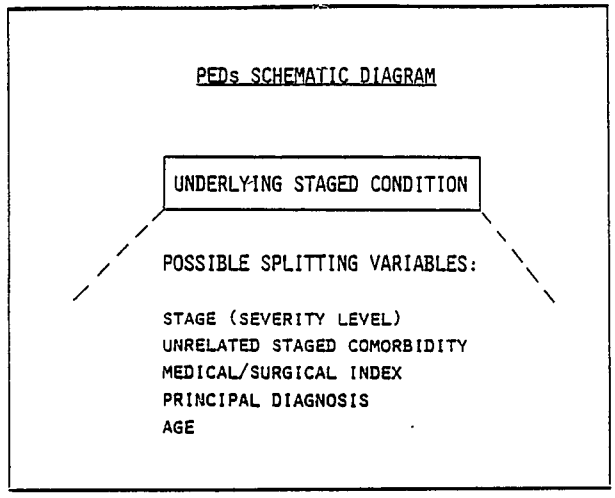
Based on this development process, PEDs has distinct advantages over DRGs for reimbursement of children's hospitals. Besides using solely pediatric data, actual cost instead of length of stay was used to measure resource consumption. The incorporation of disease severity and relationships among comorbid conditions results in a more accurate profile of the health care needs of each patient. Meaningful age splits for the pediatric population are incorporated. Diseases typical to children, such as respiratory distress and cystic fibrosis, are given greater emphasis. Finally, the associated relative costliness weights of each group are based on the combined cost experience of the five children's hospitals.

Implementation

PEDs provides the basis of semi-annual case mix adjustments for Medi-Cal per-diem reimbursement of the CACH hospitals. Each hospital's adjustment factor compares the case mix intensity of patients discharged during the previous six months to those discharged in the base comparison year. The hospital receives an additional payment from the state if it has treated sicker patients. Conversely, the hospital provides the state a partial refund of monies already paid if the hospital has treated less severely ill patients during the interval. A safety net of 4% in either direction protects both parties from random fluctuations in case mix.

Future Enhancements

CACH and SysteMetrics are planning enhancements of the present system to make it more responsive to the needs of the hospitals and the state. A continually updated database of 85,000+ patients



is being collected from CACH to allow more statistically rigorous refinement of the current system. Besides capturing new additional clinical descriptors for each CACH patient, the compilation of a stronger list of major surgeries rather than using the Class I procedure list is planned. Additional medical/surgical and age splits have recently been incorporated into a second version of the system having 257 groups. Further refinements, however, are anticipated. Applications to per discharge and capitation based methods of reimbursement are also being examined.

Conclusion

The Pediatric Diagnostic Classification System has proven to be a useful tool in the Medi-Cal reimbursement of children's hospitals in California. Additional study and experimentation must be undertaken to evaluate the impact of its potential implementation in other states, as well as the feasibility of its application to pediatric patients treated in the general acute care hospital setting.

References

Elia, E., and Mills, L.M.: The AUTOGRP User's Manual: A Basic Reference. Yale University, Institution for Social and Policy Studies, Center for the Study of Health Services, Working Paper No. 24, April 1975.

Gonnella, J.: Clinical Criteria for Disease Staging, Santa Barbara: SysteMetrics, Inc. 1983.

Gonnella, J., Hornbrook, M., Louis, D.: "Staging of Disease: A Case-mix Measuresment". Journal of the American Medical Association, Vol. 251, No. 5, February 3, 1984.

Mills, R., Fetter, R.B., Riedel, D.C., Brauer, L.D., Carlisle, J., Averill, R., Adler, D., and Mills, L.: AUTOGRP: An Interactive System. Yale University, Institution for Social and Policy Studies, Center for the Study of Health Services, Working Paper No. 23, 1973.

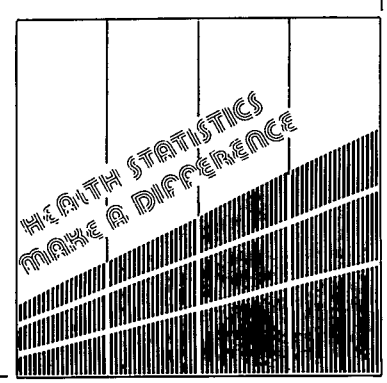
Morgan, J. N. and Sonquist, J.A.: "Problems in the Analysis of Survey Data and a Proposal", Journal of the American Statistical Association, 58, 415-434, September 1963.

SAS Institute Inc. "SAS Users Guide: Basics" Version 5 Edition. Cary, NC: SAS Institute Inc., 1985.

SysteMetrics, Inc. and Ernst & Whinney: Comparison of Resource Consumption Patterns Between Pediatric and General Acute Care Hospitals - Phase I, Report on Findings and Conclusions, Santa Barbara and San Francisco, April 2, 1984.

Session O

Current Topics in Mortality



RECORD LINKAGE OF AIDS SURVEILLANCE AND DEATH CERTIFICATE FILES: Changes in Premature Mortality Patterns in New York City

Alan R. Kristal, New York City Department of Health

INTRODUCTION

This paper describes the techniques and some results of linking records in the New York City Acquired Immunodeficiency Syndrome (AIDS) surveillance registry with New York City death certificates. The purpose of this research is to better understand AIDS epidemiologically, and to describe the impact of AIDS on mortality patterns in New York City.

Reliance upon either disease surveillance or death certificates exclusively as sources of information on AIDS is misleading. In New York City, where AIDS surveillance in 1983 was estimated to be about 90% sensitive (1), follow-up of incident cases for additional diagnoses and ultimately death, is incomplete. On the other hand, there is no single international classification of disease (ICD-9) rubric specific and sensitive to the AIDS. The World Health Organization designated ICD 279.1 for the reporting category for AIDS in January 1983. Still, AIDS deaths may be reported as a pneumonia or cancer, and other immunologic dysfunctions may be coded justifiably to rubric 279.1.

Our approach to more accurately describing the impact of AIDS on mortality is to link all reported cases in the surveillance registry to their certificate of death. This allows accurate analyses of survival patterns after diagnosis of AIDS, and reliable calculations of cause-, race-, age-, and sex-specific mortality rates. Finally, we can better quantify the public health impact of AIDS by examining changing patterns of mortality over time.

METHODS

Morbidity Surveillance

New York City maintains an active surveillance program for AIDS, using voluntary physician reporting and active hospital surveillance. Once notified of a new case, staff of the AIDS surveillance unit confirm the diagnosis and, if no obvious risk for AIDS (male homosexual activity, intravenous drug use, hemophilia, or history of transfusion) is reported, initiate a thorough investigation. A description and evaluation of the New York City Department of Health's AIDS surveillance program has been published elsewhere (1). A detailed analysis of New York City AIDS surveillance data for the years 1980-1984 has also been published elsewhere (2).

Death Certificates

New York City is unique in that it is the only city in the U.S. which maintains an independent vital registry. Computer access to death certificates is available within one week of filing. All deaths are coded using ICD-9 underlying cause of death of criteria (3). In addition, deaths in which narcotics use is confirmed by the medical examiner or narcotism is reported as an underlying or contributory cause of death are classified as "narcotics related." All the analyses reported here are those restricted to deaths occurring in New York City to persons aged 15-64 during the years 1980-1984.

Record Linkage

We used a two-step record linkage procedure to match AIDS surveillance records with death certificates. Two records with identical sex, last name, and first three letters of the first name, and either a) underlying cause of death ICD-9 codes 279.1 (immune dysfunction), 136.3 (pneumocystosis), or 173.9 (neoplasm of the skin, unspecified) on the death certificate, or b) additional correspondence of age, race, and date of death, were accepted as matched records. Additional matched records were obtained by: a) comparing all deaths coded ICD-9 279.1, 136.3, or 173.9 to a list of persons on the AIDS registry; b) comparing all persons on the AIDS registry known dead to a list of all deaths in persons aged 15-64 in New York City; and c) examining all sex and name matches with underlying cause of death not ICD-9 279.1, 136.3, or 173.9. Also, all persons on the AIDS registry with duplicate last names and first three letters of the first name were separated and matched carefully to correct death certificates using additional discriminatory information.

Rates were calculated using 1980 census data. Because there are insufficient data to calculate the incidence of AIDS among gay men specifically, the New York City Department of Health uses an ecological approach: health areas (the smallest geographic area available for health-related analyses) in which the average yearly incidence of pharyngeal and rectal gonorrhea between 1978 and 1982 was greater than

TABLE 1:
Construction of Linked AIDS Surveillance and NYC Death Certificate File.
Techniques Used to Link Records

Source of Merge	Number	(Percent)
DIRECT, COMPUTER MATCH	1303	(83)
HAND MATCH:		
Name match, not AIDS ICD	77	(5)
ICD 279.1, 136.3, 173.9 to registry	21	(1)
Registry deaths to certificates	130	(8)
Registry duplicate names to certificates	47	(3)
Total	1578	

Note:
Computer match = sex, last name, first 3 letters of first name, and
1) ICD 279.1, 136.3, or 173.9, or
2) match on age, race, and date of death

TABLE 2:
Comparison of AIDS Deaths in Surveillance Registry to NYC Death Certificate Files, 1980-1984.

DEATH CERTIFICATE	AIDS REGISTRY Yes Alive	Yes Dead	No Match	Total
Match	162	1400	---	1562
ICD 279.1	98	1004	237	1339
No Match	---	188	---	---
Total	---	1588	---	---

500 per 100,000 males aged 15-54 are selected as health areas with large gay male populations. Rates for never-married males aged 15-64 residing in these health areas are used to approximate the incidence of disease among gay men. Details of this approach were published elsewhere (4).

RESULTS

A total of 1578 linked AIDS surveillance registry-death certificate records could be created (Table 1). The majority could be linked by computer: only 17% of matches required examination of records by hand.

Table 2 shows an evaluation of the matching procedure. Of the total 1,588 deaths reported in the AIDS surveillance registry, 88% could be matched to death certificates. Of the total 1,339 deaths with the underlying cause of 279.1, 82% could be matched to the registry. Reasons deaths on the AIDS registry could not be matched to death certificates include: a) persons died outside of New York City, and b) names were incorrect or missing. Reasons deaths coded to 279.1 could not be matched to the registry include: a) cases were not reported; b) cases were reported without names (10.2% of all cases reported in 1984); c) names were reported incorrectly; d) deaths did not fit the AIDS diagnostic criteria of the Centers for Disease Control (5); and e) deaths were not due to AIDS.

Table 3 shows the trends in underlying cause of death coding for males dying of AIDS. In this and all analyses which follow, a death due to AIDS is defined as a death in which the underlying cause was coded 279.1 or a death which could be linked to the AIDS surveillance registry. Before 1982, AIDS deaths were attributed primarily to pneumonia and pneumocystosis. In 1983 the classification of AIDS as 279.1 became predominant, though other causes, particularly pneumonia, pneumocystosis and neoplasm of the skin, continue to be used.

The percent of all male deaths aged 15-64 which could be matched to the AIDS surveillance registry within each ICD-9 rubric is shown in the right-hand column of Table 3. Classifications 130, 136.3, and 279 were the most specific for AIDS.

Trends in AIDS mortality between 1980-1984 are shown in Table 4. The rates in males aged 15-64 has increased from less than 1 per 100,000 to 42 per 100,000, accounting for 6.2% of all mortality in this age group. Among women the trends are similar though the rates are far lower.

By 1984 AIDS had become one of the five leading causes of death in each five-year age group for males aged 20-49 and for females aged 25-34 (Table 5). In this table, suicide includes deaths to external causes where the medical examiner could not determine self-inflicted or accidental injury.

The effects of AIDS on cause-specific mortality in the two largest risk groups, intravenous drug users and gay men, are shown in Figures 1 and 2. The number of deaths classified as narcotics-related has increased from 554 to 1240 between 1980 and 1984. The increase can be attributed to AIDS and pneumonia, primarily. In

TABLE 3:
Underlying Cause of Death for Males, Ages 15-64, Dying of AIDS in New York City, By Year of Death

	80	81	82	83	84	% 5yr Total
Tuberculosis, 10-18	1	0	4	2	1	2.2
Bacterial, 31-41	0	1	4	1	2	4.8
Viral, 45-79	1	2	4	1	1	6.2
Mycoses, 110-118	0	2	6	3	5	24.7
Toxoplasmosis, 130	0	1	3	0	1	62.5
Pneumocystosis, 136.3	1	10	24	16	29	84.2
Cancer, Skin, 170-173	3	3	29	18	17	11.1
Lymphomas, 200-208	1	0	6	4	1	0.6
Immune Disorder, 279	0	1	31	366	813	83.0
Drug Dependence, 304	0	1	9	4	5	0.9
Nervous Disorder, 320-389	0	0	4	1	2	1.0
Cardiovascular, 320-389	0	0	2	10	12	0.0
Pneumonia, 480-485	0	13	13	23	14	3.6
Chron Liver Dis, 571-573	1	2	1	2	6	0.2
Suicide, 950-959 980-989	0	0	0	0	2	0.1
Others	2	4	8	17	12	0.1

TABLE 4:
AIDS Mortality in New York City, Ages 15-64, By Sex, 1980-1984

	1980	1981	1982	1983	1984
Males					
Rate/100,000	0.46	1.83	6.76	21.38	42.16
% Total Deaths	0.07	0.28	1.03	3.28	6.18
Females					
Rate/100,000	0	0.14	0.55	2.88	6.03
% Total Deaths	0	0.04	0.15	0.75	1.58

TABLE 5:
Leading Causes of Death and Proportion of Age-Specific Total Mortality in New York City, by Age and Sex, 1984.

MALES										FEMALES			
20-24		25-29		30-34		35-39		40-44		45-49		25-29		30-34	
Homicide	40	Homicide	28	AIDS	17	AIDS	13	CHD	25	CHD	25	Homocide	14	Cancer	15
Suicide	17	Suicide	15	Homicide	13	Homicide	12	Cancer	13	Cancer	21	Cancer	13	AIDS	11
Accident	7	AIDS	12	Suicide	11	Cirrhosis	11	Cirrhosis	12	Cirrhosis	11	Suicide	10	Ciorrhosis	9
Cancer	7	Drugs	8	Drugs	11	CHD	10	AIDS	12	AIDS	7	AIDS	9	Suicide	8
AIDS	4	Cancer	6	Cirrhosis	9	Cancer	9	Homicide	7	Homicide	6	Drugs	9	Drug	7

CHD = Coronary Heart Disease

1984 AIDS accounted for 27% and pneumonia for 10% of all narcotics-related mortality, where previously they were rarely observed in this population.

Among never married males aged 15-64 living in health areas with large gay male populations there was a 34 percent increase in total mortality, from 580 to 776 per year (Figure 2). The entirety of this increase is explained by AIDS, accounting for 33 percent of total mortality in 1984.

DISCUSSION

The impact of AIDS on patterns of premature mortality is profound. Among the two highest risk groups, gay men and intravenous drug users aged 15-64 in New York City, AIDS is the single leading cause of death and accounts for more than a third of all deaths. Even among groups not at high risk of AIDS, i.e. young women, AIDS has become one of the leading causes of premature mortality.

Accurate analysis of AIDS mortality is not feasible from AIDS surveillance or death certificates alone. Surveillance data underestimate both the number of AIDS cases and the number of cases reported that have died subsequently. Death certificate data are incomplete, because no single ICD underlying cause of death code, as none is highly sensitive and specific for AIDS. Thus, while the ICD-9 code is not due for revision until 1995, it may be possible to adapt our

FIGURE 1:

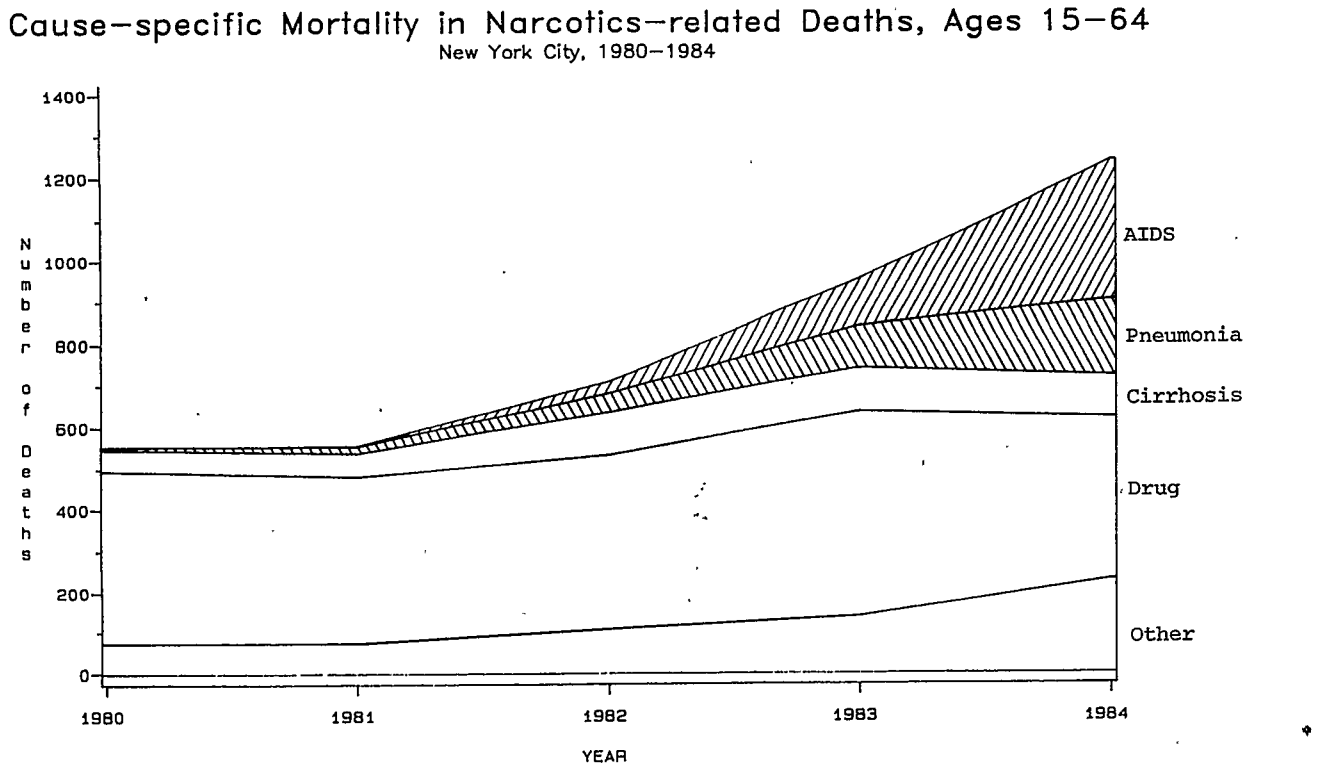
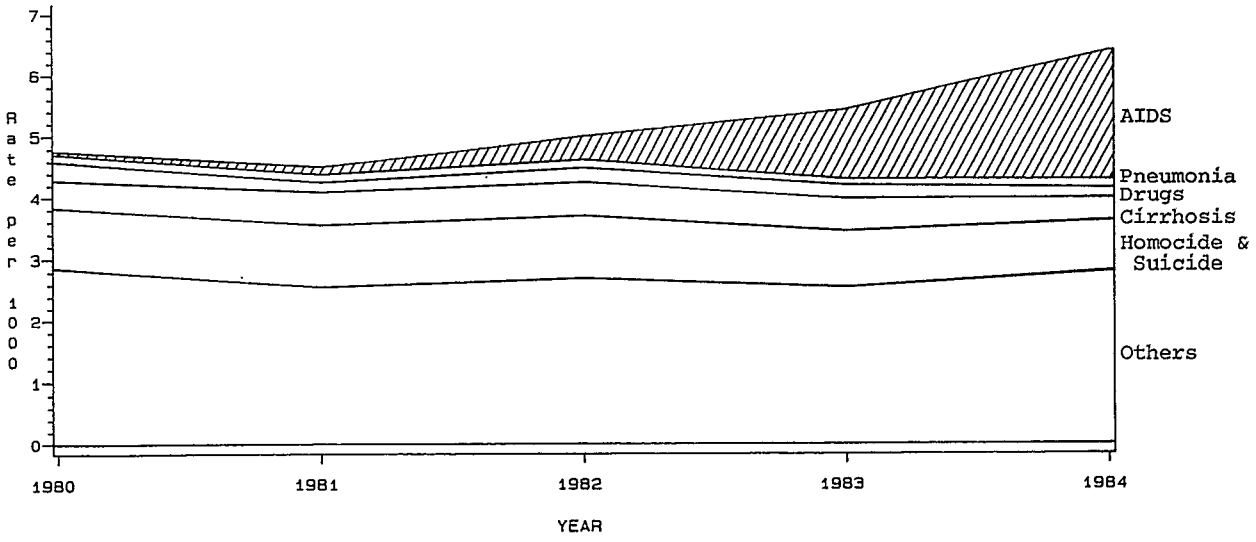


FIGURE 2:

Cause-specific Mortality in Never Married Males Living in Health Areas with Large Gay Male Populations*

per 1,000 ages 15-64, New York City 1980-1984



* Health areas with mean male rectal and pharyngeal gonorrhea rates over 500/100,000 males aged 15-54, 1978-1982

present systems of vital records analysis to reflect accurately the impact of AIDS. This probably requires additional research similar to the linkage reported here, and careful education of physicians and nosologists in the processing of AIDS death certificates.

References

1. Chamberland ME, Allen JR, Monroe JM, et al. Acquired immunodeficiency syndrome, New York City: Evaluation of an active surveillance system. JAMA 255:383-87, 1985.

2. New York City Department of Health Surveillance Team, The AIDS epidemic in New York City, 1981-1984. Submitted for publication.

3. International Classification of Diseases, 9th Revision. U.S. Dept. of Health and Human Services, DHHS Publication No. (PHS) 80-1260. Washington, DC: U.S. Government Printing Office, 1980.

4. Kristal AR, Milberg JA, Rutherford GW, et al. The epidemiology of viral hepatitis in New York City. Submitted for publication.

5. Jaffe HW, Bregman DJ, Selik RM, Acquired immune deficiency syndrome in the United States: The first 1,000 cases. J Infect Dis 148:339-45, 1983.

A MULTIPLE CAUSE-OF-DEATH ANALYSIS OF HYPERTENSION-RELATED MORTALITY IN NEW YORK STATE, 1968-1982

Edward Jow-Ching Tu, New York State Health Department

The present study uses mutiple cause-of-death data, i.e., records of all medical conditions listed on death certificates, to monitor the total contribution of hypertension to mortality in New York State over the period 1968-1982. Mortality rates for ischemic heart disease (IHD) and stroke will be presented since they are the major underlying causes of death that have received nationally, and it will be of interest to show how the trends in these mortality rates parallel the trends in the contribution of hypertension to total mortality.

Mortality data in the period 1968-1982 were coded to the 8th and 9th revision of ICDA. Hypertension is represented in the coding scheme under the special category for hypertensive disease in the 8th revision (codes 400-404) as well as specific four-digit codes under the broad three-digit categories for IHD and stroke(all codes 410.0 to 414.0 and 430.0 to 438.0 with a fourth digit of 0, as well as 412.1 and 412.2, two codes indicating hypertension that were added after initial publication of the ICDA). In the 9th revision (codes 401 to 405) there is no separate category for malignant hypertension. The new fourth digits for hypertensive disease specified as malignant or benign. The 8th revision fourth digits denoting the presence of hypertension in IHD and stroke no longer exist, making it impossible to show hypertension with those conditions for underlying cause tabulations. Furthermore, due to the lack of the proper comparability ratio for all mentions of hypertension between 8th and 9th revision, it makes the analysis of hypertension-related mortality for the period 1968-1982 more difficult. Instead of developing a proper comparability ratio for all mentions of hypertension, we divide 1968-1982 into two periods: 1968-1978 and 1979-1982 for which mortality data were coded under the same revision of ICDA and present two independent analyses on changes in the contribution of hypertension to mortality over these two periods.

FINDINGS

Hypertension

Figures 1 and 2 show the trend across the first 11-year and second 4-year periods for the age groups 40-44 and 75-79. The rate of hypertension mentions for non-white males ages 40-44 dropped from 105.5 in 1968 to 76.3 in 1978 and for non-white females from 103.0 to 70.5 over the same period. The drop was smaller for the period 1979-1982. The largest decline for ages 40-44 during the period 1968-1978 was experienced by white males (57.0%), followed by non-white females (46.1%) and non-white males (38.3%). White females of ages 40-44 showed a slightly increase. For age group 75-79, the largest decline was experienced by white females (59.5%), followed bt white males and non-white females. Non-white males experienced an increase. For the period 1979-1982, the largest decline for ages 40-44 was non-white of both sexes, followed by white females. White males showed an increase. For ages 75-79, whites of both sexes experienced a small decline, while non-whites showed increases. In general, the differences in the rate of hypertension mentions between races is larger than between sexes and the difference between races is larger at ages 40-44 than at 75-79.

Figure 3 shows that the age-adjusted death rate of hypertension mentions for white females tends to be higher than white males except 1975 for the 1st period. However, the pattern is reversed itself for the 2nd period. The rates of white males were higher than white females. For non-whites, the rate of males was always higher than females except 1981. The rate for non-whites was always higher than that of whites regardless of sex. The largest decline for the 1st period was experienced by white females, followed by white males, non-white males and non-white females; for the 2nd period, white males experienced 10.7% decline, white females 10.4%, however, for non-whites of both sexes the patterns were different, both groups experienced a 2 to 3 percent increases in the age-adjusted death rates for hypertension mentions.

Figures 4 to 7 show the age- and sex-specific rates of hypertension mentions for 1968,1978,1979, and 1982. The rates were greater for non-whites than for whites except males ages 75+ in 1968,80+ in 1978, females ages 80+ in 1968 and 85+ in 1978. It is also noted that male mortality was higher than females at younger ages but crossover at older ages. The race crossover is not observed in 1979-1982, but the sex crossover is still shown. That is, between 1979-1982, the rates of non-whites were always higher than whites.

Dramatic declines were occurred to younger white males, followed by younger non-white females, older non-white males and older white females. This pattern did not continue for the period 1979-1982. The largest decline was experienced by ages 54-59 of white males, with actual increases occurring at ages 40-44. Modest declines were registered in the older age groups.

The striking differences between the white and non-white rates are also evident in Figures 8 and 9, which present the ratios of non-white to white death rates for all mentions of hypertension. These ratios were largest for females at younger ages and decreased steadily with age for both sexes and both periods, with larger elimination in the 1st period than in the 2nd period. Between 1968 and 1978, the age-specific non-white/white female ratios decreased at ages 54 and below, but increased above ages 54, resulting in a 16.3% increase in the total age-adjusted ratio. The age pattern of male ratios were more uneven, with decreases at ages 55-59,65-69, and 80+. The total age-adjusted ratios for both sexes increased between 1968 and 1978, again between 1979 and 1982. The increase in ratio was smaller for males than fe-

males in both periods, that is, while non-white females lost ground relative to white females, non-white males improved slightly relative to white males during both periods.

IHD

Hypertension is a condition that is considered to be a common precursor of both IHD and stroke. Therefore, data for temporal trends in the underlying cause mortality rates for IHD and stroke are presented so that their similarity to and differences from the trends for hypertension mentions can be evaluated.

In Figures 10 and 11, which show the time trends in IHD death rates for ages 40-44 and 75-79,it can be seen that there is a greater similarity in the absolute levels of the rates by sex than by race, with females having lower rates than males. This pattern was clearer in 1968-1978 than in 1979-1982.

Whites of both sexes tended to have higher age-adjusted IHD death rates than non-whites in the 1st period, with a slightly higher rates of white females than non-white males. During the 2nd period, the rates were greater for males of both races than females. This pattern of the 1st period was quite different from that of the rates of hypertension mentions, which showed non-whites of both sexes having higher rates than whites (Figure 12).

The largest decline was occurred to non-white males, followed by non-white females, white females, white males for the 1st period. Like the changes in hypertension mentions for the 2nd period, the total age-adjusted IHD death rates for non-whites showed increases, with higher percentage. Total age-adjusted declines in IHD were less than hypertension declines for whites in both periods, but greater for non-whites in the 1st period.

Underlying cause death rates for IHD were higher than rates of hypertension mentions for all age, race/sex groups and entire period (1968-1982). Like hypertension, the rates were greater for non-whites than whites at younger ages (Figures 13 to 16). The race crossover started at younger ages for males than for females. Unlike hypertension mentions, male IHD mortality was always higher than females. There is no sex crossover effect at all.

As was observed for hypertension mentions, there were major differences in age-specific declines between race/sex groups. Dramatic declines in IHD death rates, unlike the declines in hypertension mentions, were uniformly concentrated in the younger ages, with non-white females experiencing the largest age-specific decline among all four race/sex groups, 101.3% at ages 50-54. In general, all four race/sex groups showed small to moderate decline at all ages except non-white females ages 85+ during the 1st period. The magnitude of declines were greater than that of declines in hypertension mentions during the 1st period. For the 2nd period, whites of both sexes declines in IHD death rates, similar to the declines in hypertension mentions, were concentrated in the younger ages, with white males experiencing the largest age-specific declines for any of the four race/sex groups. Comparing to the hypertension mentions, the declines were less, but the age pattern was very similar.

Like the non-white/white mortality ratio for hypertension mentions, the ratio for IHD declined with age for both sexes and in both periods. The ratio was slightly larger for females than males, but the difference was not as large as that in hypertension mentions. (Figures 17 and 18). Between 1968 and 1978, the age-specific male ratios decreased except ages 40-44,60-64, and 75-79. For female ratios, only four age groups showed decreases. The age-adjusted ratios for males and females indicated declines. The ratios for 1979 and 1982 indicated large increases in most age groups as well as in age-adjusted ratios.

Stroke

Absolute levels of rates of stroke as an underlying cause of death were more similar to mentions of hypertension than IHD rates, generally being slightly lower than the rates of hypertension mentions at younger ages and higher at older ages. In Figure 19, which shows the 15 years of rates for stroke at ages 40-44, the race difference in stroke mortality is evident. At ages 75-79 (Figure 20), there was a substantial convergence between the races in stroke rates. In fact, like IHD rates, the difference by sex was greater than by race.

Figure 21 indicates that age-adjusted stroke death rate for white males was less than for white females, and for non-white males, greater than non-white females except 1976 and 1978. This pattern was very similar to that of hypertension mentions in the 1st period. However, the pattern became very different in the 2nd period. For race comparison, whites had higher rates than non-whites except 1970,1972 and 1980-1982. This pattern was different from that of hypertension mentions, but similar to IHD.

Total age-adjusted declines for stroke were the largest for any of the three diseases in both periods. Differences were larger compared with IHD. Males experienced greater declines for stroke than females in the 1st period. White males and females experienced larger declines in the 2nd period. Unlike hypertension mentions and IHD, non-white males and females experienced declines, with non-white females experiencing the least declines among all four race/sex groups.

Figures 22 to 25 show that the rates were greater for non-whites than whites for almost all age groups except older ages. Underlying cause death rates for stroke were lower than rates of hypertension mentions for most age and race/sex groups except few older age groups.

Dramatic declines were occurred to younger ages in the 1st period. Declines by ages were generally more uniform during the 1st period than the 2nd period, with white males and non-white females having greater declines than white females and

non-white males. White male was the only group with an age pattern of stroke declines more consistent with those for hypertension mentions and IHD, with largest declines below ages 55, moderate declines at 55-79 and less declines at the oldest ages. During the 2nd period, the largest decline for white male was occurred to ages 50-54, for white females, ages 40-44, for non-white males, ages 80-84, for non-white females, ages 45-49. Like hypertension mentions, white females was the only group experienced decline at every ages. All other three groups show increases in stroke mortality rates in some age groups.

Figures 26 and 27 indicate that the age-adjusted non-white/white mortality ratios for stroke were smaller than that of hypertension mentions, but greater than that of IHD. The age pattern was similar to both hypertension mentions and IHD. The age-specific percent changes in the ratios between 1968 and 1978 indicated that males ages 60-64 was the only group showing a decline, for females ages 40-54 and 60-64 showing declines. This resulted in a 1.5% and 7.2% increases in the total age-adjusted ratio for males and females, respectively. The age-adjusted ratios increased to 9.5% for males and 10.3% for females between 1979 and 1982. The age pattern of changes in ratios was very uneven. Overall, the percentage changes in the age-adjusted ratios between 1968-1978 and 1979-1982 were very similar to those of hypertension mentions.

Data presented here strongly suggest that there has been a major reduction in the contribution of hypertension mentions to mortality in New York State over the 15-year period 1968-1982.

Declines in the mutiple cause hypertension death rates were generally more comparable to declines in underlying cause stroke mortality than IHD over the same period, especially in the 1st period. The declines in the age-adjusted hypertension rates were greater than declines in IHD but less than declines in stroke mortality. In this study, white females showed the largest total age-adjusted decline of all race/sex groups for hypertension mentions, white males for stroke, and non-white males for IHD during the 1st period. White males showed the largest total age-adjusted decline for all three diseases in the 2nd period. It may be that whites responded more favorably to reductions in blood pressure than non-whites. Of course, in New York State, non-white category is less homogeneous than that in most other states.

This study has presented multiple cause information on the total mentions of hypertension on New York State death certificates over the period 1968-1982. It is difficult to study secular trends for the period of 1968-1982 unless one can develop proper comparability ratios for condition codes instead of only for underlying causes.

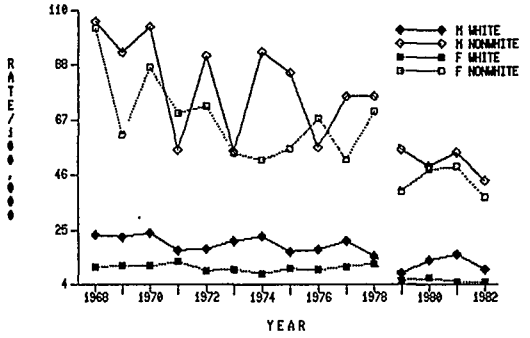


FIGURE 1. DEATH RATE FOR HYPERTENSION, NYS 1968-1982, AGE 40-44

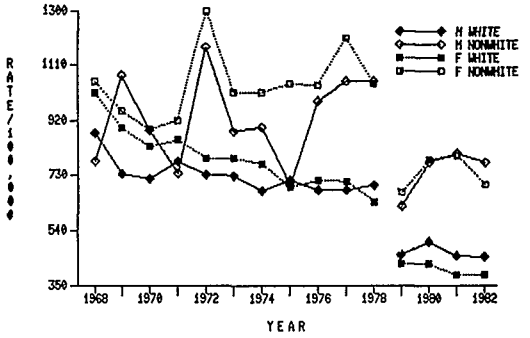


FIGURE 2. DEATH RATE FOR HYPERTENSION, NYS 1968-1982, AGE 75-79

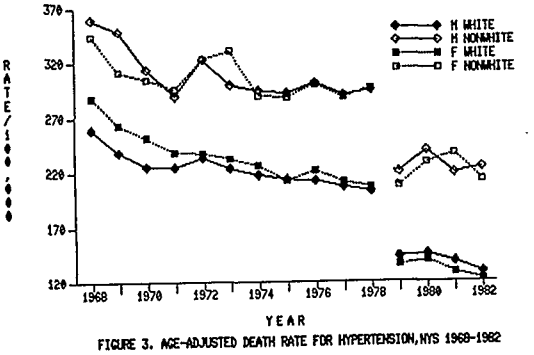


FIGURE 3. AGE-ADJUSTED DEATH RATE FOR HYPERTENSION, NYS 1968-1982

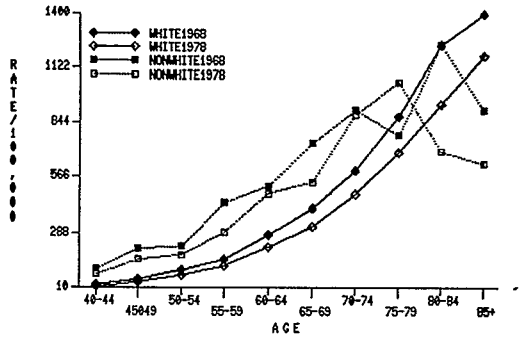


FIGURE 4. MALE AGE-SPECIFIC HYPERTENSION RATES NYS 1968 & 1978

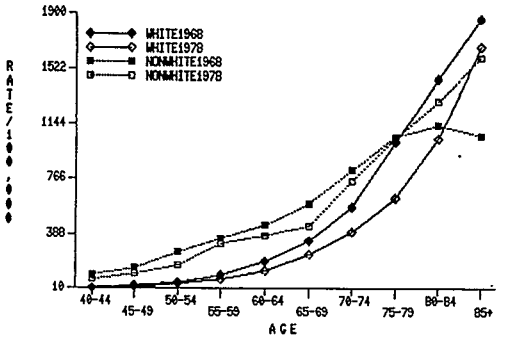


FIGURE 5. FEMALE AGE-SPECIFIC HYPERTENSION RATES NYS 1968 & 1978

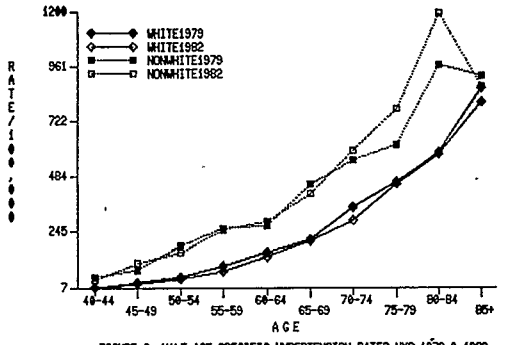


FIGURE 6. MALE AGE-SPECIFIC HYPERTENSION RATES NYS 1979 & 1982

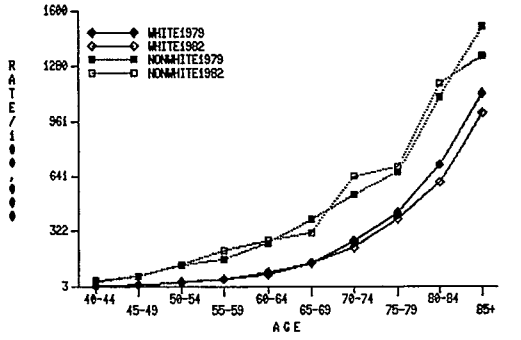


FIGURE 7. FEMALE AGE-SPECIFIC HYPERTENSION RATES NYS 1979 & 1982

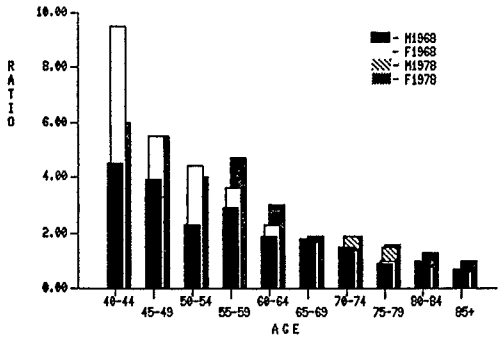


FIGURE 8. AGE AND SEX SPECIFIC NW/W RATIOS FOR HYPERTENSION NYS 1968 & 1978

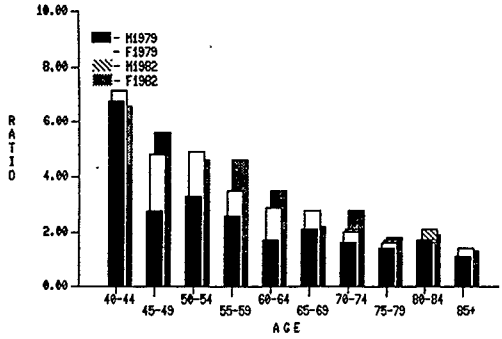


FIGURE 9. AGE AND SEX SPECIFIC NW/W RATIOS FOR HYPERTENSION NYS 1979 & 1982

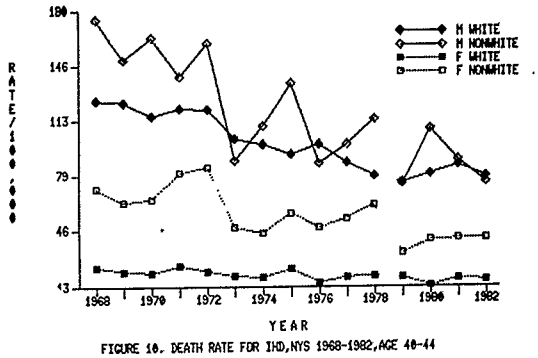


FIGURE 10. DEATH RATE FOR IHD,NYS 1968-1982,AGE 40-44

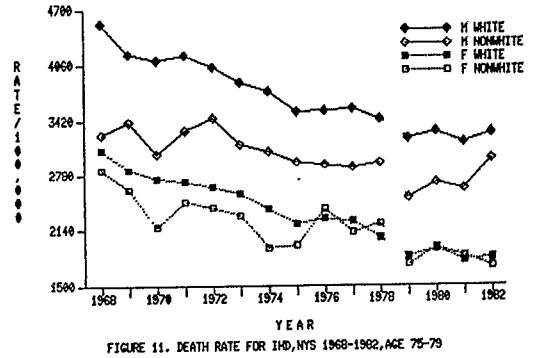


FIGURE 11. DEATH RATE FOR IHD,NYS 1968-1982,AGE 75-79

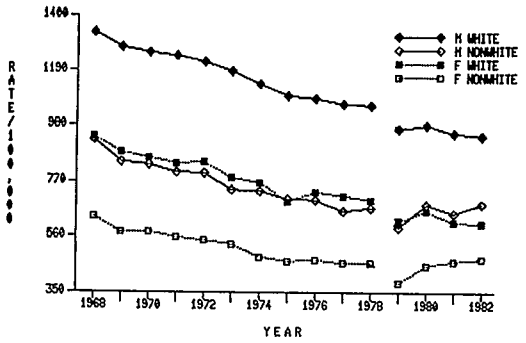


FIGURE 12. AGE-ADJUSTED DEATH RATE FOR IHD,NYS 1968-1982

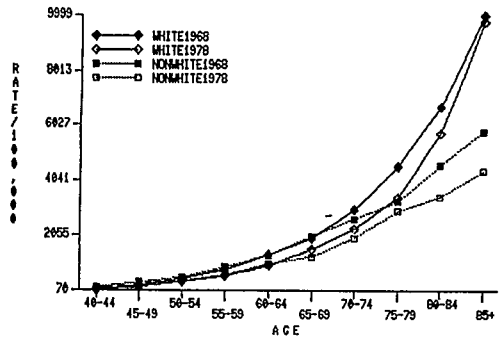


FIGURE 13. MALE AGE-SPECIFIC IHD RATES NYS 1968 & 1978

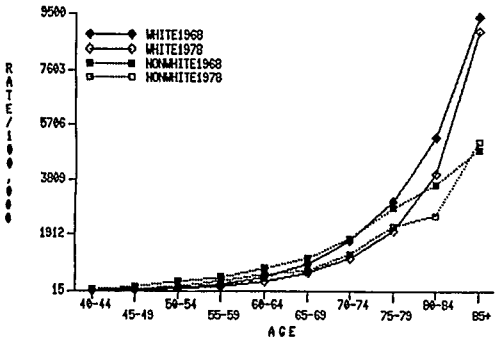


FIGURE 14. FEMALE AGE-SPECIFIC IHD RATES NYS 1968 & 1978

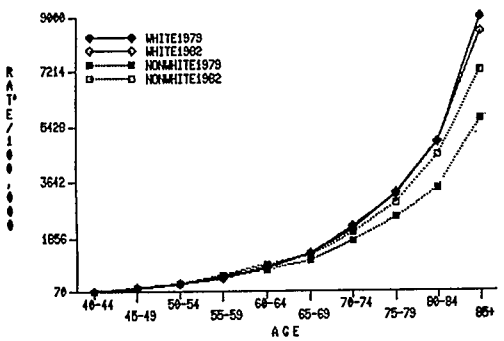


FIGURE 15. MALE AGE-SPECIFIC IHD RATES NYS 1979 & 1982

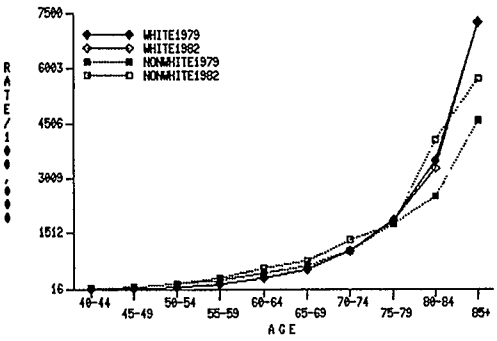


FIGURE 16. FEMALE AGE-SPECIFIC IHD RATES NYS 1979 & 1982

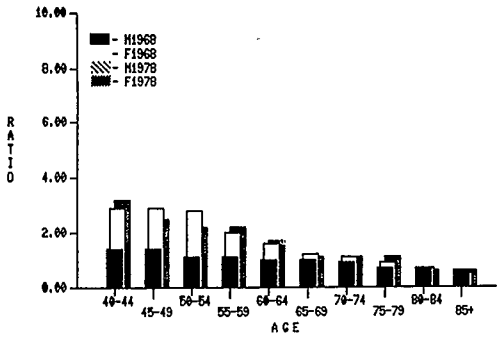


FIGURE 17. AGE AND SEX SPECIFIC NW/W RATIOS FOR IHD NYS 1968 & 1978

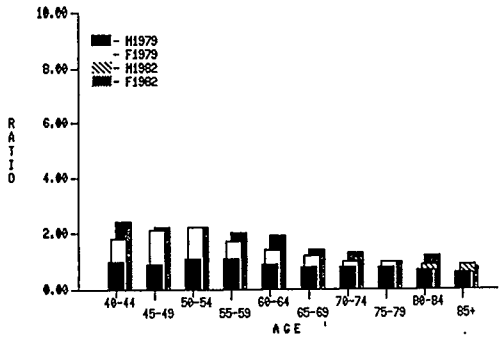


FIGURE 18. AGE AND SEX SPECIFIC NW/W RATIOS FOR IHD NYS 1979 & 1982

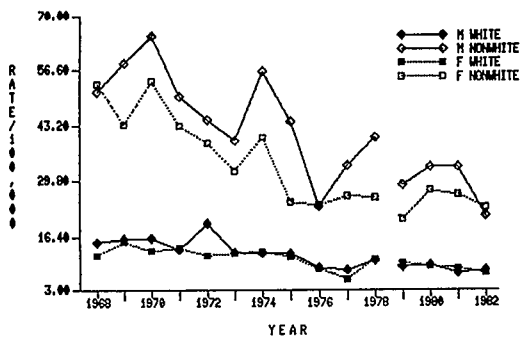


FIGURE 19. DEATH RATE FOR STROKE,NYS 1968-1982,AGE 40-44

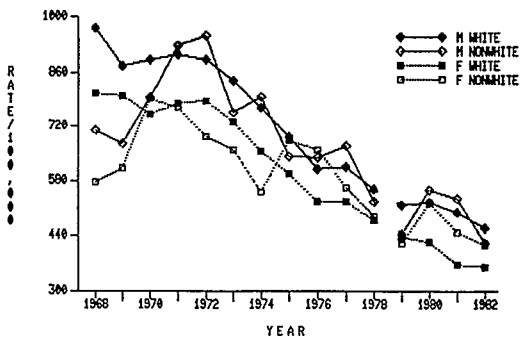


FIGURE 20. DEATH RATE FOR STROKE, NYS 1968-1982, AGE 75-79

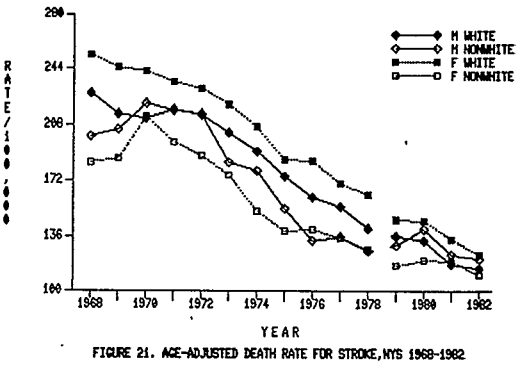


FIGURE 21. AGE-ADJUSTED DEATH RATE FOR STROKE, NYS 1968-1982

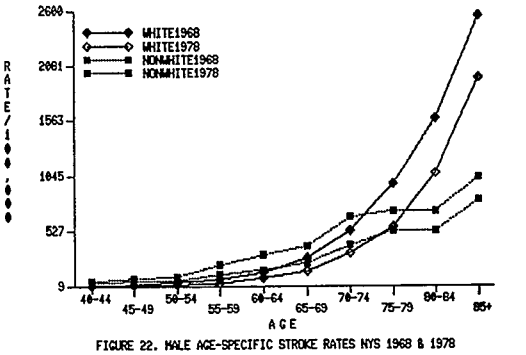


FIGURE 22. MALE AGE-SPECIFIC STROKE RATES NYS 1968 & 1978

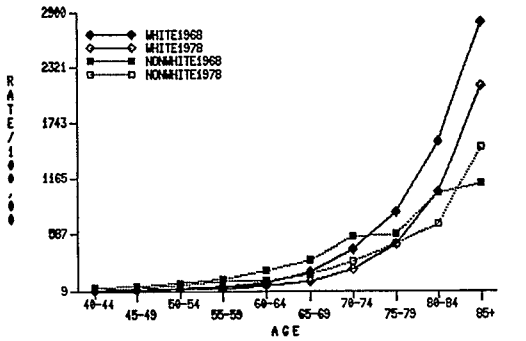


FIGURE 23. FEMALE AGE-SPECIFIC STROKE RATES NYS 1968 & 1978

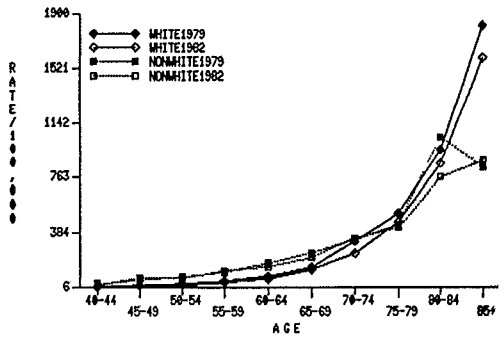


FIGURE 24. MALE AGE-SPECIFIC STROKE RATE NYS 1979 & 1982

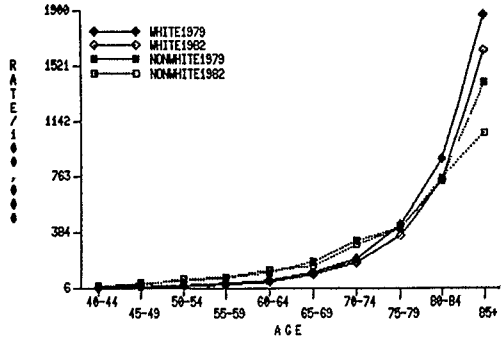


FIGURE 25. FEMALE AGE-SPECIFIC STROKE RATES NYS 1979 & 1982

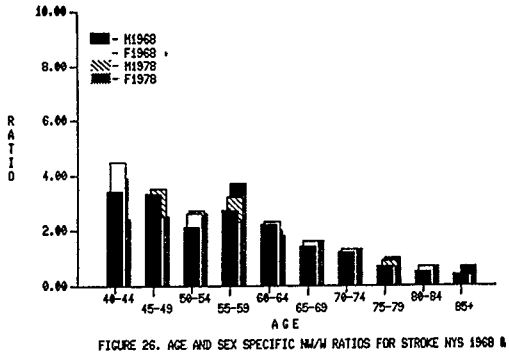


FIGURE 26. AGE AND SEX SPECIFIC NW/W RATIOS FOR STROKE NYS 1968 & 1978

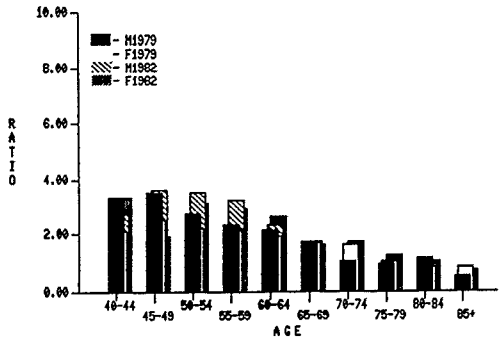


FIGURE 27. AGE AND SEX SPECIFIC NW/W RATIOS FOR STROKE NYS 1979 & 1982

MORTALITY PATTERNS AND PROJECTIONS BY EDUCATIONAL ATTAINMENT: UTAH, 1978-1982 AND 1990

Barry Nangle, John E. Brockert, and Marvin Levy, Utah Department of Health

Very little is known about the relationship between overall mortality patterns and levels of socioeconomic status. While there is some reason to believe that socioeconomic differentials exist with respect to mortality, there has never really been a national mortality data set in the United States with sufficient socioeconomic indicators to perform an analysis of death rates by socioeconomic status.

Existing data does suggest some hypotheses in this regard, however. We know that disadvantaged minorities, blacks, for example, have lower life expectancy than whites generally, and higher age specific mortality rates. Also, many occupations which carry risks for premature mortality are clustered at lower socioeconomic levels. Thus, while there is reason to hypothesize an inverse relationship between socioeconomic status and mortality, a thorough analysis of this topic for the United States awaits a mortality data set which has variables directly measuring socioeconomic status. The primary obstacle to the development of such a data set is the fact that the only source of mortality data is death certificates registered in the states, and only a few record any socioeconomic information.

Utah has recorded decedent's educational level since 1976 and can serve as a site for an initial exploration of the relationship between socioeconomic status and mortality. While we feel there is a need for basic research in this area, the analysis reported here has primarily a policy orientation. That is, from a public health point of view, the significant thing about Utah's educational distribution is that currently and in the near future it will be changing dramatically in the age groups most at risk for mortality. For example, among residents 75 and older in 1980, the majority (about 55%) had less than a high school education. In the age group right behind them, persons 65-74, only 41% were without a high school diploma, and for people 55-64 only 27% had completed less than 12 years of education.

To some extent, these rather large educational differentials by age reflect a growing importance of educational certification in society which occurred some years ago, and are not necessarily related changes over time in peoples life chances. At the same time we suspect that changing educational distributions do reflect some degree of change in the occupational mix, or class structure, of these cohorts. As such, these differentials should be related to the mortality patterns of these cohorts. Further, given the magnitude of these educational differentials over time, as these cohorts age we should see an impact on death rates in the state generally, both in terms of the the overall level of mortality and, more interestingly perhaps, in the mix of causal factors contributing to mortality rates. From a policy perspective then, the analysis of mortality by educational level should suggest trends in mortality rates due to rising educational levels, and causes of death likely to be of future significance.

The first step in the analysis was the calculation of education-specific death rates. Table 1 shows these data broken down by age, sex and cause of death for Utah residents 20 years and older, 1978-82. These rates were calculated using the educational distributions of decedents from Utah Death Certificates, and population by age, sex and education from the 1980 Public Use Sample tape distributed by the U.S. Bureau of the Census.

TABLE 1
Average Annual Death Rates
By Age, Sex and Years of Education
Residents 20 Years and Older: Utah, 1978-82

	Completed Years of Education				
Sex/Age	7 or less	8-11	12	13 or more	All Educ. Levels
ALL CAUSES					
Males					
20-44	244.7	287.6	232.2	129.0	179.6
45-54	555.5	662.9	653.0	449.9	556.0
55-64	1202.9	1284.8	1739.7	1213.1	1412.2
65-74	2049.9	2965.4	4395.0	3258.6	3407.6
75 +	6303.4	9050.8	15508.7	9022.2	9776.7
Females					
20-44	179.9	113.4	76.	65.0	76.3
45-54	363.2	329.4	322.	262.4	305.2
55-64	487.5	621.8	855.9	723.2	753.1
65-74	1009.4	1452.9	2286.5	1727.8	1796.9
75 +	4935.1	6972.8	10766.8	6367.2	7389.3
HEART DISEASE					
Males					
20-44	*	25.5	19.5	14.2	17.0
45-54	147.9	262.4	284.8	175.5	225.1
55-64	501.7	611.9	817.2	608.7	677.2
65-74	1004.7	1402.9	2217.0	1707.7	1698.9
75 +	3430.4	4905.0	8287.9	4986.8	5302.8
Females					
320-44	*	19.1	6.7	5.3	7.5
45-54	111.4	84.0	84.8	40.8	71.1
55-64	188.7	231.3	276.4	209.5	244.1
65-74	432.8	703.6	1018.0	802.9	825.8
75 +	2879.2	4217.0	6441.8	3840.8	4439.4

TABLE 1 (Continued)

Average Annual Death Rates
By Age, Sex and Years of Education
Residents 20 Years and Older: Utah, 1978-82

Sex/Age	Completed Years of Education 7 or less	8-11	12	13 or more	All Educ. Levels
MALIGNANT NEOPLASMS					
Males					
20-44	*	19.3	12.9	17.2	15.8
45-54	106.8	111.2	120.2	92.8	105.4
55-64	272.6	255.5	371.9	295.1	311.2
65-74	395.1	678.5	928.0	721.4	743.5
75 +	890.6	1311.0	2383.4	1411.8	1462.8
Females					
20-44	*	12.0	18.0	17.8	17.4
45-54	*	104.8	126.4	116.1	118.4
55-64	116.2	202.3	307.4	285.8	271.1
65-74	224.7	337.6	635.8	468.0	471.6
75 +	449.4	733.8	1317.1	809.8	844.0
MOTOR VEHICLE ACCIDENTS					
Males					
20-44	*	81.0	65.0	28.8	45.3
45-54	*	37.0	26.6	32.9	31.9
55-64	*	27.5	45.5	27.7	32.7
65-74	*	33.2	40.9	29.0	32.6
75 +	*	38.6	61.0	59.5	47.8
Females					
20-44	*	20.4	14.5	13.8	14.7
45-54	*	*	15.2	9.0	12.6
55-64	*	*	18.0	9.7	12.9
65-74	*	18.0	33.8	43.7	29.1
75 +	*	34.9	41.3	30.8	32.2
NON-MOTOR VEHICLE ACCIDENTS					
Males					
20-44	47.3	31.9	34.1	22.2	27.4
45-54	*	24.8	35.7	34.8	32.3
55-64	91.5	31.2	40.5	33.4	37.7
65-74	*	46.4	53.5	42.9	47.0
75 +	80.9	156.1	304.5	146.1	168.5
Females					
20-44	*	8.9	4.1	2.6	3.9
45-54	*	*	9.8	9.7	10.5
55-64	*	*	16.4	12.4	13.3
65-74	*	16.5	26.2	28.3	23.1
75 +	134.8	134.0	229.5	134.6	154.5
INFLUENZA AND PNEUMONIA					
Males					
20-44	*	*	2.5	*	1.7
45-54	*	*	8.8	*	8.6
55-64	*	18.3	25.5	*	17.2
65-74	65.2	57.6	59.1	68.3	61.3
75 +	505.8	461.2	697.3	436.2	505.6
Females					
20-44	*	*	*	*	0.9
45-54	*	*	*	*	4.9
55-64	*	*	13.2	10.6	11.7
65-74	*	31.3	50.9	35.5	39.3
75 +	309.2	355.2	420.0	238.6	334.1

TABLE 1 (Continued)

Average Annual Death Rates
By Age, Sex and Years of Education
Residents 20 Years and Older: Utah, 1978-82

Sex/Age	Completed Years of Education 7 or less	8-11	12	13 or more	All Educ. Levels
SUICIDE					
Males					
20-44	*	57.2	44.8	19.3	31.2
45-54	*	*	40.2	28.8	29.4
55-64	*	33.1	44.9	32.3	36.9
65-74	*	42.3	46.4	33.9	41.1
75 +	*	53.3	97.8	*	44.0
Females					
20-44	*	9.7	6.2	6.3	6.7
45-54	*	*	10.1	16.2	11.6
55-64	*	*	6.2	22.8	10.2
65-74	*	*	*	*	7.0
75 +	*	*	*	*	*
DIABETES					
Males					
20-44	*	*	4.5	4.2	4.2
45-54	*	*	*	7.1	7.5
55-64	*	18.4	19.3	22.8	20.9
65-74	*	48.1	60.1	70.0	55.4
75 +	99.9	151.1	296.9	157.9	171.1
Females					
20-44	*	*	2.5	2.7	2.4
45-54	*	*	*	*	*
55-64	*	14.1	26.2	28.6	25.4
65-74	73.1	61.4	68.9	49.7	61.9
75 +	123.1	194.0	327.7	141.1	200.8
CHRONIC LIVER AND CIRRHOSIS					
Males					
20-44	*	6.9	6.6	1.3	3.7
45-54	*	62.1	34.7	14.6	31.2
55-64	83.0	54.6	60.7	40.8	52.8
65-74	*	44.7	66.3	40.2	48.9
75 +	*	25.4	49.1	*	31.4
Females					
20-44	*	7.8	2.5	1.1	2.4
45-54	*	34.2	15.9	9.9	17.2
55-64	*	24.0	33.6	16.5	26.2
65-74	*	12.0	24.6	*	16.1
75 +	*	14.9	42.0	*	19.6
CHRONIC OBSTRUCTIVE LUNG					
Males					
20-44	*	*	*	*	0.6
45-54	*	23.8	8.6	*	9.7
55-64	*	75.1	81.5	23.2	57.8
65-74	212.5	243.5	353.4	159.4	254.3
75 +	293.6	523.5	896.9	295.4	499.3
Females					
20-44	*	*	*	*	*
45-54	*	*	*	*	3.9
55-64	*	31.5	31.1	19.2	26.6
65-74	87.9	48.8	78.6	48.3	61.9
75 +	85.7	75.5	131.8	78.6	89.4

TABLE 1 (Continued)

Average Annual Death Rates
By Age, Sex and Years of Education
Residents 20 Years and Older: Utah, 1978-82

Sex/Age	Completed Years of Education 7 or less	8-11	12	13 or more	All Educ. Levels
RESIDUAL					
Males					
20-44	110.0	59.6	41.0	21.1	32.6
45-54	103.9	104.0	86.9	54.0	74.9
55-64	83.9	159.4	232.7	122.5	167.9
65-74	234.9	368.5	570.2	385.7	424.6
75 +	909.4	1425.8	2433.8	498.0	1543.3
Females					
20-44	84.9	32.3	21.0	14.7	20.2
45-54	*	53.1	55.1	53.1	55.0
55-64	74.5	89.9	127.4	108.0	111.7
65-74	143.0	215.9	342.7	230.6	261.1
75 +	908.6	1213.5	1815.8	1073.3	1270.6

Note: Education not stated distributed.
* Rates not computed on less than seven events.

Two patterns predominate in the data in Table 1. The first is found in deaths to younger persons, 20-44 year old males and both 20-44 and 45-54 females. Here mortality rates are clearly inversely related to educational attainment. This was the pattern we expected to find throughout the age cohorts. As age increases, however, a second pattern emerges in which mortality rises with educational level, peaking at high school education, then declining in the more than high school group. Aside from accidents and suicides there are very few exceptions to these two general patterns found in these data.

One hypothesis for these unexpected results is under reporting of lower educational levels on Death Certificates. Cases in which education is not stated were distributed for the calculations in Table 1, and for some cohorts as much as 15% of decedents did not report educational level. If much of the non-reporting was at lower education, and this seems to us quite possible, then the seemingly direct relationship between mortality rates and education could be an artifact. We may be able to address this question soon, since reporting of education an death certificates has improved markedly over the past few years, from about 20% not stated in 1976 to less than 10% in 1983.

The second phase of the analysis was projection of the causal mix of mortality for 1990 based on the educational differentials found in 1980 mortality. The projection method employed was essentially a cohort survival technique designed to reveal the cause specific mortality rates which should prevail given the aging of more educated cohorts. Specifically, we:

(1) Survived the age-education cohorts of 1980 ten years to arrive at 1990 population by age, sex and education. Age-sex-education specific death rates for 1980 applied to five year age groups were used to calculate the number of 1990 survivors.

(2) Assumed that 1980 age-sex-education-cause specific death rates would be the same in 1990.

(3) Estimated 1990 deaths based on these rates and the new, more educated population.

(4) Summed over age groupings to get 1990 cause-specific mortality rates by sex.

The results of these calculations are displayed in Table 2. Shown separately are the effects of simply aging the population and aging the population within educational attainment groupings. Note that the overall projection is, in most cases, for a much reduced mortality rate. The exceptions are suicide and motor vehicle accidents, both sources of mortality projected to increase in significance by 1990. Note also the separate effects of changes in age and education are for the changed age structure to lower death rates substantially, and for this change to be offset somewhat by the changed educational distribution. Death rates will necessarily decline in Utah in the near future simply because we are a very young state with a high birth rate, and a growing youthful population not at high risk for mortality. Indeed, the effect of this changing age structure tends to overwhelm the projected changes in mortality patterns due to our changing educational distribution.

The effect of rising educational levels raising expected death rates in nearly every cause is clearly due to the high number of reported deaths to high school graduates that was discovered in our earlier analysis, and the expected growth of this educational group by 1990. Of course, if the education-specific death rates arrived at in this analysis are significantly colored by non-reporting of education by lower socioeconomic level decedents, then the age/education related projections reported in Table 2 may be misleading.

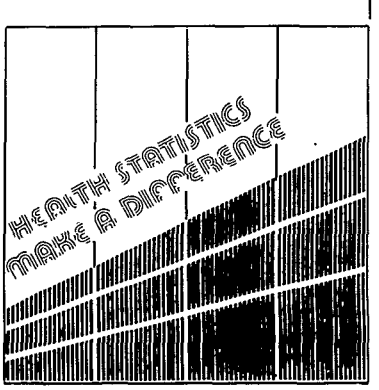
This analysis underscores the importance of accurate and complete reporting of vital records data. As more states consider the addition of socioeconomic indicators on certificates, it is important to keep in mind that the more completely the items are reported the more useful is the data for subsequent health policy analyses. In Utah, we are hopeful that as we improve our collection of educational attainment data on Death Certificates we will be able to make a contribution to the understanding of the effect of socioeconomic differentials on mortality patterns.

TABLE 2
Cause-Specific Mortality Rates, Actual and Projected, Based on 1980 Education-Cause-Specific Rates, and Projected Changes in Age and Education Distributions. Residents 20 Years and Over: Utah, 1978-82 and 1990

Cause	1978-82 Deaths	Actual Rate	1990 Projection Based on: Age Change Deaths	Age Change Rate	Age/Educ Chg. Deaths	Age/Educ Chg. Rate
MALES						
All Causes	20,339	969.7	5,420	814.9	5,469	822.4
Heart Disease & Stroke	9,295	443.2	2,393	359.9	2,425	364.6
Malignant Neoplasms	3,553	169.4	919	138.2	930	139.8
Motor Vehicle Accidents	865	41.2	280	42.1	286	43.0
Non-motor Vehicle Accidents	755	36.0	229	34.4	231	34.7
Influenza and Pneumonia	584	27.8	155	23.2	153	23.1
Suicide	688	32.8	216	32.4	220	33.0
Diabetes	349	16.6	95	14.2	97	14.5
Chronic Obstructive Lung	960	45.8	246	37.0	240	36.0
Chronic Liver and Cirrhosis	363	17.3	95	14.2	91	13.7
Residual	2,927	139.6	793	119.2	797	119.8
FEMALES						
All Causes	16,623	754.3	4,567	651.2	4,677	667.0
Heart Disease & Stroke	8,123	368.5	2,239	319.3	2,295	327.2
Malignant Neoplasms	3,219	146.0	853	121.6	884	126.0
Motor Vehicle Accidents	363	16.5	114	16.2	115	16.5
Non-motor Vehicle Accidents	357	16.2	102	14.5	104	14.8
Influenza and Pneumonia	555	25.2	156	22.2	157	22.4
Suicide	169	7.7	52	7.4	53	7.5
Diabetes	500	22.7	128	18.3	131	18.7
Chronic Obstructive Lung	312	14.2	79	11.3	81	11.5
Chronic Liver and Cirrhosis	204	9.3	54	7.7	54	7.7
Residual	2,821	128.0	790	112.6	805	114.9

Special Session 1 (NIMH)

Analysis of National Data on Mental Health Services in the United States: An Update



ANALYSIS OF NATIONAL DATA ON MENTAL HEALTH SERVICES IN THE UNITED STATES: AN UPDATE

Ronald W. Manderscheid, Michael J. Witkin,
Marilyn J. Rosenstein, and Rosalyn D. Bass
National Institute of Mental Health

Summary

This symposium examined trends in mental health services for the United States as a whole, based upon data collected by the National Institute of Mental Health (NIMH). The configuration and characteristics of specialty mental health organizations and the patients they serve were examined for the period between 1970 and the present. For each of these areas, specific uses of the data and improvements in the design and content of national data collections were highlighted. Changes in the locus of mental health services were discussed, with attention to proposed new data collections. All information was derived from the National Reporting Program (NRP) in mental health statistics, a system based upon voluntary reporting of data, operated by NIMH in close collaboration with the State mental health agencies.

Organizational Trends and Characteristics (Michael J. Witkin)

This presentation discussed trends in the delivery of mental health services over the past 25 years, with emphasis on the period between 1971 and 1983. The data were derived from information collected through biennial inventories of all specialty mental health organizations and general hospitals with separate psychiatric services. Findings presented include comparisons of the distribution of patient care episodes, staffing, and expenditures.

Prior to 1960, inpatient services in State mental hospitals were the primary setting for the care of psychiatric patients. In 1955, 77 percent of patient care episodes occurred in inpatient settings. With the advent of the community mental health center movement in the early 1960s, the growth in outpatient and day treatment episodes proliferated. In 1983, outpatient care episodes comprised 74 percent of all episodes. Between 1971 and 1975, outpatient care episodes increased from 55 to 70 percent of all episodes. However, between 1975 and 1983, the proportion of total episodes that were in outpatient settings leveled off between 70 and 74 percent.

With regard to staffing, a steady increase occurred in the proportion of full-time equivalent (FTE) staff who were in the four-core mental health disciplines (psychiatrists, psychologists, social workers, registered nurses). In the aggregate, these disciplines comprised 28 percent of total FTEs in 1983, as compared with 19 percent in 1971. By contrast, other patient care staff decreased from 46 percent to 42 percent of all FTEs in this period.

Expenditures by mental health organizations increased from $3.8 billion to $13.2 billion between 1971 and 1983. During this same time span, expenditures by State and county mental hospitals as a proportion of total expenditures decreased steadily from 62 percent in 1971, to 47 percent in 1979, to 42 percent in 1983.

Patient Trends and Characteristics (Marilyn J. Rosenstein)

Patients served in the specialty mental health organizations surveyed by the NRP were the focus of this presentation. NIMH has periodically collected data about patient characteristics through a sample survey program which began in 1969. The most recent patient sample surveys, completed in 1980-1981, collected data on admissions to inpatient psychiatric services. This presentation compared 1980 and 1970 data on the patient characteristics of sex, race, age, diagnosis, and length of hospital stay (LOS) for inpatient admissions to State and county mental hospitals, private psychiatric hospitals, and the separate psychiatric services of public and private non-Federal general hospitals.

The surveys of State and county mental hospitals and private psychiatric hospitals centered on a sample of admissions from 1-month who were followed for an additional 3-month period. The general hospital survey centered on discharges during a 1-month period. All data were inflated to represent annual estimates, adjusted to known totals. Because the LOS in general hospitals is short, the characteristics of admissions and discharges during any 1-year period were essentially the same. Thus, patients from all three surveys were referred to as admissions.

As the mental health service delivery system evolved over time, many changes took place in the location and type of services provided to patients. In parallel, the presentation compared the characteristics of the patients to see what changes have occurred in the types of people actually being admitted for care. Over the 10-year period between 1970 and 1980, increases occurred in the number of admissions to private psychiatric hospitals and the separate psychiatric services of private general hospitals, and decreases occurred in the number of admissions to State and county mental hospitals and the separate psychiatric services of public general hospitals, such

that, by 1980, the number of admissions to private general hospitals exceeded that of State and county mental hospitals. The trend is inpatient admissions over the 10-year period was a shift away from public facilities toward private facilities.

Between 1970 and 1980, several differences were observed between public and private facilities. In general, higher percentages of males, minorities, and patients diagnosed with schizophrenia were admitted to public facilities. However, the composition of the 1980 incoming caseloads of the separate psychiatric services of public general hospitals was more similar to private facilities, so that State and county mental hospitals differed more from the private facilities than did the public general hospitals.

Of particular interest was a comparison of the amount of time that inpatients received careonce admitted. Over the 10-year period, only minor changes occurred in the LOS for admissions to private psychiatric hospitals and the separate psychiatric services of general hospitals, but a considerable change took place in State and county mental hospitals. While State and county mental hospital admissions had the highest median LOS in both years, LOS decreased from 41 days in 1970 to only 23 days in 1980. Many factors contributed to these differences in LOS. The remainder of the presentation discussed the relationship of patient age and diagnosis to LOS. Because the major change in LOS occurred in State and county mental hospitals, the analysis highlighted these facilities.

Patient age appeared to be related to LOS. In 1970, children and youth had the longest median LOS. Although the median LOS decreased over time for all other ages, LOS remained constant for those in the 65 and older group, LOS remained constant, so that by 1980 the elderly had longer stays than children and youth. Diagnosis also appeared to be related to length of stay. Of the three major diagnostic groupings (schizophrenia, affective disorders, and alcohol-related disorders), those admissions with schizophrenia had the highest median LOS in 1970 and 1980; however, the median LOS decreased for all three diagnostic groups over the 10-year period. A comparison of the median LOS for different age groups with schizophrenia indicated an interaction between age and diagnosis with respect to median LOS. The median LOS of the under 18, 18-24, and 25-44 age groups decreased over the 10-year period, whereas, the median LOS for the 45-64 age group remained about the same, and the median LOS for the 65 and older group actually increased.

These data showed some changes in the composition of patients admitted to mental health inpatient facilities, but they also indicated a fair degree of stability over time in the characteristics of patients admitted to the different types of facilities.

In order to examine similar data for patients from other types of facilities, for patient groups other than admissions, and for services other than inpatient care, the presentation also reported on NIMH plans to conduct a survey in 1986 that will sample patients from inpatient, outpatient, and partial care programs of all the specialty mental health organizations covered by NIMH. Admissions, terminations, and patients under treatment during a 1-month period will be included in this survey. Data from this survey effort will enable us to begin to address issues of services to patients in a wide range of mental health programs.

Unmeasured Development in Services
(Rosalyn D. Bass)

The NRP of the NIMH currently focuses on the mental health resources, services, and persons treated in organized mental health settings, i.e., in the specialty mental health sector. In spite of the many different types of mental health organizations reporting into the NRP on a voluntary basis, this program falls short of reflecting the Nation's de facto mental health service delivery system because mental health services today are also provided by trained mental health professionals in settings other than the specialty mental health sector. For example, they are provided in the educational sector, the criminal justice sector, the military, in industrial settings, in community residential facilities (CRFs), etc.

Two factors have operated historically to limit the NRP perspective on the Nation's mental health service delivery system: (1) legislative restrictions on the Federal Government with respect to collecting mental health data from individuals in the community; and (2) the decentralization and differentiation of the mental health service delivery system.

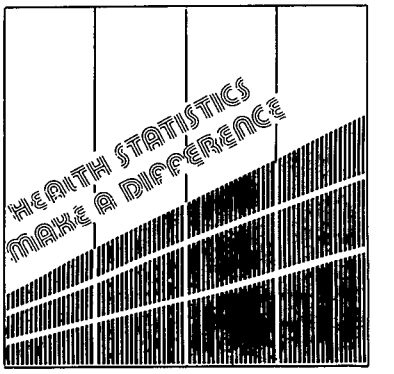
The history of the NRP can be traced back to the decennial census of 1840 which sought to identify "insane" and "idiotic" persons among those being counted. Legislation enacted in 1902 limited census collection of data on the mentally ill and retarded to those residing in institutions.

Although the mental health service delivery system was largely a monolithic system of institutions in 1902, it is not so today. Since then, the mental health service delivery system has been decentralized and differentiated not only into different types of mental health organizations (e.g., outpatient clinics, day/night facilities for the mentally ill, community mental health centers, etc.) and different types of mental health treatment (e.g., inpatient, partial, outpatient, residential, and emergency care), but also it has begun to develop in sectors such as criminal justice, education, CRFs, etc.

NIMH has undertaken to try to broaden the current scope of the NRP by seeking to obtain data on mental health services and resources located in the criminal justice system and in CRFs. The most immediate plans call for a survey of mental health services and resources in State adult correctional facilities, with longer range plans for a similar survey of county and city jails. Current plans also include a first step in preparing to survey CRFs by developing a taxonomy of CRFs from which it would be possible to delineate CRFs that would be considered part of the Nation's mental health service delivery system. Longer-range plans for a survey of CRFs and for a survey of mental health services and resources located in the universities and colleges of the Nation are currently being developed.

Session P

The Use of Personal Computers in the Analysis of Health Data



The Micro-to-Mainframe Connection: Accessing Mainframe Health Care Data for Individual Microcomputer Use

Allan M. Miller, Inland Counties Health Systems Agency

Micro-mainframe communications is one of the newest and most rapidly developing areas in the Information Systems Management field. The purpose of this paper is to discuss the uses of this new technology in accessing and processing statistics for health planning/issue analysis activities.

I. Why Make the Micro-Mainframe Connection?

The most obvious reason for linking-up to mainframes is that health care data sets are usually quite large and are only available on mainframe computers. Discharge data, hospital financial and utilization data and vital statistics are generally compiled and stored on magnetic tape at central computing facilities. In order to electronically access these data sets, you must somehow "plug into" a mainframe or minicomputer system.

A natural question at this point would be: "if the data is only available on mainframes, then why use microcomputers anyway?" For users of health statistics, understanding this is the key to understanding the uses of micro-mainframe communications technology.

The "microcomputer revolution," by lowering the price of processing power, has made computer technology accessible to millions. Microcomputers allow users the benefits of:

* immediately accessible processing capabilities;
* powerful, yet easy-to-operate software packages designed for use by non-computer professionals;
* the ability to store, process and retrieve information, literally at will.

No longer must users of health statistics wait for costly batch jobs to be run by programmers at remote computing facilities. Microcomputers allow for direct control over the data processing environment. Using microcomputer software packages such as spreadsheets and databases with query capabilities, the analyst may now process and re-process health care data sets many times with the results being available almost instantaneously.

While microcomputers have advantages in obtaining control over a data processing environment, they also have their limitations in speed, storage and processing power. For example, discharge data sets are so large and complex that processing them with microcomputers would be literally impossible.

Significant benefits are thus obtained by using micros and mainframes together, within one data processing environment. The process may best be described as one where information is "squeezed" from a large data set on a mainframe system and then electronically transferred to a microcomputer where it can be easily and quickly processed.

III. Technical Aspects of Micro-Mainframe Communications

Because of the tremendous variety of computer systems and the different purposes to which micro-mainframe communications may be used, it is impossible to discuss here the detailed technical aspects of linking up micros to mainframes. What follows below is a brief discussion of the technical issues to be considered when attempting micro-mainframe communications.

Physically connecting a microcomputer to a mainframe may be as simple as hooking up a cable from one system to the other; however, setting up a connection where two systems can actually "communicate" and operate in a useful fashion can often be technically complicated and generally difficult to achieve. Perhaps the first question that should be asked is: what is the basic purpose in making the micro-mainframe link?

Using an office microcomputer to work and store results on a mainframe located at another facility is an example of using a microcomputer as a "remote terminal workstation." This type of micro-mainframe connections is often referred to as "remote terminal emulation," in a sense the mainframe computer is made to "believe" that the microcomputer is nothing more than another one of it's on-site terminals. The more useful type of connection allows the "downloading" or "uploading" (i.e. transfer back and forth) of data files between the two systems. In this case, the microcomputer is able to send and receive files to and from the host, using it's own processing capabilities interactively with the mainframe computer.

Once the physical link is properly established, correct communications protocol must be used. Protocol, as used in computer communications, is nothing more than a set of rules for how information is exchanged over a communications line. The parameters that must be determined include the transmission speed (usually 1200 baud), signals signifying the starting and stopping of transmission, the length of transmission "words" and methods for error-checking and validation of signals.

There is no standard protocol used by all computers. However, there are several that are widely used on most mainframes such as "Xmodem" and "Kermit" (named after the famous "Sesame Street" character). Kermit was developed at Columbia University and is widely available in the public domain for most computer systems. Using file transfer programs such as these allows users of micros to send files back and forth to mainframe or mini systems, or for file transfer between microcomputers themselves.

IV. Issues in Micro-Mainframe Use of Health Statistics

The above issues in micro-mainframe communications are best brought out using real-world examples. For example, discharge data may be analyzed on a microcomputer used as a remote workstation. Or an on-site hospital financial and utilization database may be set up from a statwide mainframe data set. In either case, the steps would be as follows:

First, the relevant data subset must be "stripped out" from the larger data set. Figure I shows how discharge data are stored on a typical mainframe system on magnetic tape. Information for each discharge record from an acute care facility includes:

* hospital facility number;
* patient age, sex, race and zip code;
* length of stay;
* principal procedure and diagnosis;
* Diagnostic Related Group;
* amount of charge, and
* source of payment.

Each piece of information occupies a uniform, specified field within each record, making it readily available for processing. The same type of format is also used for the financial and utilization data set, including such information as:

* patient days;
* capital expenditures;
* gross patient revenue;
* expense per patient day, and
* total operating expense.

A typical first step might be to abstract all of the discharge and hospital records for a particular county, for example, to form a basic data subset that is more manageable in size. This might be accomplished using a high-level programming language such as FORTRAN or a mainframe statistical package such as SAS. In either case, the most important factor to consider is the processing time necessary to complete the job, which sometimes can be expensive (the California Hospital Discharge data set for 1983 contains over nine million records alone!).

Second, the storage media for the resulting data subset must be chosen. Cost issues may also govern how much and where to store a data subset. Often times, the discharge records for one large or several small counties can be prohibatively expensive to store on hard disk at the remote (mainframe) site, and too large to store on a micro. In this case, magnetic tape must be used once again as the primary storage medium. A smaller data subset may be stripped out and stored on disk at the remote site or on the microcomputer. The advantage to this type of storage is that the data sets are more accessible than magnetic tape which must be mounted at the remote site every time it is used.

It sometimes pays to store frequently used data sets (such as population statistics) right on your microcomputer system. On the other hand, as in the case of the discharge data, magnetic tape may be the best media for storage with processing being done in batches running several jobs with each pass-through of a tape.

Up to this point, the processing of the mainframe data as described has been done exclusively by means of remote terminal emulation and batch processing from a microcomputer. The other, often more useful way of processing mainframe data is to put it in a form where it is usable directly by a microcomputer itself.

Setting up a hospital financial and utilization database using an on-site microcomputer would initially involve the same steps as above. The relevant records must be stripped from a data tape run on a mainframe system.

Unwanted fields might then be removed from the data subset itself to make the working data file more manageable. Using a file transfer program based on the Xmodem or Kermit protocols, the resulting file could be downloaded to a microcomputer in preparation for inputting it into an on-site microcomputer spreadsheet or database program. In order to do this, the data must be put into the proper format for loading it into the program. For example, programs such as Multiplan and Lotus 123 have specific internal formats for storing their worksheet files. The technical specifications for getting the downloaded data into the microcomputer spreadsheet or database program are usually contained in the software instruction manual.

Originally an embedded fragment of a large and inaccessible mainframe data set, the hospital financial and utilization data would now be directly accessible to the microcomputer user. Further data processing would be possible on the micro, with "what-if" scenarios and particular data queries being almost immediately accessible. All because of successful use of the micro-mainframe connection.

ON-LINE DATA ACCESS - THE MEDICAID WORKSTATION

Embry M. Howell, SysteMetrics
David K. Baugh, HCFA
Penelope Pine, HCFA
Anthony Pepitone, SysteMetrics

Many health care data bases, particularly health care claims and discharge abstract files, are extremely large and difficult to manipulate. Analysts may be required to wait for days or weeks to receive computer analyses which they have requested, due to the time lag which is often involved in processing such large data sets. The advent of micro-computer technology has provided the tools to quickly process small data sets, but it has been difficult to take advantage of this low-cost, new technology for many health applications due to the volume of data which must be analyzed.

A joint project of SysteMetrics, Inc. and the Health Care Financing Administration has led to the development of the Medicaid Workstation which links a micro-computer to a mainframe computer. Some information is processed on the micro-computer, while large volumes of data are processed on the mainframe. Data are transmitted between the two systems via telephone lines. The Medicaid Workstation is a spin-off of the Hospital Workstation which SysteMetrics developed for processing hospital discharge abstract data and other large hospital data files.

The data base which the Medicaid Workstation is designed to analyze has been constructed at the federal level by the Office of Research of the Health Care Financing Administration. It contains complete Medicaid claims, eligibility, and provider files from five states (California, Georgia, Michigan, New York, and Tennessee) for several years beginning in 1980. The data have been provided voluntarily by the states to HCFA for use in research. The project is known as the Medicaid Tape-to-Tape project.

Medicaid files are extremely voluminous, especially for the largest states. For example, in 1982 over 26 million claims were received from Michigan, over 58 million from New York, and over 87 million from California. The timely analysis of such a large data base presents great challenges. The purpose of developing the Medicaid Workstation was to provide Medicaid researchers with a means of analyzing this enormous data set quickly, without long data processing delays.

Several steps were taken to reduce the volume of data initially. Records were aggregated into "person records". Individual claims records were summarized to provide counts of visits, stays, and dollars spent by service type for each person eligible for Medicaid during the year. In addition, since the institutionalized population is of high interest to HCFA, an extract of persons in one state who were ever in institutions during the year was prepared. This file became the "base" file for initial analyses. This file contains 50,000 records and resides on disk storage at the mainframe computer site. Storage on disk provides the opportunity for more rapid turnaround, since the need for mounting tapes is avoided.

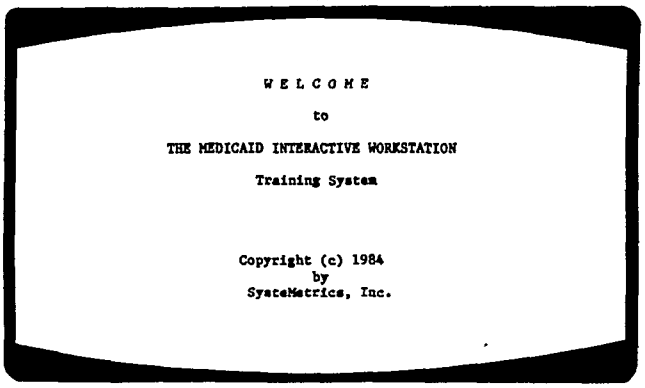
The Workstation operates as follows: a set of screens are stored on the micro-computer. These are called up and critical information is requested from the analyst about the files which are to be analyzed, the variable names, any recodes which are necessary, and the types of analyses to be produced. The information in these screens is translated by the micro-computer into SAS statements. These statements are transmitted to the mainframe computer. The data are processed (using SAS) on the mainframe and results are downloaded to the micro-computer for viewing and storage. The aggregated results can be further analyzed using micro-based software such as LOTUS.

There are two types of screens which request information from the analyst. The first is a menu, in which a selection is made from a list. The second is a form which must be filled in. A typical session on the Workstation involves dialing up to the host computer (this is quite easy because all critical information such as the telephone number, user code, etc., is stored on the micro-computer). The analyst may then want to pass through the base file to further subset the file for analysis. The smaller the file which is used, the shorter the turn-around time will be for receiving results. Then a series of tables may be requested from this smaller file.

These steps are illustrated below in a sample session which illustrates an analysis of psychiatric care for children. After logging on, the base file of all institutionalized persons is subsetted to include only persons under 21 years of age. An age variable is created which categorizes people by five-year age intervals. A table is then requested which displays psychiatric expenditures by sex and age category. This table is downloaded to the micro-computer. This sample session takes about 10 minutes including the time required to pass the file and tabulate data.

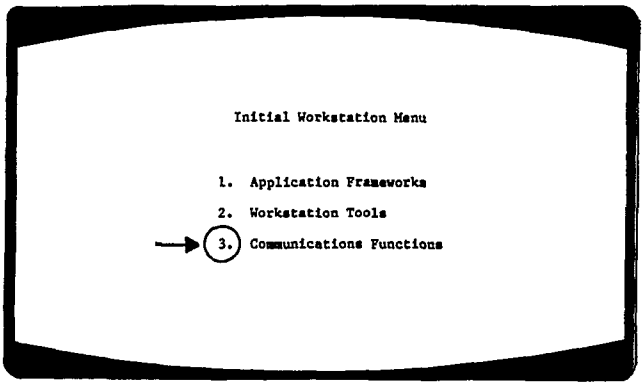
SAMPLE SESSION

(1)



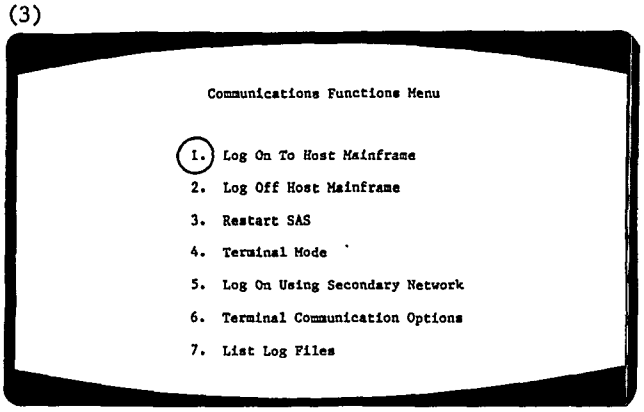
AUTOMATICALLY PROCEED TO INITIAL WORKSTATION MENU.

(2)



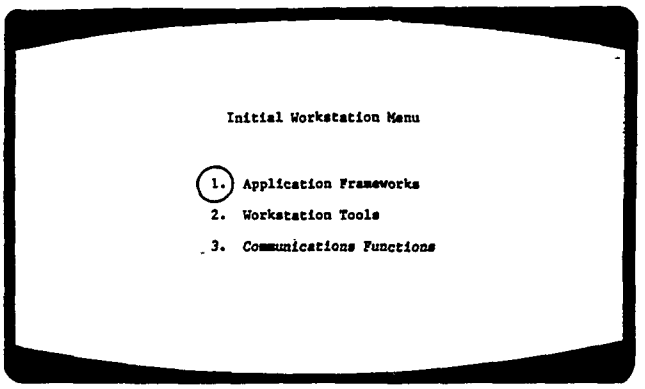
USER SELECTS #3 TO BRING UP SCREEN WITH OPTIONS FOR COMMUNICATING BETWEEN MICRO-COMPUTER AND HOST MAINFRAME.

(3)



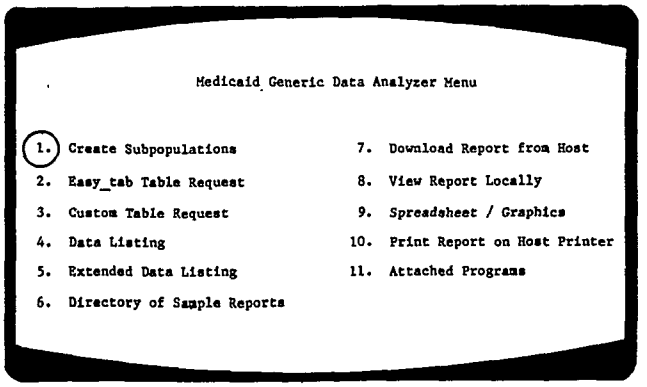
USER SELECTS #1 TO LOGON TO HOST MAINFRAME.

(4)



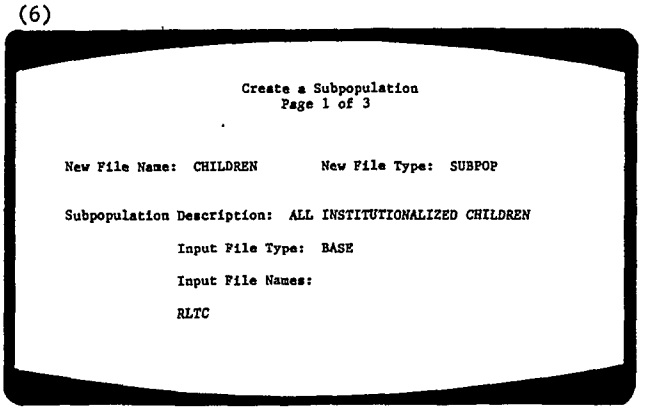
USER SELECTS #1 TO BRANCH TO MEDICAID DATA ANALYZER.

(5)



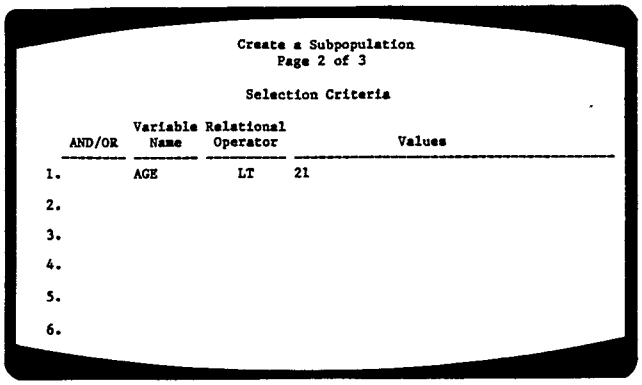
USER SELECTS #1 TO SUBSET THE BASE DATA FILE. FORMS WILL AUTOMATICALLY APPEAR ON THE SCREEN WHICH REQUEST INFORMATION NEEDED IN SUBSETTING THE FILE.

(6)



IN THIS SAMPLE PROBLEM, AN ANALYSIS FILE HAS BEEN PREVIOUSLY CREATED CONTAINING UTILIZATION AND EXPENDITURES INFORMATION FOR ALL INSTITUTIONALIZED RECIPIENTS (CALLED THE "BASE" FILE). IN THIS STEP THE FILE WILL BE SUBSETTED TO INCLUDE ONLY INSTITUTIONALIZED CHILDREN.
THIS FORM HAS THREE PAGES. THE FIRST PAGE OF THE FORM ENTERS THE NEW FILE NAME AND FILE TYPE, AND THE INPUT FILE NAME AND FILE TYPE.

(7)



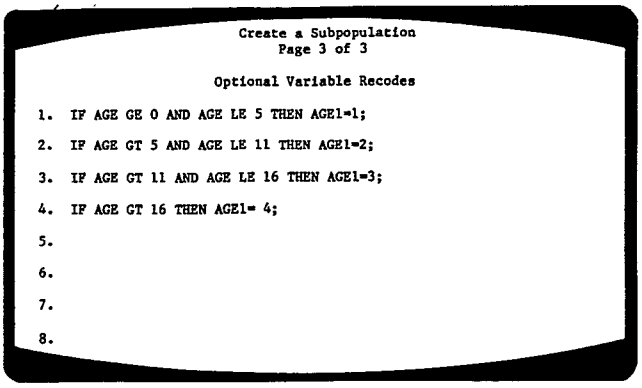
Create a Subpopulation
Page 2 of 3

Selection Criteria

	AND/OR	Variable Name	Relational Operator	Values
1.		AGE	LT	21
2.				
3.				
4.				
5.				
6.				

IN THE SECOND PAGE OF THE FORM, ALL PERSONS UNDER 21 YEARS OF AGE ARE SELECTED FOR ANALYSIS.

(8)



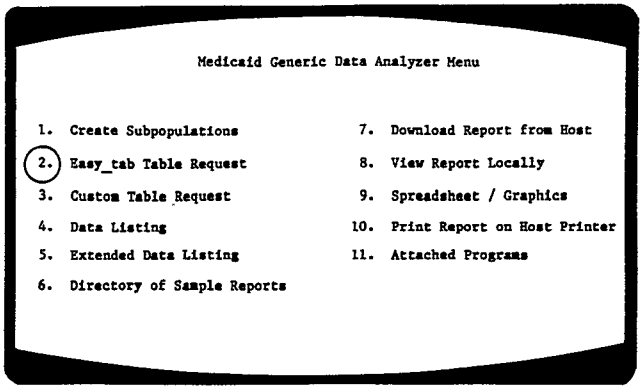
Create a Subpopulation
Page 3 of 3

Optional Variable Recodes

1. IF AGE GE 0 AND AGE LE 5 THEN AGE1=1;
2. IF AGE GT 5 AND AGE LE 11 THEN AGE1=2;
3. IF AGE GT 11 AND AGE LE 16 THEN AGE1=3;
4. IF AGE GT 16 THEN AGE1= 4;
5.
6.
7.
8.

IN THE THIRD PAGE OF THE FORM, A CATEGORICAL VARIABLE FOR AGE IS CREATED. THIS FORM IS THEN TRANSMITTED TO THE HOST COMPUTER FOR PROCESSING. AFTER PROCESSING, THE USER RETURNS AUTOMATICALLY TO THE MEDICAID DATA ANALYZER MENU.

(9)

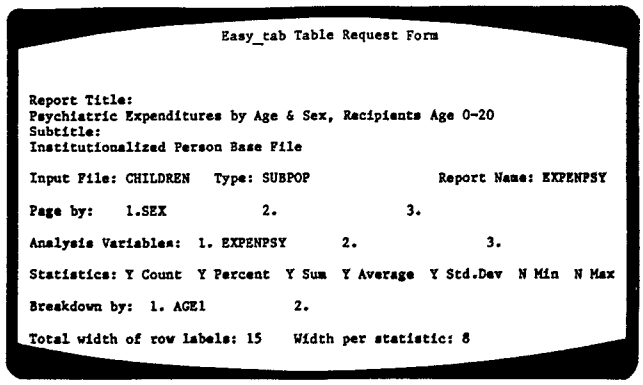


Medicaid Generic Data Analyzer Menu

1. Create Subpopulations
2. Easy_tab Table Request
3. Custom Table Request
4. Data Listing
5. Extended Data Listing
6. Directory of Sample Reports
7. Download Report from Host
8. View Report Locally
9. Spreadsheet / Graphics
10. Print Report on Host Printer
11. Attached Programs

USER SELECTS #2 TO PROCEED TO FORMS WHICH SPECIFY SIMPLE DATA TABLES FROM THE SUBSETTED FILE.

(10)



Easy_tab Table Request Form

Report Title:
Psychiatric Expenditures by Age & Sex, Recipients Age 0-20
Subtitle:
Institutionalized Person Base File

Input File: CHILDREN Type: SUBPOP Report Name: EXPENPSY

Page by: 1.SEX 2. 3.

Analysis Variables: 1. EXPENPSY 2. 3.

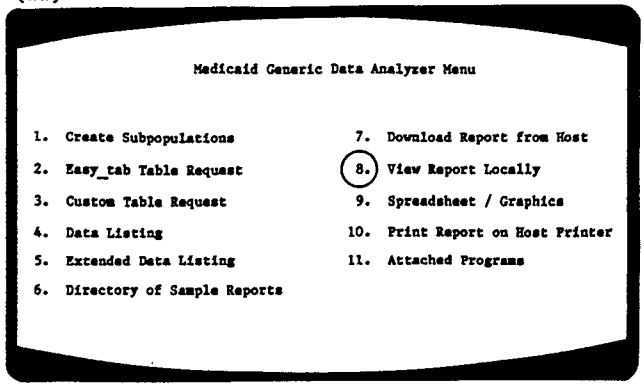
Statistics: Y Count Y Percent Y Sum Y Average Y Std.Dev N Min N Max

Breakdown by: 1. AGE1 2.

Total width of row labels: 15 Width per statistic: 8

THIS FORM REQUESTS A TABLE OF PSYCHIATRIC EXPENDITURES BY AGE AND SEX FOR CHILDREN. USER RETURNS AUTOMATICALLY TO THE MEDICAID DATA ANALYZER MENU.

(11)



Medicaid Generic Data Analyzer Menu

1. Create Subpopulations
2. Easy_tab Table Request
3. Custom Table Request
4. Data Listing
5. Extended Data Listing
6. Directory of Sample Reports
7. Download Report from Host
8. View Report Locally
9. Spreadsheet / Graphics
10. Print Report on Host Printer
11. Attached Programs

SELECT #8 TO EXAMINE THE REPORT WHICH HAS JUST BEEN CREATED.

Psychiatric Expenditures by Age & Sex, Recipients Age 0-20
Institutionalized Person Base File
SEX 1 = MALES

	INPATIENT PSYCH COSTS(AGE 0-20)				
	N	PCTN	SUM	MEAN	STD
AGE1					
0- 5 YRS OLD	60	3.2	249652	4160.9	17090.5
6-11 YRS OLD	378	19.9	6235033	16494.8	23975.9
12-16 YRS OLD	751	39.5	9435440	12563.8	16517.0
17-20 YRS OLD	710	37.4	3047546	4292.3	8932.3
ALL	1899	100.0	18967671	9988.2	16885.9

Psychiatric Expenditures by Age & Sex, Recipients Age 0-20
Institutionalized Person Base File
SEX 2 = FEMALES

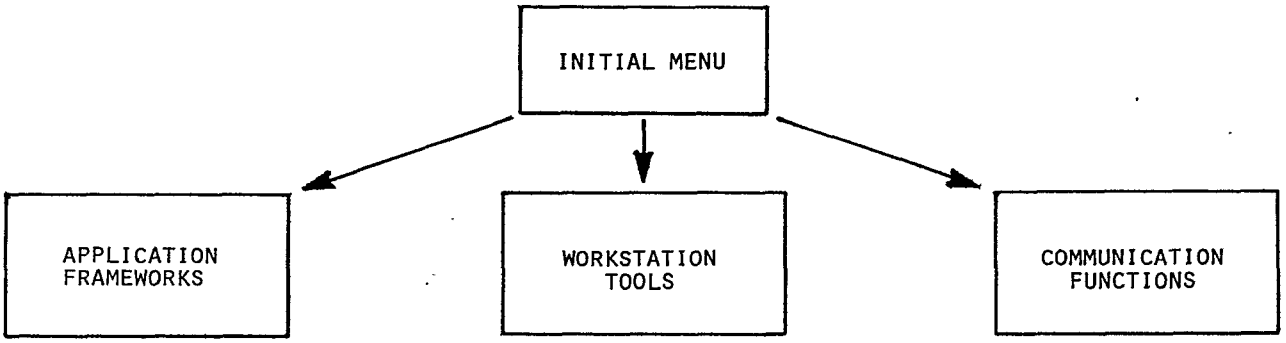
	INPATIENT PSYCH COSTS(AGE 0-20)				
	N	PCTN	SUM	MEAN	STD
AGE1					
0- 5 YRS OLD	29	2.2	0	0.0	0.0
6-11 YRS OLD	166	12.8	1907840	11493.0	22163.9
12-16 YRS OLD	550	42.3	6621046	12038.3	15504.6
17-20 YRS OLD	555	42.7	2291865	4129.5	8402.5
ALL	1300	100.0	10820751	8323.7	14501.5

THE REPORT SHOWS:

- THERE ARE MORE MALES (1899) THAN FEMALES (1300) RECEIVING PSYCHIATRIC CARE.
- MEAN EXPENSES ARE HIGHER FOR MALES ($9988) THAN FEMALES ($8324) FOR THE YEAR.
- CHILDREN IN THE MIDDLE AGE GROUPS HAVE HIGHER MEAN EXPENSES THAN THE YOUNGEST AND OLDEST GROUPS.

USER RETURNS AUTOMATICALLY TO MEDICAID DATA ANALYZER MENU.

SYSTEM OVERVIEW



1. MEDICAID DATA ANALYZER
2. MARKET SHARE
3. COST REPORT
4. EXTERNAL DATABASES

1. CONVERT TABLE TO LOTUS
2. LOTUS SPREAD SHEET
3. CHARTMASTER GRAPHICS
4. WORD PROCESSING
5. TEXT EDITOR
6. SYSTEM MANAGEMENT
7. FILE MAINTENANCE

1. LOG ON
2. LOG OFF
3. RESTART SAS
4. TERM MODE
5. LOG ON SECOND NETWORK
6. TERM COMMUNICATIONS
7. LIST LOG

THIS IS AN OVERVIEW OF THE WORKSTATION MENUS. ONLY THE MEDICAID DATA ANALYZER AND LOG-ON FUNCTIONS ARE ILLUSTRATED IN THIS SAMPLE SESSION.

ON-LINE ACCESS: AN APPROACH TO COPING WITH INFORMATION OVERLOAD

Cynthia E. Burghard and Elliot M. Stone
Massachusetts Health Data Consortium, Inc.

Among other characteristics, the microcomputer has a schizophrenic personality which includes its use as a communication terminal as well as a stand-alone central processing unit. The communications capacity of the microcomputer allows direct access to mainframe computer data bases over telephone lines. In addition, the microcomputer is a stand alone computer system that can be used to store, analyze and graph data that have been "downloaded" or transmitted from the mainframe computer or inputted manually into the microcomputer.

The Massachusetts Health Data Consortium has a multiyear data base of approximately three and a half million billing and discharge records from acute care hospitals. The Consortium utilizes a variety of ways to access the database. These analytic tools include a series of approximately 60 batch reports to analyze hospital market share, charges, migration patterns and case mix intensity. The organization has also experimented with transmitting subsets of the data base on floppy discs to clients and found that the capacity and limitations of floppy discs made this impractical.

Custom designed processing for individual clients is another way to access the data base. We have found this too expensive and too resource-intensive. The Consortium has made available data on magnetic tapes for the users. This access mode merely transfers the burden of accessing the data and generating the reports to the client.

With the ever pressing needs of clients to have more direct access in a timely manner to the data, the Consortium collaborated with Data Resources, Inc. (DRI) of Lexington, Massachusetts to develop ON-LINE ACCESS.

The Consortium has placed a portion of its massive data base on the DRI computers and have allowed clients to access these data through their own microcomputers. Clients have the choice of either doing their analysis directly on the DRI interactive system or downloading data into their microcomputers to do their analysis locally. Data Resources Inc. has developed a computer language called RETRIEVE which is a data base management package with English language commands that allow the non-data processing user to access the data with relative ease. The Health Data Consortium and DRI charge clients at an hourly rate for time sharing.

The data base which is available on ON-LINE ACCESS includes the following data elements:

- Hospital Name
- Patient Residence
- Age
- Sex
- Days
- Disposition
- The expected principal payment source
- The Diagnostic Related Group (DRG)
- The Major Diagnostic Category
- The Clinical Specialty
- Unique Physician Identifier
- Total of All Charges for Hospital Stay
- Routine Charges
- Special Care Charges
- Ancillary Charges

ON-LINE ACCESS was conceived and first introduced into the market in June of 1984 with an initial client base of both hospital and non-hospital clients. The Consortium and DRI were faced with the problem of how to train traditionally noncomputer-literate managers and analysts to properly use the new technologies or how to cope with information overload. Clients are taught to use the traditional batch reports as baseline data, to help them focus their questions and generate additional questions. Clients are then taught to query the interactive systems to pull subsets of the data and to answer the questions raised by the batch reports and to further refine the questions and analyses. Once the subset has been identified using the interactive system, clients are then instructed on how to download these data into their microcomputer where they can be stored, analyzed and graphed locally.

Using an actual analyses problem, I will step you through the use of batch reports, ON-LINE ACCESS and a microcomputer: A marketing manager for a suburban Boston hospital has been asked to investigate the possibility of adding a vascular surgery service to the hospital.

The marketing manager begins the analyses using a batch report showing patient origin of the hospital's case load. The hospital needs to use this report to identify its primary market, i.e., from which towns does it draw most of its patients? Fifteen (15) towns make up approximately 85% of the hospital's business. The batch reports have been used to identify the towns that need to be analyzed. The Marketing Manager can then go on to the ON-LINE ACCESS system to continue his analyses.

Using a microcomputer as a terminal, and a communications software package called Smartcom, the Marketing Manager logs onto the ON-LINE system using a series of commands. Once onto the DRI computer, it takes only four or five commands to define the subset for analyses and determine format for the report.

In this example, the Marketing Manager wants to define the subset of the data by first SELECTing the 15 towns that make up the hospital's primary service area; the hospitals--both local hospitals as well as Boston area where patients were treated; and the patients who were classified under the sub-specialty of vascular surgery. This establishes the criteria for the records to be selected.

Next the user has to decide which CONCEPTS or which data elements need to be included in the analyses. In this case, the Marketing Manager wants to look at the individual Diagnosis Related Groups (DRG) that are included as part of the subspecialty of vascular surgery; which hospital patients went to, their principal payment source, the town that they resided in and their length of stay (LOS). The concepts are hierarchal in nature, so that the first concept to be displayed would be DRG; then within each DRG, the individual hospitals would be listed; then within each individual hospital, the payors and so forth.

Clients than have a choice of either AGGREGATing their reports so that they are two or three dimensional cross-tabulations or they can LIST the actual records. In this example, there are approximately 400 records selected, small enough to be stored on the microcomputer. The Marketing Manager is going to LIST the records so that he can have the raw data base reside on the microcomputer. Once all of the criteria listed above has been input, the users need to use the command RUN, give the file a name, and then in approximately four to five minutes, the data will be retrieved.

Because this is a relatively small file of approximately 400 records, the Marketing Manager will download this subfile. To obviate the need for clients to develop their own protocols for a microcomputer-to-mainframe connection, ON-LINE ACCESS includes a series of preprogrammed commands which require the client to invoke the single command of DOWNLOAD and press the carriage return a couple of times to successfully transmit the data file from the mainframe to the microcomputer.

Unlike many other downloaded data sets, ON-LINE ACCESS allows clients both to download text as well as numbers so that the file can be easily manipulated. Having all the 400 records resident on the microcomputer, the connection to the mainframe computer can be severed and no additional time sharing charges will be incurred. The types of analysis and graphics that can be done at this point are virtually endless.

The Marketing Manager analyzed the differences in market share between the local hospitals and the Boston area hospitals. The data indicates a fairly even market share split between the local hospitals and the Boston area hospitals with the exception of some emergency cases, like amputation, or other "less serious" surgical procedures like vein stripping which one would expect to be done locally.

Analysis of the difference in length of stay patterns in Boston hospitals and local hospitals was also undertaken. This would help to uncover whether there were differences in the possible resource needs of patients in Boston hospitals versus local hospitals. Again, the Marketing Manager saw fairly consistent length of stay patterns with the Boston hospitals, slightly higher in some cases. However there appeared to be a wide variation in one of the DRG's, DRG 120, Other Operating Procedures on the Circulatory System. The local hospital had a length of stay of approximately thirty (30) days and the Boston hospitals length of stay was only five (5) days.

At this point, if the Marketing Manager were working with batch reports he/she would have to go back to the data processing department and: a) see if there was an error in coding, and/or b) possibly have some additional runs to be able to understand or explain the variations in LOS. Another option for the Marketing Manager would be to go back to the ON-LINE ACCESS system and prepare another data run which would incur expenses of probably, in this case, of $100 - $150. However, because the data base is resident on the microcomputer, he/she can go back and review the records included in the Average Length of Stay calculation and determine what might be causing the variations.

The Marketing Manager determined that local area hospitals included many cases with very long lengths of stay. There were cases with lengths of stay of eighty (80) to eighty-five (85) days, one with fifty-four (54) days--all of which contributed to an average length of stay of thirty (30) days. The Boston area hospitals had a total of three (3) cases, all of which had low lengths of stay.

Because the data was resident on the microcomputer, the Marketing Manager could quickly and with no additional expense find that the answer to the length of stay problem was due to several cases with very long lengths of stay.

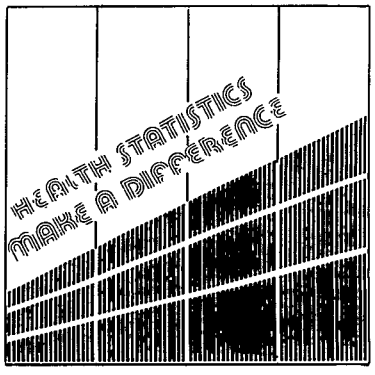
For the Marketing Manager who is trying to make decisions about the possibility of vascular surgery in the area, he/she may want to do further analysis to understand whether the long lengths of stay in the local hospitals was due to perhaps the inavailability of long term care placement or perhaps the practicing patterns of some of the local physicians. In any case, the Marketing Manager was able to manage the analyses process utilizing a combination of batch reports, an interactive system and microcomputer all for a cost of only $125 for one on-line data pull.

One of the other features of the microcomputer is that it provides you the capacity to generate graphics using standardly available software packages. ON-LINE ACCESS clients have been able to generate a wide range of graphics from their downloaded files. Where previously graphic artists had to be hired or analysts and managers had to wade through long tables, they are now able to view the graphics providing a more effective presentation medium.

One rule of thumb when trying to cope with information overload to consider using the approaches outlined in this paper. Do not try and take on all of the technologies at once, but use a phased approach. Start with your batch reports to generate baseline information and to answer the first level of research questions. If you have access to the interactive system, use that to help you answer the questions generated from the baseline data and to subset the data you need to complete your analysis and then use your microcomputer to actually do the analysis. Using this system, I think you will be able to efficiently and effectively utilize the abundant data bases which are now available to health care policy analysts.

Session Q

Statistical Resources and Program Approaches



THE NEW PERSON-BASED, NATIONAL MEDICAID STATISTICAL REPORTING SYSTEM

Donald N. Muse, Ph.D., Health Care Financing Administration
Richard L. Bale, Ph.D., Health Care Financing Administration
Richard H. Beisel, Health Care Financing Administration

PURPOSE

The purpose of this paper is to introduce the design concept for the new, computerized Medicaid Statistical Reporting System, called MEDSTAT.

BACKGROUND OF MEDICAID STATISTICAL REPORTING

Medicaid statistical reporting began in 1967, two years after passage of Title XIX. States were required to submit to HCFA a report containing key monthly data currently known as the "Monthly Statistical Report on Medical Care," or the HCFA-120. Similarly, a more detailed annual report known as the "Statistical Report on Medical Care: Recipients, Payments, and Services" (HCFA-2082) has also been required. Both of these reports are used to collect and compile data on the Medicaid program both at the State and National levels. Federal policy makers, including the Congress, have relied primarily on these reports for information concerning the management and future of the Medicaid program.

Given the dramatic cost increases in the Medicaid program since 1980, HCFA has reconsidered the adequacy of State Medicaid statistical reporting. Current reports have been found to be inadequate in terms of level of detail and accuracy. In addition, new forms of service delivery and financing of Medicaid at the State level is not adequately captured by the current statistical reporting system. Examples of changes in the program due to cost containment initiatives include capitation financing arrangements, HMOs, and DRG-related provider reimbursement schemes. Current statistical reports do not capture sufficient information about these new developments. Further, current statistical reporting does not provide a good basis for conducting research on the Medicaid program in terms of person-based files that would reveal what types of eligibles are associated with high turnover rates, utilization of expensive services, and the effects of changes in program policies.

Effective October 1, 1983 (FFY 84) HCFA issued revised and expanded Medicaid statistical reporting requirements for States to follow (Transmittal No. 29, September, 1984 -- Revision to Section 2700 of the State Medicaid Manual). These new reporting requirements eliminated the monthly HCFA-120 report and expanded the annual HCFA-2082 report. The expanded HCFA-2082 report requests more information on institutionalized recipients, dual eligibles, very young and old age groups, and participation in capitation programs. The new reporting requirements were effective with FFY 84.

Along with the revised reporting requirements States were given the option of forwarding to HCFA, in lieu of submitting hard copy reports, standardized computer tapes of their eligibility, claims payment, and provider files. This tape reporting option (referred to as the MEDSTAT System) will reduce State reporting burden and at the same time provide HCFA with person-based Medicaid service usage and expenditure profiles. From the MEDSTAT data, HCFA will produce the new annual 2082 report. In the future, additional information may also be submitted under the MEDSTAT System so that HCFA can also generate other Federal statistical reports now required of the States. HCFA researchers, actuaries, and other users will have, for the first time, access to data at a level of detail that will greatly improve their ability to monitor the Medicaid program and better understand its dynamics.

The quarterly person-based files will be available for actuarial research and forecasting Medicaid expenditure trends, basic research on the characteristics of the Medicaid population, evaluation of the impact of HCFA demonstration projects and waivers, and policy analyses of how changes in eligibility, reimbursement, and coverage policies may affect State Medicaid programs. The data will have the advantages of uniform definition and format across States, high reliability, being person-based, and being current (as opposed to being over a year old when they become available).

OVERVIEW OF THE MEDSTAT SYSTEM

The MEDSTAT system is designed to provide HCFA with information needed to manage and analyze the Medicaid program. States that report Medicaid statistical data under the MEDSTAT System will be submitting to HCFA, on a quarterly basis, five tape files comprised of claims data, eligibility data, and provider data. The MEDSTAT system will receive these files, verify the accuracy of the data, store the data in a database management system for easy retreival, and conduct standardized analyses of the data for HCFA. Figure 1 provides an overview of the MEDSTAT general system design.

Participating States will submit Medicaid data to HCFA using fixed format records. The data will be submitted in five separate files:

(1) Paid claims for inpatient hospital care file (CLAIM-IP);

(2) Paid claims for long term institutional care file (CLAIM-LT);

(3) "Other" paid claims file containing claims that do not fall into the above two categories (CLAIM-OT);

(4) Eligibles file containing basic information on all eligibles (ELIGIBLE);

(5) Provider file containing basic information on all providers (PROVIDER).

Data for four of these files will be submitted to HCFA on a quarterly basis, with the PROVIDER file being submitted annually. The three tables on the following pages present lists of the variables included in each of the files. Table A shows variables included in each of the three paid claims files, Table B shows the variables for the ELIGIBLE file, and Table C shows the variables for the PROVIDER file.

As illustrated in Figure 1, the MEDSTAT System may be seen as being comprised of four major processing components: (1) Receipt, (2) File Processing, (3) Media Production and Control, and (4) Reports.

The Media Production and Control (MPC) component is really controls all the other components of the MEDSTAT system. MPC stores information on data files received from States, generates notification letters to States, tracks which jobs in the File Processing component have been run and need to be run, generates the JCL necessary to run a job, and keeps track of the results of each job that has been run, plus all the tape and disk files created by a job.

The Receipt component logs tape files received from States into the MEDSTAT System's MPC subsystem for tracking. This process was computerized because MEDSTAT will be receiving and processing about 1,000 reels of tape from States each year when it is in full operation. These 1,000 reels of tape will be processed into more numerous tape and disk files.

The File Processing component performs several tasks for MEDSTAT, including: running it through the Validation module, which checks every field in every record against a set of error detection specifications and error tolerance standards; and producing a Quality Assurance Report.

If a file passes the error tolerance specifications a complete backup of the original file is created for tape archival and the file is written out in an extended format that prepares it to be loaded into the database management system.

Acceptable files are then loaded into Model 204 (the database management system used to store the data for access and most analyses.) The Model 204 files are extensively keyed to allow instantaneous selection of individual or groups of records based on one or more keyed fields.

The Reports component of MEDSTAT is being designed to produce a series of standard reports required by HCFA, with the annual HCFA-2082 being the first in line. Several other Medicaid reports now now required of States in hardcopy form probably will be added in the near future.

In addition to these types of statistical reports, the data will be used extensively for actuarial analyses. The MEDSTAT System will eventually contain a module to help other users find their way through the database to easily and quickly produce ad hoc analyses.

The MEDSTAT System is in its early stages at present. Approximately ten states will be submitting Medicaid statistical data to HCFA through the MEDSTAT System for FFY 1985, and we expect the number of states to increase at the rate of about ten each year.

Future plans for MEDSTAT data include developing a variety of special purpose files of smaller, more efficient sizes to be used for analyses to answer policy and basic research questions. Within the constraints of personal privacy, we hope to make several types of files available for use to researchers and analysts outside of HCFA somtime within the next couple years.

Figure #1
System Overview

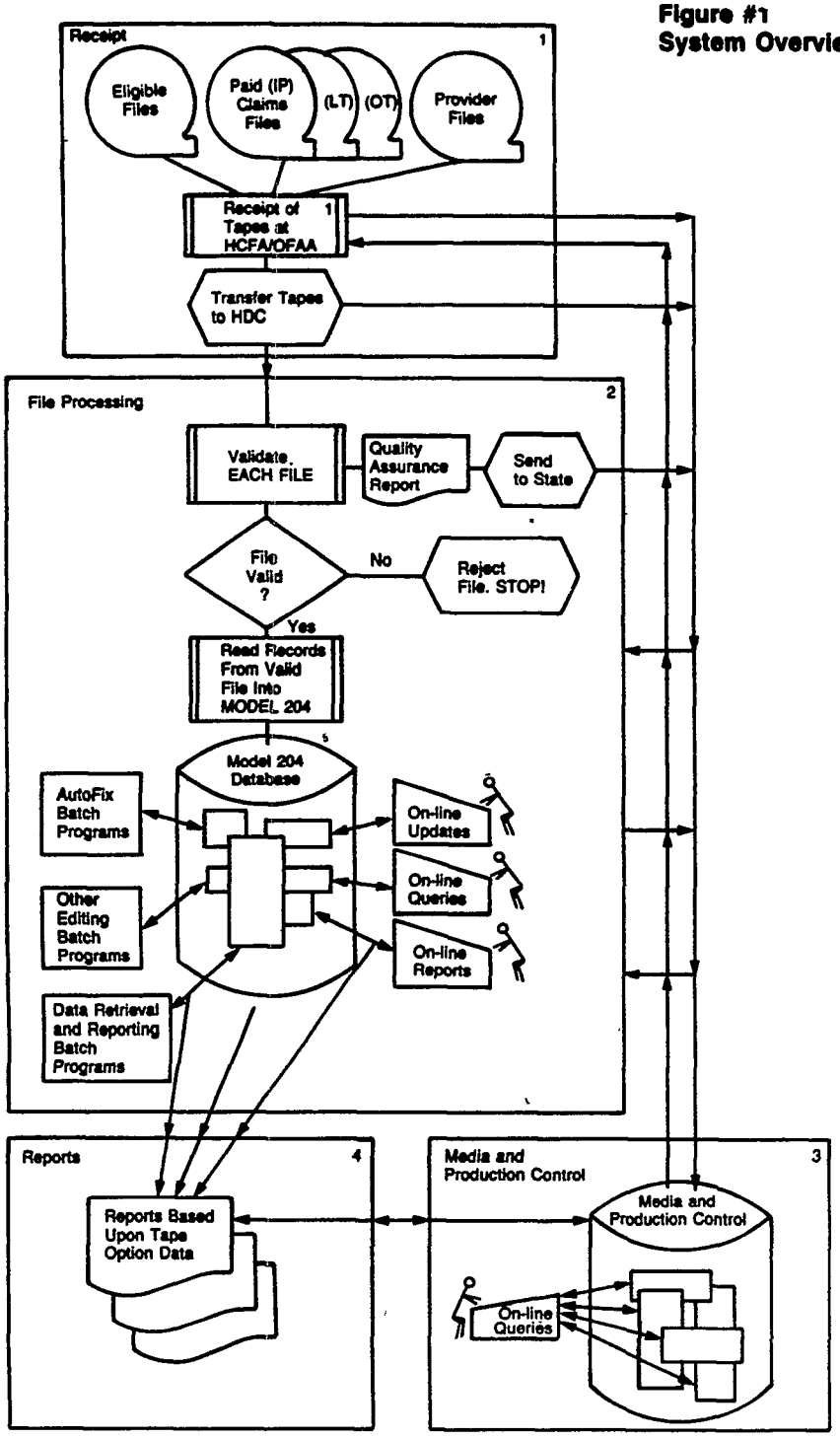


Table A

VARIABLES FOR THE PAID CLAIMS FILES

Var #	Field Name	CLAIM-IP Inpatient	CLAIM-LT Long Term	CLAIM-OT Other
01	Recipient ID	x	x	x
02	Date of Birth	x	x	x
03	Sex Code	x	x	x
04	Type of Coverage	x	x	x
05	Type of Service	x	x	x
06	Claims Adjustment Indicator	x	x	x
07	Payment/Adjustment Date	x	x	x
08	Medicaid Amount Paid	x	x	x
09	Beginning Date of Service	x	x	x
10	Ending Date of Service	x	x	x
11	Provider ID	x	x	x
12	Amount Charged	x	x	x
13	Other 3rd Party Payment	x	x	x
14	Medicare Deductible	x	x	x
15	Coinsurance Payment	x	x	x
16	Diagnosis Code	x	x	x
17	Place of Service	x	x	x
18	Medicare Covered Days/ Quantity	x	x	x
19	Admission Date	x	x	NA
20	Discharge Status	x	x	NA
21	Principle Procedure Category	x	NA	x
22	State-Specific Prin. Procedure Code	x	NA	x
23	State-Specific Procedure Flag	x	NA	x
24	State-Specific Code Modifier	x	NA	x
25	Principle Procedure Date	x	NA	NA
26	State-Specific Secondary Procedure Code	x	NA	NA
27	State-Specific Secondary Procedure Code Flag	x	NA	NA
28	Secondary Procedure Code Modifier	x	NA	NA
29	Secondary Diagnosis Code	x	NA	NA
30	Accommodation Charges	x	NA	NA
31	Ancillary Charges	x	NA	NA
32	Skilled Care Days	NA	x	NA
33	Intermediate Care Days	NA	x	NA
34	Leave Days	NA	x	NA
35	State-Specific Drug Code	NA	NA	x
36	Reason for Denial of Claim	x	x	x
37	Date of Claim Denial	x	x	x
38	Date of Claim Receipt	x	x	x

Table B

VARIABLES FOR THE ELIGIBLE FILE

Var #	Field Name
01	Eligibles ID#
02	Date of Birth
03	Date of Death
04	Sex Code
05	Race/Ethnicity Code
06	Social Security Number
07	County Code
08	Zip Code
	The following variables have values for each of the 12 months in a year
09-20	Days of Eligibility (Months 1-12)
21-32	State Specific Eligibility Group (Months 1-12)
33-44	Maintenance Assistance Status (Months 1-12)
45-56	Basis of Eligibility (Months 1-12)
57-68	Health Insurance Coverage (Months 1-12)
69-80	HMO/Capitation Coverage (Months 1-12)
81-92	EPSDT Flag (Months 1-12)

Table C

VARIABLES FOR THE PROVIDER FILE

Var #	Field Name
01	Provider ID
02	Provider State Code (Practice Site)
03	Provider County Code (Practice Site)
04	Provider Zip Code (Practice Site)
05	Provider State Code (Billing Address)
06	Provider County Code (Billing Address)
07	Provider Zip Code (Billing Address)
08	Medicare Provider #
09	Number of Certified Beds
10	Capitation Flag
	Type of Service Checklist: (19 types of service)
11	Inpatient Hospital
12	Mental Hospital-Aged
13	SNF/ICF Mental Health-Aged
14	Inpatient Psychiatric Facility-Age < 22
15	ICF-MR
16	ICF-Other
17	SNF
18	Physicians
19	Dental
20	Other Practitioners
21	Outpatient Hospital
22	Clinic
23	Home Health
24	Family Planning
25	Lab & X-Ray
26	Prescribed Drugs
27	EPSDT
28	Rural Health
29	Other
30	Physician Specialty or Other Practitioner Category #1
31	Physician Specialty or Other Practitioner Category #2
32	Physician Specialty or Other Practitioner Category #3
33	Total Title XIX Payments

THE CONNECTICUT NURSING HOME PATIENT DATA SYSTEM

Christine Pattee, Connecticut Dept. of Health Services

INTRODUCTION

Since 1977 the Connecticut Department of Health Services has collected demographic data on every patient admitted to each of the almost 300 SNFs and ICFs in the state. In 1982 we added patient functioning levels including ADLS, continence and readmission status. The data is reliable in that our reporting instrument has been consistent through the years, every facility report is individually checked and coded by in house staff, and the entire data set is subjected to substantial computer editing for quality control.

Data are reported by the facilities and though we believe the information is valid, we have never done a field survey to validate the information reported. However, the questions asked are simple with little room for subjective choices. There is no known motivation to misreport any data especially since there is no connection between Health Department data collection and Department of Income Maintenance payments.

The data collection system is efficient and relatively inexpensive. Findings are used extensively by planners and budget developers. Ordering information for detailed data findings and data collection methods is found at the end of this article.

MEASUREMENT OF LENGTH OF STAY

This paper will concentrate on length of stay (LOS) and the different patient populations found within a nursing home. The only other LOS study drawn from a large data base was developed from the 1977 National Nursing Home Survey, in which researchers had to construct estimated lengths of stay based on life table analysis of cross sectional data (1-4). Because Connecticut has annual reporting on individual patients, we are able to measure LOS directly.

Measurement of nursing home LOS, where stays are often well over a year, is very different from the same measurement in a hospital where LOS is a matter of days. In a nursing home, LOS must be calculated from date of admission to date of discharge. Furthermore, it should be measured in two ways, at discharge (complete LOS) and on a specified census date (incomplete LOS).

LOS may be summarized as either a mean or a median (Fig. 1). Mean or average LOS is substantially greater than median LOS beause the average is skewed by a small number of very long stays (e.g. ten or more years.) For graphic presentation and tabular summaries, I believe that median LOS is the more appropriate figure and it is used in the following charts. Mean LOS must be used in formulas relating total days of stay to patient or bed counts (i.e. volume measurements).

CENSUS, DISCHARGE AND ADMISSIONS POPULATIONS

Measurement of discharge (complete) and census (incomplete) LOS represent two very different populations. The median complete length of stay, 107 days, reflects a population that turns over relatively quickly for a nursing home. There is duplication in that the same patient can be discharged and readmitted to the facility and will be

COMPARATIVE MEASUREMENTS OF LENGTH OF STAY
CONNECTICUT SNFS AND ICFS 10-1-82 to 9-30-83
ALL PATIENTS

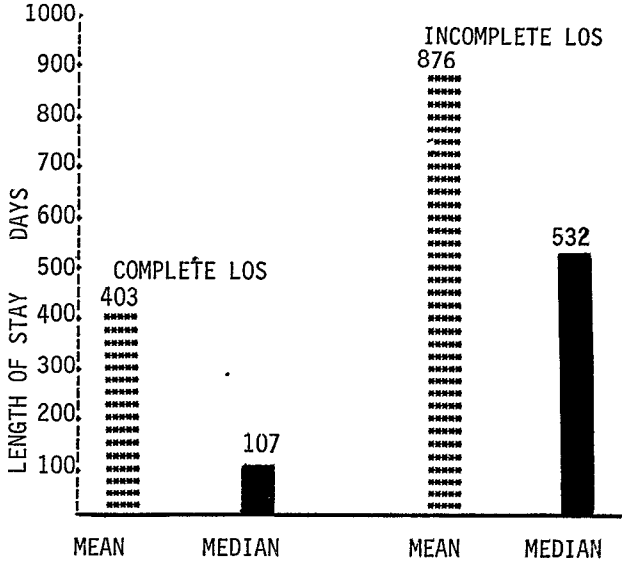


FIG. 1

LENGTH OF STAY-FIRST ADMISSIONS ONLY COMPARED
TO ALL PATIENTS CONNECTICUT SNFS AND ICFS
10-1-82 TO 9-30-83

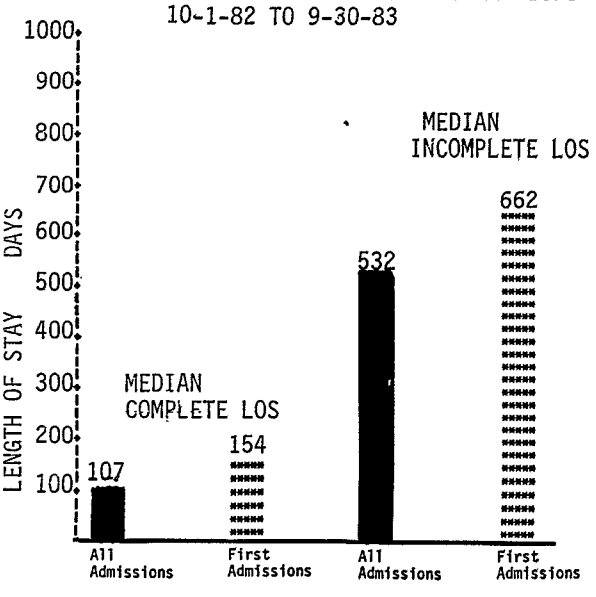


FIG.2

counted again in the data set. The median incomplete or census LOS of 532 days, about one and a half years, is much longer than the discharge LOS even though it is not yet complete. The census population is an unduplicated count of patients and represents mainly the long stayers. It is interesting to note that discharge and census LOS have remained quite stable in Connecticut over the four years that we have been measuring them. In the 79-80 reporting year, discharge LOS was 92 days and census LOS was 515 days, just slightly less than the more recent totals.

There is a third identifiable population-the admissions cohort within a reporting year. In general the profile of the admissions population is similar to the discharge population, with some significant exceptions in payment source. In Connecticut, in the 82-83 reporting year, the discharge population was 21,551 and the mutually exclusive census population, on 9-30-83, was 25,665. The number of admissions during the year was 21,909. Admissions are included in both the discharge and the census group and include readmissions of the same patient.

It is important to be able to separate readmissions from first admissions because length of stay on a discharge and subsequent readmission tends to be relatively short. As Fig. 2 shows, there is an increase in LOS for first admissions only, to 154 days for discharge length of stay and to 662 for census length of stay. One third of all admissions to Connecticut nursing homes are readmissions of the same patient in the same year (Fig. 3). Of the unduplicated census population, 14% had a history of readmissions during the immediately preceding reporting year (Fig. 4).

READMISSION STATUS OF PATIENTS ADMITTED
10-1-82 TO 9-30-83

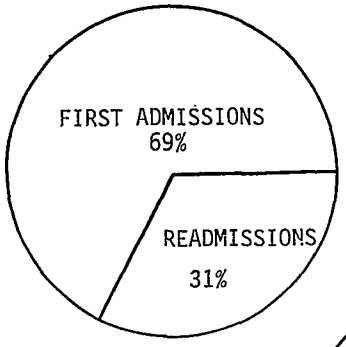


FIG. 3

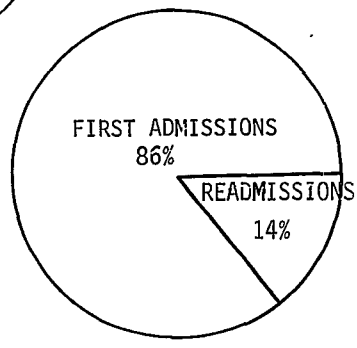


FIG. 4

READMISSION STATUS OF CENSUS POPULATION
9-30-83

AGE.

Figure 5 shows the expected age distribution with the highest proportion in the older age groups. However, Fig. 6 shows that median incomplete LOS is longest for the under 65 group, who are only about 10% of the total population.

DIAGNOSIS.

Circulatory conditions, 22% plus stroke 11%, are the most common diagnosic group. Mental diagnoses are the primary diagnosis of 20% of the population: 11% with a mental but not psychiatric diagnosis, 7% with a psychiatric diagnosis and 2% with mental retardation. Comparing census counts (Fig. 7) with median incomplete LOS (Fig. 8), the mentally retarded population has by far the longest LOS, about three and a half years. The next longest LOS is for patients with a psychiatric diagnosis. Neoplasms and respiratory conditions have the shortest LOS.

CENSUS - AGE

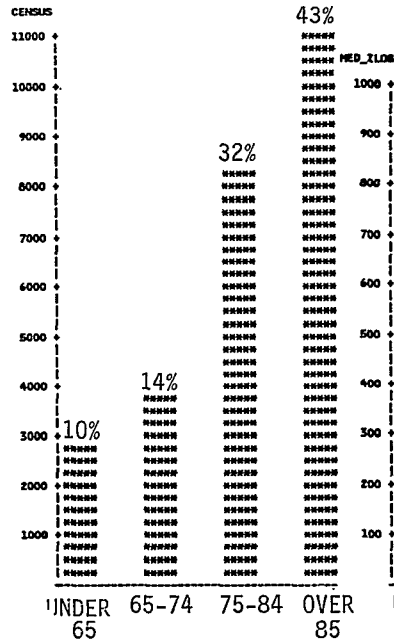


FIG. 5

LOS - AGE

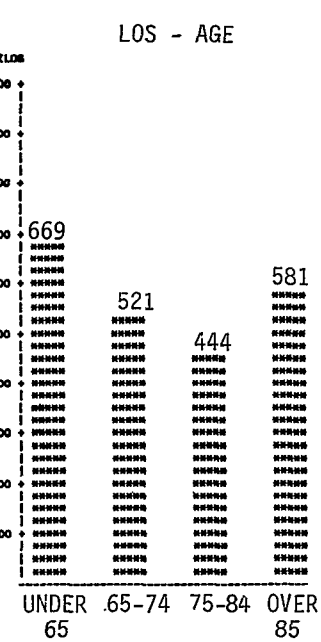


FIG. 6

CENSUS - DIAGNOSIS

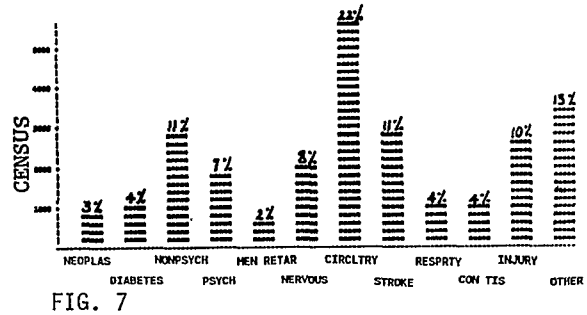


FIG. 7

CENSUS LOS - DIAGNOSIS

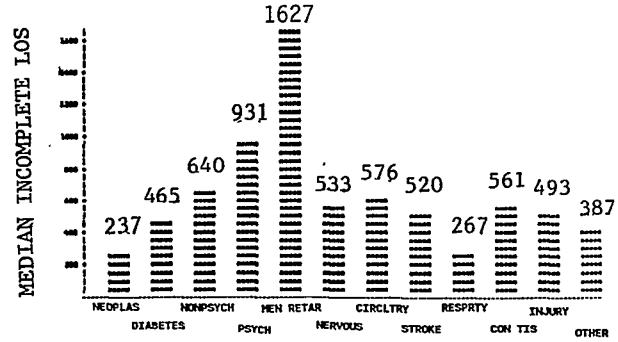


FIG. 8

PRIMARY DIAGNOSIS

SOURCE OF PAYMENT.

The most noticable difference among the census, admission and discharge populations is in their source of payment. (Fig. 9a-c Note: in these figures, 'Other', which is 2% of the population, is not included.) Twelve percent of the population were on Medicare when admitted to the facility during the 82-83 reporting year. This is notable since patients receiving Medicare are generally considered a very small part of the nursing home population. A little over one third of the admissions were Medicaid eligible. Since a portion of this admission population is actually being readmitted, this includes patients who may have entered the nursing home as private payors, were discharged to a hospital for a spell and then re-entered the nursing home and went on Medicaid at a later point than is apparent in these data. Almost half of the admissions are private pay (Fig. 9a).

Significant changes are evident in the payment sources of discharged patients. The Medicare population has dropped and the Medicaid population has increased (Fig. 9b). In the census population, almost two thirds of the population are on Medicaid, but one third are still paying out of their own pockets (Fig. 9c).

The median incomplete LOS for the census population has significant bugetary implications when categorized by payment source (Fig. 10). The as yet incomplete LOS of the Medicaid population is about 2 years, whereas the private pay LOS is under a year.

SOURCE OF ADMISSION

Over half of the census population entered the nursing home from a general hospital (Fig. 11). However, the admissions from the hospital have the shortest length of stay, and once again patients who started out in a mental hospital are filling up years of bed space (Fig. 12).

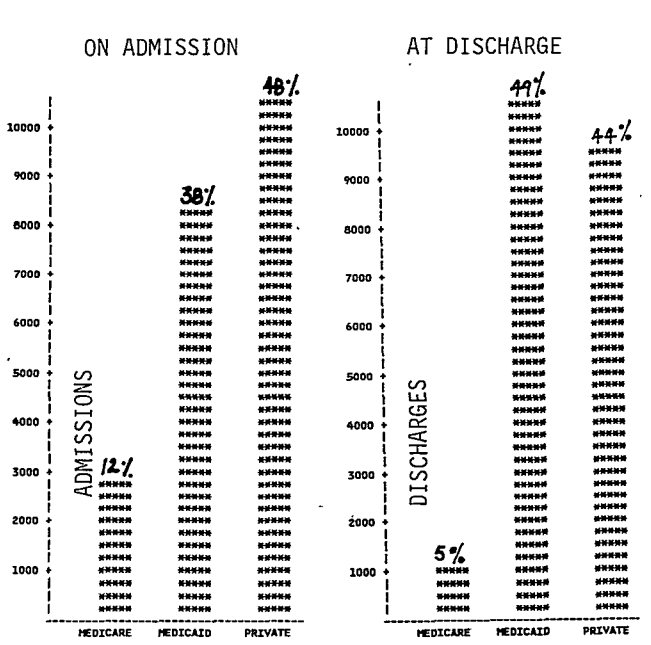


FIG. 9a

FIG. 9b

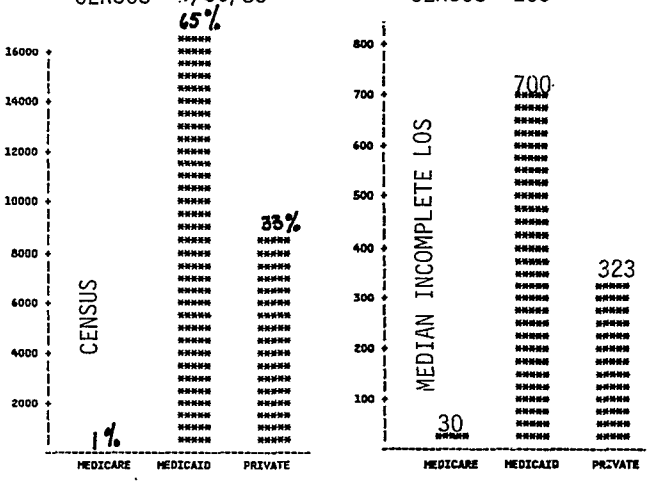


FIG. 9c

FIG. 10

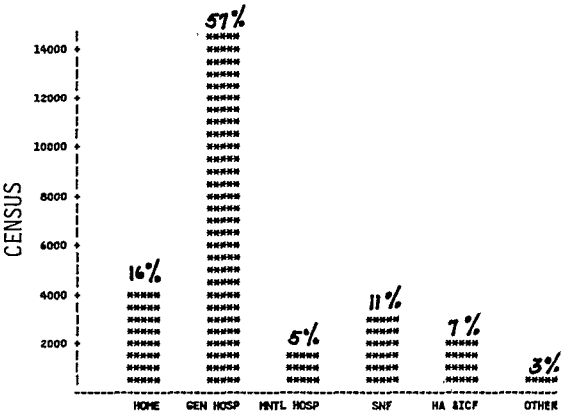


FIG. 11

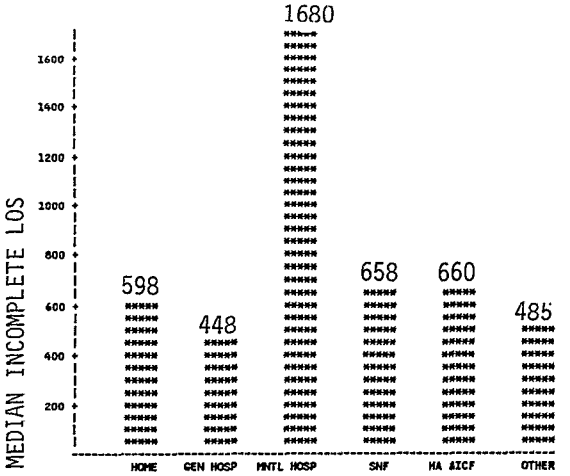


FIG. 12

DESTINATION ON DISCHARGE

Forty four percent of the patients are discharged to a general hospital. A significant but unknown proportion of them will die there, and a hospital discharge is often an interim period before a nursing home readmission (Fig. 13). There is very little variation in LOS prior to any destination on discharge except for discharge home which has a median LOS only 33 days. Of course

DISCHARGES - DESTINATION

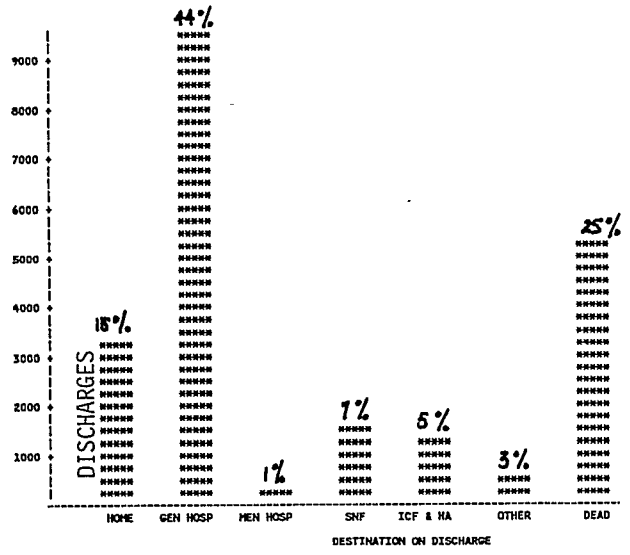


FIG. 13

LOS - DESTINATION

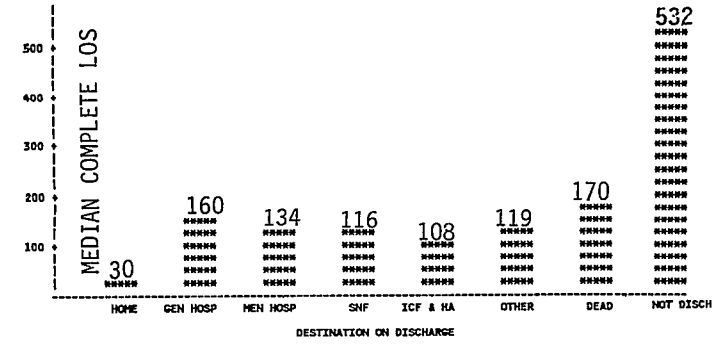


FIG. 14

the as yet incomplete LOS of the undischarged population far outstrips any discharge LOS (Fig. 14).

PATIENT FUNCTIONING LEVELS

1982-83 was the first year that Connecticut collected data on patient functioning levels. For four activities of daily living (ADLs), facilities were asked to indicate whether the patient was independent, sometimes dependent, or dependent. For ability to transfer, e.g. from bed to wheelchair, approximately a third of the patients fall into each category (Fig. 15a). A higher proportion, 43%, are able to ambulate without assistance (Fig. 15b). Fully two thirds of the patient population are able to feed themselves (Fig. 15c). The smallest number of patients, 29%, are able to dress themselves, whereas 44% are dependent in this function (Fig. 15d).

Fig. 16a and b report bowel and bladder continence in the census population.

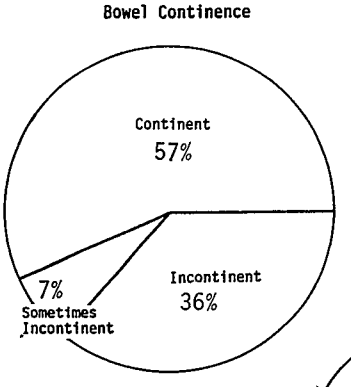


FIG. 16a

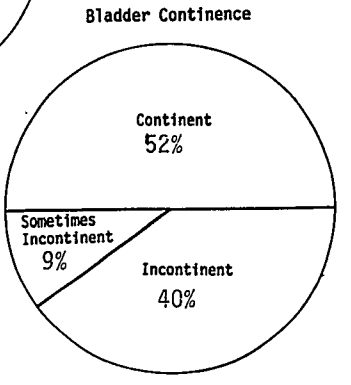


FIG. 16b

ADL SCORE

A composite ADL score was constructed by assigning 1 for 'independent', 2 for 'sometimes dependent', and 3 for 'dependent'. Numerical values for each of the four ADLs were summed so that a score of 4 indicates independence in all four ADLs and a score of 12 means dependent in all four functions. Over a quarter of the census population is independent in all four ADLs (Fig. 17). ADL score has been examined by facility

ACTIVITIES OF DAILY LIVING

TRANSFER

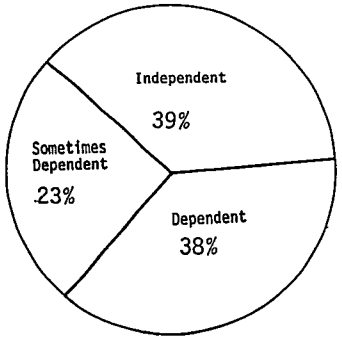


FIG. 15a

AMBULATION

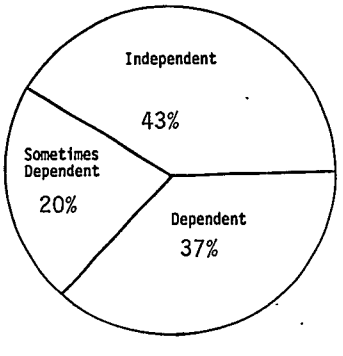


FIG. 15b

ABILITY TO FEED SELF

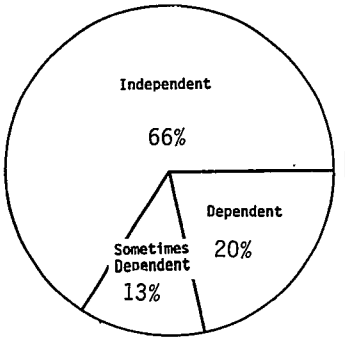


FIG. 15c

ABILITY TO DRESS SELF

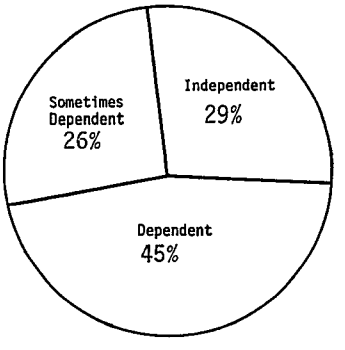


FIG. 15d

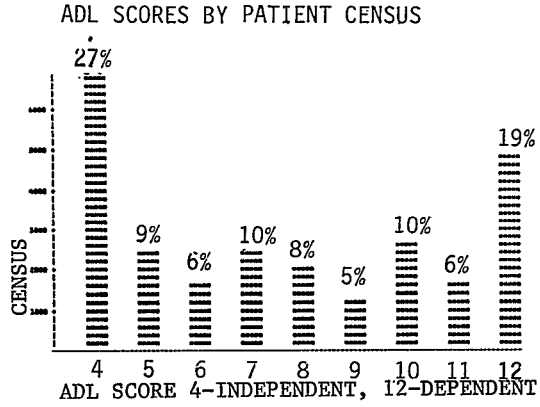


FIG. 17

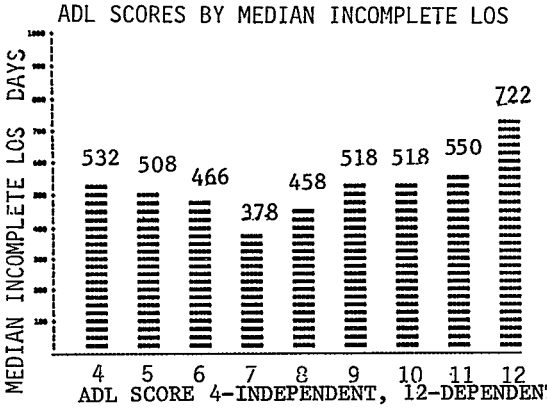


FIG. 18

level, and although functioning level is definitely higher in ICFs, there is a remarkably high proportion of independently functioning patients in SNFs also.

There is a slight U shape in the distribution of lengths of stay for the census population of patients with different ADL scores, with longer LOS's at the independent and dependent ends of the spectrum of scores (Fig. 18).

REFERENCES

1. Keeler, B. Emmett, Robert L. Kane, David H. Solomon; 'Short and Long-Term Residents of Nursing Homes'; Medical Care, March 1981, 363-369.
2. Liu, Korbin, Yuko Palesch; 'The Nursing Home Population: Different Perspectives and Implications for Policy'; Health Care Financing Review, December 1981, 15-22.
3. Liu, Korbin, Kenneth G. Manton; 'The Length-of-Stay Pattern of Nursing Home Admissions'; Medical Care, December 1983, 1211-1222.
4. Liu, Korbin, Kenneth G. Manton; 'The Characteristics and Utilization Pattern of an Admission Cohort of Nursing Home Patients(II)'; The Gerontologist, 24:1,1984, 70-76.

CONNECTICUT NURSING HOME DATA BOOK 1982-83

There is a data book containing comparative statistics on all patients in Connecticut's SNFs and ICFs between 10-1-82 and 9-30-83. Contents include:

A. Patient data (for admissions and resident census population)
1. Demographic-age, diagnosis, admission from, destination, payment source, patient origin, length of stay for discharges and resident census, readmission status.
2. Functioning Level-ambulation, transfer ability, ability to dress self, ability to feed self, ADL score (sum of above), bladder continence, bowel continence, continence score (sum of above), catheterization, confusion status.

B. Facility Data-occupancy rates, bed count and changes in number of beds, medicaid rate and days, HSA and town location.

This information is aggregated in about 80 pages of cross tabulations in which each characteristic is compared with every other characteristic for the entire nursing home population. Additionally, there is a 400 page supplement in which statistics on each of the characteristics above are listed for each of the 294 SNFs and ICFs, grouped by HSA. Other sections include a detailed explanation of data collection procedures, quality control, and reporting instructions to the facilities.

Cost: Summary Tables $5.00
Summary Tables + individual facility listing $20.00

Note: The book is in press and will be available in late 1985.

Make check payable to: "Department of Health Services"

Mail to: Statistical Analysis Unit-Data Book
Department of Health Services
150 Washington Street
Hartford, Connecticut 06106

MEDICAID HEALTH MAINTENANCE ORGANIZATIONS: USE AND AVAILABILITY OF UTILIZATION, ENROLLMENT AND HEALTH STATUS DATA

David Spivack and Karla J. Keith, Mount Sinai Medical Center of Greater Miami

INTRODUCTION

The rising cost of the Medicaid program has become a major concern for the federal government as well as State governments. Beginning in 1981 with implementation of the Omnibus Budget Reconciliation Act, Public Law 97-35, States were given increased flexibility to deal with the problem (1). This increased flexibility was a direct result of changes in the freedom of choice provision that had historically required States to offer Medicaid recipients freedom to obtain services from any qualified provider. Based on these changes and the perceived benefits of prepaid systems, an increasing number of States have chosen to enter into risk contracts with health maintenance organizations and other health care providers to provide prepaid case managed care to Medicaid recipients in a defined geographic area. Interest in prepaid health plans is based on their potential for providing comprehensive, cost-effective health care to an enrolled population and the demonstrated success of these plans in controlling costs and reducing the use of inpatient hospital services. In addition to these State initiatives, several proposals to reform the Medicaid program into separate state and federal programs have proposed that a federal program of primary care be funded through a prepaid, capitation system as a means of controlling costs (2).

Although the development of these types of prepayment alternatives to traditional fee-for-service Medicaid reimbursement offers States an excellent opportunity to control costs, while at the same time retain or expand benefits, the success of these ventures is severely jeopardized by the lack of appropriate data by which prospective providers can participate in the rate setting process, assess the feasibility of risk contracts and plan for their implementation. National and state data currently available to develop payment systems and capitation rates is generally limited to historic utilization and cost data from services provided in a fee-for-service delivery system in the absence of case management. The experience of the Mount Sinai Medical Center in negotiating a risk contract for a Medicaid HMO in Miami Beach demonstrates the need for a national Medicaid HMO database that would include information on health status, utilization by age and sex, enrollment and disenrollment rates and length of enrollment. The availability of such a database will not only improve the planning, development and evaluation of Medicaid prepaid plans, but also encourage new providers to participate in these initiatives.

BACKGROUND

Mount Sinai Medical Center is a 699 bed non-profit, voluntary teaching hospital in Miami Beach, Florida. The Medical Center's primary service area consists of Miami Beach as well as other parts of Dade County. Miami Beach contains one of the largest elderly populations in the country with close to 52% of all persons over 65 years of age and close to 29% over 75 years of age. Reflective of this age composition, close to 70% of all inpatient days at Mount Sinai are Medicare patient days. Medicaid and indigent admissions represent approximately 9% of all inpatient days.

The Miami Beach Medicaid population is comprised of approximately 57.4% SSI-Medicaid enrollees and 42.6% AFDC-Medicaid enrollees. This distribution represents a much higher proportion of SSI enrollees in Miami Beach given that SSI enrollees represent only 30% of all Medicaid enrollees statewide. The high proportion of SSI eligible enrollees in the Miami Beach Medicaid population is explained by the large elderly population in Miami Beach coupled with the fact that approximately 18.4% of all Miami Beach residents 65 years and older live below the poverty line. This is of particular concern in the "South Beach" area with approximately 25.6% of all persons over 65 living below the poverty line. Within the Miami Beach SSI population, approximately 39% have Medicaid only, 23% have Medicaid and Medicare Part B and 38% have Medicaid and Medicare Parts A and B.

Mount Sinai's interest in developing a Medicaid HMO for Miami Beach Medicaid enrollees stemmed largely from the observation that the Medical Center was providing over 85% of all inpatient days utilized by Miami Beach Medicaid enrollees. Institutional data revealed that Mount Sinai was also providing a large volume of services to Medicaid enrollees in the outpatient department and emergency room. Given that the Medical Center was viewed as the primary source of care for the vast majority of Medicaid enrollees on Miami Beach, Mount Sinai initiated the Medicaid HMO project as a means of providing a more comprehensive range of services to these patients. In addition, it was felt that many of the patients receiving free care were eligible for Medicaid based on State eligibility criteria. It was hoped that mechanisms established as part of a Medicaid HMO enrollment process would facilitate Medicaid eligibility determination for many of these patients.

With these objectives in mind, Mount Sinai initiated negotiations with the State of Florida Medicaid Office in May of 1984. The initial proposal included an overview of the proposed HMO service delivery system including policies and procedures, the

administrative and organizational structure responsible for managing the program, and the Medical Center's capability to respond to a number of reporting and data collection requirements. Once these components were approved by the State, the final component of contract negotiations was the business proposal. As part of the business proposal, Mount Sinai was required to project utilization and unit costs and propose its own capitation rates within State ceilings. Mount Sinai was also required to project HMO enrollment for both the SSI and AFDC eligible groups.

MEDICAID HMO UTILIZATION PROJECTIONS AND THE DEVELOPMENT OF CAPITATION RATES

From the State's perspective, the objective of the institution's rate setting process was to ensure that Mount Sinai could provide HMO enrollees with needed services given the institution's costs of providing those services. Risk contracts negotiated with the Florida State Medicaid Program offer providers a maximum capitation equal to 95% of per recipient expenditures under the fee-for-service system. Unlike average adjusted per capita cost (AAPCC) capitation for Medicare beneficiaries provided for under TEFRA, Medicaid capitation is only grossly adjusted on the basis of national origin and welfare status - SSI and AFDC. To assist in this rate setting process, the State Medicaid Office provided historic utilization and expenditure data for all Medicaid enrollees for fiscal year 1983-1984. This data was based on Medicaid enrollee case months to account for both users and non-users. The State also provided utilization assumptions developed from the experience of large Medicaid populations in other States.

The utilization rates derived from the historic Miami Beach Medicaid enrollee data were much higher than comparable national data for Medicaid enrollees and utilization rates for other HMO populations, e.g. employed HMO populations and Medicare HMO demonstration populations. Of particular concern was the significant difference between the Miami Beach SSI utilization rate for inpatient hospital days of 6,048 days per 1,000 enrollees compared to the initial experience of 2,880 days per 1,000 enrollees for the Medicare HMO Demonstration population. Another striking difference between these two groups was in the number of prescription medications per member per year. The Miami Beach SSI group had a rate of 28 prescriptions per year compared to the Medicare HMO enrollee rate of 10 prescriptions per member per year (3). Differences between Miami Beach AFDC historic utilization rates and utilization rates for non-Medicare HMO populations were not as striking. A comparison of the historic and projected utilization rates for both Miami Beach AFDC and SSI Medicaid enrollees is contained in Table 1. A summary of the initial utilization experience for Medicare HMO Demonstration enrollees is contained in Table 2.

The historic utilization rates for the Miami Beach groups were also much higher than the utilization assumptions provided by the State. A summary of these assumptions is contained in Table 3. Although higher than the rates established in the initial experience of the Medicare HMO demonstrations, these utilization assumptions were also considerably lower than the Miami Beach historic utilization rates.

Despite the apparent limitations of the historic utilization data, it served as a general indication of the utilization patterns of the prospective HMO population and as a starting point in projecting future HMO utilization and capitation rates. However, it was clear that the extent of utilization revealed in the historic data could not be accomodated within the 95% ceiling proposed by the State. Faced with this task, several assumptions were made regarding the projected impact of the prepaid, case managed system on the utilization by Medicaid HMO enrollees. These assumptions included:

- hospital inpatient days would be reduced
- physicians inpatient visits would be reduced
- emergency room visits would be reduced
- number of prescription medications per enrollee would be reduced
- hospital outpatient visits would increase
- outpatient physician visits would increase
- home health visits would increase

Limited experience from other Medicaid HMO programs have shown a reduction in inpatient hospital days and prescription medications (4,5,6). Similar reductions in inpatient hospitalization have also been observed for both employed and Medicare HMO populations. Increased use of ambulatory care services has also been demonstrated in these HMO populations (7,8).

The other assumptions were based on the proposed objectives and anticipated outcomes of case management, and the use of restrictions on the use of emergency room services for non-emergency care. Where applicable, the utilization guidelines provided by the State were used as an indication of the low end of a range of possible utilization projections. Based on the assumptions described above and the State utilization guidelines, Mount Sinai was able to propose SSI and AFDC capitation rates within the State set ceiling that will allow for a relatively high level of utilization given the experience of other HMO populations. These assumptions were reviewed by the State Medicaid program officers and

accepted as a rationale for Mount Sinai's proposed utilization rates.

LIMITATIONS IN THE USE OF HISTORIC DATA AND THE NEED FOR A NATIONAL MEDICAID HMO DATABASE

Although the historic utilization data for Miami Beach Medicaid enrollees was useful as a general indication of the utilization patterns of the prospective HMO population, there were several inherent limitations in using this data to project utilization for an HMO enrolled population. The most serious limitation with this historic data was that it was based on utilization in a fee-for-service system with quite different incentives for both patients and providers than those inherent in a prepaid, capitation system. In many cases fee-for-service Medicaid reimbursement has encouraged the use of in-hospital as opposed to ambulatory care services for both patients and providers.

A second limitation was that, in the absence of information on the health status of the Miami Beach Medicaid population, it was virtually impossible to discern to what degree historic utilization rates were the result of a fragmented delivery system with incentives for inpatient care or an accurate reflection of the health status of the Miami Beach Medicaid population. Although prior utilization has been proposed as a possible health status adjustment in determining Medicare AAPCC rates, the use of prior utilization as a proxy for health status has been criticized in that it may reflect the practice patterns of a particular provider or system and not the actual need for services (9,10). A third limitation inherent in the historic data was that it was based on a more restrictive benefits package than that proposed for the Medicaid HMO. A fourth and final limitation was that the historic data was based on Medicaid expenditures and did not reflect the use of services that were not reimbursed through the Medicaid program.

The need to project HMO enrollee utilization as part of the contract negotiation process and to assess the financial feasibility of risk contracts, in light of the data limitations described above, is the primary justification for development of a national Medicaid HMO database. The availability of national data would provide information on utilization rates for comparable Medicaid HMO populations, as well as observed trends in utilization as a result of improved case management and cost containing incentives. In the absence of such a yardstick for evaluating utilization projections, many prospective HMO providers may over estimate the potential of a prepaid, case managed system to reduce utilization.

A national database would also provide more accurate information on the health status and utilization patterns of Medicaid enrollees who choose to join HMO plans, and the length of their enrollment. The availability of such a database would enable providers to anticipate and plan for the influence of health status, preferred and/or adverse selection and enrollment patterns on utilization and enrollment that have been observed in other HMO populations. Although similar types of information is increasingly available for other HMO populations, the applicability of these findings to Medicaid HMO enrollees is yet to be determined.

Data from other HMO eligible populations and initial data on Medicaid HMO enrollees in Michigan have tended to show that those who enroll in HMOs tend to be lower utilizers and are at lower risk of utilization than those who choose to remain in the fee-for-service system. These observations of "preferred selection" have been explained in part by the unwillingness of sicker patients to sever existing physician relationships in order to join an HMO plan (11). However, in the case of providers like Mount Sinai who already serve as the primary source of care for those they seek to enroll in a Medicaid HMO, this observations may not be valid. In fact, the established relationships of many Medicaid recipients with outpatient department physicians and other community-based physicians who will be included in the HMO may serve to enroll a disproportionately higher risk group. Consistent with this concern, additional findings suggest that "preferred selection" is more likely to occur in the case of enrollment in prepaid group practice plans and less likely to occur in the case of enrollment in individual practice association plans (IPAs). These findings were based on the observation that IPAs were more likely to allow enrollees to maintain an existing physician-patient relationship (12).

A national Medicaid database would provide data not only on the health status and utilization of the HMO enrolled population, but also an ongoing comparison of this group with Medicaid enrollees in the fee-for-service system. The availability of such data may allow for the development of adjustments that would protect potential providers from adverse selection. National data would also prove useful in projecting rates of voluntary HMO enrollment among Medicaid eligibles.

A final issue that may potentially jeopardize the soundness of Medicaid HMO utilization projections is the impact of high turnover and short length of enrollment in the HMO. Data from employed HMO populations suggests that the use of medical services changes with the duration of enrollment with utilization of provider visits and well care being higher in the first months of enrollment. These findings also suggest that utilization of these services decreases and stabilizes after the first year of enrollment (13). If this observation is also true for Medicaid enrollees, utilization projections made in the

absence of accurate enrollment data may not accurately account for the impact of enrollment patterns in this population. A national database would potentially provide information on the length of enrollment of Medicaid HMO enrollees and the relationship between length or enrollment and patterns of utilization.

CONCLUSION

The Mount Sinai experience of negotiating a risk contract for the development of a Medicaid HMO in Miami Beach demonstrates the need for a national database on Medicaid HMO enrollees. Based on this experience, such a database would include information on the health status and demographics of HMO enrollees, utilization by age, sex and welfare status, enrollment and disenrollment rates and length of enrollment. Although many of the providers negotiating these types of risk contracts have had experience in the planning and management of prepaid plans, this experience is not readily applicable in the planning of similar plans for low income groups. The availability of national data on the enrollment of Medicaid recipients in HMOs, as well as their experience in these plans, would greatly enhance the decision-making and planning capabilities of providers entering into these contracts. Unless better data is made available, it may become difficult to encourage providers to enter into risk contracts to serve Medicaid enrollees and many programs may fail as a result of inaccurate projections and planning. The availability of a national database will also facilitate the evaluation and improvement of States' alternative delivery system initiatives. The success of these programs will have a major impact on the nation's ability to continue to provide quality health care to low income persons.

FOOTNOTES

1. Federal Register, Volume 46 (October 1, 1981): 48524-29.

2. Bloom, M. Medicaid Overhaul Recommended. Modern Healthcare 14(March):146, 1984.

3. Leutz, W.N., et al. Changing Health Care for an Aging Society. Lexington, Massachusetts: Lexington Books. 200-08, 1985.

4. Johnson, R.E and Azevedo, D.J. Comparing the Medical Utilization and Expenditures of Low Income Health Plan Enrollees Having Medicaid Eligibility. Medical Care 17(9):953-66, 1979.

5. DesHarnais, S.I. Enrollment In and Disenrollment From Health Maintenance Organizations by Medicaid Recipients. Health Care Financing Review 6(3):39-50, 1985.

6. Rabin, D.L., Bush, P.J. and Fuller, N. Drug Prescription Rates Before and After Enrollment of a Medicaid Population in an HMO. Public Health Report 93(1):16-23, 1978.

7. Luft, H. HMO Performance: Current Knowledge and Questions for the 1980s. The Group Health Journal 1(1):34-40, 1980.

8. Greenlick, M.R., et al. Kaiser-Permanente's Medicare Plus Plan: A Successful Medicare Prospective Payment Demonstration. Health Care Financing Review 4(4):85-97, 1983.

9. Thomas, J.W., et al. Increasing Medicare Enrollment in HMOs: The Need for Capitation Rates Adjusted for Health Status. Inquiry 20(3):227-39, 1983.

10. McClure, W. On the Research Status of Risk-Adjusted Capitation Rates. Inquiry 21(3):205-13, 1984.

11. Ibid. DesHarnais.

12. Beebe, J., Lubitz, J. and Eggers, P. Using Prior Utilization to Determine Payments for Medicare Enrollees in Health Maintenance Organizations. Health Care Financing Review 6(3):27-38, 1985.

13. Yesalis, C.E. and Bonnet, P.D. The Effect of Duration of Membership in a Prepaid Group Health Plan on the Utilization of Services. Medical Care 14(12):1024-36, 1976.

TABLE 1

HISTORIC AND PROJECTED UTILIZATION
FOR SSI AND AFDC MEDICAID ENROLLEES
(Rates per 1,000 per Year)

	SSI		AFDC	
	HISTORIC	PROJECTED	HISTORIC	PROJECTED
Hospital Inpatient Days	6,048	3,600	1,171	600
Physician Visits	29,400	12,000	10,800	4,000
Hospital Outpatient Visits	1,657*	4,500	653*	3,000
Emergency Room Visits	200*	400	150*	200
Outpatient Laboratory Tests	3,315	6,000	2,176	4,500
Outpatient Radiology Tests	830	1,300	218	800
Outpatient Prescription Drugs	28,570	18,000	13,600	7,000
Dental Services	412	412	1,200	1,200
Eyeglass Prescriptions	540	540	280	280
Hearing Services	50	50	1	1
Home Health Visits	571	800	400	400
Transportation	1,100	1,000	70	70
Screening Visits	-	100	540	750
Family Planning Visits	-	1	40	40
Independent Lab/X-ray Tests	682	250	200	150

*Estimated based on aggregate historic data for hospital outpatient and emergency room visits.

TABLE 2

INITIAL EXPERIENCE OF HMO MEDICARE DEMONSTRATION PROGRAMS
(Units per 1,000 Members per Year)

Acute Hospital	1,700 - 2,880 days
Outpatient Physician Services	5,730 - 5,875 visits
Outpatient Nonphysician Services	1,890 - 2,300 tests
Laboratory	5,470 - 6,920 tests
X-Ray	1,130 - 1,490 tests
Pharmacy	9,010 - 10,730 prescriptions
Refractions or Eyeglasses	560 - 660 sets
Hearing Aids	106 aids
Home Health	300 - 450 visits

Source: Walter N. Leutz, et al. Changing Health Care for an Aging Society, Lexington, Massachusetts, Lexington Books, p.204.

TABLE 3

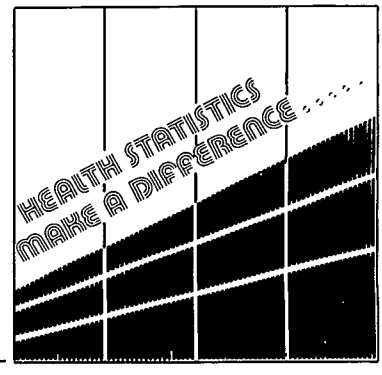
FLORIDA MEDICAID PROGRAM UTILIZATION ASSUMPTIONS
(Rates per 1,000 Enrollees per Year)

	SSI	AFDC
Hospital Inpatient Days	3,300 - 4,100	500
Physician Visits	5,100 - 7,100	4,000
Emergency Room Visits	75 - 200	125
Laboratory Tests	3,500 - 5,600	2,500
X-Ray Tests	550 - 1,000	400
Prescription Drugs	9,300 -15,900	4,000
Home Health Visits	300 - 500	25

Source: Florida Medicaid Program - Alternative Health Plan

Session R

Issues in the Quality of Vital Statistics Data



SOURCES OF DATA ERRORS ON BIRTH CERTIFICATES AS DETERMINED BY HOSPITAL CHART REVIEW IN VERMONT

Gail Rushford, Vital Statistics Vermont

Today I want to tell you about some of the things that I learned during my first few months as a Field Representative with Vermont's Division of Public Health Statistics. What I learned in the field surprised the Vital Statistics staff and caused us to make several changes in the way we work with birth certificate data. For instance, we were surprised to find that two different people working from the same medical chart, can come up with two different answers to the question, "In which month of the pregnancy did the mother begin prenatal care?" We learned that the fact that one hospital consistently reports a lower number of prenatal visits than others do does not necessarily mean that that hospital is actually serving mothers who received less prenatal care. I'll explain these findings as I go on, as well as tell you what we did about them.

We learned that Vermont had these and other similar problems as a result of a survey that I did in November and December of 1984. For the past several years, Vermont's field program had focused primarily on registration issues and the field representative's time was spent with the local registrars. I joined the staff in 1984 to deal with statistical issues and data quality. I began with the birth certificates. Most Vermont births (98%) occur in hospitals, as they do nationally, so I knew that I would be doing most of my work with the hospitals. However, before addressing the problems in the hospitals that affect data quality, I needed to identify those problems and their causes.

Vermont is a small state. We have just 14 hospitals with maternity wards. Our largest hospital delivers approximately 2000 babies per year, or one quarter of all Vermont births, and our smallest has around 30 deliveries per year. No matter how you look at it, we are never dealing with large numbers. Our size makes it possible for us to do projects that would be much more costly and time consuming for other states.

STUDY DESIGN

The purpose of this project was not to gather scientifically reliable numbers and statistics regarding the completion of birth certificates, but to get a sense of what was required to complete them and how reliable we could expect them to be. In Vermont, hospital personnel prepare the birth certificate for the attending physician's signature. Most of our hospitals make use of the worksheets provided by the Health Department. The worksheet is a duplicate of the birth certificate form and is used to gather the information that will be recorded on the certificate. The original birth certificate is filed with the Clerk of the town where birth occurred. The Town Clerk registers the birth and forwards a copy of the certificate, along with the confidential section, to the Health Department.

I made a field visit to each of the 14 hospitals to conduct a data quality survey. At each hospital, I interviewed the personnel responsible for completing the birth certificates, usually medical records or nursing staff. I asked them to describe the procedures and data sources used in completing the certificates. I also completed a hospital chart review to determine the accuracy of a sample of birth certificates. Statewide, 334 certificates were sampled. I compared the information that I found in the charts with what was recorded on the hospital work-sheet and the original birth certificate. Any time the information differed on one of these three sources, a discrepancy was counted.

Vermont's birth certificate contains all of the items that are found on the U.S. standard certificate as well as items for "weeks of gestation" and "dates of prenatal blood tests". (Figure 1) I looked for discrepancies in all legal items and all confidential items except for "dates of blood tests" and the "complications" section.

GOOD DATA

While we were most interested in finding and fixing problem areas, it is worth noting that the quality of reporting was very good on most items. The legal section was generally good, with only 32 total discrepancies counted out of 334 records with 27 legal items per record. Most of these can be attributed to typographical errors or discrepancies within the medical records.

Of the medical items, only 18 discrepancies were found in reported birthweights and 12 in the Apgar Scores. Most of these can be attributed to discrepancies in the charts. There were 7 discrepancies in the legitimacy item, most of which reflected a difference between the information on the hospital's worksheet and what was ultimately recorded on the birth certificate. It is difficult to judge the accuracy of the self-reported items for race and education of the parents. For those cases where information was available in the charts, there were 2 discrepancies in the father's race and 8 in his education. I noted no discrepancies in the mother's race and 7 in her education.

PROBLEM IDENTIFICATION

In spite of those encouraging results, I found that there is more potential for error on the confidential portion of the birth certificate than we had anticipated. I counted a total of 205 discrepancies in the confidential sections. (This does not reflect the number of records that had discrepancies.) There are several items on the certificate which currently require individual calculations. Varying methods of calculation result in inconsistent reporting on the statewide level. The data items where these inconsistencies are most apparent are "weeks of gestation", "number of prenatal visits", and "month prenatal care began". Another area where discrepancies were noted was the pregnancy history section. This is due to confidentiality issues. I'll explain to you why we believe that these items have not been as accurate as they

could be and what we are doing in Vermont to make them more exact.

As I mentioned, the Vermont birth certificate asks for both the date of the last menstrual period and the weeks of gestation. Our assumption, and our intention, was that we were getting a physician's estimate of the baby's gestational age based on an examination of the baby. This would supplement our data regarding weeks of gestation as calculated by dates. We learned that what we were often getting instead was just another calculation of the weeks of gestation by dates.

Procedures for obtaining this information included:

1. Medical records personnel calculating the weeks by the date of the last normal menses and the date of birth. Some of them would calculate from two weeks after the last menses to the date of birth.

2. Delivery room personnel calculating the weeks by dates and recording the number in the mother's chart.

3. The obstetrician determining the weeks of gestation by exam.

4. The pediatrician determining the weeks of gestation by exam.

I found that often the person completing the birth certificates did not know if the gestational age recorded in the medical records represented a calculation of the dates or a physician's determination based on examination of the infant. The lack of definition has made this data unusable.

The "month that prenatal care began" is another item that required medical records personnel to calculate the correct answer based on dates. The instructions for that item indicate that the month of the pregnancy that the mother had her first prenatal visit should be recorded. The month that care began is calculated from the date of the last normal menses to the date of the first prenatal visit. Some people count the months of pregnancy by rounding off the dates and counting whole calendar months and others count the months in four week increments. That leads to discrepancies. In doing my survey, I found 34 such discrepancies (10%) out of the 334 records reviewed. Let me give you an example of this sort of discrepancy. If a woman began her last normal menses on May 20, 1985 and had her first visit for prenatal care on July 8, 1985, you could calculate that she had her first prenatal visit in her second month of pregnancy by counting from May 20 to June 20 to July 8 or in her third month by counting the whole months of May, June and July.

The Handbook that is published by NCHS for use by the hospitals in completing birth certificates does not specify whether or not the months should be rounded off. The birth certificate contains no instructions, including the fact that the length of the pregnancy is measured from the date of last menses and not from the date of conception.

The "prenatal visit" data is also a problem, due mainly to the lack of defined standards for counting visits. We were not able to determine the accuracy of this data on the statewide level because it was not comparable among the hospitals. Since most doctors send their prenatal care records to the hospital two or more weeks before the expected delivery date, several visits can get "lost" if just the number of recorded visits are counted. Generally, the doctors do not update the prenatal data in the hospital file and the mother often does not remember how many visits she made to the doctor. Most of the people that I interviewed told me that in calculating the number of visits made for prenatal care they would assume that a certain number of visits were made between the date of the last recorded visit and the delivery and then adjust the number of visits accordingly. However, at two of our hospitals, this is not done. Only the visits that are recorded on the prenatal record are counted and recorded on the birth certificate. This variance in standard procedures can cause these two hospitals to appear to be serving mothers who are receiving less prenatal care than the mothers who gave birth in the rest of Vermont hospitals. (Figure 2)

The standards vary for determining what visits are counted among hospitals. Visits that are made for lab work and ultrasounds may or may not be counted in the total number of visits made for prenatal care. Some hospitals counted all visits and some disregarded visits that did not appear to involve physical examination. There are no guidelines in the NCHS handbook regarding what type of visits should be counted or disregarded.

One of our most difficult problem areas, due to the sensitive and private nature of the information, is the pregnancy history. These items are the same on the Vermont birth certificate as they are on the U.S. Standard Certificate. The following questions are asked about the mother's pregnancy history: How many children born alive to her are still living? How many are now dead? What was the date of the last live birth? How many spontaneous and induced terminations of pregnancy has she had before 20 weeks of gestation? How many after 20 weeks? What was the date of the last other termination of pregnancy?

What I discovered during the interviews and the chart review is that a conflict often arises between accuracy and confidentiality. The mother's chart may contain information indicating that she had a child who was given up for adoption or that she had a miscarriage or abortion. Most of our hospitals have a procedure for verifying the information with the mother. When the mother denies this information or refuses to have it recorded on the birth certificate, the person responsible for completing the birth certificate is in a difficult position. The policy at most of the hospitals is to respect the mother's wishes regarding the pregnancy history data that will be revealed on the birth certificate. Some of the people that I interviewed have stated that their main concern is to protect the confidentiality of their patients' records and they will not release information that the patient does not want released or is not aware is being released.

Adding to the problem is the fact that the

chart may not be complete, making it necessary to ask the mother for the information. The mothers often can not remember the months of their last live births and months are often not given in the chart. If the chart contains information about terminated pregnancies, the dates may be missing altogether or only the year might be given. There may not even be a space in the prenatal record for the month of the occurrence; although, there is usually a space for the year.

The chart review yielded 27 discrepancies on the item for "number of terminations before 20 weeks". This represented the largest number of discrepancies in the pregnancy history section. All of these 27 discrepancies represented under-reporting on the birth certificate. While these 27 discrepancies affected 8% of all records reviewed, they affected 28% of the records of the 96 women whose medical records indicated a pregnancy that had ended before twenty weeks of gestation.

One final point that I would like to make regarding discrepancies in general, which was brought home to us by this survey, is that completeness can not and should not be the only measure of data quality. Many of those 205 discrepancies that I found were caused by simple human error -- carelessness, poor record keeping and oversight. For example, the "date of last normal menses" resulted in 37 discrepancies (11%) that were caused by mistakes or omissions.

SOLUTIONS

One we had identified these problem areas we began to work to solve them. First we developed a manual to be used by the hospitals and midwives in the completion of birth certificates. The manual contains new procedures for the completion of several items. These procedures reflect our efforts to eliminate the need for individual calculations and to develop clearly defined standards to be used when completing the certificates. We have instructed the data providers to begin using the new procedures right away. Any necessary language changes will be incorporated into the next revision of our certificates.

For the "weeks of gestation" item, we specify that we want an estimate of the baby's gestational age based only on a physical examination of that baby by a physician or midwife. If that information is not available, the item is not to be left blank, but the words "unknown" or "not given" should be filled in. We will not query this item as long as we have the date of the last menses so that our computer can calculate the weeks of gestation by dates. The birth certificates and worksheets will be changed to read "physician's estimate of infant's gestational age".

We are now asking for the date of the mother's first prenatal visit instead of the "month of pregnancy prenatal care began". The month will be calculated by computer. To get a more accurate count of the "total number of prenatal visits", we are asking for the total number of recorded, verifiable visits as well as the date of the last recorded visit. On our next revision we will have two separate boxes for these entries. We have not yet decided whether visits for lab work and ultrasounds should be counted.

In response to the confidentiality issue, we have encouraged a procedural change. The confidential sections of our birth certificates are detachable from the carbon copy of the legal section. As I mentioned earlier, the hospitals previously forwarded the entire certificate intact to the local town clerk who, in turn, forwarded the carbon copy and the confidential section to the vital statistics office. We are now encouraging the hospitals and midwives to forward the confidential section directly to us when the legal section is sent to the town clerk for registration. This is a small change which we hope will demonstrate our sincere commitment to preserving confidentiality and lessen resistance to providing personal information on the birth certificates.

We are no longer querying unknown months for pregnancies that ended more than three years ago. We hope that this will also help to make completion of the pregnancy history section a little less difficult.

As I mentioned, these changes are included in our new birth certificate manual. The manual was based on the NCHS handbook regarding the completion of birth certificates, but our manual contains more detail and is geared specifically to Vermont. We sent the manual out to all of our hsopitals. We also held a workshop on the changes for everyone who is responsible for completing birth certificates. The meeting was well attended and those people who didn't attend received a personal visit from me. I also attended a meeting of a midwives group. I put a lot of effort into communicating these changes that we had made because we would not see any improvement in our data quality if nobody knew about all of these changes.

We see ongoing communication as the key to ensuring the quality of our data. We are maintaining contact with the hospitals through our field program. As field representative, I will be making use of quarterly newsletters and regular field visits throughout the year in order to keep both the hospitals and Vital Statistics office well-informed. We also plan to make an annual event out of our first successful joint meeting between hospital and Vital Statistics representatives to share ideas and concerns regarding vital records.

Two other improvements that we hope to make in the future are development of a prenatal worksheet and computerization of the birth registration process. The worksheet would be distributed to the obstetricians' offices to be completed by the expectant mothers. Most of the information that is presently on the birth certificate would be on the worksheet so that the parents' full names, ages, places of birth, race, and education, and the mothers' pregnancy histories would be in the mothers' charts and ready to enter on the birth certificates. This would eliminate the need to ask all of the questions of the mothers after the birth occurs. We hope that this worksheet will benefit us in two ways. One, it should make the job of completing the birth certificates a bit easier and less time-consuming. Second,

it might make confidentiality less of an issue. If the mothers are providing the Health Department with their pregnancy histories themselves it is less threatening than having that data reported by a third patry. Regarding computerization, we hope to eventually join some of the other states who are currently receiving birth records directly from the hospitals via computer diskettes. This would help to improve timeliness and accuracy as most editing would be done at the hospital level.

CONCLUSION

When I was considering what thoughts I would like to conclude with today, I realized that I have begun using a very basic philosophy in my work with the hospitals. In order to get the most accurate and useful data possible, we need to make the job of providing the data as simple as possible. We have tried to do that in Vermont through the methods that I have just described to you. We changed the definitions of any items that required calculation so that now the information requested can simply be transferred from medical records to the certificate. This makes the job a bit easier. We provided a manual which defines all items and preferred procedures clearly. Those changes and the proposed worksheet and computerization, should also help to save time, which is always an important factor in simplification. We have also instituted procedural changes to make confidentiality less of an issue.

We realize that completing vital records is not the sole priority on our data providers' lists. Therefore, we will continue to look for ways in which to simplify the job. We are confident that this will aid us in our goal of obtaining even more accurate and complete data.

BIBLIOGRAPHY

1. Carucci, Peter M.: Reliability of Statistical and Medical Information Reported on Birth and Death Certificates. New York State Department of Health Monograph No. 15, Albany, N.Y., May, 1979.
2. Hospital Handbook on Birth Registration and Fetal Death Reporting. US DHEW Publication No. (PHS) 78-1107, Hyattsville, MD., February, 1978.
3. Querec, Linda J.: Characteristics of Births, United States, 1973-1975. US DHEW Series 21, No. 30, PHS 78-1908, Hyattsville, Md., September, 1978.
4. Querec, Linda J.: Comparability of Reporting Between the Birth Certificate and the National Natality Survey. US DHEW Series 2, No. 83, PHS 80-1357, Hyattsville, Md., April, 1980.
5. Taffel, Selma: Prenatal Care, United States: 1969-1975. US DHEW Series 2, No. 33, PHS 78-1911, Hyattsville, Md., September, 1978.

ATTENDING PHYSICIAN, MIDWIFE, OR HEAD OF FAMILY shall complete this certificate and file it within 10 days after birth with the clerk of town/city in which birth occurred (Title 18, V.S.A. Chapter 103, §5071)

DH-PHS-1-83

No. 23802

VERMONT DEPARTMENT OF HEALTH

CERTIFICATE OF LIVE BIRTH

LOCAL FILE NUMBER

STATE FILE NUMBER 144

TYPE, OR PRINT IN PERMANENT BLACK INK PRESS FIRMLY

CHILD

CHILD—NAME (First, Middle, Last)	DATE OF BIRTH (Month, Day, Year)	HOUR
1	2a	2b m

SEX	THIS BIRTH—SINGLE, TWIN, TRIPLET, ETC (Specify)	IF NOT SINGLE BIRTH—BORN FIRST, SECOND, THIRD, ETC (Specify)	COUNTY OF BIRTH — VERMONT
3	4a	4b	5a

CITY, TOWN OF BIRTH	FACILITY—NAME (If not in facility give street and number)	IF HOSPITAL
5b	5c	☐ Inpatient ☐ Enroute 5d

MOTHER

MOTHER—MAIDEN NAME (First, Middle, Last)	AGE (At time of this birth)	TOWN AND STATE OF BIRTH (If not in U.S.A. Name country)
6a	6b	6c

MOTHER'S MAILING ADDRESS (Including Street or Rural Number and Zip)

7

RESIDENCE—IN WHAT CITY OR TOWN DOES 8a MOTHER ACTUALLY LIVE	STATE	COUNTY	CITY, TOWN
		8b	8c

FATHER

FATHER—NAME (First, Middle, Last)	AGE (At time of this birth)	TOWN AND STATE OF BIRTH (If not in U.S.A. Name country)
9a	9b	9c

INFORMANT	RELATION TO CHILD
10a	10b

CERTIFIER

I certify that the above named child was born alive at the place, date and with the facts stated above

ATTENDANT: ☐ MD-DO ☐ RN, LPN, PA, EMT ☐ FAMILY MEMBER ☐ CNM ☐ LAY—MIDWIFE ☐ OTHER

11a SIGNATURE	11b	DATE SIGNED (Month, Day, Year)
		11c

CERTIFIER—NAME (TYPE OR PRINT)	MAILING ADDRESS	CITY OR TOWN, STATE, ZIP
11d	11e	

REGISTRAR

REGISTRAR—Signature	DATE RECEIVED BY LOCAL REGISTRAR (Month, Day, Year)
12a	12b

TO BE SIGNED BY REGISTRAR ON COPY ONLY

TRUE COPY CLERK—Signature	TOWN	DATE (Month, Day, Year)
13a ATTEST	13b	13c

INFORMATIONAL COPY ONLY

CONFIDENTIAL INFORMATION FOR MEDICAL AND HEALTH USE ONLY. NOT FILED AS PART OF PUBLIC RECORD

No. 23802

ATTENDANT THIS SECTION MUST BE COMPLETED (Title 18, V.S.A. Chapter 103, §5071)

HEALTH DEPARTMENT PERSONNEL: DETACH FOR STATISTICS AND DESTROY

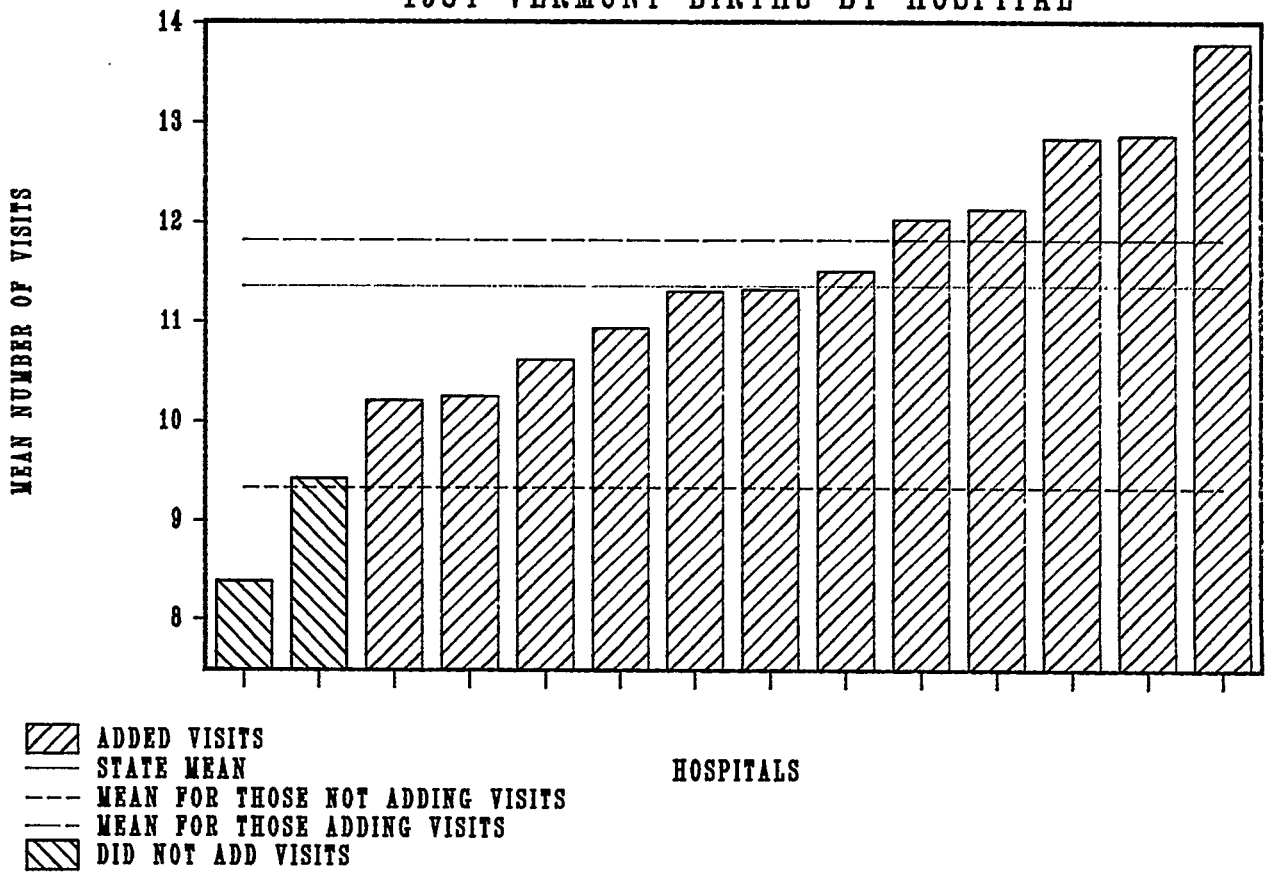
FATHER RACE	EDUCATION—SPECIFY HIGHEST GRADE COMPLETED		PREGNANCY HISTORY (Complete each section)	
WHITE, NEGRO, AMERICAN INDIAN, ETC. (Specify)	Elementary or Secondary (0-12)	College (1-4 or 5+)	LIVE BIRTHS (Do not include this Child)	OTHER TERMINATIONS (Spontaneous and Induced)
15	16		17a Now living / 17b Now dead	17d Before 20 weeks / 17e After 20 weeks
MOTHER RACE	EDUCATION—SPECIFY HIGHEST GRADE COMPLETED		Number ___ None ☐ / Number ___ None ☐	Number ___ None ☐ / Number ___ None ☐
WHITE, NEGRO, AMERICAN INDIAN, ETC (Specify)	Elementary or Secondary (0-12)	College (1-4 or 5+)		
18	19			

DATE LAST NORMAL MENSES BEGAN (Month, Day, Year)	WEEKS OF GESTATION	MONTH PREGNANCY PRENATAL CARE BEGAN FIRST, SECOND, THIRD, ETC (Specify)	PRENATAL VISITS TOTAL NO (If none, so state)	DATE OF LAST LIVE BIRTH (Month, Year)	DATE OF LAST OTHER TERMINATION (as indicated in d or e above) (Month, Year)
20	21	22a	22b	17c	17f

DATES OF BLOOD TESTS (STS) FIRST TEST Month, Day, Year	SECOND TEST Month, Day, Year	IF NONE, GIVE REASON	WAS MOTHER MARRIED
24a	24b	24c	25 ☐YES ☐NO

COMPLICATIONS OF PREGNANCY (Describe or write "none")	BIRTHWEIGHT	WAS SILVER NITRATE OR OTHER SUITABLE PROPHYLACTIC USED IN BABY'S EYES?
26		☐YES ☐NO
CONCURRENT ILLNESSES OR CONDITIONS AFFECTING THE PREGNANCY (Describe or write "none")	30	23
27		
COMPLICATIONS OF LABOR AND OR DELIVERY (Describe or write "none")	DELIVERY ☐VAGINAL ☐CAESAREAN	
28	31	
CONGENITAL MALFORMATIONS OR ANOMALIES OF CHILD (Describe or write "none")	APGAR SCORE 1 min / 5 min	IS INFANT LIVING AT TIME OF REPORT?
29	32	33 YES ☐ NO ☐

PRENATAL VISITS FOR INFANTS 40-42 WEEKS
1984 VERMONT BIRTHS BY HOSPITAL



AN IMPROVED SYSTEM FOR REPORTING CONGENITAL MALFORMATIONS ON THE BIRTH CERTIFICATE

Stephen Minton, M.D. and Robert E. Seegmiller, Ph.D.

Information on state birth certificates pertaining to congenital malformations allows states to monitor the distribution and changes in incidence of these disorders and to study their associated risk factors. Such information can be used in epidemiological studies to detect new syndromes and to educate the public regarding causation and prevention. However, the reliability of the birth certificate in terms of completeness and accuracy of recorded medical and health information is often uncertain, and therefore, the usefulness of this document is limited.

Several investigators have attempted to measure the completeness of reporting congenital malformations on birth certificates by examining either records at a few hospitals such as the studies by Lilienfeld (1), Montgomery(2), Oppenheimer (3), and Mackeprang (4), or by examining records in certain selected malformations in the Milham (5), or Bock (6) studies. These studies have shown underreporting of congenital malformations on birth certificates in widely varying degrees from 0-75%.

In a retrospective study of seven Utah hospitals published in 1981, Seegmiller et al. (7) compared 1968-1972 birth certificates against hospital records and noted that congenital malformations were inaccurately recorded on the birth certificate. As seen in Table 1, the four smaller hospitals were better than the three larger hospitals in terms of completeness and accuracy of reporting. Utah Valley Regional Medical Center (UVRMC), with the largest number of births in the study (7158 in 1970, 1972), reported only 12% of the total malformations and only 44% of marker malformations. We concluded that although birth certificates may be useful for determining rates and associated factors of certain marker malformations, they should not be used in their present state to provide a complete picture of the occurence of congenital malformations.

In 1980 we initiated a study to evaluate UVRMC's method of reporting congenital malformations. We found that the major cause for the discrepancy between the data reported on birth certificates and the actual incidence of congenital malformations was related to the reporter. In many instances the deliverer of the baby had the responsibility of completing and signing the birth certificate while the infant exam was performed by the nurse and/or a different physician. In addition, if the chart work was not completed at the time of discharge, malformations were often missed when the birth certificate was completed retrospectively. These findings confirmed that of Hay (8) and Mackeprang (4) that the procedure for documenting and reporting congenital malformations is poor. If the birth certificate was to be used reliably in determining malformation rates and associated risk factors, a system for improved completeness and accuracy was needed.

The present study was undertaken to establish a more accurate system of reporting congenital malformations on the birth certificates of liveborn children.

METHODS AND PROCEDURES

This study was performed in two parts.

Part I: In early 1981, two procedural changes for recording and reporting congenital malformations on the birth certificate were initiated at Utah Valley Regional Medical Center. These changes were 1) the responsibility for recording congenital malformations was transferred, by hospital policy, from the deliverer of the baby to the baby's physician, and 2) a congenital malformation reporting sheet was included in the front of each newborn chart. The baby's physician was instructed to fill it out during the hospitalization and review it at the time of discharge. An extensive orientation to the worksheet was carried out in Obstetric, Family Practice and Pediatric Department Meetings as well as written instruction was given to each physician in those departments. The congenital malformation worksheets were collected from May-September 1981 and their corresponding newborn charts were reviewed.

Part II: The second part of the study involved the appointment of a single, centralized medical record's person to review the newborn charts (physician and nurse exams) and the congenital malformation reporting sheet and to personally complete the birth certificate on all babies born at UVRMC in 1982. Photocopies of the completed birth certificate for each child born during 1982 at UVRMC was reviewed by date of birth.

To evaluate the accuracy and completeness of birth certificate reporting, UVRMC newborn patient files were retrospectively evaluated. These files consisted of charts of a) all children with any congenital malformation(s) indicated on the birth certificate; b) all babies which had been transferred to the Newborn Intensive Care Unit (NBICU), exclusive of the above; and c) every tenth of the remaining charts pulled in order of birth, exclusive of the above-mentioned children.

Each of the above patient medical files was reviewed with specific attention given to a) the "Nursing Assessment" (performed in the delivery room by the attending obstetrical nurse); b) "Neonatal Nursing Admission Assessment" (performed at the time of baby's admission to the nursery); c) "Physician's Record of Newborn"(performed after the physician examination); and d) "Top Sheet" (final diagnostic physician summary record). This information was then compared with its corresponding birth certificate to determine the accuracy of reporting and the likelihood of a malformation being reported on the birth certificate.

As a standard for classifying congenital malformations, we used the Eighth Revision International Classification of Diseases, Adapted for Use in the United States, (1965, ICDA). It was not determined whether the malformations required corrective surgery or therapy, or resulted in severe physical or mental handicap; thus we did not classify malformations as major or minor as Regemorter (9) or Mackeprang (4). "Conditions of the newborn" such as birth asphyxia, prematurity,

Respiratory distress and transient heart murmurs, in addition to conditions which were of little clinical significance (e.g. hip click when the physician felt no therapy was indicated, simple birth marks, skin taps, masses and pilonidal dimples) were not regarded as congenital malformations in the present study. Where an infant had three or more anomalies the individual was classified as having multiple malformations.

RESULTS

Part I: In the congenital malformation worksheet study in 1981, 2361 congenital malformation worksheets were filled out by the pediatrician or family practitioner and compared to the corresponding patient record. Ten worksheets had no malformation when in fact a malformation was recorded on the chart. Thus, there was a 99.58% accuracy of reporting a malformation on the worksheet as compared with the medical record.

Part II: There were 4949 live births at UVRMC in 1982. There were 236 children, or 4.8 percent of the total live births, with one or more malformations listed on the birth certificate.

A total of 1015 UVRMC newborn patient files (20.5% of the total live births) were then retrospectively evaluated. These files consisted of a) the 236 charts of all children with any congenital malformation indicated on the birth certificate; b) the 286 charts of all babies which had been transferred to the Newborn Intensive Care Unit (NBICU), exclusive of the above; and c) the 493 charts selected by every tenth of the remaining charts pulled in order of birth, exclusive of the 522 above-mentioned children.

The comparison of the birth certificate and hospital record found errors of three major types (Table 2).

A). The reporting of malformations inaccurately (Table 2-A). Five babies had inaccurate classification, i.e., four babies reported having heart murmurs actually had congenital heart disease, and one baby reported to have multiple malformations had only hydrocephalus. One child had an inaccurate diagnosis of congenital hip dislocation when it was actually a well baby.

B). The reporting of malformations incompletely (Table 2-B). A child diagnosed as having hypospadias and gastroschisis was reported as having only gastroschisis; two children diagnosed as having meningocele and hydrocephalus had one of the malformations omitted; and a child diagnosed as having multiple anomalies was reported as having only clubfoot.

C). The reporting of non-malformations as malformations (Table 2C). On the birth certificates of 120 children, 126 conditions of the newborn such as growth retardation, transient heart murmur, hip click, pilonidal dimple, birth marks, skin tags, etc. were inaccurately reported as either the sole malformation(s) listed or were listed in association with a true malformation. The more frequently listed non-malformations were heart murmur, hip click, pilonidal dimple, and pigmented nevi.

Thus, there were a total of 136 errors, only ten of which were due to recording or transfer errors (six inaccurate, four incomplete) and 126 were due to non-malformations being reported by the physician as malformations.

Of the original 236 birth certificates, 116 had a true malformation listed, 12 had a malformation and a non-malformation listed in association, and 108 had a non-malformation as the sole malformation listed (Table 3). This means there were 128 children or 2.6% of the total live births with 148 malformations and 120 children with 126 non-malformations.

DISCUSSION

If one looks at the transference of the physician's and nurse's observations to the birth certificate, this system improved the accuracy and completeness of reporting from 26% in 1970, 1972, to >98% in 1982. Although there were ten errors related to accuracy and completeness of reporting in the 128 malformed children, not one of the 128 children went unreported, i.e., at least some mention was made of the malformations in each case. The comparison of 1015 UVRMC newborn records revealed no further congenitally malformed children than had already been ascertained directly from the birth certificate. Thus, the considerable underreporting of malformations on birth certificates seen in other studies was not seen in this system.

If one looks at the accuracy of the system as a reflection of actual congenital malformations reported by the physician or nurse and defines non-malformations reported by the physician or nurse as errors in reporting, then this system has only a 45% accuracy in reporting. The major deficiency noted in the present study was the significant reporting of non-malformations as malformations. One-hundred-twenty-six non-malformations were listed as malformations. Other researchers have noted a similar dilemma. Non-standardization of definitions and terminology has lead to confusion on what to report. Some reports have listed some of these under the insignificant category or trivial category. However, usually these were still greatly under-reported versus major or minor malformations or almost completely ignored.

It appears that the inclusion of the congenital malformations worksheet in the front of the chart raised the level of compliance on the part of the attending physician such that even the insignificant and non-malformations were reported.

This lead us to develop a "Classification of Congenital Malformations" for the physicians and the specified medical records reporter which not only listed true congenital malformations but addressed "non-malformations" such as heart murmur or hip click. For example, when a heart murmur is listed, the recorder uses the following algorhythm to establish the correct completion for the birth certificate (Figure 1). She reviews the records to see if the murmur was present or absent at discharge, whether the patient was transferred to another institution for evaluation of possible congenital heart disease or if a diagnosis of congenital heart disease was made. She then follows the appropriate pathway to determine what to record. This removes the guesswork of the worker in deciding whether a condition is a true malformation or not and refers it back to a physician.

In summary, the institution of the baby's physician as the reporter, a congenital malformation worksheet included in the front of the newborn chart, and a single centralized medical record's recorder for filling out the birth certificate improved the completeness of reporting congenital malformations on birth certificates from 26% (1970, 1972) at UVRMC, to <98% (1982). Nevertheless, the following problems still remained: a) 6 newborns of the 106 reported with a congenital malformation had an inaccurate congenital malformation listed on the birth certificate. b) 4 newborns' birth certificates listed one congenital malformation but missed a second malformation. c) 120 newborns had non or insignificant congenital malformations listed as malformations, i.e., heart murmur, pilonidal dimple, etc. An education process with algorhythms was instituted with the centralized medical records recorder after the study to eliminate the reporting of non-congenital malformations. It is anticipated that this system can be a very accurate surveillance mechanism for monitoring the incidence of congenital malformations and can easily be used by other hospitals to improve their reporting.

RECOMMENDATION

A) Physician of baby is Recorder
B) Centralized Reporter
C) Congenital Malformation Worksheet
D) Expanded Classification of Congenital Malformation Guide

TABLE 1

COMPARISON OF THE BIRTH CERTIFICATE AND HOSPITAL RECORD FOR COMPLETENESS OF REPORTING TOTAL MALFORMATIONS 1968-1972

County and Institution	Years Examined	Live Births	Total Malformations: Hospital Record	Total Malformations: Birth Certificate	Percent Reported on Birth Certificate
Duchesne County					
Duchesne County Hospital	1968-1972	1,322	51	21	41
San Juan County					
San Juan County Hospital	1968-1972	788	38	22	58
Monument Valley Hospital	1968-1972	1,152	66	29	44
County Total		1,940	104	51	49
Uintah County					
Uintah County Hospital	1968-1972	1,173	25	15	60
Utah County					
American Fork Hospital	1970,1972	1,235	41	14	34
Payson City Hospital	1970,1972	1,109	40	16	40
Utah Valley Hospital	1970,1972	7,158	358	44	12
County Total		9,502	439	74	17
Total		13,937	619	161	26

TABLE 2
ERRORS IN REPORTING CONGENITAL MALFORMATIONS ON THE BIRTH CERTIFICATE
UVRMC--1982

Birth Certificate	Hospital Record	No.
	PART A INACCURATE REPORTING	
heart murmur	cong. heart disease	3
heart murmur	cong. heart dis./club foot	1
multiple	hydrocephalus	1
cong. hip dislocation	well baby	1
total inaccurate		6
	PART B INCOMPLETE REPORTING	
gastroschisis	hypospadias/gastroschisis	1
hydrocephalus	meningocele/hydrocephalus	1
meningocele	meningocele/hydrocephalus	1
clubfoot	multiple	1
Total incomplete		4
	PART C NON-MALFORMATIONS	
heart murmur	heart murmur	51
hip click	hip click	23
pilonidal dimple	pilonidal dimple	17
birth mark (pigmented nevi)	birth mark (pigmented nevi)	16
tags	tags	8
skin growth	skin growth	3
partial skull closure	partial skull closure	2
cerebral atrophy	growth retardation	1
foot slightly diverted	foot slightly diverted	1
questionable musculoskeletal	questionable musculoskeletal	1
hemangioma	hemangioma	1
cystic mass	cystic mass	1
left abdominal mass	left abdominal mass	1
Total non-malformations		126
Total Errors in Reporting		136

TABLE 3

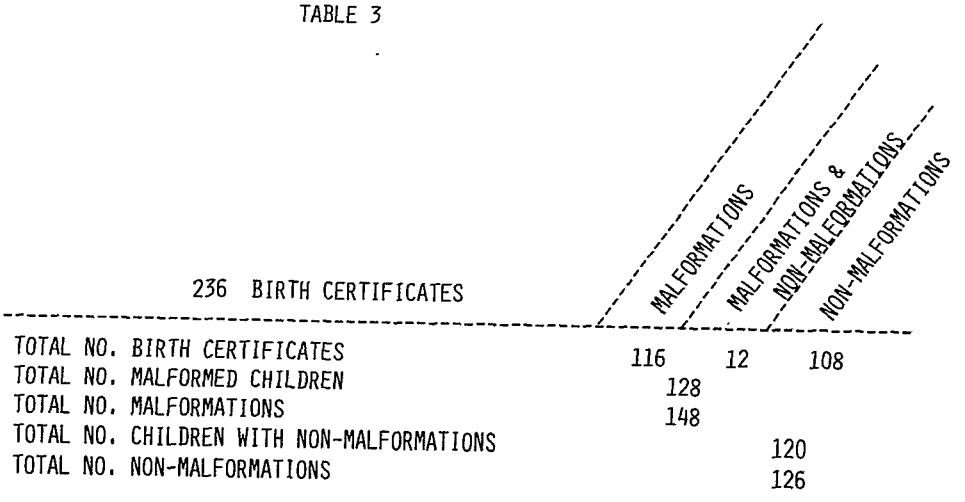
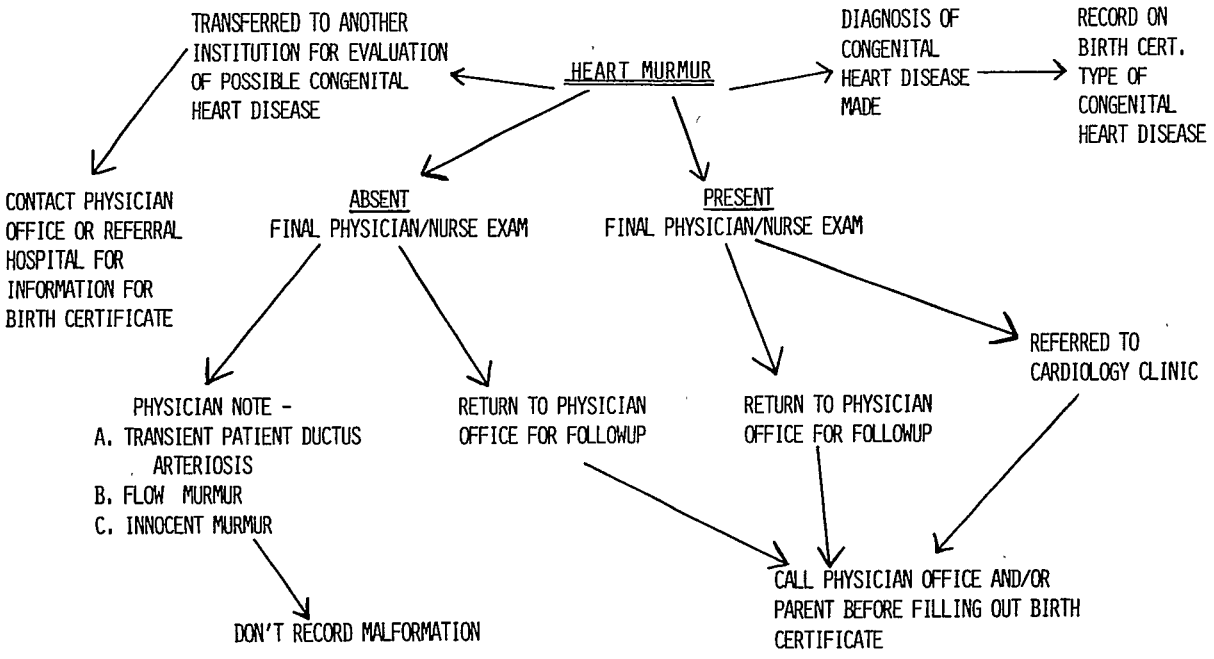


FIGURE 1



IF ANY QUESTION - CALL PHYSICIAN'S OFFICE TO DETERMINE IF CONGENITAL MALFORMATION AND WHAT KIND.

BIBLIOGRAPHY

1. Lilienfeld, A.M., Parkhurst E., Patton R., et al.: Accuracy of supplemental medical information of birth certificates. Public Health Rep. 66:191-198, February, 1951.
2. Montgomery, T.A., Lewis, A., and Hammes L.: Live birth certificates-evaluation of medical and health data in California. California Med. 96:190-195, March, 1962.
3. Oppenheimer, E., et al.: Evaluation of obstetric and related data recorded on vital records and hospital records: District of Columbia, 1952. National Office of Vital Statistics--Special Reports, Selected Studies. Vol. 45, No. 13, U.S. Government Printing Office, Washington, D.C., Nov 20, 1957.
4. Mackeprang, M., Hay S., Lunde, A.S.: Completeness and accuracy of reporting malformations on birth certificates. HSMHA Health Reports. Vol 87, January, 1972, pp. 43-49.
5. Milham, S., Jr.: Underreporting of incidence of cleft lip and palate. Am. J., Dis. Child. 106:185-188, August, 1963.
6. Bock, H.B., Zimmerman, J.H.: Study of selected congenital anomalies in Pennsylvania. Public Health Reports. Vol. 82, No. 5, May, 1967,pp. 446-451.
7. Seegmiller, R.E., Swenson, M.W., Skinner, G.C., et al.: Reporting of congenital malformations on Utah birth certificates. Utah state Department of Health, Salt Lake City, Utah, April, 1981.
8. Hay, S., Barbano, H.: Independent effects of maternal age and birth order on the incidence of selected congenital malformations. Teratology. 6:271-279, 1973.
9. Regemorter, N.V., Dodion, J., Druart, C., et al.: Major congenital malformations in 5448 newborns: Comments on genetic counseling and prenatal diagnosis. Acta. Paediatr Belg. 34:73-81, 1981.

COLLECTION OF MEXICAN-ORIGIN VITAL AND HEALTH STATISTICS IN SOUTH TEXAS COUNTIES ALONG THE UNITED STATES-MEXICO BORDER

Rumaldo Z. Juarez, Department of Sociology, Pan Am University

INTRODUCTION

The purpose of this paper is to identify: (1) progress in the collection of Hispanic Vital and Health Statistics, with a specific focus on the population of Mexican origin; (2) some problems that still persist; and (3) some suggestions for addressing these problems. The paper is based on information obtained from two sources. One is a visual examination of 623 birth certificates and 504 death certificates selected from three registrar offices in two U.S.-Mexico border counties in Texas. The other is information gathered by an Hispanic National Committee on Vital and Health Statistics supported by the Ford Foundation in 1983. The Committee reflected local, state, and national perspectives about the problems of collecting and reporting vital and health statistics of persons of Hispanic origin. The membership of this Committee also reflected a number of dimensions: (1) Geographic: U.S.-Mexico border, West Coast, Midwest; (2) Occupational: vital statistics recorder, mid-wife, funeral director, university Hispanic health services researchers, social scientists in the National Center for Health Statistics; (3) Government: city, county, national; and (4) Hispanic ethnics: Puerto Rican, Mexican, and Central or South American.

In the following paragraphs, I shall address myself to three major areas: (1) Health Statistics, (2) Natality and (3) Mortality inclusive of Infant Mortality.

HEALTH STATISTICS

To date, the most comprehensive data set available on the health status of Mexican-origin population is the widely publicized Hispanic Health and Nutrition Health Examination Survey. Its data collection activities began in 1980 and the first release of public-use tapes for the Southwestern states may be possible in December 1985. As far as data for border counties in the state of Texas, data should be available for two counties, El Paso and Cameron.. Since these data are comprised of information obtained from both a household interview and a physical examination, it will be the first data set in the history of health surveys that should be able to reflect a comprehensive health status assessment of these populations along the U.S.-Mexico border.

These two counties, however, are not exactly representative of border counties; thus their findings should be interpreted with great caution if the concern is generalizability to other border counties. El Paso and Cameron counties are relatively urbanized populations with industrial bases quite different than most of the other border counties. As examples, both Hidalgo and Starr counties in the Rio Grande Valley are more rural in nature with agriculture being a primary industry. As such, Hidalgo and Starr counties have one of the largest migrant agricultural labor populations in the country, a population that is well known for its poor health status. In addition, the rates of poverty in these two counties are also among the highest in the country and their levels of unemployment, especially during some parts of the year, are especially high. In the case of Starr county, its typical unemployment rate ranges from 30 to 50 percent during the year. Currently, it stands at about 39 percent. Compounding these factors is the lack of health care services in the region. When they are available, their services may not be accessible due to income limitations and no health insurance.

Unfortunately, there is no other comprehensive health status assessment data set for border counties like Cameron in the Lower Rio Grande Valley. While the U.S.-Mexico Border Health Association, through the efforts of the Pan American Health Organization, makes an attempt to collect data on diseases of epidemiological interest, it does not have the resources to collect information that would be needed for health status assessments. On occasion, as in the case of the Lower Rio Grande Valley Development Council's Area Agency on Aging 1983 Elderly Needs Assessment, the health of a specific population (in this case, the elderly age 60 and over) may be studied (Juarez, Lopez, and Garcia, 1984; Juarez and Lopez, 1984 and Juarez and Lopez, 1985). These types of studies, however, are severely limited in their scope and generalizability.

Unlike some states that have morbidity data collection systems, the state of Texas does not have a population-based morbidity data collection system. This greatly restricts the various counties and the state in their ability to: develop short and long-term health plans, have effective resource allocation plans, and make appropriate financial decisions (Texas Statewide Health Coordinating Council, 1985:8). In the absence of these data, the local and state agencies rely heavily on data obtained from local programs and health clinics. While it is better than no data at all, it does have its limitations in that these data do not accurately represent the counties' populations, is subject to variability of reporting methods, and suffers from a lack of continuity in its collection as well as its reporting. In this regard, it is at about the same level of data quality provided by the various county health departments.

Currently, the county health departments in the border counties suffer from a severe lack of funds, a health reporting system that does not receive the cooperation of the local physicians, and county health departments that are more concerned with political considerations and internal personnel conflicts than with maintaining the public's health. A county health department with no "teeth" in their health data collection activities is completely powerless to have a decent vital and health statistics data collection system.

Recommendations: (1) Similar health surveys to the HHANES need to be conducted in all of the border counties; (2) the state needs to develop a population-based morbidity system; (3) the state needs to emphasize on the counties the importance of health data collection and provide the necessary resources to support the data collection activities.

NATALITY

Prior to 1978, natality data on this county's Hispanic population was next to a disaster. Since 1978, however, a new era of Hispanic natality is in progress. Largely due to the works of Stephanie J. Ventura and Robert L. Heuser of the Natality branch in the National Center for Health Statistics, much more is known about Hispanic births than before 1978 (Ventura and Heuser, 1981; Ventura, 1982, 1983, 1984, and 1985). For the first time Hispanics have been able to learn more about their natality patterns, the interethnic differences, the effects of educational levels on birth rates and the need for prenatal care programs in Hispanic communities. Of particular significance of these reports is that they clearly reveal the Hispanic populations' major role in the growth of the U.S. population and the value of having finally included the Hispanic identifiers in State birth certificates. Needless to say that these data are also a tribute to the cooperative efforts of states who are now reporting Hispanic births.

Equally as impressive in the collection of natality data of Hispanics has been the gradual success in improving its quality. In 1978 only 17 states were using the Hispanic identifier covering about 40-50% of the Hispanic births but hindered by a substantial 12.1 percent who lacked an Hispanic origin response for mother's origin. By 1982, 23 states were reporting Hispanic birth statistics, covering about 95 percent of Hispanic births and only 3.8 percent were not reporting Hispanic origin for the mother, Table 1:

Table 1. Hispanic Natality Reporting Progress

Yr.	No. States	% Nat'l Est. of Hisp. Births Represented	% Hisp. Origin not reported	
			Mother	Father
78	17	40-50	12.1	20.2
79	19	60	9.6	18.1
80	22	90	7.0	16.4
81	22	90	6.4	15.9
82	23	95	3.8	13.9

The reporting of natality statistics about Hispanics in Texas has been long overdue, slow in coming, turbulent to say the least, but nevertheless marked with significant progress in the last five years. Eventhough the state has been keeping vital records since 1903, it wasn't until 1980 that Texas adopted the use of ethnic identifiers in both birth and death certificates. Prior to 1980, these certificates contained only a "Color or Race" item typically completed as either "White" or "Black." Consequently, all of the natality and mortality studies conducted on Hispanics in Texas born before 1980 were based on Spanish surname. To name a few of these: Bradshaw and Fonner, 1977; Ellis, 1959, 1962; and Roberts, 1973. The addition of the ethnic identifier provides us with a strong ray of hope that better days are ahead in regards to the quality of Hispanic natality data in Texas. There are, however, some problems that still remain to be overcome.

An examination of a sample of 25 birth certificates (all Spanish surname) from a pool of 843 for the year 1982 of one of the city registrar's office had 15 certificates that did not respond to the Spanish origin item. This amounts to a 60 percent nonresponse rate. In this same office, however, a review of 298 birth certificates from 1985 indicated a complete reversal. That is, an estimated less than 5 percent were not answering both the Spanish origin and the ethnic origin item correctly. This drastic improvement in identification is attributed largely to the individual efforts of the local clerk who began to monitor the certificates closer as they were coming from the local hospital. This closer monitoring was stimulated by the involvement of the clerk in a project that was addressing the needs of more accurate classification of ethnic origin information. This outcome demonstrated two very important points. One is the importance of maintaining communications between the local hospital that completes the certificate and the registrar. The other is the importance of training. In fact, it is relatively easy to tell when a new records clerk has been hired at the local hospital because the quality of the certificates begins to deteriorate, at least temporarily.

A similar review of 75 birth certificates from 1982 and 75 from 1985 for each of two border counties did not reveal any problems of nonresponse or misclassification in the ethnic origin items. This was a bit puzzling since all three registrar's offices (one city and two county) are in the same counties. In other words, why would there be error rates higher in the city records but not in the county? The main reason is the trained personnel. The births for the city were being delivered in the local city hospital. On the other hand, the births being recorded in the county were being delivered at other larger hospitals in the area whose clerks were apparently better trained.

Correct completion of the ethnic origin item, however, should not be the only item of concern in the birth certificates. As was found to be the case in one of the county offices birth records for 1982, the item pertaining to the month prenatal care began was incorrect in a number of the certificates. Instead of writing in "first, second, etc.," the actual month, and in some cases the date, in which prenatal care began was written in, e.g., "February or October 12, 1982." In a few instances, the word "yes" was entered. Overall, though, the number of certificates with this type of response was not overwhelming. Of concern here, nevertheless, is the variability in accuracy that can be introduced when changing health records personnel, e.g., when the usual records clerk is on vacation or replaced.

Having included the Hispanic identifier in the birth and death certificates in Texas only resolves part of the problem of Hispanic data. Still needed is the publication of these data in the state's annual reports. Collection of the data is of limited value if the results are not made known to the public in the usual state reporting mechanisms. This is not to say that the data are not available. A printout can be easily obtained from the state department of health upon request. The point is that states which have Hispanic data should not wait for special data requests before releasing Hispanic results. Rather, it should be available as a routine reporting mechanism in a similar form to that of the vital statistics reports from NCHS but with an emphasis on county and regional comparisons.

The problems of local registrars in border counties, and perhaps in other counties as well, go well beyond the identification of Spanish/ethnic origin and the item on prenatal care. Some of the other problems that local registrars have to contend with are:

1. Formal training of registrars:
   Little or no formal training is afforded to some of the registrars. While the Texas Department of Health provides annual conferences in Austin, not all registrars are provided the opportunity to attend. Consequently, the guidelines and directives issued by the State are followed to this minimum level. Also, since most registrars normally have more duties in their place of work than just recording vital events, little time is given them to keep up with recent State directives, or for that matter, having the time to monitor closely the quality of records being brought into the office. A related problem is not having the time to follow up on incorrect records as soon as errors are noted, e.g., calling the errors to the attention of the hospital clerks.
2. Registrar's status in the respective organizations.
   The County Registrar's Office is considered vital to the function of the respective counties. However, municipalities seem to have other priorities, thus depriving this office of the needed status and funding to accomplish its objectives. It is not uncommon to find record books in need of considerable filing updating and organization.
3. Procedural reporting problems.
   Eventhough the procedures for reporting vital records information from the local areas to the state appear to be well defined, those pertaining to local and county governments are not. Some counties require that all reporting entities (e.g., city, municipality, hospital, etc.) be sent to the county clerk's office -- others do not. Also, at the local level, there seems to be some confusion of where the events should be recorded, especially when geographic boundaries are not well defined. Many times the determining factor of whether a birth is recorded in the county or the city office is largely a function of where the birth took place, i.e., at home, or at a particular hospital. It is quite possible that under these kinds of circumstances some of the births could "fall through the cracks" and go unreported.
4. Reporting problems unique to the U.S.-Mexico Border counties.
   Cities directly adjacent to Mexico have continuous difficulties in registering births by midwives. Since this is a relatively inexpensive form of delivery in these areas when compared to the inaccessibility of preferred health care services from local clinics and physicians, a significant proportion of the deliveries to low-income families are of this type. Unfortunately, reporting by mid-wives ranges from excellent to poor, depending on the midwife's experience, qualifications, and training. Some midwives leave the registering up to the parents and do not follow-up to see that the births are recorded. In some instances, the local registrars have a close working relationship with some of the midwives and this helps to insure appropriate registration. In others, it is left to "la voluntad de Dios."
5. Fraudulent records/reporting activities.
   Fraud encountered in registering:
   a. registering by midwives.
   b. registering by registrars/deputies.

   Fraud encountered in reproduction of records:
   a. copies made in registrar's office. It is not uncommon to find as many as six individual requests at different times for a birth certificate. Needless to say that local registrars question the person's ability to hang on to their birth certificates since they appear to "lose" them quite easily. But they nevertheless have to honor the person's request.
   b. copies duplicated and printed from an original record by the public.

c. copies created for an individual by the public.
d. using birth certificates of children who have died either by duplicating these records or obtaining legal copies.

Recommendations: (1) Eventhough there are currently 23 states and the District of Columbia reporting births by Spanish origin, there still remain some very important states to be added to the list. These states with significant Hispanic populations are Connecticut, Michigan, Wisconsin, and Washington. (2) The quality of vital records is largely determined by the resources invested in the training of recorders at the local level, thus a greater emphasis should be placed on the continuous training of records clerks. (3) The vital and health statistics problems along the U.S.-Mexico border counties are believed to be so interrelated that a standing U.S.-Mexico Commission on Vital and Health Statistics should be established to address these problems and develop joint solutions. Among the problems that this Commission needs to address are adult and infant mortality rates, records exchange, and record fraudulency. (4) The state of Texas needs to routinely publish annual natality reports by race/ethnicity for each of the counties. (5) The state of Texas also needs to conduct an evaluation of the quality of vital records, particularly those in the border counties.

MORTALITY

While we can point to the great progress that has been made in natality statistics, the same cannot be said about mortality. We have yet to see any national reports on mortality statistics among Hispanics at either the national level or in the state of Texas. Apparently, this is not a priority area of concern at the federal level, in spite of there already being mortality data from about 22 states over the last five years. As for the state of Texas, it has mortality data on Hispanics since 1980 and is able to produce mortality statistics on Hispanics, on special requests. As in the case of birth statistics, it does not routinely publish the mortality statistics by ethnicity in their annual reports.

In all fairness to those who work with the mortality statistics, part of these differentials in progress between natality and mortality statistics may be a function of the different kinds of problems encountered in death records as opposed to the birth records. Foremost among these problems is the accuracy in reporting of Spanish origin identity. In contrast to the high rate of correct response to the ethnic identifier on the birth certificate, the response to this item on the death certificate, at least in the border counties, is mortal. An examination of the 504 death certificates from three registrars offices reveals an error rate as high as 50% in two of the three locations and in the third one it was approximately 25-30%. A typical response pattern is to answer "yes" to the question on Spanish origin and then follow it up with one of the following responses to the followup question on ethnicity: Hispanic, Caucasian, Spanish Origin, Mexican American and even a "yes" or a "no". Why the drastic difference in the quality of ethnic reporting is not clearly known other than the realization that it is a substantial problem. Records from one county did reflect a certain pattern of identification, depending on the location of death. Deaths that occurred at one of the larger hospitals were usually recorded correctly, i.e., indicated the appropriate Spanish origin identification followed by the appropriate specific ethnic origin. Deaths that occurred at another smaller hospital, however, were usually recorded as "Hispanic" only with no specification of whether the person was of Mexican, Puerto Rican, or Cuban origin. A similar pattern was observed for these deaths that occurred at home and whose death certificates were completed by the Justice of the Peace. In the final analysis, what this points to is a great need for evaluating the source and magnitude of the problem and the need for drastic training measures.

Another problem area, as can be expected, is in recording the cause of death. It is not unusual, especially among the older population (age 60 or over), to simply have recorded "natural causes" or "old age" as the cause of death. This response pattern is particularly apparent in death certificates completed by the Justice of the Peace. Further complicating this problem is the lack of autopsy information, which is more the exception than the rule, i.e., very few autopsies are ever conducted on Hispanics in this region, unless foul play is suspected. This may be due to two major factors. One is cultural and the other one is economics. The typical autopsy can run from $300 to $600 dollars, a fee that few families can afford.

Recommendations: (1) Most of the ones made for Natality above. (2) Training of funeral directors and local Justices of the Peace needs to be intensified.

CONCLUSIONS

Overall, much progress has been made in obtaining Hispanic vital and health statistics but much remains to be done, especially in the area of mortality. If there is one area that needs to be emphasized in the vital records, it is the need to maximize the appropriate use of the ethnic identifiers. This item is critical if Hispanic vital and health statistics are to become a reality. Unlike in past years when Spanish surname was the typical mode of analysis to Hispanic vital events, this approach is rapidly becoming obsolete. As relations between ethnic groups continue to improve, the rates of marriage structural assimilation can expect to also increase. In fact, this pattern has been quite visible beginning with the 1950's up to the 1970's (Murguia, 1982). This pattern is also visible in the proportions of Hispanics

who are not identified as being of Spanish origin, the further they are from the border regions (Hout, 1982). Further complicating this method is the increasing number of surnames that are typically "Anglo" but are held by persons who come from ethnic intermarriages and still regard themselves as being of Spanish origin. In the end, it is going to be self-identification that will play a major role in the "appropriate" ethnic identification. This in turn, is going to be largely dependent on the proper training and education of those persons who complete the vital records.

Paper presented at the 20th National Meeting of the U.S. Public Health Conference on Records and Statistics, August 13-15, 1985, Washington, D.C. Conference sponsored by the U.S. Department of Health and Human Services, Public Health Service, Office of the Assistant Secretary for Health, National Center for Health Statistics.

REFERENCES

Bradshaw, Benjamin S. and Edwin Fonner, Jr.
1977 Mortality of Mexican Americans: 1969-71. Paper presented at the Conference on Demography of Racial and Ethnic Groups. Austin, Texas, August 19.

Ellis, John M.
1959 Mortality Differences for a Spanish-Surname Population Group. Southwestern Social Science Quarterly, 39:314-321.

1962 Spanish-Surname Mortality Differences in San Antonio, Texas. Journal of Health and Human Behavior, 3:125-217.

Hout, Michael
1982 What Current Estimates of Hispanic Mortality Really Tell Us. Department of Sociology, University of Arizona, Tucson. Mimeograph.

Juarez, R.Z., M. M. Lopez, and L. L. Garcia.
1984 Preliminary Findings of A South Texas Elderly Needs Assessment Survey: A Rural-Urban Comparison of Hispanic and Non-Hispanic Elderly Family Support. Paper presented at the annual meetings of the Rural Sociological Society, College Station, Texas, August 24.

Juarez, R.Z. and M. M. Lopez
1984 Health Status of Mexican American and Anglo Elderly in South Texas. Paper presented at the annual meetings of the American Public Health Association, Annaheim, California, November 12.

1985 A South Texas Rural-Urban Comparison of Mexican American and Anglo Elderly Housing Conditions and Needs. Paper presented at the annual meetings of the Southwestern Sociological Association, Houston, Texas, March 22.

Texas Statewide Health Coordinating Council
1985 Texas State Health Plan 1985. Austin, Texas: Texas Department of Health.

Ventura, S. J. and R. L. Heuser.
1981 Births of Hispanic Parentage, 1978. Monthly Vital Statistics Report, Vol. 29, No. 12 Supp. DHHS Pub. No. (PHS) 81-1120. Public Health Service. Hyattsville, Md., Mar. 20.

Ventura, S. J.
1982 Births of Hispanic Parentage, 1979. Monthly Vital Statistics Report, Vol. 31, No. 2 Supp. DHHS Pub. No. (PHS) 8201120. Public Health Service. Hyattsville, Md. May.

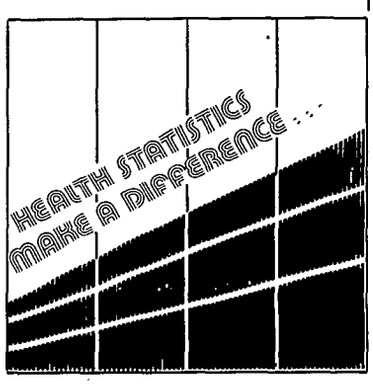
1983 Births of Hispanic Parentage, 1980. Monthly Vital Statistics Report, Vol. 32, No. 6 Supp. DHHS Pub. No. (PHS) 83-1120. Public Health Service. Hyattsville, Md. Sept.

1984 Births of Hispanic Parentage, 1981. Monthly Vital Statistics Report, Vol. 33, No. 8 Supp. DHHS Pub. No. (PHS) 84-1120. Public Health Service. Hyattsville, Md. Dec.

1985 Births of Hispanic Parentage, 1982. Monthly Vital Statistics Report, Vol. 34, No. 4 Supp. DHHS Pub. No. (PHS) 85-1120. Public Health Service. Hyattsville, Md. July 23.

Special Session 2 (AMA)

American Medical Association's Data Resources



THE AMERICAN MEDICAL ASSOCIATION'S PHYSICIAN MASTERFILE

John D. Loft and George A. Ryan
American Medical Association

Introduction

The American Medical Association's Physician Masterfile is the most comprehensive source of physician information available in the United States. The Masterfile contains demographic, educational, and current practice information about all of the nearly 520,000 physicians in the country, non-members as well as members. It is the basis of the AMA's data resources and is used for a wide variety of purposes including verifying physician credentials, distributing scientific information to the medical community, promoting membership services, and monitoring social and economic trends in medical practice.

In order to assure the completeness and accuracy of the Masterfile's data, the AMA devotes a considerable amount of manpower and resources to the compilation and maintenance of the File. Each month, an average of 12,700 physicians' records are updated with new information of one sort or another. Because more than one data element is updated for each of these physicians, over the year, the updating process involves nearly 800,000 transactions with the File.

Historical Development

As a census of the physician population in the United States, the Masterfile dates from 1905, when the AMA began a card index of licensed physicians in preparation for the publication of the first American Medical Directory[1]. Prior to this time, the AMA had kept Membership rosters since it was formed in 1847. Other medical directories were in circulation, but were incomplete in that they contained the names of only those who paid a fee to be listed. Moreover, information contained in these early directories was not verified and was quite often fraudulent.

In 1905, the House of Delegates of the AMA, recognizing the lack of a comprehensive and accurate list of physicians in the U.S., decided to establish a biographical record of physicians and publish its own medical directory. The first edition of the American Medical Directory was published in 1906. It listed the full name of each physician, year of birth, medical college, year of graduation, year of licensure, and office address for a total of 128,173 physicians. Subsequent editions were published every two or three years until World War II disrupted regular publication. More recently, the Directory has been published every four years following a quadrennial census of physicians conducted by the AMA.

From its initiation, an important feature of the AMA's Directory was that information regarding medical college and year of graduation, date of licensure, and membership in medical societies would be verified with official sources. Prior to the AMA's Directory, the only records of graduation were alumni lists maintained by each medical school. The value of an alternative compilation of such information became apparent as records were inadvertently destroyed and as medical schools closed and records lost. When the U.S. mobilized for the First World War, the AMA provided a valuable service to the armed forces by verifying the records of thousands of physicians who entered the military service. The AMA continues to provide the Physician Profile Service for use by State Licensing Boards, Hospitals, Medical Schools, State, County, and Specialty Societies, and other health-related organizations including government agencies.

Initially, the Masterfile was a card index system used principally to produce the American Medical Directory and for membership and mailing purposes. Record-keeping procedures were not designed for statistical aggregation or data analysis and all entries were narrative. In 1948, electronic accounting machines were installed for the use by the AMA Bureau of Medical Economic Research for the analysis of sample survey data. In 1958, the entire card index system was converted to machine-readable form as the AMA initiated the use of computers to maintain physician data.[2] By this time, the File contained information on year of birth, sex, medical school and year of graduation, year of licensure, National Board certification, American Specialty Boards, primary and secondary specialties, location and type of practice, sources of income, principal employer, membership in specialty societies, and professorial appointments. With the exception of current practice information, which was obtained directly from the physician through a mail questionnaire, all information on the File continued to be obtained or verified through institutional sources.

As the needs of Masterfile users have changed, the amount and type of information maintained on the File have also changed. An important development was implemented in 1968. In collaboration with the National Center for Health Statistics, the Division of Demographic Studies, and the Bureau of the Census, the AMA Department of Survey Research developed a new classification of physicians' professional activities. The new classification differed from the previous method in two important features.[3]

First, the earlier classification was based on the physician's financial arrangement (private practice or not) rather than activity. As noted in the 1964 critique prepared by the U.S. National Committee on Vital and Health Statistics, the former classification did not allow the identification of all physicians directly involved in patient care a critical gap in medical manpower

studies.[4]

Second, prior to 1968, physicians classified themselves into response categories describing professional activities, principal employer, primary and secondary specialty. The National Committee commented that physician manpower could not be properly allocated among different activities until some objective measure of workload--such as hours worked or number of patients seen--was made available.

The 1968 reclassification addressed both of these points by developing a Type of Practice-Professional Employment categorization system based on the numbers of hours per week spent in various arrangements. The concept of classifiying physicians according to number of hours worked was also extended to primary and secondary specialties. A new questionnaire, the "Record of Physicians' Professional Activities" (PPA), was developed based on the new classification and implemented in 1968.

The PPA has continued to evolve. Most recently, in preparation for the 1985 Census of Physicians' Professional Activities, the AMA's Division of Survey and Data Resources undertook a thorough evaluation of the PPA. The form used in earlier censuses was reviewed internally and by several external committees including an advisory committee of Federal government researchers and statisticians. This review resulted in a number of modest, though important revisions. A question was added to ask about office address when the preferred professional mailing address is not a practice address. An item to obtain office telephone number was added to the form. A category for hours in fellowship programs was added to the section on professional activities. A new definition was developed for medical research activities. The classification of specialty was modified by adding some specialty codes and deleting others. Definitions for office practice categories were refined and expanded in order to assure comparability with other data collection systems operated by the AMA. Items to record the name of the hospital where the physician admits most patients and the number of hours worked in that hospital have been added to the form. The new form asks for the physician's race and ethnicity in order to enhance studies of physician manpower by inhouse researchers and external data users.

Current Data Collection Procedures

Conceptually, each Masterfile record consists of an historical data section and current data section. The historical section contains demographic, educational, and permanent professional information. Features of each physician's current practice arrangements are maintained in the current portion of the File. Current data are subject to constant change as physicians move from location to location, alter their professional activities, or change employment arrangements. This section of the File is updated continuously through an intensive monitoring process.

Medical schools provide the information used to initiate a record in the Masterfile: the name and address of each student, sex, birthdate and birthplace, name of the medical school, and expected date of graduation. These data are stored in a separate "Student" File. A unique record identifier, the Medical Education Number, is assigned to the student record when the individual enters and can remain unchanged for the course of a physician's career. Students are tracked as long as it takes them to complete their undergraduate medical education.

As students graduate and enter residency or fellowship programs, their records are shifted from the Student File to the Physician Masterfile. Additional data describing graduate medical training programs--type of program, date entered and date completed--are added to the historical portion of each record as the physician completes his or her graduate medical education. These data are obtained through a yearly Census of Graduate Medical Training Programs. Data from the National Residency Matching Program also play an important role in tracking first year residents. Records for graduates of foreign medical schools are established as they enter residency programs. Background information about foreign medical graduates is supplied to the AMA by the Educational Commission for Foreign Medical Graduates (ECFMG).

Completion of the National Board of Medical Examiners examination, the date and state of each licensure (including disciplinary action indications received from the Federation of State Medical Boards), dates of certification by specialty boards, indicators of government service, and professional affiliations with state and county medical societies are added to the record during the physicians professional career. Information about deceased physicians is permanently maintained on the File for security and verification purposes.

The current section of the record contains the physician's preferred professional mailing address, type of practice, professional employment, and primary, secondary, and tertiary specialty. These data are obtained directly from the physicians through a mail questionnaire, the Record of Physicians' Professional Activities (PPA). A change in a physician's current status on any of these variables may be signalled by a number of sources that are monitored constantly: AMA mailings and publications, commercial mailings sent by the ten addressing companies licensed by the AMA, physician correspondence, correspondence from hospitals, government agencies, medical schools, medical societies, specialty boards, and licensing agencies. Any indication of a change in address or activity triggers the mailing of a PPA questionnaire. In addition, every four years, the PPA questionnaire is mailed to the entire physician population in the Census of Physicians' Professional Activities. The 1985 Census is currently in the field with an expected completion date of mid-1986.

The Quality of the Physician Masterfile Data

In 1977, Goodman and Eisenberg[5] reviewed

several papers that evaluated the quality of the AMA's Physician Masterfile. More recently, the AMA's Department of Data Planning and Evaluation has conducted several reliability studies[6]. Studies evaluating the Masterfile have used two different approaches to the issue of reliability. In the first approach, a sample is selected from the Masterfile and comparable data are obtained from the sample and checked with the Masterfile record. In the second approach, data from a sample of Masterfile records are compared with data from other lists.

Results of these reliability studies show some variation in the reliability of the items and general improvement in the quality of Masterfile data over the years. Studies generally focus on the reliability of three key items: address, specialty, and primary activity. The reliability of the physician's location is of obvious importance in attempts to contact physicians. Specialty and activity are often used as selection criteria in drawing samples and their reliability is critical in minimizing sample bias due to inaccuracies in the sample frame.

Theodore and Sutter[7], using a sample of 2,833 physicians selected from the 1966 Masterfile, showed that specialty in the Masterfile record matched the respondent's answers in 88.1 percent of the cases. This measure of specialty reliability ranged from 99.2 percent among pediatricians to 88.0 percent among general practitioners. The sample was drawn from office-based records in the Masterfile and 6.4 percent of the cases were found to be in other professional activities, providing a partial measure of the reliability of activity.

Cherkin and Lawrence[8] compared 1974 Masterfile data on 6,001 physicians in Washington state with similar data obtained from the state's Division of Professional Licensing (DPL) and from a sample survey of 300 physicians. According to the DPL File, 5,467 physicians were licensed and practicing in Washington at the same time the AMA File was selected; 14.0 percent (836 cases) of the 6,001 physicians in the AMA Masterfile were not on the state's list; 5.5 percent (302 cases) of the 5,467 physicians on the DPL File were not on the Masterfile; and 5,165 cases were listed on both files. Using the Masterfile data from other states, the authors were able to resolve all but six of the discrepancies by locating physicians in states other than Washington.

The survey data in the Cherkin and Lawrence study were used to examine the reliability of Masterfile data on birthyear and birthplace, place of medical education, professional activity, and specialty. For each of these variables, agreement between the Masterfile data and survey data was over 90 percent.

Since 1980, the AMA's Department of Data Planning and Evaluation has conducted annual validation studies of the information contained in the Physician Masterfile. These studies examine the reliability of twenty variables on the File. In the most recent study[6], a representative one-percent sample of the physician population (n = 5,188) was mailed a questionnaire and asked to confirm or correct preprinted data from the physicians' Masterfile records. Three waves of mailing were used to achieve a response rate of 72 percent.

The study evaluates the accuracy of the address information in the Masterfile and the accuracy of individual data elements. The effectiveness of the mailing address --addressability--is indicated by the rate of "undeliverable" questionnaires that the AMA received after the first wave of the survey. In 1984, the overall rate of addressability was calculated to be 99 percent.

The comparison of individual data elements generally showed a high rate of agreement between the Masterfile data and the survey data, with some variation in the reliability by element and by certain physician characteristics. The average rate of agreement across all of the 20 variables was 94.0%. Agreement on background and education information was quite high: 99.1 percent agreement on medical school (differences between the survey data and Masterfile are most often due to variant names of medical schools particularly foreign medical schools); 98.0 percent on year of graduation (differences appear to be due to memory effects in the survey data); 97.0 percent on birthplace; and 97.4 percent on physician's name (differences were due to name changes through marriage and variant spelling of the same name). Agreement on current professional activities was slightly lower: 94.1 percent agreement on primary specialty and 88.5 percent on type of practice.

Questions about specialty differ in the PPA and in the Validation Surveys, which accounts for some of the discrepancy. The PPA asks for number of hours work in a specialty while the Validation Survey asks the respondent to verify primary, secondary and tertiary specialty. Respondents to the Validation Survey often reverse primary and secondary specialty.

The lower rate of agreement on type of practice is not surprising as practice characteristics can change often during a physician's career. We know from the volume of address changes that annually about 20 percent of the physician population move, often changing employment characteristics as they change location. Differences between the Masterfile data and the Validation Survey data in these variables usually represents the timespan between locational changes and resulting data collection and updates to the Masterfile.

Not surprisingly, rates of agreement are correlated with age of physician. The average rate of agreement is slightly lower for physicians in training and early practice (92.1 percent and 92.7 percent, respectively) and higher for those in established practices and those retired (95.6 percent for both groups). The average rate of agreement for AMA members was somewhat higher than among non-members (by a difference of 6.7 percent points). Physicians in patient care activities had higher rates of reliability than physicians in other professional activities (d = 5.2 percentage points).

Validation studies completed to date have

demonstrated that the Masterfile is a comprehensive and accurate source of physician data. The reliability of historical data is excellent--very close to 100 percent. Reliability of current practice data can never be 100 percent because of the inherent problems of capturing information from a constantly changing universe. Physicians will continue to move and change professional activities and there will always be some lag between the change and recording the new status on the File. Through extensive data collection efforts, the AMA is able to achieve the high levels of reliability indicated in the Validation Studies.

Uses of the AMA Physician Masterfile

The Masterfile's original purpose--a data file for use in the preparation of the American Medical Directory--is one which continues today. The American Medical Directory is published every four years. It contains an alphabetical listing of all physicians in the country and state-by-state lists. Each entry records a physician's name, address, primary and secondary specialty, medical school of graduation, type of practice, year licensed in current state of professional address, and certification by specialty boards.

The AMA's credentialing services, based on the Masterfile data, are widely recognized and used in the medical community. Hospitals, medical schools, and specialty societies all use these services as physicians apply for admitting privileges, faculty positions, or society memberships. In 1984, the Association provided 136,000 physician profiles to help validate physicians' qualifications to practice. Beginning this year, the Association initiated a similar service for the Veterans Administration which has asked the AMA to verify the credentials of as many as 94,000 physicians. The VA has supplied the AMA with magnetic tapes containing information collected by that agency; the AMA uses its computers to match records from the VA tapes and Masterfile and verify that the physician meet the VA's standards for employment. Recently, the AMA has responded to similar requests from the U.S. Army, Navy and Air Force.

The AMA's most prominent mission is the representation of the medical profession. In doing so, the Association relies heavily on Masterfile data to track historical trends and monitor their impact on its constituency. Examples of such trends are the growth and development of group practice, the changes in the proportion of physicians who are graduates of foreign medical schools, and the career paths of young physicians and women physicians. As these sectors in the physician population grow, their particular needs and concerns must be considered in the development of national health policy.

In collecting and disseminating socioeconomic data, the Masterfile is useful in two ways. First, the File in its entirety is used to produce yearly monographs describing "Physician Characteristics and Distribution in the U.S."[9] This publication contains historical and current data on age, sex, specialty, national board certification, and country of medical graduation. Tables are available for both federal and non-federal physicians and for regional, county, and metropolitan area breakdowns. Licensure statistics are available in a separate volume, also published annually.[10] These population based statistics are very useful to physicians and other health care providers as they seek locations for potential practices.

Second, the Masterfile is also used as a sampling frame for a number of in-depth sample surveys that would be inappropriate to administer to the entire population of physicians. The sample survey methodology is an effective, efficient means of collecting valid and reliable data on topics that are too broad or too sensitive to address in the on-going Masterfile data collection systems.

Chief among the sample surveys supported by the Masterfile is the Socioeconomic Monitoring System (SMS). This is a telephone survey fielded four times a year with separate samples selected from Masterfile records identified as belonging to patient care physicians. The basic topics of the survey are income, practice costs, and practice patterns. SMS data are used to develop quarterly statistical profiles of physicians' practices in terms of these variables. In addition, interviewing time is available in each round of the survey to address topics of special interest to the Association.

In order to remain a strong and effective organization, the AMA relies on its membership for support. Because the Masterfile contains information on non-members as well as members, it is a critical resource in membership development. Masterfile data are used to prepare profiles identifying types of non-members as a basis for establishing programs that will encourage participation in the Association. Lists of non-members areprepared for special marketing programs designed to increase membership. Finally, environmental analyses, possible through the Masterfile and other data resources, help the AMA to be more responsive to its membership in particular and in general to the entire physician population.

Our interests in trends in American medicine are shared by many other parties that require accurate and current information about physicians and their professional activities. Masterfile data have been used extensively by various agencies within the federal government. Special tapes with identifying information deleted from each record have been supplied to the Department of Health and Human Servies for use in sample surveys and in physician manpower studies based on population statistics. Similar anonymous tapes are also provided to academic researchers.

Masterfile data are also used in the development of the Area Resource File which combines data on health providers with demographics descriptors of each county in the United States. The ARF supports numerous policy studies at the national, state, and local level.

As with the AMA's Socioeconomic Monitoring System, federal agencies and academic researchers employ the Masterfile as a sampling frame for a number of national sample surveys which provide data that are critical to the development of national health policy.

In this paper we have described the history and development of the AMA's Physician Masterfile, demonstrated its current structure and quality, and illustrated its broad range of potential uses. In its leadership role in the medical profession, the AMA requires understanding of the complex trends in contemporary American medicine. As a historical File, the Masterfile is an important tool for monitoring these trends. It is a unique sampling frame for more detailed studies of particular issues. External users within all levels of government rely on Masterfile data to inform health policy decisions. The Masterfile is a vital resource for the Association in fulfilling its responsibilities to the profession and to the public.

REFERENCES

[1]F.V. Cargill: "The American Medical Directory" pp. 1170-1179 in M. Fishbein's A History of the American Medical Association 1847 to 1947. W.B. Saunders, Philadelphia, 1947.

[2]Proceedings of the House of Delegates, American Medical Association, Clinical Session. Minneapolis, Minnesota, December 2-4, 1958. pp. 50-51.

[3]C.N. Theodore, et al: Reclassification of Physicians, 1968. Center for Health Sciences Research and Development. American Medical Association, Chicago, 1971.

[4]U.S. National Committee on Vital and Health Statistics: United States Statistics on Medical Economics. U.S Department of Health, Education and Welfare, PHS Publication No. 1125. Washington, D.C. 1967.

[5]L.J. Goodman B.S. Eisenberg: "The Quality of Physician Data", Public Data Use, Volume 5, Number 3, May, 1977 pp. 37-44.

[6]A.V.V. Daigle: "Physician Record Validation Survey 1984, Summary of Findings." Department of Data Planning and Evaluation, Division of Survey and Data Resources, American Medical Association, December, 1984.

[7]C.N. Theodore and G.E. Sutter: "A Report on the First Periodic Survey of Physicians", "Journal of the American Medical Association" Volume 202, Number 6, November, 1967 pp. 516-524.

[8]D. Cherkin and D. Lawrence: "An Evaluation of the American Medical Association's Physician Masterfile as a Data Source-One State's Experience," Medical Care, Volume 15, Number 9, September, 1977, pp. 767-779.

[9]M.A. Eiler: Physician Characteristics and Distribution in the U.S., 1983 edition. Department of Data Release Services. Division of Survey and Data Resources, American Medical Association, 1984.

[10]A.V.V. Daigle: U.S. Medical Licensure Statistics and Licensure Requirments: 1983-1984. Department of Data Planning and Evaluation, Division of Survey and Data Resources, American Medical Association, 1985.

MEDICAL GROUP PRACTICES: THE AMA DATA BASE

Penny L. Havlicek and Mary Lou S. White
American Medical Association

Since 1965 the American Medical Association (AMA) has periodically conducted census surveys of medical groups in order to monitor changes in the characteristics of such practices. The AMA defines group practice as "three or more physicians formally organized to provide medical care, consultation, diagnosis and/or treatment through the joint use of equipment and personnel and with income distributed in accordance with methods predetermined by the group." This definition was derived by the Medical Group Management Association (MGMA), the American Group Practice Association (AGPA) and the AMA and was adopted by the AMA House of Delegates in 1964.

Trends in Medical Groups

The AMA conducted census surveys of U.S. medical groups in 1965, 1969, 1975, 1980, and 1984. These census surveys determined that the number of groups and the number of physician positions[1] have grown substantially since 1965 (Table 1). In 1965, 4,289 groups were identified in the U.S. By 1984, less than 20 years later, there was almost four times that number of groups. The last four years alone have shown a growth rate of about 44 percent or an annual increase of 9.5 percent. The number of physician positions grew by almost 400 percent between 1965 and 1980 and nearly doubled over the last four years. Part, but not all, of this growth may be due to improved data collection methods.

Table 1: Growth in the Number of Groups and Physician Positions

Year	Number of Groups	Physician Positions
1965	4,289	28,381
1969	6,371	40,093
1975	8,483	66,842
1980	10,762	88,290
1984	15,485	140,213

These census surveys also determined that the practice of group medicine is being increasingly conducted in single specialty groups. The percentage of groups that are single specialty has increased steadily from 49.7 percent in 1969 to 70.0 percent in 1984 (Table 2). While the proportion of family or general practice groups has remained fairly stable over the years the proportion of multispecialty groups appears to be decreasing. Although the number of multispecialty groups may in fact be decreasing, another explanation of this apparent decline is that multispecialty groups may be merging in order to attract the business of prepaid health plans or to increase market share.[2]

Table 2: Specialty Composition of Groups, 1965-1984

	Specialty Composition			
Year	Single Specialty %	Multi-Specialty %	Family/ General Practice %	Total N
1965	50.3	34.4	15.2	4,289
1969	49.7	38.0	12.3	6,371
1975	54.2	35.1	10.7	8,483
1980	57.2	33.0	9.8	10,762
1984	70.0	18.3	11.7	15,186[a]

[a]Excludes 299 groups whose specialty compostion was unknown.

Medical groups have also increased in size. Although most groups had three or four members in both 1980 and 1984, the mean number per group has increased from eight to nine. The largest increase occurred in groups with 50 or more members. The number of such groups doubled from 146 in 1980 to 306 today.

The growth in larger groups illustrates the increasing complexity of medical practice. The number of groups has grown and larger groups are growing faster than smaller groups. In addition, our data show that professional corporations have emerged as the most popular legal form of organization. Professional corporations represented 16 percent of all groups in 1969, but 73 percent of all groups in 1984.

Past Data Collection Efforts and the Creation of the Group Practice Data Base

Each of these census surveys was an independent survey effort requiring the reidentification of the population frame. However, in light of the increasing numbers of group practices that these surveys revealed and their growing significance to the practice of medicine, the AMA created a data base of group practices to be maintained on a continuous basis beginning with the 1984 Census of Medical Groups. In order to best describe the AMA group practice data base, the methodology of the 1984 census is outlined below.

Methodology of 1984 Census

The 1984 Census of Medical Groups was initiated in 1983 and completed in the Spring of 1984. As in previous census surveys, the population frame was assembled using internal and external sources. The first step in

assembling the population frame was the identification of medical groups surveyed in the 1980 census. This list was updated through the AMA's Census of Physicians' Professional Activities (PPA). The PPA is a quadrennial census of the U.S. physician population with targeted mailings to physicians who indicate a change in their practice as well as other selected subpopulations in non-census years. This survey collects information regarding type of practice, employment and specialty. If a medical group was mentioned on the PPA census that was not on the list of groups surveyed in 1980, it was added to that list. Finally, lists of groups were obtained from the MGMA, the AGPA and other relevant health care organizations and incorporated into the AMA list. The resulting list of groups formed the initial data base.

The data were collected over four mailings. The first mailing was conducted in September of 1983, and the last mailing was conducted in March of 1984. A longer, more detailed questionnaire was used in the first two mailings, and an abbreviated questionnaire was used in the last two mailings to facilitate and encourage response. A response rate of 86.9 percent was obtained representing a total of 15,485 groups.

Data Elements

The group practice data base includes information collected on the 1984 Census of Medical Groups. This information includes:

- o whether the group met the AMA's definition of group
- o the size and specialty of the group
- o the group's legal form of organization
- o the percentage of total care provided that is prepaid
- o the group's relationship to a hospital (such as operated by a hospital, renting space from a hospital, etc.)
- o whether the group has a medical director
- o whether the group employs a business manager, group administrator, or health plan manager
- o whether the group has the following facilities on-site: pharmacy, clinical laboratory, routine radiology, routine electrocardiology, audiology, vision testing
- o whether the group owns a videocassette recorder for use in patient education or continuing medical education
- o whether the group currently uses and plans to use a computer for business transactions, clinical records, medical (nonpatient) data retrieval
- o whether the group pays for memberships for its physcian members to metropolitan/county medical societies, the state medical association, the AMA or national medical specialty societies and for subscriptions to medical journals
- o whether the group consists of a parent group with one or more branch or satellite clinics, and if so their names and addresses

These data were collected to describe certain dimensions of medical groups. Group size and specialty data provide basic background information. Legal form of organization, the percentage of total care that is prepaid, the group's hospital relationships and the management personnel employed tap the organizational complexity of medical groups. The on-line facilities not only indicate the range of services a group may provide but also the degree of independence a group may have from other providers. Videocassette recorder ownership and computer usage reflect the group's sophistication regarding practice management.

Data Base

Not all of the information collected from groups is represented on-line. The on-line data elements comprise a single screen of information for each group that is called up via specification of the group ID number. The on-line data elements include:

- o group ID, name, and address
- o date the group was entered on to the data base
- o date of most recent address update
- o parent group ID
- o active/inactive flag
- o date group inactivated
- o whether the group meets the AMA definition
- o former or cross-reference group ID
- o specialty composition
- o the percentage of total care that is prepaid
- o group size
- o date of most recent group size update

These data elements can be updated on-line while access to the other information that is collected on groups is through conventional rectangular data sets stored on disk or tape. The date entered, date of address change, date inactivated, and date of size change are generated automatically when the appropriate on-line field is updated.

Another segment of the group practice data base is the group physician file. This file enables the AMA to link physicians to groups as it contains the identification numbers of all physicians who reported on the PPA census to be affiliated with a medical group and the identification number of that group. The group physicians file is updated on a periodic basis in conjunction with the PPA census. An extensive update of this file will be conducted next year toward the completion of the census.

Data Base Maintenance

Currently the group practice data base maintenance involves 1) identifying groups that need to be added to the data base and identifying groups that have dissolved, 2) updating group addresses, and 3) updating the data elements.

New groups are primarily identified through the PPA census which asks physicians if they practice in group arrangements. New groups are also identified in various print media such as

group practice newsletters, journals, etc. Marketing materials obtained from health maintenance organizations (HMOs) are used to identify groups affiliated with HMOs. Dissolved groups are identified through the PPA census and group practice surveys to update data elements.

Group addresses are also updated through the group practice surveys and through the PPA census. In addition, the AGPA notifies the AMA of the address changes of their journal recipients.

Data elements are updated periodically. When new groups are added to the data base or when an address change occurs, groups are automatically surveyed. Another census to all groups is scheduled for late 1986.

Data Base Use/Users

The data has internal and external uses. Internally, the data have been used to develop a membership program for group physicians. Externally, medical product and supply companies have used the data to analyze their markets. The data are made available to external commercial users through licensed addressing companies to which the AMA provides updated group practice files on a quarterly basis.

Future Plans for Group Practice Data and Data Base

The AMA has several plans for the group practice data and data base. These plans include publication of the results of the 1984 Census of Medical Groups, reorganization of the group practice data base and collection of additional information on groups.

Publication

In September the results of the 1984 Census of Medical Groups will be published. Many of the major survey findings relate to the specialty composition of the group. Some of these findings are now described.

Groups were classified as single specialty, multispecialty, and family or general practice. Of the 15,186 groups that could be so classified about two-thirds were single specialty, 18 percent were multispecialty and 12 percent were family or general practice. Multispecialty groups are larger than either single or family or general practice groups. The average size of multispecialty groups was 26.6 physicians, the average size of single specialty groups was 5.8 physicians and the average size of family or general practice groups was 5.7 physicians.

Table 3 highlights other differences among these three types of groups. The percentages of each type of group with various characteristics are reported. The numbers in parentheses represent the total number of cases on which the percentages are calculated.

Table 3: Percentage of Medical Groups with Selected Characteristics by Specialty Composition, 1984[a]

	Type of Group					
	Single Specialty		Multi-Specialty		Family/ General Practice	
Characteristics	%	N	%	N	%	N
Prof. Corp.	78	(7,941)	63	(1,881)	60	(1,205)
Bus. Manager	52	(7,434)	66	(1,468)	62	(1,101)
Group Admin.	16	(5,828)	58	(1,465)	23	(875)
Med. Director	23	(7,671)	44	(1,809)	24	(1,164)
Health Plan Man.	2	(5,403)	15	(1,043)	2	(768)
Hosp. Assoc.	58	(7,618)	43	(1,807)	37	(1,152)
Pharmacy	16	(6,450)	47	(1,653)	29	(1,092)
Clin. Lab.	47	(6,806)	84	(1,792)	92	(1,183)
Routine Radiology	46	(6,977)	78	(1,769)	70	(1,169)
Routine Electrocardiology	32	(6,495)	87	(1,746)	94	(1,170)
Audiology	27	(6,299)	66	(1,659)	68	(1,127)
Vision Testing	30	(6,247)	75	(1,651)	87	(1,115)
VCR	32	(7,817)	36	(1,844)	23	(1,194)
Providing Prepaid Care	22	(6,980)	36	(1,634)	27	(1,074)
Computers for Bus. Trans.	58	(6,986)	76	(1,708)	50	(1,062)
Computers for Med. Nonpatient Data Retrieval	21	(6,092)	33	(1,465)	17	(956)
Computers for Clin. Records	16	(6,113)	16	(1,437)	9	(973)

[a]This table is based on the 11,243 groups that responded to the long form of the questionnaire. The 4,242 groups that responded to the short form were not asked these questions. The reported Ns vary due to item nonresponse.

Although over 50 percent of each type of group are professional corporations, multispecialty groups and family or general practice groups are more likely to assume other legal forms. Although not shown in the data presented, family or general practice groups are more likely than either multispecialty or single specialty groups to be organized as partnerships.

Business managers, group administrators, medical directors or health plan managers are more common in multispecialty groups than single specialty or family or general practice groups. Mulltispecialty groups are also more likely than single specialty or family or general practice groups to provide prepaid care. The larger size and more administrative support staff indicates the greater complexity of multispecialty groups.

Single specialty groups are less likely than multispecialty or family or general practice groups to have any of the selected facilities on-site. Some of these facilities may not be relevant to some single specialty groups. In addition, single specialty groups appear to have greater access than multispecialty or family or general practice groups to these facilities through their hospital relationships. Close to two-thirds of all single specialty groups maintain some kind of relationship to a hospital, but less than half of multispecialty and family or general practice groups do.

Multispecialty groups are more likely than single specialty or family or general practice groups to use computers in the practice. Multispecialty groups are more likely than single specialty or family or general practice groups to use computers for business transactions and medical nonpatient data retrieval, but are equally likely as single specialty groups to use computers for clinical records. Again the larger size of multispecialty groups and the greater volume of business generated appears to require greater technological support in administering the practice.

Data Base Reorganization

In addition to publishing the survey results, the AMA plans to bring on-line all data currently collected on groups. This will facilitate both data entry and data retreival. The AMA also plans to make the separate data bases on groups and physicians in groups interactive so that access and use of these will be more efficient.

The AMA is now in the process of reorganizing the data base so that the horizontal and vertical integration taking place among groups are better reflected. For example, we would like to be able to isolate networks of groups and the physicians affiliated with them so that they may be called up on screen. This reorganization should be completed sometime next year.

Omnibus Survey

Finally, the AMA will be collecting additional data on groups so that we can better describe the variance in their complexity. For example, we would like to know the number of groups and the number of physicians affiliated with HMOs, Preferred Provider Organizations and other emerging practice arrangements. An omnibus survey to a sample of group practices is scheduled for later this year to collect such information.

Discussion

The AMA census surveys have revealed a growing number of medical group practices. This growth in medical groups may be occurring for a number of reasons. The number of medical groups may be growing because the environment of medicine is becoming increasingly competitive, and groups are better able to compete due to their superior ability to generate capital. Some of this capital may be used to market medical services, a growing and expensive trend among health care organizations. The number of medical groups may also be growing because given the increasing cost of medical care, group practices are less costly to each physician to establish than solo practices, and may permit economies in the delivery of medical services. Although group practices have always presented certain potential practice advantages to physicians such as facilitating referrals and better, more flexible hours, some groups may also be able to absorb the costs of malpractice insurance, a renewed concern among physicians today. Finally, increasing numbers of medical groups may be forming in response to the increased numbers of physicians and the resulting competition among them.

Recognizing the growing importance of group practices to the delivery of medical care, the AMA has established a data base of group practices. Maintained on a continuous basis, the data base will be used to investigate the impact of medical groups on the practice of medicine. Such analyses will have implications for health policy issues such as access to medical care.

Footnotes and References:

1. Physician positions reflect the slots within groups. Because physicians may occupy positions in more than one group, the number of physicians in groups may be overstated.

2. Dan Richmond: "Groups hope mergers will attract business of prepaid health plans," Modern Health Care, August 2, 1985, pp. 67-68.

FOREIGN MEDICAL GRADUATES: 1983 PROFILES

Mary Ann Eiler, American Medical Association

INTRODUCTION

Historically, the foreign medical graduate (FMG) has played an important role in US medicine. FMGs have been the focus of both divergent and complex legislation motivated by alternating cycles of perceived physician shortage and oversupply. The liberal international exchange programs and legislation in post World War II and the 1960s gave way to the more restrictive policies of the 1970s and new examination requirements in the 1980s. Changes in the demography and growth rate of the US population, shifts in migration patterns and concerns of "brain drain" in developing countries motivated discussions about the total US physician supply and the increasing numbers of FMGs.

In 1971, FMGs numbered 62,214 to represent approximately 18% of a total MD count of 344,304. By 1976, the FMG complement included 85,626 MDs or 20.9% of all physicians (409,446) in the US. Five years later (1981), FMGs in the US numbered 102,762 out of a total MD stock of 467,679 -- 22%. One-fifth (21.6%) of all 519,546 physicians (112,005 MDs) in the US in 1983 had received their medical education in schools outside the US and Canada.

PURPOSE OF STUDY

Although FMGs as a percent of total MDs has remained fairly stable since the early 1970s, this apparent stability belies important trends and variations within the FMG population. In response to these variations, the AMA is publishing Foreign Medical Graduates: 1983 Profile which provides detailed statistics from the AMA Physician Masterfile on FMGs by type: foreign national FMGs (FNFMGs), US citizen FMGs (USFMGs), and exchange visitor J-Visa FMGs (EVFMGs). US medical graduates (USMGs) and Canadians are also included for comparison. Although the volume discusses FMGs in historical and current contexts, data are presented for each physician type primarily in three major sections: TRENDS, PROFILES, and LOCATION.

TRENDS

The Trends section provides statistics on activity and specialty choices. Age and sex distributions and country of graduation are also presented as are physician population ratios. Some major findings reveal that although FMGs increased numerically by 80% between 1971-1983 their rate of growth declined from 37.6% (1971-1976) to 20% (1976-1981). Figure 1 illustrates percent changes for MDs by type. The cumulative growth rate of US medical graduates in 12 years was almost 45% - matching that of USFMGs (45.3%). The growth rate of FNFMGs between 1971-1983 was dramatic (99%), exceeding that of the other groups, with the highest percent occurring between 1971-1976.

Trend analysis also reveals that the proportion of FNFMGs (includes EVFMGs) out of total MDs doubled in 12 years: 1971 (40,222), 1983 (80,044) from 11.7% to 15.4% of all MDs. USFMGs remained at about 6% of total MDs in the 12 years but gained 9,969 physicians. FNFMGs continued to dominate the total FMG cohort, accounting for nearly two-thirds in 1971 with USFMGs at slightly more than one-third. In 1981, FNFMGs comprised over 70% but USFMGs declined to slightly over one-fourth. (Table 1)

FIGURE 1 -- CHANGES IN PHYSICIAN POPULATIONS -- 1971-1983

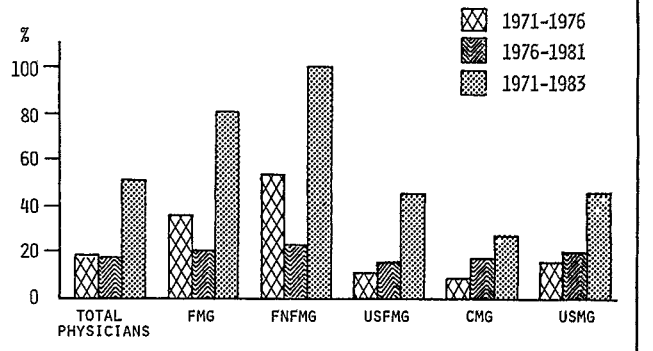


Table 1 - Federal and Non-Federal FMGs, FNFMGs, and USFMGs for Selected Years 1971-1983

Year	TOTAL FMGs	FN-FMGs*	US-FMGs
1971	62,214	40,222	21,992
1976	85,626	61,456	24,170
1981	102,762	74,914	27,848
1983	112,005	80,044	31,961
1971	100.0	64.7	35.3
1976	100.0	71.8	28.2
1981	100.0	72.9	27.1
1983	100.0	71.5	28.5

*FNFMGs include EVFMGs.

Activity/Specialty

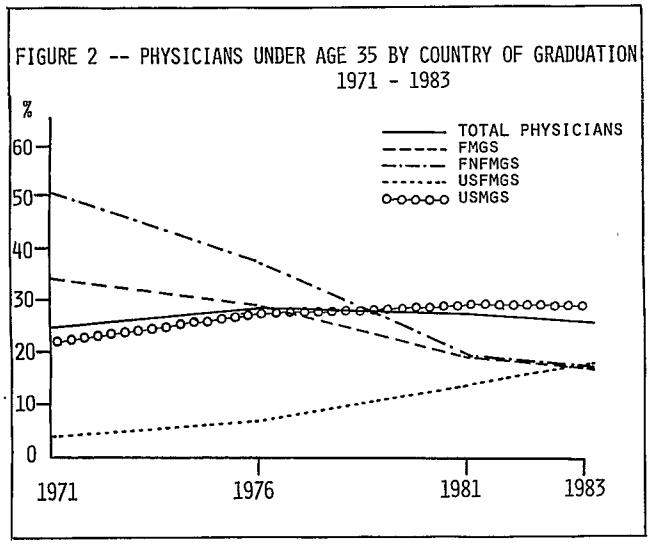
FMGs in Patient Care grew by 76% between 1971-1983. In this interim, FMGs in Office Based practice increased by 38,322 for a growth of 162% while Hospital Based practice grew by only 4.5%. In 1971, Office Based practice comprised 45% of all FMGs in Patient Care but by 1983 about 68%. FIVE specialties accounted for one-half or more of all Patient Care FMGs in each trend year (Table 2). FNFMGs (includes EVFMGs) in Patient Care comprised 11.7% of all Patient Care MDs in 1971 and about 15% in both 1981 and 1983. In both 1971 and 1983 USFMGs accounted for approximately 6% of all Patient Care MDs.

Table 2 - Specialties Ranked by Size For FMGs In Patient Care In Selected Years 1971-1983

1971		1976		1981		1983	
IM	15.4%	IM	16.3%	IM	15.9%	IM	17.1%
GP/FP	12.8	GP/FP	12.9	GP/FP	12.0	GP/FP	11.9
GS	11.1	GS	10.4	GS	8.6	PD	8.7
P	9.1	P	9.1	P	8.3	P	8.4
PD	6.9	PD	7.8	PD	8.2	GS	8.2
Total	55.3	Total	56.5	Total	5.3	Total	54.3

Age/Country of Graduation

Male FMGs grew from 16.5% of all male MDs in 1971 to one-fifth in 1983 while female FMGs decreased from 36% in 1971 to nearly 32% all female MDs in 1983. A more complete age profile of total MDs, US medical graduates and foreign nationals suggests correlations with developments in US medical education, immigration policies and legislation enacted in the 1970s that affected FMGs. In 1971, one-half of all FNFMGs (includes EVFMGs) were under 35, suggesting the influx of young MDs from abroad who came to the US for residency training. By 1976, the proportion decreased to 37.9%. Five years after PL 94-484 about 1-in-5 FMGs were under 35 and in 1983 17.8% were so. In contrast, USFMGs under 35 have steadily risen -- from nearly 5% in 1971 to 18% in 1983. Figure 2 illustrates trends for MDs under 35 by type between 1971-1983.

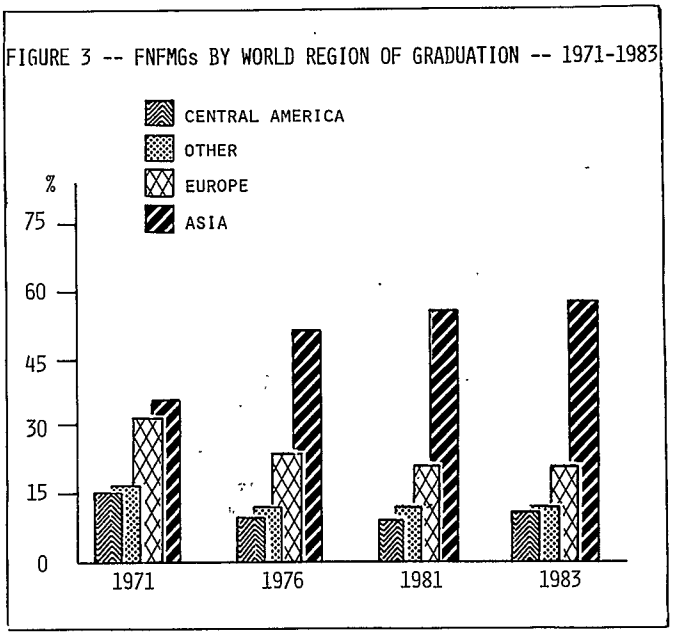


The proportion of FMGs under 35 in 1971 exceeded that of USMGs -- 34.3% vs. 23.7%. While in 1976 the proportions for these groups were at near parity -- 29.7% (FMGs) and 27.4% (USMGs), by 1981 the proportion of USMGs under 35 exceeded that of FMGs by 10 percent -- 29.4% (USMGs) and 19.3% (FMGs) with greater disparity by 1983 -- 29.1% (USMGs) and 17.9% (FMGs).

Although country and region of origin in the literature on FMGs often refers to an immigrant's last permanent residence, in Foreign Medical Graduates: 1983 Profile country and region refer specifically and only to country of medical education based on current political boundaries.

Since 1971, Asia, Europe, and Central America have consistently contributed the largest percentages of all FMGs in US medicine. Asian schools represented 45-49% of all FMGs in the US for each trend year (except 1971 - 39.2%). Central American schools grew in representation from 10% of all FMGs in the early/mid '70s to 14.6% in 1983. Europe, for the same interim showed steady declines: 40.7% (1971) vs. 27.2% (1983).

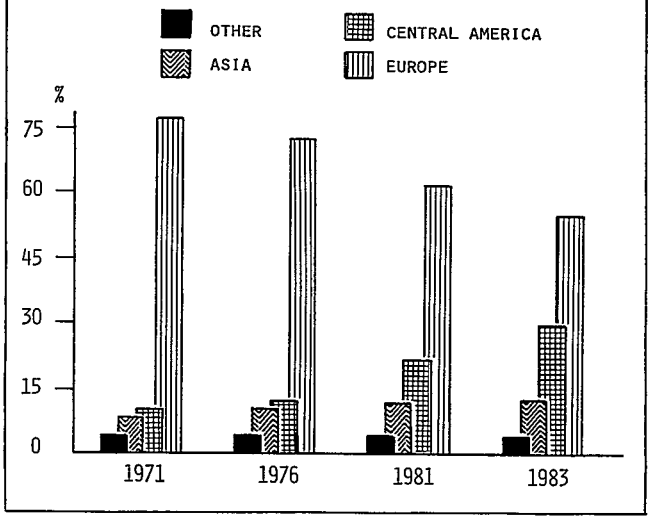
Figures 3 and 4 present percentage breakdowns by region of graduation for FNFMGs* and USFMGs and show similarities as well as striking divergencies. FNFMGs (Figure 3) from Central American schools declined from 15.4% in 1971 to 11.4% 12 years later. European representation also diminished -- from about 30 to 20%. Asian contribution, however, accelerated from 36.3% of all FNFMGs in 1971 to over one-half (55%) in 1983.



While FNFMGs from Central American schools declined, USFMGs (Figure 4) from this region nearly tripled since 1971 -- from 10.6% of all USFMGs in 1971 to 29.3% in 1983. Like FNFMGs, USFMGs from Europe decreased but their proportionate representation initially and through the trend years was larger -- over three fourths of all USFMGs in 1971 were European graduates while only slightly over one-half were

so in 1983. Similar to statistics for FNFMGs, Asian data shows increases of USFMGs between 1971-1983 but their proportionate percentages were significantly lower than those for FNFMGs -- 8.9% of all USFMGs in 1971 and 12.5% in 1983.

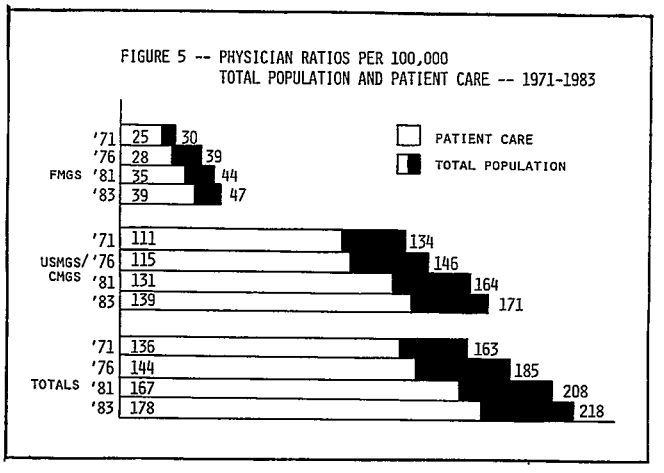
FIGURE 4 -- USFMGs BY WORLD REGION OF GRADUATION -- 1971-1983



Ratios

Although physician-population ratios are not intended as definitive indicators of manpower surplus or shortage, they are general guidelines to compare the distribution of MDs by type over time. Figure 5 presents ratios for physicians per 100,000 total population and in Patient Care.

FIGURE 5 -- PHYSICIAN RATIOS PER 100,000 TOTAL POPULATION AND PATIENT CARE -- 1971-1983



In 1971 there were 163 MDs per 100,000 total population; in 1983 the count was 218. Between 1971-1983 total MDs grew from 344,304 to 519,546 for 50.9% growth. The ratio of all MDs in Patient Care for the same population improved from 136 (1971) to 178 (1983). Patient Care physicians increased from 286,733 in 1971 to 423,361 at a percent change of 47.6.

Ratios for US and Canadian medical graduates per 100,000 population rose from 134 to 171. US and Canadian MDs gained 125,451 physicians -- from 282,090 to 407,541 -- at a rate of 44.5%. Patient Care ratios for US and Canadians grew from 111 to 139. In 1971, USMGs and Canadians totaled 234,539 but in 1983, 331,574 for over 41% growth.

The percent growth of FMGs nearly doubled that of US and Canadians -- 80% from 62,214 to 112,005 in the 12 years. FMG ratios increased from 30 to 47. In 1971, FMGs in Patient Care numbered 52,194; in 1983, 91,787 for a 75.9% growth with ratio growth from 25 to 39 per 100,000 population.

Table 3 displays the TEN HIGHEST states ranked by size for FMGs and the ratios for each state in 1971 and 1983. In both years, New York was highest in absolute counts -- 16,985 and 21,462. Its physician population ratio of 93 per 100,000 population was highest among the 10 states in 1971 followed by Maryland (61) and New Jersey (50).

Table 3 -- States with HIGHEST COUNTS of FMGs by RATIO Per 100,000 Total Population for 1971 and 1983

	---1971---			---1983---	
State	Total FMGs	Phys/ Pop. Ratios	State	Total FMGs	Phys/ Pop. Ratios
New York	16,985	93	New York	21,462	122
Illinois	5,006	45	California	10,269	41
Ohio	3,832	36	Illinois	8,622	75
New Jersey	3,605	50	New Jersey	7,419	99
California	3,130	15	Florida	7,260	68
Pennsylvania	3,045	26	Ohio	5,656	53
Michigan	2,766	31	Pennsylvania	5,200	44
Maryland	2,457	61	Texas	5,100	32
Massachusetts	2,188	38	Michigan	4,689	52
Florida	2,036	28	Maryland	4,022	94

PROFILES

The profile section represents AMA Masterfile data as of December 31, 1983 on key professional characteristics of FMGs such as board certification status, year of graduation, region and country of medical education, practice specialty preferences of USMGs, Canadians, and FMGs, and specialty of residency training.

Board Certification

In 1983, 38% of all FMGs were certified by a specialty board, 31.5% by a board corresponding to their specialty, 1.5% by a corresponding board and by other boards, and 5% by a non-corresponding board. Three-fifths (62.1%) of the FMG population were not board certified. FNFMGs (including EVFMGs) were more likely to be board certified than their US counterparts: 41.1% of FNFMGs were certified by at least one specialty board while 30% of the USFMGs were board certified. Canadian MDs were more likely to be board certified (51.1%) than FMGs. USMGs had the highest percentage of board certified physicians (56.2%) -- 15.1% higher than FNFMGs and 26.2% higher than USFMGs.

Year of Graduation

Foreign national FMGs (includes EVFMGs) tended to be more recent graduates than USFMGs,

Canadian, or US Medical graduates. This is clearly seen in Table 4, which displays the cumulative percentages for each decade of graduation. Nearly all FNFMGs (95.5%) graduated after 1950 while only about three-fourths of MDs in each of the other groups graduated after the same year. This difference suggests the influx of FNFMGs following World War II. Table 5 presents a more discrete breakdown of graduation for physicians by type. A relatively high proportion of USFMGs graduated before 1940, over 14 percent compared with less than 10 percent of the USMGs and only about 1 percent of FNFMGs. These data may reflect the fact that the early part of this century saw the evolution of American medical education from a collection of mostly developing medical schools to its current position of pre-eminence.

Table 4 - Cumulative Percent Distribution by Year of Graduation for Foreign, Canadian and US Graduates, December 31, 1983

	---Year of Graduation---				
Group	After 1970	After 1960	After 1950	After 1940	After 1930-
FMG	36.4	70.6	89.2	95.2	98.8
FN-* FMGs	37.0	79.4	95.5	98.9	99.6
US-FMGs	34.6	48.4	73.3	85.9	96.6
CMGs	28.7	52.1	74.8	87.4	95.6
USMGs	44.9	63.6	78.6	90.7	97.5

*Includes EVFMGs

Table 5 - FMGs, US and Canadian Graduates by Year of Graduation, December 31, 1983

Group	Total	Prior 1940	1940 1949	1950 1959	1960 1960	1970 1979	1980 1983
FNFMGs	100.0	1.0	3.4	16.1	42.4	34.1	2.9
USFMGs	100.0	14.2	12.6	24.9	13.8	22.7	11.9
USMGs	100.0	9.3	12.1	15.0	18.7	29.3	15.6
Canadians	100.0	12.7	12.6	22.7	23.4	24.8	3.9

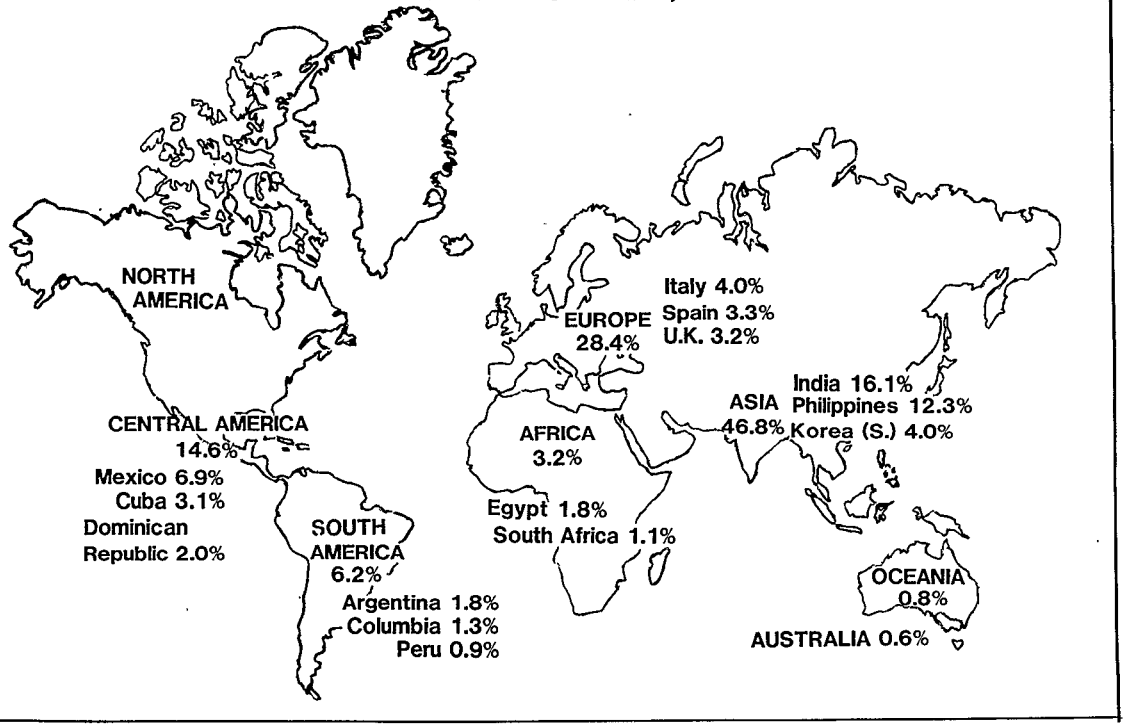
*FNFMGs include EVFMGs.

Country/Specialty

Figure 6 illustrates countries of graduation that contributed highest counts of FMGs to US medicine as of December 31, 1983. India was the largest supplier of FMGs - 16.1% - followed by the Phillipines - 12.3% - and Mexico 6.9%. The 15 countries in Figure 6 accounted for over 60% (62.4) of all FMGs in the US in 1983. Figure 6 also displays the prercent of FMGs out of total FMGs by region of graduation.

The Profile section also includes data on age and country of graduation for each of the ten specialties most popular among FMGs as a whole and among USFMGs and FNFMGs. The country of graduation for these data is based in each case on the 30 highest countries specific to each specialty. Thus, the 30 countries listed for FMGs in Internal Medicine are not completely synonymous with those in Pediatrics or any

FIGURE 6 -- FMGs BY WORLD REGIONS AND SELECTED COUNTRIES OF GRADUATION -- DECEMBER 31, 1983



other of the ten as the complete study demonstrates.

For each of the populations, Internal Medicine was the top ranking specialty. Table 6 displays the percent of total FMGs, FNFMGs, and USFMGs in Internal Medicine as well as those under 35 by the five highest countries of graduation ranked by size for the discipline.

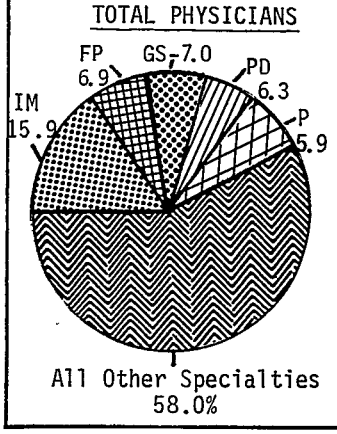
One-fifth of all FMGs in Internal Medicine as of 1983 were graduates of Indian schools. This high rank of India among FMGs in Internal Medicine was due largely to its popularity (28.2%) among FNFMGs (includes EVFMGs). India was particularly dominant (43%) among FNFMGs in Internal Medicine in the under 35 group. The second rank of the Phillippines for FMGs in Internal Medicine was due largely to FNFMGs. Mexico's third rank among FMGs may be attributed to its high representation among USFMGs (20.8%). Among those under 35 in Internal Medicine, Mexico accounted for approximately 5% of FNFMGs but for over one-third of the USFMGs in this group.

Table 6 - FMGs in Internal Medicine by FIVE Highest Countries of Graduation, December 31, 1983

----Internal Medicine----					
Country	% FMGs	Country	% FNFMGs	Country	% USFMGs
		All Ages			
India	20.7	India	28.2	Mexico	20.8
Phillippines	11.3	Phillippines	14.1	Italy	13.3
Mexico	7.8	Pakistan	3.7	Dominican Rep.	7.3
Italy	4.7	Korea (S)	3.3	Spain	6.1
Dominican Rep.	3.0	Taiwan	3.2	Switzerland	5.5
		Under 35			
India	28.6	India	43.0	Mexico	34.4
Mexico	14.9	Phillipines	6.3	Dominican Rep.	15.6
Dominican Rep.	6.7	Mexico	4.8	Italy	9.4
Phillippines	6.0	Pakistan	4.2	Grenada	9.1
Italy	4.4	Taiwan	2.6	Phillippines	5.5

Specialty distributions may also be compared for total MD population, USMGs, Canadians and FMGs. Five specialties illustrating highest representations for each of these populations are depicted in Figure 7. Internal Medicine was the top choice with near parity percentages in each population. Anesthesiology was represented for FMGs at 6.5% of total FMGs but did not rank in the top five for all MDs or for all USMGs and Canadians. Family Practice did so for these groups but not for FMGs. Obstetrics/Gynecology was among the highest choices for USMGs and Canadians but absent from the top list for total physicians and FMGs.

FIGURE 7 - TOP SELECTED SPECIALTIES FOR PHYSICIANS, December 31, 1983



	US & CMGs	FMGs
OS	57.9%	55.8%
IM	16.0%	15.6%
FP	7.6%	--
GS	7.0%	7.0%
PD	5.9%	7.7%
P	--	7.4%
OBG	5.6%	--
AN	--	6.5%

Specialty of Residency Training

According to preliminary data TEN specialties of residency training were responsible for about 70% of all US physicians in graduate medical education in 1983: Internal Medicine, General Surgery, Family Practice, Pediatrics, General Practice, Psychiatry, Obstetrics/Gynecology, Anesthesiology, Orthopedic Surgery, and Pathology. US medical graduates accounted for nearly three-quarters or more of residents in these specialties except in Anesthesiology (63%) and Pathology (69%). USMGs accounted for over 80% of residents in Family Practice (87%), General Practice (82%) and Orthopedic Surgery (89%).

Nearly one-fifth of all residents in Anesthesiology (19%) were FNFMGs while Pediatrics (13%) and Pathology (14%) were second and third with highest proportions of FNFMGs. Almost 9% of all residents in Psychiatry were USFMGs, representing the largest USFMG constituency across the ten specialties. The percent of EVFMG residents exceeded minimally that of USFMG residents in five specialties: Internal Medicine (5.5 vs 5.8%), General Surgery (5.8 vs 5.5%), Pediatrics (7.5 vs 5.2%) Anesthesiology (9.0 vs 7.5%) and Pathology (8.7 vs 6.9%).

These same residency specialties except for Orthopedic Surgery and Cardiovascular Diseases matched those for highest percents of total FMGs (64%) in graduate medical education with a variation in rank order for size. Orthopedic Surgery appeared only on the total MD list and Cardiovascular Diseases appeared only on the FMG list. FNFMGs comprised nearly 40% or more of all FMGs in residency specialties except in General Practice and Family Practice where USFMGs exceeded FNFMGs. Approximately one-half of all FMG residents in Anesthesiology and Pediatrics were FNFMGs. EVFMGs and FNFMGs demonstrated near parity proportions of all FMG residents in General Surgery -- nearly 40% for each group.

LOCATION

Location data focuses on census division and world region of graduation, state of residency training and state of practice, state of location and activity, among other variables.

State of Location

The state as an areal unit of analysis is appropriate for a variety of research efforts that include public finance policy making and graduate medical education assessments. Physician population ratios within the state however may vary widely. Also interstate differences may reflect an underlying urban/rural distribution modified by state policy variables, climate and regional influences, as well as economic and professional factors.[1]

In 1983, the FIVE states of New York, California, Illinois, New Jersey, and Florida cumu-

latively accounted for nearly one-half (49.1%) of all FMGs, 47.2% of all FNFMGs, and one-half (54.1%) of all USFMGs. These same states comprised over one-half (56.1%) of all FMGs in residency training.

State of Residency Training

Five states (as indicated above) accounted for nearly one-half of all FMGs in the US. Preliminary 1983 state of residency data indicates that New York and Illinois demonstrated the highest retention of FMGs who did graduate medical education in these states.

Of all FMGs located in New York in 1983, 85.2% did residency training in the state, approximately 7% did so in the contiguous states, and about 7% trained in other US states. Nearly 73% of FMGs in Illinois trained in Illinois while only approximately 4% did so in states contiguous to Illinois. About 23% of the FMGs in Illinois trained in other US states.

FMG retention statistics for California and Florida reflect the geographical movement of civilian population growth in the sun belt. Over 60% of all FMGs located in California in 1983 did residency training in states other than California and its contiguous states. About 38% trained as residents in the state. Similarly, two-thirds of all FMGs (66.4%) in Florida and its contiguous states while only slightly over 30% (31.6%) trained in Florida.

Census Division Location

Representation of physicians by type may also be viewed within census division boundaries. Table 7 illustrates the percent of total MDs, FMGs, FNFMGs, USFMGs, and USMGs and Canadians for each census division. Largest proportions of USFMGs in 1983 were concentrated in the Middle Atlantic - 36% while the lowest percentage was indicated for the East South Central area (1.6%). The Middle Atlantic also had the highest percent of constituency of FNFMGs -- 28.2% while the Mountain division (1.6%) had the least. Slightly over one-fourth (26.2%) of all Canadians were located in the Pacific division while only 3.0% were so in the East South Central. US medical graduates demonstrated largest percentage constituencies in the South Atlantic (16.8%) with an only 5.2% concentration in the East South Central.

Table 7 - Physicians by Census Division of Location and Country of Graduation, December 31, 1983

Census Division/ Area of Location	Total Physicians	FMGs	FNFMGs*	USFMGs	Canadians	USMGs
Total Physicians	100.0	100.0	100.0	100.0	100.0	100.0
New England	6.8	5.5	5.1	6.3	11.8	7.1
Middle Atlantic	19.1	30.4	28.2	36.0	15.2	16.0
East North Central	15.2	19.2	21.1	14.3	12.0	14.2
West North Central	6.2	4.0	4.5	2.9	5.6	6.8
South Atlantic	16.6	16.1	16.9	14.2	11.4	16.8
East South Central	4.6	2.3	2.6	1.6	3.0	5.2
West South Central	8.5	6.1	7.0	4.0	8.1	9.2
Mountain	4.7	1.8	1.6	2.2	5.2	5.5
Pacific	16.6	10.5	10.6	10.3	26.2	18.1
Possession	1.3	2.7	1.0	7.1	0.1	0.9
Address Unknown	0.6	1.3	1.4	1.1	1.4	0.4

*Includes EVFMGs

Table 8 presents a somewhat different spatial dispersion of FMGs by illustrating FNFMGs and USFMGs by percentages of total FMGs in each census division. Foreign national FMGs accounted for 71% of all FMGs in the US and Possesions in 1983; USFMGs did so at 29%. Within the 9 census divisions, FNFMGs accounted for the highest share of FMG populations in the East and West South Central divisions -- 81% in each case. USFMGs had highest percents of total FMGs in the Mountain (35%) and Middle Atlantic Divisions (34%).

Table 8 -- FNFMGs and USFMGs as Percentages of Total FMGs by Census Division, December 31, 1983

Census Division	Total FMGs	FNFMGs	USFMGs
Total U.S. & Possessions	100.0%	71.0	29.0
New England	100.0%	67.0	33.0
Middle Atlantic	100.0%	66.0	34.0
E. North Central	100.0%	79.0	21.0
W. North Central	100.0%	80.0	20.0
South Atlantic	100.0%	75.0	25.0
E. South Central	100.0%	81.0	19.0
W. South Central	100.0%	81.0	19.0
Mountain	100.0%	65.0	35.0
Pacific	100.0%	72.0	28.0

CONCLUSION

FMGs are not a homogeneous group. Ideally, their distribution could be further studied, like other topics of physician supply, within the context of the social, economic, and political infrastructures[2] of the US and donor countries. Future research might address these areas and consider a variety of complex, possibly interdependent, patterns within the American medical system itself: (1) the correlation, if any, between FMG spatial dispersion in metropolitan and non-metropolitan areas and country of medical education (2) the relation between residency and practice specialty (3) the demographic and social characteristics of USFMGs and their subsequent career paths; and others.

Foreign Medical Graduates: 1983 Profile recognizes the need for such additional studies and provides a descriptive base from which many such efforts might proceed.

*These data exclude EVFMGs, Not Classified, Inactive, and Address Unknown.

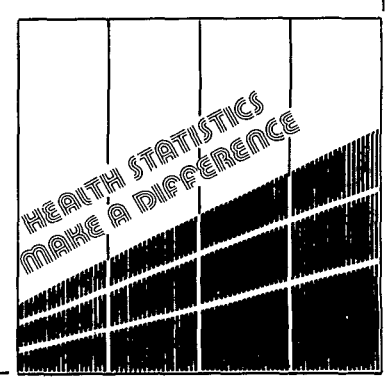
REFERENCES

1. Eisenberg, Barry S., and Cantwell, James R. The Spatial Distribution of Physicians: A Literature Review. Center for Health Services Research and Development, American Medical Association, June 1975.

2. Kindig, David and Taylor, Charles. Growth in the International Physician Supply, JAMA (Vol. 253, No. 21), June 7, 1985.

Third Plenary Session

Statistics Make a Difference—
The Local Perspective



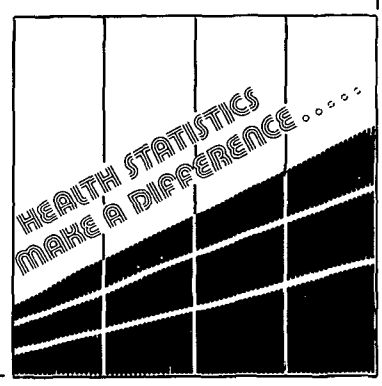
Statistics Make A Difference -- The Local Perspective

David J. Sencer, M.D., M.P.H.
Commissioner of Health
New York City Department of Health

(Not available for publication)

Session S

Local Geographic Information Systems



GEOGRAPHIC INFORMATION SYSTEMS
SERVING LOCAL POLICY DECISIONS

Alan B. Humphrey, University of Rhode Island

INTRODUCTION

The current proliferation of hardware and software is leading to new concepts in the analysis and presentation of health statistics. The use of computers has evolved from tabbing equipment (Electronic Data Processing, EDP), to the generation of routine reports for managers (Management Information Systems, MIS), to hands on inquiry of databases (Decision Support Systems, DSS). The transition from one phase to the other has been gradual and usually perpetuated by the release of new computing equipment and software. Most impressive from the standpoint of quantity are the MIS available today. They range from the user specific to the general, but they all have one characteristic in common - they are database specific. The input and output formats are well defined, and the timing of the reports are fairly specific. In most cases the audience for which the reports are intended are middle and upper management with well defined problem specifications. These characteristics have changed as the audience requesting the information has shifted from the manager to the policy maker.

This policy orientation has associated with it partially defined problems, changing analysis and output requirements, and a need for results at non specific time intervals. In addition, the user generally has a minimum knowledge of statistics. To meet this need, DSS have evolved. They vary in their complexity and are often extensions of existing MIS.

The purpose of this paper is to explore some of the major components of a DSS to serve local policy decision making and to suggest alternatives for their development and implementation. The framework around which such a DSS could be constructed is related to geography and the recording of events and characteristics of individuals residing on specific pieces of land.

The primary components of a DSS are 1) Database, 2) Geographic Base, 3) Database Management System, 4) Graphics System, 5) Model Base, and 6) Dialog Generation Interface. Each of these must be capable of operating as stand alone systems as well as linked systems.

DATABASE

The composition of the database should be considered not in terms of data sources, but rather in terms of data types. For example, the Census is the primary source of demographic data, but there is also demographic information found in health interview surveys and vital records. Similarly, health status information can be found in both health interview surveys and vital records. This can be thought of as a data source by data type matrix where the user is more concerned with the data type than the source. Therefore, the four primary categories of the database would be something like the following from the users perspective:

- DEMOGRAPHICS
- HEALTH STATUS
- HEALTH RESOURCES
- MEDICAL CARE UTILIZATION

Each of these categories could have sub-categories, such as environmental characteristics within the Demographics group, etc.

While the database need not contain data linked to geography, this discussion will deal only with data that is geographically coded. Many data systems have geographic coding inbedded in the data collection procedures. However, for those that do not, there are methods available for inserting the necessary geographic information. For example, vital records, hospital discharge records and health manpower data contain addresses that can be geographically coded. This can be accomplished with such computer tools as the ADDMATCHER that will geocode addresses to geographical entities (e.g. census tracts, city blocks, city block faces).

The database can also contain smaller data sets that are aggregated to different geographic levels. For example, if census tract aggregates were used extensively, certain key variables with specific applications could be maintained in special data files. This would reduce the access time as well as the computer costs.

GEOGRAPHIC BASE

The geographical unit to which the data is coded should not be the same unit that the final analysis will eventually use. The flexibility and ultimate uses of a database are greatly enhanced if very small, hierarchically consistent units are used. This allows the user to define the larger areas for analysis and is not tied to existing boundaries. Hence, the optimum size for primary coding is a square foot or possibly smaller. For most this is impractical, but the point to be made is that the units should be as small as possible. All too often data is collected at the county or state level because that is the level of final interest.

Also, the smaller the primary unit the greater flexibility there is in combining data from various sources. A number of geographic data bases exists that are not directly related to

health affairs, but could be used in analyzing and interpreting health related data. This includes ground water characteristics, archeological data, soil types and vegetation.

The units should also be as uniform in size as possible. This is not as important as size, but the more uniform the size the easier resultant maps will be to interpret.

Perhaps the most widely used geographic data base is that developed by the U.S. Bureau fo the Census, i.e. the GBF DIME File. Others have been developed and are increasing in their popularity such as those based on the U.S. Postal Service's ZIP Codes. The problem with these geographic identifiers is that they change over time and the variability within an area can be larger than the variability among areas.

DATA BASE MANAGEMENT SYSTEM

Since the database will contain many different types of files covering various temporal and spatial dimensions, it is imperative that a system be used that can manipulate large data files quickly and easily. There are many alternatives on the market that operate on all sizes of computers from micros to main frames. Some now have the capability to interface between the various types of machines.

FOCUS is a database management system that was originally designed for use only on main-frame computers, but has recently been adapted to the micro computer. A user can be working on a PC in FOCUS and access data from the main-frame when necessary with a minimum of effort.

SAS is a powerful and easily learned statistical package that many of us use to edit, analyze and display our results. For many, using a main-frame computer, this is the selection of choice. It also has the capability to handle geographic data sets.

ARC/INFO is a geographic database management system that has received acclaim in the environmental and land management area and is being used in a number of states throughout the country. It operates on medium sized computers (MINIS), creates maps readily and data sets that can be used by other software packages.

GRAPHICS SYSTEM

While graphics have become an integral part of most analytical systems, its inclusion still needs to be stressed. Not only should a graphics system be included, but it should be easily accessible and linked to the other components of the DSS.

Of primary importance is a mapping system that is easy to use and fully integrated with the other components of the system. A combination of the flexibility of CALFORM and the ease of use of SAS/GRAPH would be ideal.

In addition to being able to chart, plot and map the results of an analysis, it should have the capacity to draw user defined graphics. Many situations arise for which schematic diagrams would make the interpretation of a complex analysis much easier.

MODEL BASE

The Model Base contains instructions that manipulate and analyzes the data from the data base. These models generate the results that are used in the decision making process and can vary from the very simple to the complex.

The Model Base compliments the database. The database provides the input to the model base while the model base defines the data needed for the database.

Many of the instructions included in the model base are those that would be used very often while others are very specific and would be used rarely. One set of instructions for example might calculate age-sex adjusted death rates. The user would be requested to indicate the variables to use, the algorithm would calculate the rates and output a new data set with the age adjusted rates. Additional models might be a health status indicator profile for specified areas or a resource allocation model based on various budget constraints.

DIALOG GENERATION INTERFACE

Central to the DSS is the Dialog Generation Interface, (DGI). It serves many functions including the linkage of the DSS components, generating computer instructions, listing a work sessions' activities, generating specific data sets from the database, and retaining the instructions from a work session for later use and modification.

The DGI would operate from the users standpoint from a primary menu with a list of options. The selection of an option might lead to yet other options which would result in specific tasks being accomplished. The user could request that all the instructions be saved. If, at a later session, the user wanted to repeat some, but not all, of these instructions, the previous work could be edited and the n re-run. The implications of this for simulation and sensitivity analyses should be obvious.

Another important feature of the DGI is the creation of instructions to be included in the Model Base. Once the user has completed a set of instructions they could be saved under a specified name and run with new data or with modified inputs.

DESIGN STRATEGIES

There are essentially two ways to approach the construction of a DSS. The first is patterned after the construction of a Management Information System. Obtain all the information

on inputs and outputs, configure the system for optimal performance, write the code, and then implement the system. This is very difficult to do for a DSS since the outputs, by definition, are ill defined or missing at the inception of the project. Also, the inputs can and should change as time progresses and the system is used. Consequently, a more evolutionary approach is needed.

This approach does not imply that definitions and file specifications are not made, nor that the components are loosely or haphazardly thrown together. Rather, it implies that once an overall scheme has been defined, the components are developed, tested, used, and modified. Each component should be able to stand alone as well as being capable of working with other components.

The first consideration is the definition of the users of the system and how they will use it. They are not program managers receiving routine reports. Rather, they are policy makers that will be using the system on an ad hoc basis. For the most part they will not be technically sophisticated but will want to play a number of what-if types of games. Thus, the system will have to remember instructions that have been submitted at an earlier date. This critical requirement leads to the Dialog Generation Interface.

IMPLEMENTATION STRATEGIES

In order for the DSS to be truly effective it must be used and shared by many individuals. The construction and maintenance of the database alone can be costly and time consuming. However, if this activity is shared by several agencies it can be manageable.

Through the efforts of the National Center for Health Statistics and the Cooperative Health Statistics System, many of the guidelines needed for the building and sharing of databases have already been established. In particular, the Uniform Hospital Discharge Data Set efforts provide insights on how a number of different program needs can be met with one, well thought out set of data items, how the data can be collected and distributed to a number of users, and how the results can be used.

An equally important effort to that of creating the database is the sharing of software and statistical techniques for analysis and display. The use of graphcis and especially the use of mapping has come a long way since the early Census Use Studies. There are several mapping packages available today that run on many different size computers.

CONCLUSIONS

Many of the components discussed here have been developed and are being used in a variety of contexts. The Massachusetts Health Data Consortium, under the direction of Elliot Stone, has developed a user oriented data retrieval package for Hospital Discharge Data. This package allows the user to create a small database for analyses and downloading to a micro comptuer. Dr. Jeffrey Gould has developed a census data retrieval package using SAS that allows the user to select specific census variables. These variables are stored in a SAS database for later analysis and mapping. SysteMetrics, Inc., has developed a comprehensive decision support stystem for hospital planning and evaluation. It includes most of the components discussed here with the primary exception being the Dialog Generation Interface.

While none of these systems are directed toward issues dealing with public policy, the adaptation capabilities are certainly there.

With these advances in computer hardware and software, mechanisms need to be established for the sharing of information. There are many alternatives available which include newsletters and electronic bulletin boards. A newsletter, possibly under the auspices of the National Center for Health Statistics, could include brief articles and letters regarding new applications, innovations, and techniques that are not ready for journal publication. This could also be a place to publish negative results and commentaries. An electronic bulletin board could be established for the exchange of computer software, data, and techniques.

Regardless of the details, the time has arrived for the states and localities to join with the National Center for Health Statistics to form a partnership for the advancement of information systems that will serve local policy decisions.

MICROCOMPUTER-BASED GEOGRAPHIC INFORMATION SYSTEMS FOR LIMITED AREA ANALYSES

Debora Sidor Stewart, Loyola M. Caron, James W. Merchant and
P. Andrew Norton, University of Kansas Space Technology Center

Introduction

Environmental management and policy decisions must, almost always, be based upon examination and analysis of the interplay of many different factors which may bear upon a particular issue. Data involved in analyses usually have a locational, or geographic component.

Maps are important sources of geographic data. However, significant problems can be encountered in using them:

- They are frequently out of date and are costly and time consuming to update;
- They are often at the wrong scale or in the wrong format for a particular need;
- It is typically too expensive or time-consuming to produce cartographic products (e.g., maps at different scales) to address different needs since such work usually involves recompiling and/or redrafting the maps; and
- Maps are difficult to compare and overlay in order to discern important spatial interrelationships, especially when the problems noted above exist, or where it is necessary to overlay, differentially weight and compare three or more maps.

During the last decade, great advances have been made in the use of automated techniques to store, manipulate, compare and display geographically-referenced data. A *geographic information system* (GIS) is a system in which all data are spatially referenced so that multiple themes of data can be registered and analyzed in concert[1/]. A computer-based GIS allows a "map" user to store "map" data in digital (numerical) format and automatically retrieve, display, overlay and update data. Data may come from topographic maps, land use maps, soils surveys, hydrologic records, the census or a myriad of other sources. Virtually any data that are, or can be, mapped (i.e., geographically-referenced) can be "digitized" and stored in the computer. Once stored, these data can be automatically extracted, reconfigured, updated, analyzed, mapped in a format and at a scale designed to meet a specific need, and used for many types of decision-making.

Some of the advantages of using a GIS include:

- Low cost of analyzing various scenarios and relationships once data are entered;
- Rapid analysis and output;
- Maps and accompanying statistics can be generated for specific applications; and
- Data can be easily updated and expanded.

The focus of our current work is on agricultural counties. We are working with the American Farmland Trust, a non-profit organization dedicated to the preservation of agricultural lands and the promotion of farming opportunities. Their interest is in implementing geographic information systems in rural counties, to assist them in identifying prime agricultural lands and to aid them in developing defensible policies relating to their preservation. They are especially concerned with rural counties adjacent to metropolitan areas which are experiencing rapid growth.

Objectives

The goal of our study is to facilitate the adoption of GIS techniques by local jurisdictions, for use in limited area analyses. Specific objectives are to:

(1) Stimulate the market to encourage the development of affordable geographic information systems; and

(2) Provide technical assistance to local-level users.

Although our current efforts focus on implementing GIS techniques for planning and management of natural resources, the same techniques can be applied to health statistics or any other discipline requiring spatial (i.e., geographic) data.

Nature of the Problem. In the past, the implementation of geographic information systems techniques in non-urban local jurisdictions has not proceeded as quickly as it has in other arenas. This slow adoption of GIS technology is largely due to a number of characteristics of local jurisdictions that, we believe, have impeded progress in this area. These include:

- Small population (low density) of the administrative or management area;
- Limited budgets, due to a relatively small tax base;
- Relatively few planning/ natural resources management professionals;
- Lack of computer staff to implement the technology; and
- Administrators and policymakers who control the purse-strings are often unfamiliar with technical problems of planning and natural resources management, and consequently are not willing to support a long-term investment in a geographic information system intended to alleviate the situation.

Clearly, these characteristics have slowed the adoption of GIS techniques in local jurisdictions. Equally important, however, is the effect these factors have had on the vendors and developers of geographic information systems. Vendors respond to demand. To date, the greatest demand has come from large metropolitan areas concerned with transportation and zoning issues, or other urban problems requiring (in some cases) a sophisticated system of very high spatial resolution. Unfortunately, systems developed for this market are far too expensive for jurisdictions

working with limited budgets. Moreover, these systems are typically more sophisticated than a local area would require, at least initially, for its needs.

It is probable that a local jurisdiction could justify the purchase of a geographic information system that can be installed on a relatively inexpensive microcomputer, or personal computer. Until recently, however, such microcomputer-based geographic information systems have not been well packaged. Although affordable, they have been:

(1) Focused primarily on digital image processing, with a poorly developed GIS component;

(2) Designed primarily for automated mapping (i.e., they lack analytical and/or overlay capabilities); or

(3) Developed for educational purposes.

In addition, these microcomputer-based systems have not offered adequate graphic capabilities, and have not lent themselves to enhancement or upgrading through the addition of hardware or other software modules.

In summary, we have observed that the adoption of GIS techniques by non-urban local jurisdiction is influenced by:

- GIS vendors/developers,
- Local users (e.g., county planners), and
- Local-level administrators and policymakers.

Vendors are responding to the demands of highly urbanized areas and engineering firms requiring sophisticated (and consequently, expensive) systems. Although local users frequently recognize the value of these systems, and may wish to acquire them for planning and management efforts, they have found it difficult to convince administrators and policymakers to invest in the technology, especially when existing systems are either too expensive or are not packaged to meet their specific needs.

Until vendors develop more affordable systems, or re-package existing systems, the local user cannot demonstrate the benefits of GIS technology to the administrators controlling the purse-strings.

Scope of Work

The overall strategy for facilitating the adoption of GIS techniques involves a three-phase (four-year) program. During the first phase of our work (the subject of this paper), we identified and documented geographic information systems software. The purpose of this effort was to ascertain the availability of existing software, and to characterize each package based on its capabilities and operating environment.

During the second phase, one of these packages will be implemented in a (prototype) rural county. Through this process, we will:

(1) Assist rural planners and managers to create and use the system, and provide other technical assistance, as required;

(2) Provide feedback to vendors regarding those aspects of the system not well suited to meet their needs (i.e., are not well packaged), so that necessary modifications can be made;

(3) Identify new applications of this technology brought to light by the active utilization of the system by local-level users knowledgeable of specific situations existing at the local level; and

(4) Define cost/benefit relationships. This type of information has not been well documented in the past, but is essential if administrators and policymakers are to be convinced to invest in the technology.

In the third phase of the project, geographic information systems that have been repackaged to reflect the specific needs of local users will be installed in up to five additional counties selected throughout the United States. During this phase, technology transfer between peers (e.g., county planner to county planner) will, we believe, facilitate the adoption of the systems.

Characteristics of a Model GIS. It is our belief that a model GIS suitable for use by non-urban local jurisdictions should have at least the following characteristics:

(1) Full GIS Capabilities - The system should not just be designed for automated mapping, but should allow overlaying of multiple data sets, spatial modeling, and area measurement. All functions should be generic (i.e., they should not be restricted to any particular application). The system should also have sufficient data storage capability to handle "county size" data sets at operationally required spatial resolution.

(2) Low cost - A fully functional system (hardware and software) should sell for no more than $40,000, and should provide, minimally, digitizing capabilities, data storage, interactive color display, and hard copy output capabilities.

(3) Packaged as an integrated system - The system should be capable of being sold "off the shelf" as a unit (including hardware and software), and should require no special expertise for assembly or operation. Maintenance and service should be easily available. No unusual operating environment should be required.

(4) Flexible -The system should lend itself to enhancement, through the purchase of additional software modules and/or hardware, implementation of specialized models, and networking.

(5) High quality graphics - The system should possess graphics capabilities (color and black and white) sufficient for day-to-day decision-making, for publication, and for public and administrative presentations.

(6) User friendly operation - The system should not require a background in computing to be used effectively.

Results: Phase I

During Phase I, a survey of geographic information systems software was conducted. Over 2,000 brief questionnaires were mailed to individuals believed to be possible contributors of information (e.g., known vendors, computer firms, members of organizations and societies currently using GIS techniques). The purpose of this questionnaire was three-fold: (1) To identify developers or sellers of GIS software; (2) To request assistance in identifying other individuals who might be developing GIS software; and (3) To inform individuals of the study, in the event that they might wish to acquire a summary of its findings.

This questionnaire identified 84 vendors or developers of GIS software. Detailed information about their software was acquired through a follow-up survey, designed to address the following aspects:

- General information on their software (whether it was currently available or under development, the cost of the basic GIS software or turnkey system, "user-friendliness", and the availability and cost of software support and training);
- Operating environment needed for the software (what computers it had been or could be installed on, the operating system, what type of storage was needed, and whether special peripheral hardware and software was required, or if it was optional); and
- Functional capabilities of the software (e.g., data entry, editing and updating, data analysis and map and graphic output)

Fifty-six follow-up surveys were returned. Of these, three software packages appear to meet several of the criteria set forth earlier. That is, they have full GIS capabilities, are packaged as integrated systems, are flexible, user-friendly, and relatively low in cost. Graphics offered by these systems can be described as adequate to good.

The three systems are (alphabetically, by vendor):[2/]

1. GIS-100 by Aeronca Electronics, Inc.
2. AUTOGIS by Autometric, Inc.
3. ERDAS GIS and Image Processing System by Earth Resources Data Analysis Systems.

Although several of the other systems appear to offer very attractive capabilities and operating environments, they are not included here because they are either:

1. Not yet released or under development,
2. Can handle only limited amounts of data,
3. Were developed solely as educational tools, or
4. When coupled with the cost of the host computer, are quite expensive, and consequently, beyond the reach of most non-urban local jurisdictions.[3/]

Brief descriptions of the three systems that best meet our criteria for implementation at local levels are included below.

GIS-100[4/]

Aeronca Electronics, Inc. offers the GIS-100 for the IBM PC, IBM PC/XT and compatibles.

> "This low cost software package was developed for project level data manipulation, map display and business and statistical data display. The software produces high resolution color graphic displays, automatic business type graphics (i.e., PIE charts of map statistics), and the whole system is driven by a MOUSE for easy interaction of map data and screen graphics. The software not only provides color images, but scaled graphic printer or plotter output and polygon line or point input through the MOUSE or table digitizer. The analytical capabilities include defining proximities (locational analysis), suitability, siting, impact assessments, pair-wise combinations, rescaling, and much more."

AUTOGIS[5/]

> "The Automated Geographic Information System (AUTOGIS) is a software package that has been designed for land management and military agencies in the United States. It provides the basic functions of data capture, storage, retrieval, analysis, modelling and display of spatially referenced data.
>
> "AUTOGIS is menu or command driven and supports batch mode operations; abbreviated and concatenated commands are available for the experienced user. All sequences, whether menu or command driven, contain on-line help files. The system specifications and design were largely determined by users and potential users in the Federal Land Management and Resource agencies. AUTOGIS is therefore orientated towards the user and 'user friendliness'
>
> "AUTOGIS installation on DEC-VAX Systems operating under VMS, HP-9000 series under UNIX and Data General computers operating under AOS or AOS/VS, including the Data General DeskTop Series, is undertaken at the client's site."

ERDAS GIS and Image Processing System[6/]

"The ERDAS-PC Integrated Image Processing and GIS system utilizes ERDAS' high resolution (512 x 512 x 32 bit) image display to provide professional level capabilities for image enhancement, classification, geometric correction, GIS merger and analysis and scaled hard copy. Over 120 easy to use production tested programs produce intelligently computed default answers. The menu-driven software is user-friendly with program inputs in plain English.

"Modular design allows for acquisition of software/hardware modules on an individual basis (e.g., GIS, IMAGE PROCESSING, TAPES, COLOR SCALED HARDCOPY, POLYGON DIGITIZING AND VIDEO DIGITIZING). Optional Mass storage may include a tape cartridge disk drive. ERDAS offers complete installation and on-site applications software training. Each system comes with thorough documentation, including a 400-page USER'S GUIDE and 200-page APPLICATIONS PROGRAMMERS MANUAL. The optional SOFTWARE SUBSCRIPTION SERVICE (SSS) provides yearly updates and new releases. For advanced users, menus may be edited or bypassed and the SOFTWARE TOOL KIT enables automatic linkage of user developed programs."

"The ERDAS-PC is only one of the ERDAS family of systems which includes DEC, DG, and PRIME. Because ERDAS menus and data formats are consistent, the ERDAS-PC systems may serve as intelligent workstations to a larger computer."

Descriptions of all GIS software packages documented throughout the course of this study, have been published by the American Farmland Trust, 1717 Massachusetts Avenue N.W., Washington, D.C. 20036. The report includes several tables summarizing software availability, operating environments, and functional capabilities.

Acknowledgements

This study was supported by a grant from the Andrew W. Mellon Foundation administered by the American Farmland Trust (AFT). Their support is gratefully acknowledged. The authors would especially like to thank Robert J. Gray, AFT Director of Policy Development and Project Director, and Margaret Maizel, AFT Project Manager, for their assistance throughout the course of this study.

Footnotes

1/See, for example, "The Functional Characteristics of Geographic Information Systems," by K. C. Clarke, Contract NCAZ-OR305-201, NASA/Ames Research Center, Moffett Field, CA, 1983; and "Introduction to Computerized Land-Information Systems," by J. E. de Steiguer and R. H. Giles, Jr., Journal of Forestry, Vol. 79, No. 11, November 1981.

2/Please note that this does not constitute an endorsement of these software packages. On the basis of our study, however, they appear to be packages that would be appropriate for use by rural counties for agricultural applications.

3/Please note that our software survey did not request costs of host computers because prices for the same computer/model may vary tremendously, especially where discounts are available through large-scale purchasing agreements (e.g., state contracts).

4/Annotated from GIS Product Brief provided by Aeronca Electronics, Inc., Charlotte, North Carolina (Undated).

5/Annotated from Summary Information Abstracted from "Statement of Qualifications" provided by Autometric, Inc., Fort Collins, Colorado (Undated).

6/Annotated from Product Brief for ERDAS-PC provided by Earth Resources Data Analysis Systems, Atlanta, Georgia.

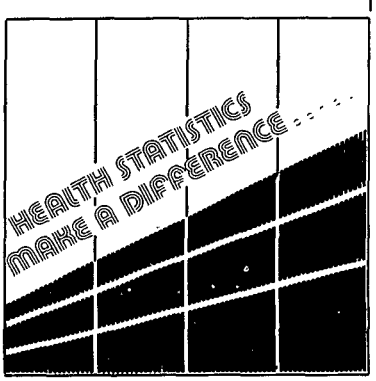
New Investments in Localized Health Statistics

Robert Lewis
School of Public Health
University of South Carolina

(Not available for publication)

Session T

Methodology for the Analysis of Infant and Fetal Mortality



USING THE FETAL LIFE TABLE IN ENVIRONMENTAL EPIDEMIOLOGY

Marilyn K. Goldhaber, Kaiser Permanente Medical Care Program
Oakland, California

INTRODUCTION

As more attention is directed toward environmental issues, the environmental epidemiologist is increasingly faced with the problem of providing well thought-out answers to public concerns. Did a suspected environmental hazard have an adverse health effect on an exposed population? If so, how can we measure it? The epidemiologist may take several approaches to answering these questions. One important approach is to study fetal survival following exposure to the hazard.

Investigating fetal survival is efficient and practical for the environmental epidemiologist. Results can be obtained in a relatively short time, as little as nine to ten months after exposure, and usually with a fair degree of objectivity. Because fetal cells multiply and differentiate rapidly, the fetus is often more susceptible to environmental insults and thus provides a more sensitive indicator of biological damage than does the infant, child or adult.

For these reasons, we find it desirable to study fetal survival or, rather, its complement, fetal loss. This is usually measured by dividing the number of spontaneous fetal losses by the total number of pregnancies at risk (excluding those ending in voluntary induced abortion).

The resulting proportion, however, may be an inadequate expression of the incidence of fetal loss since it does not take into account gestational age. The risk of fetal loss decreases dramatically with increasing gestational age. Thus, the distribution of gestational ages at first observation must be considered when comparing one population to another. A population of pregnant women followed, on the average, from the 10th week of gestation, for example, will have a higher proportion of fetal loss than one followed from the 15th week of gestation.

There is a tool available to the epidemiologist which can overcome differences in gestational age distributions among populations being compared. This is the fetal life table. The fetal life table provides an "adjusted" incidence of fetal loss as a function of gestational age, using gestational-age-specific rates determined empirically. The first question to consider when constructing a fetal life table is: from what point in pregnancy do we wish to evaluate fetal loss?

DEFINING THE EARLIEST POINT IN THE LIFE TABLE

It is currently nearly impossible to diagnose pregnancy before 10 days after conception and certainly impractical for the great majority of women before they miss a menstrual period - approximately two weeks after conception. It has been conjectured that 50% or more of conceptuses fail to survive these first two weeks[1,2,3]. Most of these pregnancies go unnoticed by the women themselves and do not contribute to estimates of the incidence of fetal loss. The earliest time that it is practical to identify pregnancy and begin to follow it up is at the time that the woman has missed her first menstrual period, approximately 14 days after conception, or 28 days from the first day of the last menstrual period. By convention, we denote gestational age from the first day of the last menstrual period. Thus, a practical statement of the research question would be: what is the expected incidence of spontaneous fetal loss from the beginning of the fifth week of gestation onward?

This definition of the research question has been used in nearly all fetal life table studies published to date[4,5,6,7,8]. It is dictated by practicality more than by anything else. The starting point of observation could be set at any gestational age at which we desire to evaluate the subsequent risk of fetal loss. In other words: given that pregnancy is at a certain gestational age (the beginning of the fifth week of gestation, for example), what is the subsequent probability of fetal loss?

DESCRIPTION OF THE LIFE TABLE METHOD

The first step in constructing a fetal life table is to partition pregnancy into intervals that are small enough to assume a constant risk throughout, usually one-week periods. The next step is to determine the rate of fetal loss during each interval, usually by dividing the number of fetal losses during a given week by the number of woman-days (pregnancy days) at risk during that week. Finally, these gestational-age-specific rates are applied iteratively over the intervals to a hypothetical population where all members are under observation from a specified gestational age (usually the beginning of the fifth week of gestation). The proportion of pregnancies ending in fetal loss in the hypothetical population is the estimated incidence of fetal loss in the study population, adjusted by the life table method. A paper by Taylor describes the fetal life table procedure[5].

To appreciate better the need for the life table method it is instructive to see how fetal loss varies with gestational age in a reference population.

FETAL LOSS IN A REFERENCE POPULATION THE KAISER PERMANENTE STUDY

There have been several large prospective studies

of the baseline incidence of spontaneous fetal loss[4,5,6,7,8,9]. I will use the most recent study to illustrate the underlying risk of spontaneous fetal loss during pregnancy.

In the mid-1970's, a study was conducted by Harlap *et al* among Kaiser Permanente Medical Care Program members in Northern California[8]. A total of 32,367 pregnant women were recruited into the study at their first prenatal visit. The women were followed through the end of the 27th week of pregnancy in order to calculate the incidence of miscarriage before that time. The probability of miscarriage before the 28th week, given that pregnancy had survived to the fifth week of gestation, was calculated as 14.4% by the life table method. Estimates of conditional probabilities of miscarriage over the full term of pregnancy, given that pregnancy had survived until the beginning of designated gestational weeks, are shown in *Table 1*. Harlap estimated from previous studies[4,5,6] (which were in good agreement as to the degree of fetal loss after 27 weeks) that about 2.0 out of 100 pregnancies that survived to the fifth week of gestation would end in fetal loss after 28 weeks. This has been incorporated into the figures in *Table 1*.

Table 1: Estimated Probability of Spontaneous Fetal Loss given survival to the beginning of Designated Week of Pregnancy.

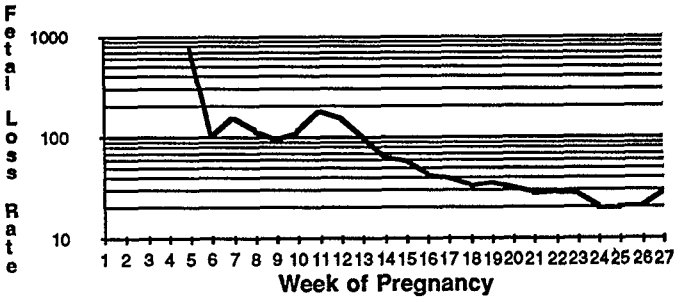
Beginning Week of Pregnancy	Estimated Probability of Fetal Loss
28th	.023
--	--
20th	.039
--	--
13th	.066
--	--
9th	.101
8th	.108
7th	.118
6th	.127
5th	.164

1. From Harlap et al, 1980 .
2. Probability estimates from Harlap et al, 1980, were adjusted upwards assuming 2.0 per 100 pregnancies which survive to the 5th week gestation would end in fetal loss after 27 weeks. Data from Harlap et al cover only up to the 27th week.

Table 1 shows that the probability of fetal loss drops dramatically with advancing gestational age. This occurs for two reasons. First, the later a pregnancy is identified, the shorter the amount of time to observe a fetal death. More importantly, the instantaneous risk of fetal death (the hazard function) is much higher in the early weeks of gestation as illustrated in *Figure 1*.

If we were to compare, without the benefit of fetal life table methodology, two communities with identical underlying risks of fetal loss but the pregnant women in one community were identified, on the average, earlier in gestation than in the other community, we might find dramatic differences in the measured proportion of fetal losses. When doing epidemiological investigations, we try to avoid this problem by assuring that pregnant women are sampled in exactly the same way in both study and control communities. Nevertheless, statistics

Figure 1: Average Daily Fetal Loss Rate Per 100,000 Pregnancies by Gestational Week



Data from Harlap et al, 1980.

generated in even the best situations can appear misleading. Often, pregnancies are identified on the average quite a bit later than the fifth week of gestation in both communities. The resulting proportions of pregnancies ending in fetal loss in both communities may then be quite a bit lower than the expected probability of spontaneous fetal loss from the beginning of the fifth week of gestation. This, at first glance, can be confusing to the investigator who is expecting about 16% of pregnancies under study to end in spontaneous fetal loss.

The example below illustrates how situations like that described above can be handled with and without the aid of the fetal life table.

EXAMPLE 1: WELL WATER CONTAMINATION

A study was recently conducted by the California Department of Health Services in response to well water contamination in Santa Clara County [10]. This was known as the "Fairchild Study" which involved an electronics plant, the Fairchild Camera Company whose underground waste solvent storage tank leaked toxic chemicals, including trichloroethane, into ground water leading to a community well. It was not known how long the leak had been occurring prior to its discovery in December, 1981. The Health Department responded to the incident by conducting a door-to-door census of persons living in the area served by the well and a nearby control area served by a separate water company. During the census, women who had been pregnant at any time during an estimated two-year exposure period from January 1, 1980 to December 30, 1981 were identified. In a follow-up telephone survey of the women with uterine pregnancies (ectopics were excluded), 228 women in the study area and 274 women in the control area responded with information about their

pregnancies.

As a matter of design, the analysis was restricted to only those women who had *conceived* uterine pregnancies during the two-year study period, thus excluding more advanced pregnancies, most of which were clearly beyond the period at greatest risk for spontaneous fetal loss at the time of first observation in January, 1980. With the restricted dataset, there were 41 spontaneous fetal losses and 4 elective abortions out of 191 pregnancies conceived in the study area during the study period, and 23 spontaneous fetal losses and 5 elective abortions out of 210 pregnancies conceived in the control area. The proportion of pregnancies ending in spontaneous fetal loss among pregnancies at risk (i.e. excluding pregnancies ending in elective abortion) was 21.9% in the study area and 11.2% in the control area, yielding a statistically significant relative risk of 2 ($p<.01$).

If the more advanced pregnancies had been included, the incidence figures in both areas would have appeared "watered down". Although there is nothing incorrect about such figures, they would have appeared misleading. If one were expecting a baseline incidence similar to that in the reference population cited earlier, the control area incidence might appear low and the study area incidence more normal. A serious environmental problem might thus go unrecognized if the control area incidence is perceived to be aberrant instead of the study area incidence.

Although no life table adjustment was done in this example, the importance of gestational age at first observation was clearly recognized by the investigators.

A life table analysis would not have corrected for the possibility of early miscarriages going undetected (by the women themselves), as in the Kaiser Permanente study[8]. However, it would have allowed for the inclusion of the advanced pregnancies without "watering down" the incidence of fetal loss. The suggested method is to enter all pregnancies conceived *during* the study period into the earliest gestational interval of the life table. Each pregnancy conceived *before* the study period is then entered into the life table in the interval corresponding to the gestational age of the fetus on the first day of the study period (January 1, 1980, in the well water example).

EXAMPLE 2: MALATHION SPRAYING

The next example shows how the fetal life table can be used to determine whether methods of identification and follow-up of pregnant women are yielding reasonable numbers of fetal losses. This can help assure investigators that their research methods are sound.

Using a multiple decrement method described previously[11], I recently had the opportunity to participate in a life table analysis in a preview of data which were collected to study the association of malathion spraying in Northern California with pregnancy outcome. Malathion, a relatively mild pesticide, was sprayed during the one-year period , July 1, 1981 through June 30, 1982, in an attempt to erradicate the Mediterranean Fruit Fly. The California Department of Health Services enlisted the cooperation of three large obstetrics and gynecology clinics in and around the sprayed areas to provide data on pregnant women whose pregnancies did not end in voluntary induced abortions. A total of 7,830 women, involving many unexposed to malathion as well as exposed, were identified at the time of their first pregnancy test or pregnancy confirmation visit during the one-year spray period. All but 350 were followed until the termination of their pregnancies. Those whose pregnancies were known to end in miscarriage, stillbirth or live birth of an infant with congenital anomalies or low birthweight for gestational age, and an equal number of control women with pregnancies ending in healthy live birth, were solicited for participation in a case-control study to determine their malathion exposures*.

The investigators were interested in assessing whether the number of fetal losses identified in the cohort was reasonable. Of the 7,830 women, 7,213 were entered into a life table analysis. (Women who were first seen at the clinics after their outcome date, or within three days of their outcome date, were excluded. This was to eliminate the bias caused when pregnancies come under observation because of threatened miscarriage - *see section at end of this paper* "PITFALLS OF THE FETAL LIFE TABLE METHOD". Also, women for whom gestational age at entry or outcome could not be determined or reasonably estimated were excluded.) Of the 7,213 women, 518 had spontaneous fetal losses of which 23 were due to ectopic pregnancy.

Using the fetal life table method, the incidence of spontaneous fetal loss from the fifth week of gestation was estimated as 15.4% including ectopic pregnancies and 14.1% excluding ectopic pregnancies. Although we cannot make any claims about the effect of malathion from these figures since both exposed and unexposed women are included, we can see that the incidences are within reasonable limits. Taylor[5], for example, demonstrated an incidence of 18.6% in 1964 versus 16.4% demonstrated by Harlap[8] a little more than a decade later. Eariler life table studies[4,6,7] showed even higher incidences of spontaneous fetal loss, due at least partially to the unintentional inclusion of voluntary abortions. Because voluntary abortions were illegal at the time, they were unlikely to be properly identified in the studies.

Notice that the proportion of the at-risk pregnancies (excluding ectopics) which ended in spontaneous fetal loss was only 6.9% (495/7190), as compared to the life table incidence of 14.1%. Whereas the 6.9% figure has little meaning, the life table incidence can be compared to figures from reference

*Data on malathion exposure, collected from women participating in the case-control study, are currently being analyzed by investigators at the University of Southern California, School of Medicine. A report on the association of malathion exposure and pregnancy outcome will be forthcoming.

populations.

EXAMPLE 3: THREE MILE ISLAND

The final example, probably the most interesting and illustrative of the life table technique, is that of the accident at the Three Mile Island nuclear power plant near Harrisburg, Pennsylvania, March 28, 1979[11]. Within three months after the discharge of radioactive materials, the Pennsylvania Department of Health, in conjunction with the Centers for Disease Control and the U.S. Bureau of the Census, conducted a census of all persons living within five miles of the plant. No control population was similarly censused. Each woman was asked whether she was pregnant during the 10-day exposure period.

A total of 479 pregnancies were identified of which 28 ended in spontaneous fetal losses, 2 in ectopic pregnancies, 13 in voluntary induced abortions and 436 in live births. In order to determine whether the incidence of spontaneous fetal loss was greater than expected, each woman was entered into a multiple decrement life table analysis at the exact gestational age (in days) of her pregnancy on March 28, 1979. Using only pregnancy experience after the Three Mile Island accident, the adjusted incidence of spontaneous fetal loss from the fifth week of gestation was found to be 16.1%, similar to or lower than incidences in four reference populations[4,5,6,8].

From the number of woman-days of observation within each gestational month, expected numbers of fetal losses were calculated using rates from the four reference studies. A small number of losses among the Three Mile Island women after 20 weeks of gestation (1 observed, 5.9 - 6.5 expected) was balanced by a clustering in the 13-16 week period (13 observed, 2.9 - 7.2 expected). It is unknown whether the age-at-death distribution was related to the nuclear accident. Such a distribution had not been suggested *a priori*. It could be that some agent such as stress or emotional trauma from the nuclear accident caused "doomed" fetusus, those that would have aborted eventually, to abort earlier. This is a hypothesis that merits further investigation in other studies of emotional stress and pregnancy outcome.

The investigators concluded that no overall excess in the number of fetal losses had occurred at Three Mile Island. Studying the total population, rather than just a sample, strengthened this conclusion.

FUTURE RESEARCH

When using the fetal life table method in environmental epidemiology, gestational age at first exposure is an important part of the model. In a more sophisticated model, gestational age at first exposure could be replaced by an exposure index for each pregnant woman (a function of chronological time and space, for example) and a fetal susceptibility index (a function of gestational age)[14]. A technique such as Cox's regression (which also takes into account important covariates such as maternal age, race, parity, etc.) could test for an effect of an environmental exposure that varied within and among individuals whose susceptibility varied with gestational age. Although this subject is beyond the scope of this paper, it is an important area for future research in environmental epidemiology.

PITFALLS OF THE FETAL LIFE TABLE METHODOLOGY

In ideal situations, the fetal life table method completely eliminates the problem of differences in gestational age distributions among populations being compared prospectively. However, for the life table to function perfectly, two assumptions must be met[12]: (1) the intervals must be small enough to assume a constant risk throughout, and (2) entry and exit (i.e. censorship) must be independent of outcome events. The second assumption is particularly troublesome because women with problem pregnancies (or higher exposures) tend to come under observation earlier than women experiencing no problem (or lower exposures). Previous life table studies have tried to remedy this situation by eliminating any pregnancy with symptoms of threatened abortion at the time of entry. It remains a challenge to correct the bias without forcing it in the opposite direction. Unfortunately, the life table is extremely vulnerable to distortions in the early intervals, where the above assumptions are most likely to be violated[5,13].

Another problem with the life table approach is communicating to the public the necessity of making the complicated manipulations that are required. As demonstrated in the well water contamination example, gestational age at entry can be considered in the analysis without using a life table. However, we can think of life table manipulations as somewhat analogous to age adjustments when comparing the incidence of diseases, such as cancer, between populations. Although the public may lag behind in comprehending new or unfamiliar methodologies, reproductive and environmental epidemiologists will find it valuable to understand the fetal life table concept and apply it when appropriate.

SUMMARY

In summary, when investigating the incidence of spontaneous fetal loss, it is essential to consider gestational age distributions of populations being compared. This can be done (1) by assuring that the populations were sampled in the same way relative to gestational age, or (2) by adjusting incidences in populations being compared by the fetal life table method.

REFERENCES

1. Hertig AT, Rock J, Adams EC, Menkin MC: *Thirty-four fertilized human ova, good, bad and indifferent, recovered from 210 women of known fertility.* Pediatrics 1959; 23:202-211.

2. James WH: *The incidence of spontaneous abortion.* Population Studies *1970;* 24:241-245.

3. Roberts CJ, Lowe CR: *Where have all the conceptions gone?* Lancet 1975; i:498-499.

4. French FE, Bierman JM: *Probabilities of fetal mortality.* Public Health Reports 1962; 77:835-847.

5. Taylor WF: *On the methodology of measuring the probability of fetal death in a prospective study.* Human Biology 1964; 36:86-103.

6. Shapiro S, Levine HS, Abramowicz M: *Factors associated with early and later fetal loss.* Advances in Planned Parenthood 1971; 6:45-63.

7. Erhardt CL: *Pregnancy losses in New York City, 1960.* American Journal of Public Health 1963; Sept:1337-1352.

8. Harlap S, Shiono PH, Ramcharan S: *A life table of spontaneous abortions and the effects of age, parity, and other variables.* IN: Hook EB, Porter I (eds): Reproductive Loss. New York: NY Academy Press, 1980;145-158.

9. Wilcox AJ, Treloar AE, Sandler DP: *Spontaneous abortion over time: comparing occurrence in two cohorts of women a generation apart..* American Journal of Epidemiology 1981 114:548-553.

10. California Deparment of Health Services: *Pregnancy outcomes in Santa Clara County 1980-1982.* .Reports of two epidemiological studies. January 1985.

11. Goldhaber MK, Staub SL, Tokuhata GK: *Spontaneous abortions after the Three Mile Island nuclear accident: a life table analysis..* American Journal of Public Health 1983; 73:752-759.

12. Abramson FD: *Spontaneous fetal death in man..* Social Biology 1973; 20:375 403.

13. Leridon H:*Intrauterine mortality.* IN: Human Fertility: the Basic Components. University of Chicago Press 1977; 48-81.

14. Goldhaber MK, Staub SL, Tokuhata GK: *Re: spontaneous abortions after the Three Mile Island nuclear accident,* Goldhaber et al, respond (letter). American Journal of Public Health 1984; 74:520.

THE COMPARISON OF INFANT MORTALITY RATES WHEN BIRTHWEIGHT DISTRIBUTION DIFFERS

Andrew M. Friede, Stan Becker, Phillip H. Rhodes, Centers for Disease Control

INTRODUCTION

Local and state health departments and federal agencies are frequently called upon to compare infant mortality rates (IMR), defined as the number of live-born infants who die during days 0-365 of life per 1,000 live births; although a risk, it is by convention called a rate. Some of these comparisons focus on the identification of those local areas where the IMR is anomalously elevated, or even rising.[1,2] Alternately, these analyses are designed to follow trends over time in a single area's IMR, and to suggest possible reasons for improvement, or lack thereof. Program planners and public health officials often need reliable summaries of these trends and their components.

A crude IMR can be thought of as having two components: the first is the birthweight distribution, and the second is the IMR's of a series of birthweight strata, called birthweight-specific IMR's. This is directly analogous to a population's crude mortality rate, which is determined both by the distribution of ages within the population and by the mortality rates of these different age groups. The comparison of crude mortality rates is a standard problem in demographic analysis. However, demographers usually view age as a "nuisance" variable, or a confounder of the comparison, and they seek to eliminate its effect, not to estimate it. By contrast, the contribution of low birthweight to differences in IMR's is of intrinsic programmatic interest. Hence, in addition to controlling the effect of different birthweight distributions in order to facilitate the comparison of IMR's, one would like to be able to estimate the contribution of birthweight to unequal IMR's.

The principal objective of this paper is to investigate the statistical behavior of several techniques which summarize the comparison of IMR's, and which analyze the components of these trends. Within this context, we make recommendations as to the choice of method(s). First, we compare the results of four different techniques which summarize the comparison of IMR's when the comparison is confounded by unequal birthweight distributions. Second, we compare the results of two different techniques that estimate the relative contributions of differences in birthweight distributions and birthweight-specific mortality rates in the comparison of unequal IMR's. We used data from Massachusetts for the purposes of illustration, as it was the problem of analyzing the decline in the infant mortality rate in Massachusetts that prompted the present inquiry.

METHODS

Data Sources and Population

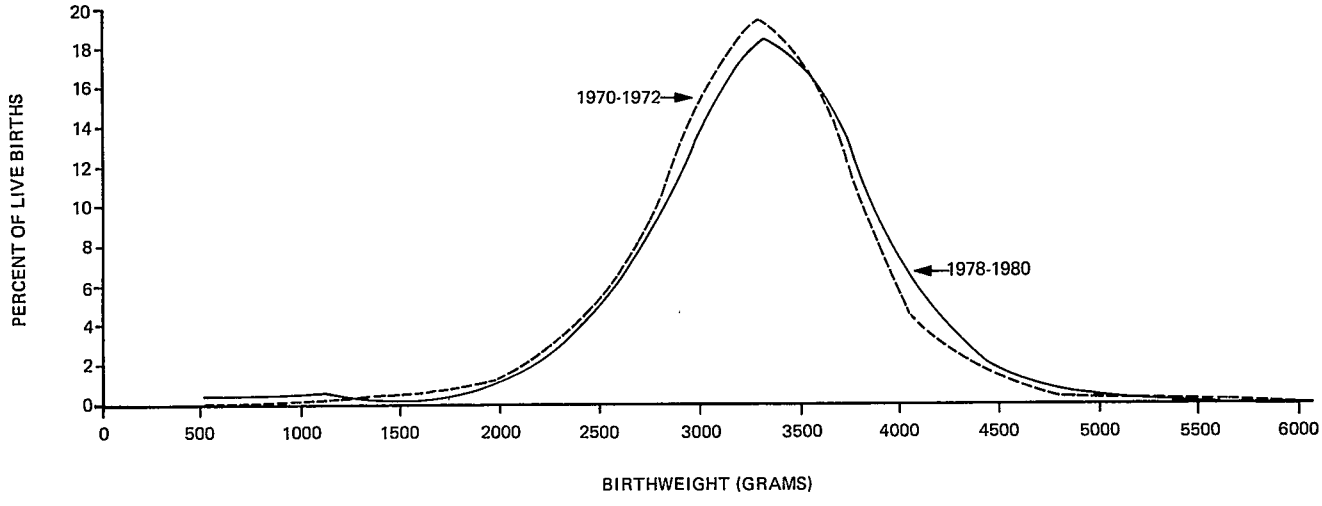
Two sources of computerized vital records were provided by the Massachusetts Department of Public Health. The file of births was used to obtain the distribution of birthweights of Massachusetts resident live births. The file of linked birth-infant death certificates was used to obtain the birthweight of each infant who died while still a resident of Massachusetts. (Copies of death certificates for infants who change their legal residence before the first birthday have been furnished to the State only since 1976. To avoid any non-comparability across time, the data were restricted to Massachusetts resident newborns who also died as residents.)

We compared the mortality experience of two birth cohorts: 1970-1972 and 1978-1980. Infants of birthweight 500-5999 grams were included in these analyses. The probability of an infant with a birthweight of less than 500 grams being considered alive at birth depends on delivery room practices, which vary greatly from region to region, and across time. Hence, hospitals and regions which intensively manage newborns who are barely alive may have paradoxically higher IMR's; to improve comparability, infants under 500 grams were excluded. Because infants with registered birthweights of 6000 grams or more (over 13 pounds) are relatively likely to have had their birthweights improperly recorded, they too were excluded.

Statistical Analysis

Relative Risks. Four methods were used to estimate a summary relative risk (RR), defined as the ratio of the risk of infant death in 1978-1980 to the risk of infant death in 1970-1972. Direct and indirect standardization were performed as described by Fleiss.[3] For the direct standardization, the birthweight distribution of 1970-1972 was used as the standard. For the indirect standardization, the schedule of stratum-specific IMR's for 1970-1972 was used as the standard. To maximize the control of the confounding effect of birthweight, eight birthweight strata were used throughout.[4] The Miettinen summary RR, which uses a maximum likelihood method with weights inversely proportional to the variance of each stratum, was calculated by the methods of Rothman and Boice.[5] The RR was also estimated by modeling the logarithm of the risks as a linear combination of birth cohort and birthweight stratum, using maximum likelihood methods.[6] All the above methods will produce an unbiased summary RR if the RR

BIRTHWEIGHT DISTRIBUTION
MASSACHUSETTS RESIDENT LIVE BIRTHS
1970-1972 AND 1978-1980



is constant across strata, i.e., if there is no effect modification, (interaction) between the RR and stratum number.

Partitioning. The difference between two crude infant mortality rates can be decomposed into two components: [7,8]

$$IMR_1 - IMR_2 =$$

$$((BW_{i1}+BW_{i2})/2)\ (MR_{i1}-MR_{i2}) +$$

$$((MR_{i1}+MR_{i2})/2)\ (BW_{i1}-BW_{i2}),$$

where IMR_1 and IMR_2 are the crude infant mortality rates in period 1 and period 2, BW_{i1} and BW_{i2} are the proportions of births in each of i strata (i.e., $\Sigma\ BW_{i1} = \Sigma BW_{i2} = 1$), and MR_{i1} and MR_{i2} are the stratum-specific mortality risks for each of i strata. The first component is the fraction of the crude difference that can be attributed to changes in the mortality risks; the second component is the fraction due to changes in the birthweight distribution. The two components sum to the crude difference. Note that this formulation applies the averages of the birthweight and mortality risks in each stratum to changes in the other. This assures that the value of each component is derived from a hypothetical mid-point during the time period under consideration, i.e., as if the changes took place continuously and concurrently.[7] This method may be most appropriate for the decomposition of changes in a single population at two points in time, as contrated with the study of differences in two distinct populations at one point in time. The ratio of two crude infant mortality rates can also be decomposed into two components:[9,10]

$$IMR_2/IMR_1 = RR_{crude} = RR_{mr} \times RR_{bw},$$

where IMR_1 and IMR_2 are the crude infant mortality rates in period 1 and period 2, RR_{mr} is that part of the decline in the crude RR that can be attributed to a change in the risk of mortality, and RR_{bw} is that part of the crude RR that can attribute to a shift in the birthweight distribution. This formula assumes that the RR_{mr} is constant across strata.

RESULTS

There was a small, symmetric upward (rightward) shift of the birthweight distribution in Massachusetts during the 1970's (Figure). However, because of the very strong association between birthweight and the risk of mortality, any comparison of the IMR's for these two time periods needs to take even this small shift into account.

The number of deaths and the associated birthweight-specific IMR for each of the two birth cohorts are given in Table 1. There were

a relatively large number of deaths in each stratum, with the possible exception of the last. This should allow a reasonably stable estimate of the RR or RD (rate difference) for each stratum. Note the very strong association between birthweight and stratum-specific IMR; the IMR's vary by a factor of close to 400.

Table 1. Infant Mortality Rates By Birthweight: Massachusetts, 1970-1972 and 1978-1980

	1970-1972		1978-1980	
Birthweight(g)	Deaths	IMR	Deaths	IMR
500- 999	846	852.8	569	737.0
1000-1499	587	437.4	209	195.9
1500-1999	504	147.9	150	63.8
2000-2499	411	33.4	188	22.3
2500-2999	478	10.0	210	6.4
3000-4499	865	4.7	451	2.8
4500-5999	16	4.9	10	2.5
Total	3707	14.6	1787	8.6

The birthweight-specific RR's and RD's are given in Table 2. The RR's were not constant across the strata. However, they were all less than one; i.e., there was no "crossover." Because the RR's were not truly constant, any summary RR's will not exactly represent the effect in every stratum. However, its general direction (less than one versus greater than one) will be correct for each stratum. In contrast to the moderate degree of constancy of the RR's, the RD varied very widely. This was a direct result of the wide variation in birthweight-specific IMR's. Hence, from the point of view of estimability, the RD's was a relatively poor choice of a parameter to employ in any summary of the mortality experience across time. No further attempts were made to summarize the RD.

Table 2. Risk Ratio and Risk Difference by Birthweight: Massachusetts, 1970-1972 and 1978-1980

Birthweight(g)	Risk Ratio	Risk Difference
500- 999	0.9	-116
1000-1499	0.4	-241
1500-1999	0.4	-84
2000-2499	0.4	-11
2500-2999	0.6	-3.6
3000-4499	0.6	-1.9
4500-5999	0.5	-2.4
Total	0.6	-6.0

Relative Risks

The four different methods used to summarize the RR produced similar estimates (Table 3). In the modeling of the log of the risk, a statistically significant interaction ($p<0.01$) (equivalently, effect modification) was noted between cohort and birthweight stratum, i.e., the RR was not constant across strata. The fact that these four different methods of summarization gave very similar results suggested that this particular pattern of effect modification affected these methods similarly.

Table 3. Estimates of Summary Relative Risk (RR) Obtained by Various Methods

Method	RR
Crude	
Direct Standardization	0.62
Indirect Standardization	0.63
Miettinen Summary	0.63
Model Log of Risk	0.70

Partitioning

The partitioning of the crude difference of -6.0 (8.6 minus 14.6) yielded a component due to a change in the birthweight-specific rates of -5.2, and a component due to a change in birthweight distribution of -0.8. The results of the partitioning of the crude risk ratio of 0.59, using the component due to the change in the rates of 0.63, yielded a component due to a change in birthweight of 0.94. The results of each analysis suggested that almost all the improvement in Massachusetts' IMR was due to an improvement in stratum-specific infant mortality rates; improvements in the birthweight distribution played a relatively minor role.

DISCUSSION

In an analysis of changes in the infant mortality rate in Massachusetts during the 1970's, we found that similar estimates of a summary RR were obtained from four different methods. The obtained RR's ranged from 0.62 - 0.70. This example had effect modification; our methods seemed to be similarly affected. We also found that two different methods gave similar estimates of the relative contributions of changes in the birthweight-specific IMR's and shifts in the birthweight distribution. Almost all of the change in the infant mortality rate in Massachusetts during the 1970's was attributable to improvements in the birthweight-specific IMR's; only a small part was due to the upward shift in the birthweight distribution.

We are unaware of a similar comparison of the effect on the choice of different analytical techniques in estimates of the RR summary or on the relative role of birthweight-specific IMR's and birthweight. However, our results of the rates of decline in mortality, and the relatively small role played by improved birthweight, are comparable to those in

California from 1970-1977 [11], and in the United States as a whole from 1950-1975.[12] With the exception of the modeling of the log of the risks, none of the methods requires a computer; hence, there are no important barriers to the verification of these results via replication in other states.

The choice of technique depends not only on ease of computation, but on availability of data, estimability of parameter, and the parameter's intrinsic interest.[13,14] Direct and indirect standardization will provide equivalent estimates of RR if the standard chosen for indirect standardization is derived from one population or the other, or a hypothetical population between the two.[15] However, when the data are sparse, indirect standardization may provide more meaningful estimates than direct standardization.[2] Finally, in the analysis of infant mortality, risk differences are unlikely to be easily summarized, although they may be of intrinsic interest when considered at the stratum-specific level.

We would like to emphasize that summarization of the RR is only intended to complement a critical examination of stratum-specific effects. This is especially important to bear in mind when there is effect modification, as there was in these data. On the other hand, program planners often need summaries for public purposes. In the case where there is no "crossover," although the exact size of the estimated RR may not reflect the effect in every stratum, its overall direction will be true to the data. If there is crossover, summarization of the RR is not appropriate.[16]

One more caveat should be noted. All these analyses assume that the birthweight distribution and the birthweight-specific IMR's vary independently. However, medical and programmatic interventions may cause them to be linked. For example, a campaign to provide improved access to care to high-risk mothers may improve both the birthweights and birthweight-specific survival. Conversely, a successful stop-smoking campaign may raise birthweights, but because babies in the new birthweight distribution at a given birthweight may be of lower gestational age, the program may result in higher birthweight-specific mortality risks for some strata. The possible linkage of the birthweight distribution to patterns of mortality should be considered in any analysis of patterns of infant mortality.

In summary, most of the decline in infant mortality in Massachusetts during the 1970's was due to a lower risk of mortality at all birthweights. Further reductions in IMR will benefit from improvements in the birthweight distribution.[17] Summaries of the RR of infant mortality, if interpreted with care, may provide health planners with useful information. However, summaries should not be substituted for a close examination of the stratum-specific effects.

ACKNOWLEDGEMENT

As part of ongoing collaborative research, the data were made available by the Massachusetts Department of Public Health.

REFERENCES

1. Foster JE, Kleinman JC. Changes in infant mortality and related rates by health service area: 1969-73 and 1974-77. Statistical Notes for Health Planners 1981; 13:1-32. Hyattsville, MD: Government Printing Office, 1981. (DHHS publication no. (PHS) 81-1237.)

2. Kleinman JC. Re: "Perinatal mortality: standardizing for birthweight is biased" (Letter). Am J Epidemiol 1984; 120:798-801.

3. Fleiss JL. Statistical methods for rates and proportions. New York: John Wiley & Sons, 1981:237-255.

4. Cochran WG. The effectiveness of adjustment by subclassification in removing bias in observational studies. Biometrics 1968; 24:295-313.

5. Rothman KJ, Boice JD. Epidemiologic analysis with a programmable calculator. Boston: Epidemiology Resources, 1982: 11-17.

6. Wacholder S. Binomial regression in GLIM: estimating the risk ratios and risk differences (abstract). Biometrics 1985; 41:582.

7. Berman SM. Movable partitions: methodological problems in assessing improvements in neonatal mortality. Am J Epidemiol 1985; 122:354-355.

8. Kitagawa EM. Components of a difference between two rates. J Am Stat Assoc 1955; 50:1168-1194.

9. Miettinen OS. Components of the crude risk ratio. Am J Epidemiol 1972; 96: 168-172.

10. Miettinen OS. Standardization of risk ratios. Am J Epidemiol 1972; 96: 383-388.

11. Williams RL, Chen PM. Identifying the sources of the recent decline in perinatal mortality rates in California. N Engl J Med 1982; 306:207-214.

12. Lee KS, Paneth N, Gartner LM, Pearlman MA, Gruss L. Neonatal mortality: an analysis of the recent improvement in the United State. Am J Public Health 1980; 70:15-21.

13. Roderick JAL, Pullum TW. The general linear model and direct standardization--a comparison. Sociological Methods & Research 1979; 7:475-501.

14. Breslow NE, Day NE. Indirect standardization and multiplicative models for rates, with reference to the age adjustment of cancer incidence and relative frequency data. J Chron Dis 1985; 28:289-303.

15. Anderson S, et al. Statistical methods for comparative studies. New York: John Wiley & Sons, 1980:113-139.

16. Wilcox AJ, Russell IT. Perinatal mortality: standardizing for birthweight is biased. Am J Epidemiol 1983; 118:857-864.

17. Guyer B, Wallach LA, Rosen SL. Birth-weight-standardized neonatal mortality rates and the prevention of low birth weight: how does Massachusetts compare with Sweden? N Engl J Med 1982; 306:1230-1233.

PREDICTING PATTERNS OF INFANT MORTALITY: MULTIPLE REGRESSION & LOGISTIC ANALYTIC MODELS

Yahya Daoud, Patricia Mullan Scalzi and Stephen Blount, Detroit Health Department, Detroit, Michigan

1. Introduction

McQueen and Siegrist's critique of existing research of social factors in the etiology of disease points out that such multivariate approaches remain "rather uncommon" in traditional social epidemiology. A more traditional approach to the study of social factors implicated in infant mortality, for example, examines the differential experience of infant mortality among groups representing different age and racial categories. Cox and Mackay caution that the analysis of mortality data in terms of the contribution of different risk factors must consider that "few of the risk factors are well-defined and truly independent and their combination is more likely to be described by an interactive rather than an additive model. Smoking, occupational and social class provide a good example of confounded factors: smoking among men is more prevalent in "blue-collar" workers than in professional and managerial classes" (Cox and Mackay, 1982: 382). McQueen and Siegrist (1982) criticize the "weakness of traditional epidemiological approach ... often in the form of tables which are bivariate with one variable controlled. Often the data are "adjusted" for variables which are conceived as simple but in reality are far more complex. Related to this bivariate correlational approach has been a failure to consider causal relationships among multiple variables". (McQueen and Siegrist, 1982: 353).

This paper presents an approach to the study of infant mortality which considers the relative contribution of coexisting characteristics of infants and their parents in producing observed patterns of reproductive outcome. This paper focuses on multiple regression and logistic analysis in the study of natal health outcomes. In this paper, the primary health outcome is the category of Infant Status at one year, which are "survival" or "failure to survive" (death). Infant "survivors" are defined as those infants born during a cohort year who survived during their first year. The common premise of both the multiple regression and logistic analytic approach is the determination to empirically test for the presence and impact of co-existing contextual factors associated with the emergence of a well-defined health outcome.

2. Data Base

This paper draws from a series of investigations examining the experience of infant mortality among the cohorts of live births in the urban Detroit community from 1976 through 1981. This section will draw on findings from these studies for the population characteristics of the cohort of infants who were born to mothers who were residents of Detroit in 1981. The analyses have been further restricted to the cases for which the birth certificate includes the information representing the risk factors used in developing the models.

The number of births for each of the major ethnic categories specified within the 1981 birth certificate data set formed the following distribution: 13,182 black, 5,578 white, 47 American Indian, 36 Filipino, 10 Chinese, 4 Japanese, 1 Hawaiian and 98 other Asian children. Plurality, the number of infants simultaneously born to the same mother, substantially determines the health of the individual infants. The overwhelming majority of the Detroit births occurred as single births. In 1981, the occasions of multiple births consisted of 463 sets of twins and 9 sets of triplets.

In terms of the frequency with which infant mortality occurred among the infants born in Detroit in 1981, 285 deaths occurred in the neonatal period. An additional 101 deaths occurred in the postneonatal period. 18,580 of the 18,966 (97.96%) children who were born survived through their first year of life.

Infant mortality rate is defined as the number of deaths occurring among children under one year of age, reported within a given year, divided by the number of live births reported during the same time interval, multiplied by 1,000. In the United States, the infant mortality rate dropped from approximately 100 deaths per 1,000 live births in 1915 (Shapiro, Schlesinger and Nesbitt, 1968), to approximately 13 deaths per 1,000 in 1980 (U.S. DHHS, 1981:102). Michigan's absolute infant mortality rate, as well as the acceleration of the rate of decline in the experience of infant mortality, is statistically concordant with that of the nation. The rate of infant mortality suffered with the City of Detroit, however, represents both a higher absolute rate of infant mortality and a slower rate of decline than the experience of Michigan and the nation. The infant mortality rate in Detroit has changed from 1976's 22.8 to 20.9 in 1980 and 21.9 in 1981. In following the rate of change continuously from 1976, Detroit has experienced only a 3.9% reduction in infant mortality rates. During this same period, the nation as a whole decreased its infant mortality rate by 21.7%.

In 1981, 97.96% of the infants born to Detroit residents survived the first year of life. In 1980, 97.93% of infants had survived to their first birthday compared to 97.75% in 1976.

In 1981, the age at which Detroit mothers gave birth ranged from eleven to forty-six years. While this pattern of maternal age is slightly negatively skewed, median maternal age and quartile distribution remained relatively invariant. In comparing the fertility rate within different age groups in the U.S. to the age-stratified fertility rates of other developed countries, a greater proportion of U.S. births are due to adolescent pregnancies (U.S. DHHS, 1980:19). Adolescent pregnancies represented 20.15% of live births in 1981, 21.17% in 1980 and 25.45% in 1976. While the rate of live births to adolescent mothers in the United States has declined over time, the rate in Detroit has remained at approximately 20%.

Birth weights recorded for infants born in Detroit in 1981 ranged from 142 grams to 7,069 grams. The distribution of birth weight categories among the cohorts of Detroit infants is presented in Table 1. Birth weight and gestational age jointly define the prematurity/immaturity status of the infant. Premature infants are defined in the occasion of a gestational age less than 37 weeks, and a birth weight less than 2500 grams. Among the cohorts of Detroit infants, the examination of the birth weight distributions presented in Table 1 indicates that the occurrence of infants with low birth weight was 11.87% in 1981, 11.01% in 1980 and 11.88% in 1976. Within the state as a whole, only 6.9% of all live births in Michigan fell into the low birth weight category.

Table 1 Percent Distribution of Birth Weight categories Among Cohorts of Detroit Infants

Birth weight in Grams	Year of Birth		
	1981	1980	1976
Less Than 500	0.43	0.38	0.23
500 - 999	0.90	0.86	0.84
1000 - 1499	0.92	1.01	1.04
1500 - 1999	2.06	2.11	2.47
2000 - 2499	7.46	6.65	7.30
2500 - 2999	22.07	22.30	22.40
3000 - 3499	37.46	36.93	37.23
3500 - 3999	21.39	21.88	21.86
4000 - 4499	5.33	5.43	5.37
More than 4499	1.06	1.20	1.14
Unknown	0.93	1.24	0.08

3. Methodology

The first set of analyses sought to estimate the probability of death for each infant using social and medical variables available on the birth and death certificates.

The estimation procedures utilized the multivariate logistic function and the 1981 data set. A maximum likelihood computational procedure was used to select sets of variables which enable this type of function to describe the data well and which could not be improved upon markedly through the addition of the variables.

In the first step infant birth weight was the dominant independent variable. The second step in the analysis was to construct a model using social and medical variables to predict the infant's birth weight.

4. Calculation of the Base-Line Models and Standard Mortality Ratios

The baseline model developed here is a multivariate logistic model which summarizes the combined mortality experience of all infants born in the year 1981 to mothers who were residents of the City of Detroit. The first model enables the probability of death to be calculated for an infant from the infant's birth weight and sex.

The logistic model used (see Cox, 1970) is of the following form:

$$P=P(\text{Death}/X_1,X_2,..,X_p)= \frac{1}{1+\text{Exp},-(B_0+B_1X_1+..B_pX_p).},$$

where $X_1,X_2,...,X_p$ are the p independent variables. X_1 is the infant birth weight, X_2 is the infant sex, B_0 is the intercept term (constant) and B_1, B_2, .., B_p are the coefficients of the p independent variables.

Two measures of how well the model fits the data will be referred to. The first is R^2 which gives the fraction of the variability in the survival or death indicator which is explained by the modeled probabilities of death. An R^2 near 1 indicates that the model is discriminating well between those infants who die and those who survive while an R^2 near 0 indicates that the model does not differentiate between the two groups. The second measure is predictive power which is the average probability which the model gives for the observed outcome for the infants in the study. The model estimates the probability for each infant's outcome after the first year as P or 1 - P, depending upon whether the infant died or survived. The average of the modeled probabilities for the observed infant outcomes within the first year of their life gives the predictive power. A predictive power near 1 implies that the model is consistently assigning high probabilities to observed outcomes, as it should. A predictive power close to 0 would imply that the model was consistently assigning high probability to the wrong outcomes and low probability to the observed outcomes.

The selection of the independent variables for the models was carried out in several stages. The first stage consisted of the development of a model using all the social and medical variables available on the birth certificates. Birth weight and sex have consistently been reported to be the most important factors in the mortality model. The following stages consisted of development of a better model with the important or semi-important variables after deleting the non-significant (p 0.05) variables from the model.

5. Results

In this section we are going to discuss the results of the three different models which had been developed.

MODEL I

Model I was the initial model developed. The model included birth weight and birth weight square as the main contributor risk factors; sex and race are also included in the model. This model was the result of our search within all available suspect risk factors to explain the variation within infant mortality.

Model I showed high significant association between infant mortality and birth weight; the partial correlation between birth weight and infant status after one year is 18.8% after adjusting to birth weight square, sex, and race. Also the same measure between birth weight square and infant mortality is 8.6% after adjusting to the other risk factors in the model. Including the birth weight and the birth weight square in the model showed that the relation between IMR and birth weight is a U shaped relation, which means that the probability of infant death, given low birth weight or very high birth weight, are higher than the probability of infant death given that their weight are around 3,500 grams. The model showed that the ideal birth weight is around 3,500 grams.

The estimates and results of Model I are contained in Table 2. Table 2 contains the risk factors included in the model, the partial correlation, the estimate of the coefficients, their standard error, and the statistics used for testing the hypothesis that no association is between the risk factor and infant status after one year. Table 3 summarizes the adjusted probability of death for each risk factor given that the other risk factors are fixed at their means, and the 95% confidence limits for the adjusted probabilities.

Table 2. Estimates of coefficient: Model I

Risk Factors	Partial Correlation (%)	estimated		Test Statistic
		Parameters	Standard Error	
Constant	11.8	4.152	0.2692	16.02
Weight	18.8	-0.005	0.0003	-21.11
Weight^2	8.6	0.721E-6	0.5356E-7	13.46
Sex				4.24
Male	0.3	0.175	0.0707	
Female	0.3	-0.181	0.0734	
Race				4.24
White	0.0	-0.049	0.1245	-0.40
Black	0.0	0.005	0.0533	0.10
Others	0.2	1.053	0.5158	2.04

Table 3. Estimated Adjusted Probability of Death Using Model I

Risk Factors	P(Death)	95% Confidence Limits	
		Lower	Upper
Overall	0.006	0.005	0.008
Sex			
Male	0.008	0.006	0.009
Female	0.005	0.004	0.007
Race			
White	0.006	0.004	0.008
Black	0.006	0.005	0.008
Others	0.018	0.007	0.048

Examining Table 2 and Table 3 showed that: sex is significant risk factor (X^2=6.10, D.F=1, p=0.02); the estimated adjusted probability of death given male is slightly higher than the estimated adjusted probability of death given female.

Although race showed a slight significant contribution to the model (X^2=4.24, D.F=2 p=0.12), Table 2 showed that the significance is due to other races, not to black and white. The estimated adjusted probability of death given black equal to the estimated adjusted probability of death given white which equal to 0.006, and the adjusted probability of death given other races is equal to 0.018, which lead us to develop Model II.

When hospital was added to the logistic model in addition to the primary factors, it was found to be not significant (p 0.1). Therefore, a hospital term was not included in the model.

For Model I the overall predictive power is 0.95, and the percent of total variation explained is 54.6. We are continuing examining the goodness of fit of this model; similar models will be developed using the 1980 data, and 1976 data, to compare the results for the three years.

MODEL II

This model is similar to Model I: the calculations of the parameters have been restricted to the two racial groups black and white (not including other races).

Birth weight, birth weight square, and sex were the significant risk factors to be included in Model II. Model II held the same relation between infant status after one year and birth weight, birth weight square and sex as in Model I. When race was added to the model in addition to the primary risk factors it was found to be not significant (X^2=0.11 D.F=1 p=0.73). Therefore, a race term was not included in Model II.

For Model II the overall predictive power is 0.952 and the percent of total variation explained is 54.1%.

Table 4 contains the risk factors included in the model, their partial correlations, the estimates of the parameters and their standard error, and the statistics used to test the hypothesis that no association between the risk factor and the infant status after one year. The estimated adjusted probability of death and their associated confidence limits are presented in Table 5.

The two models showed that birth weight is the main risk factor to infant status after one year. This finding leads us to construct Model III, which assumes that the birth weight is the outcome.

Table 4. Estimates of coefficient: Model II

Risk Factors	Partial Correlation (%)	estimated		Test Statistic
		Parameters	Standard Error	
Constant	12.1	4.146	0.2573	16.11
Weight	19.0	-0.0053	0.0003	-21.04
$Weight^2$	8.7	0.7193E-6	0.5375e-7	13.38
Sex				6.21
Male	0.3	0.1783	0.0715	
Female	0.3	-0.1844	0.0740	

Table 5. Estimated Adjusted Probability of Death Using Model II

Risk Factors	P(Death)	95% Confidence Limits	
		Lower	Upper
Overall	0.006	0.005	0.008
Sex			
Male	0.008	0.006	0.009
Female	0.005	0.004	0.007

MODEL III

Given the results of Model I and Model II, the model which has been developed in this section used infant's birth weight as the outcome and the available social/medical information to explain the variation within birth weights.

Table 6 presents the risk factors in the order that they entered into the model, the estimated parameters for each risk factor, estimated standard deviation of the estimated parameters, the fraction of explained variance (R^2), and the contribution to R^2 after adding the variable into the model.

Table 6. Estimates of coefficient and R^2: Model III

Risk factor	Estimated		R^2	Improvement in R^2
	parameters	Standard error		
Constant	-1,509.08	83.81	-	-
Gestational age	63.70	1.32	0.225	-
Apgar score 5 M	125.38	4.45	0.274	0.0490
Plurality			0.297	0.0228
single	668.36	28.05		
Named Parents			0.312	0.0146
Mother only	-49.49	10.46		
Sex			0.324	0.0135
Male	139.51	8.29		
Number of Prenatal Visits	11.05	0.94	0.331	0.0068
Race			0.336	0.0051
White	154.33	39.30		
Black	40.59	39.28		
Previous Children Delivered Now Living	36.22	3.29	0.339	0.0041

The total number of cases used to estimate the parameters is 14,588 after deleting cases with missing information (3,378 cases).

Gestational age was the first variable to enter the model; it explained 22.62% of the variance by itself. At the same time gestational age presented a large amount of missing data (1,464 cases), note that gestational age under 16 weeks and over 52 weeks are treated as missing values.

Apgar score after five minutes was the second variable to add into the model. The estimated parameter for apgar score is 124.69 which indicates that there is a positive relationship between birth weight and apgar score.

The third term entered into the model was plurality. The average birth weight for single birth will be approximately 670 grams more than for multiple births.

Following plurality "named parents" variable was entered into the model. This variable was created to measure the social support mothers were getting from the fathers during the course of pregnancy. This term reflects the presence or absence of certain information related to the father on the birth certificate (e.g. race, education and age). The model estimated that infants with both parents (married or unmarried) on the average weigh 50 grams more than infants born for single mothers.

Sex was entered into the model; the contribution to R^2 is small (0.0135). After adjusting to the other variables, the estimated birth weight for male is bigger than for female.

Number of prenatal visits risk factor was entered into the model with little contribution to R^2 (0.007). The model indicates that there is a positive linear association between birth weight and the number of prenatal visits.

The seventh risk factor included in the model was race. The improvement in R^2 after adding race into Model III was 0.005, very little improvement. Model III estimated that white infants tend to weigh more than black infants and other racial infants. When we restricted the model to the two groups white and black, race was entered into the model at step 7 again, with the same contribution to R^2.

The overall improvement in R^2 was 0.017 after adding race, previous children delivered now living, family education, previous children delivered now dead, concurrent illness or condition affecting this pregnancy, and previous deliveries born dead, which is very little improvement after adding 6 variables to the model.

References

1. Chase, H. Time Trends in Low Birth Weight in the United States, 1950-1974. In D. Reed and F. Stanley (Eds.) The Epidemiology of Prematurity. Urban and Schwarzenberg: Baltimore, 1977.

2. Cox, D.R. The Analysis of Binary Data. London, Methuen, 1970.

3. McQueen, D.V. and Siegrist, J. Social Factors in the Etiology of Chronic Diseases: An Overview. Social Science Medicine, 1982, 16:353-367.

4. Reed, D.W. and Stanley, Fiona. The Epidemiology of Prematurity. Urban and Schwarzenberg: Baltimore, 1977.

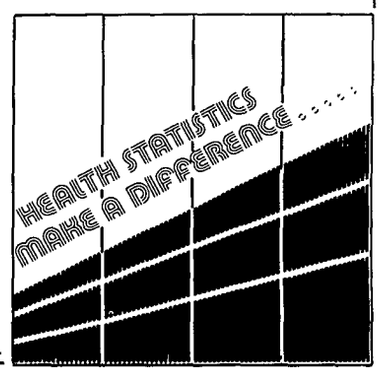
5. Shapiro, S., Schllesinger, E.R. and Nesbitt, R.E. Infant, Perinatal, Maternal and Childhood Mortality in the United States. Cambridge: Harvard University, 1968.

6. U.S. Department of Health and Human Services. Health United States, 1981. Public Health Service, National Center for Health Statistics. December, 1981.

7. U.S. Department of Health and Human Services. Health United States, 1980. Public Health Service, National Center for Health Statistics. December, 1980.

Session U

Health Statistics on Health Behavior



THE RELATIONSHIP BETWEEN MARITAL STATUS AND SMOKING BEHAVIOR

Margaret A. Cooke and Joel C. Kleinman, National Center for Health Statistics

Introduction

Over the last few years it has been suggested that separation and divorce promote health damaging behaviors including smoking (Brody, 1983 and Lynch, 1977). The association between marital status and smoking was documented in the 1950's (Haenszel, Shimkin and Miller, 1956) and was noted in the Surgeon General's first report on smoking and health which was published in 1964. The report stated that "smoking (of any kind) is most prevalent among the divorced and widowed" (Smoking and Health, 1964, page 364). In the 20 years since the Surgeon General's report was published, the rate of divorce has more than doubled, from 2.4 per 1,000 population in 1964 to 4.9 in 1984 (NCHS, 1968 and 1985). Recent data from the Census Bureau indicated that in 1984 over a quarter of all households with children were headed by a single parent (Bureau of the Census, 1985). Experts from the Census Bureau have projected that half the children born in the 1980's will spend at least part of their childhood living with one parent. Since divorce and separation are affecting an increasing proportion of both the adult and child population, the smoking habits of the separated and divorced may have important health effects.

This paper examines the prevalence of cigarette smoking by marital status using more recent data (1978-80), with particular emphasis on the combined separated and divorced group. We also consider the relationship between smoking and marital status after controlling for sex, age, and education, since these variables are related both to smoking and to marital status.

Data source

The data used for this analysis were obtained from the 1978, 1979, and 1980 smoking supplements appended to the National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics. The smoking supplement was given to a one-third sample for the last two quarters of 1978, the whole of 1979, and the last two quarters of 1980. The NHIS is a stratified household interview survey of the total U.S. noninstitutionalized civilian population (see National Center for Health Statistics, 1979 and 1981 for details of survey and sample design). The estimates in this paper are based on approximately 36,000 white and black persons age 20 to 64 years. Due to small sample size, widows were included in the analysis only for the 45 to 64 year age group.

Methods

Logistic regression was used to model the probability of being a current smoker. Initially the total data set was modeled with race, sex, age, education and marital status as independent variables. The resulting model included many higher order interactions, indicating that the variables were operating differently within different age and sex groups. Therefore, we analyzed the data separately by age and sex. The graphs presented here are based on the significant variables identified from the modelling. There were no significant differences in smoking prevalence between black and white in the total model and only a small difference among women age 25 to 44 years. Therefore the data presented here are based on black and white persons combined (other races excluded).

The graphs in figures 1 to 7 indicate the relationship of education and marital status to smoking prevalence for the various age and sex groups. Significant main effect variables were identified from the modelling and the graphs were plotted using only these significant variables. The data plotted are the weighted observed values which differed only marginally from the modelled values. Points based on less than 25 sample persons have been omitted from the graphs and the Appendix Tables.

Results

To illustrate the importance of sex, age, and education in determining smoking prevalence, Appendix Table 1 presents the prevalence of cigarette smoking for the years 1978-1980 tabulated for these three variables. Men are more likely to be cigarette smokers than women in all age and education groups. The prevalence of smoking is highest in the middle age groups for both sexes and all education groups. There is a strong inverse relationship between years of education and prevalence of smoking, except among women 45 to 64 years of age, where smoking does not vary substantially by education.

Age and sex differences in the prevalence of smoking are the result of different historical patterns. Cigarette smoking became widespread among men after the First World War and reached a peak in the 1950's. Women embraced the smoking habit somewhat later, with smoking prevalence peaking among women during the 1960's. Educated women were more likely to start smoking than were uneducated women in the early days of smoking uptake. In recent years, men have been more likely than women to stop smoking, particularly among those with higher education (Harris, 1983 and Higgins, 1984).

Appendix Table 2 shows that marital status accounts for a large additional proportion of the variation in smoking prevalence, over and above the proportion accounted for by sex, age, and education. In every age, sex and education group the separated and divorced (as a combined group) had a higher percentage of current smokers than those who were married or never married.

This higher prevalence of smoking among the separated and divorced relative to those of other marital status (generally on the order of 15 to 20 percent higher) is illustrated in Figures 1-7. There are no consistent differences in smoking prevalence between the married and the never married in the younger age groups. In the age group 45 to 64 years married men had a lower smoking prevalence than never married men. For women the reverse was true. The widowed age 45 to 64 years had a smoking prevalence intermediate between the separated and divorced and those of other marital status.

The slope of the lines on these graphs indicates the education gradient for smoking prevalence in the different age and sex groups. The education gradient declined with age for both men and women, but was always less steep for women than for men. For women age 45 to 64 there was essentially no educational

differential in smoking as indicated by the nearly flat lines in Figure 7.

From this cross sectional data it cannot be determined whether the stress of separation and divorce causes people to smoke or whether smokers should refrain from getting married. Some additional smoking variables which were examined to attempt to clarify the direction of causality were: likelihood of never having smoked, likelihood of stopping smoking, and age began smoking (See Appendix Tables 3-5).

In every age and education group a larger percentage of separated and divorced men and women had smoked at some point in their lives (Table 3). The differences were particularly striking for women. If separation and divorce increased the chances of taking up smoking, this group would have a higher proportion of "late smokers." However, Table 4 shows there was little difference between married and separated and divorced persons in the proportion who began smoking late. The higher proportion of ever smokers among the separated and divorced combined with a no later than average starting age suggest that separation and divorce did not lead to initiation of smoking. The separated and divorced, however, were least likely to be former smokers and the married most likely (Table 5). Among people age 25 to 44 years, about 15 percent fewer of those who ever smoked quit smoking among the separated and divorced compared with the married. It is therefore possible that separated and divorced people were more likely to resume smoking after having quit.

The greater prevalence of current smokers among the separated and divorced can therefore be attributed both to a high proportion of smoking uptake and a lower level of smoking cessation. Since the separated and divorced did not start smoking at an older age than the married and never married, there is no evidence that the stress of separation and divorce caused people to start smoking. However, it is possible that stress may have caused former smokers to start smoking again (data are unavailable to investigate this issue).

One final difference in the smoking behavior of the separated and divorced is that they were more likely to be heavy smokers than the married (Appendix Table 6). It is not possible to determine whether they were always heavier smokers or began smoking more heavily after separation or divorce.

Discussion

Previous research has indicated a higher prevalence of both mortality and morbidity among divorced and widowed persons relative to their married counterparts, as well as higher use of health care services, particularly inpatient services (Somers, 1979, Verbrugge, 1979, Lewis, 1984). However, we did not identify any studies which controlled for smoking behavior when relating marital status and health. Smoking has been heavily implicated in both the leading causes of mortality in the United States: heart disease and cancer, particularly lung cancer but also cancers of many other sites and the list is increasing (American Cancer Society, 1980). Some portion of these higher rates of mortality and morbidity reported among widowed and divorced persons may be attributable to the higher prevalence of smoking observed in these groups relative to married persons. From this analysis we conclude that it is necessary to control for smoking behavior when studying the relationship between marital status and morbidity, mortality and the use of health care services.

The relationship between marital status and smoking prevalence was noted two decades ago, yet little use appears to have been made of this information in the design of smoking cessation programs. These programs should target the separated and divorced and health personnel should counsel persons going through separation or divorce against restarting smoking or smoking more heavily during a potentially stressful period of their lives.

Smoking parents provide a role model for children concerning later smoking habits. A recent study found that children from one parent families were more likely to smoke than their peers and that girls were more likely to smoke if their mothers smoke and boys were more likely to smoke if their fathers smoke (Murray, Kiryluk and Swan, 1985). Since an increasing proportion of families are headed by single parents, the high smoking rates among the separated and divorced are of concern not only to themselves, but also as potentially influencing the smoking habits of the next generation. Smoking prevention and cessation programs among young people should take into account their family circumstances and the smoking habits of their parents.

References

American Cancer Society: Dangers of Smoking: Benefits of Quitting. American Cancer Society, New York, 1980.

Brody, J.E.: Divorce's stress exacts long-term health toll. New York Times. Dec. 13, 1983.

Bureau of the Census: Household and family characteristics, March 1984. Current Population Reports, Series P-20, No. 398. U.S. Government Printing Office, 1985.

Haenszel, W, Shimkin, MB, and Miller HP: Tobacco smoking patterns in the United States. Public Health Monograph No. 45. PHS Pub. No. 463. Public Health Service. U.S. Government Printing Office, 1956.

Harris, JE: Cigarette smoking among successive birth cohorts of men and women in the United States during 1900-80. JCNI 71(3):473-479. Sept. 1983.

Higgins, MW: Changing patterns of smoking and risk of disease. In: The Changing Risk of Disease in Women: An Epidemiologic Approach by Gold, EB. Collamore Press, Lexington, Mass., 1984

Lewis, FM: Marital status and its relation to the use of short-stay hospitals and nursing homes. Public Health Reports 99(4):415-424. July-Aug. 1984.

Lynch, JJ: The Broken Heart: The Medical Consequences of Loneliness. New York, Basic Books, 1977.

Murray, M, Kiryluk, S, and Swan, AV: Relation between parents' and children's smoking behavior and attitudes. Journal of Epidemiology and Community Health. 39(2):169-174, 1985.

National Center for Health Statistics: Vital Statistics of the United States, 1965: Volume 3, Marriage and Divorce. Public health Service. U.S. Government Printing Office, 1968.

National Center for Health Statistics: Mortality from selected causes by marital status, by Klebba, AJ. Vital and Health Statistics. Series 20, No. 8a and 8b. PHS Pub. No. 1000. Public Health Service. U.S. Government Printing Office, Dec. 1970.

National Center for Health Statistics: Differentials in health characteristics by marital status, United States, 1971-72, by Wilder MH. Vital and Health Statistics. Series 10, No. 104. DHEW Pub. No. (HRA) 76-1531. Public Health Service. U.S. Government Printing Office, Mar. 1976.

National Center for Health Statistics: Changes in cigarette smoking and current smoking practices among adults: United States, 1978, by Moss, AJ. Vital and Health Statistics. Advance Data, No. 52. DHEW Pub. No. (PHS) 79-1250. Public Health Service. Hyattsville, Md., 1979.

National Center for Health Statistics: Current estimates from the National Health Interview Survey: United States, 1979, by Jack, SS and Ries, PW. Vital and Health Statistics. Series 10, No. 136. DHHS Pub. No. (PHS) 81-1564. Public Health Service. U.S. Government Printing Office, April 1981.

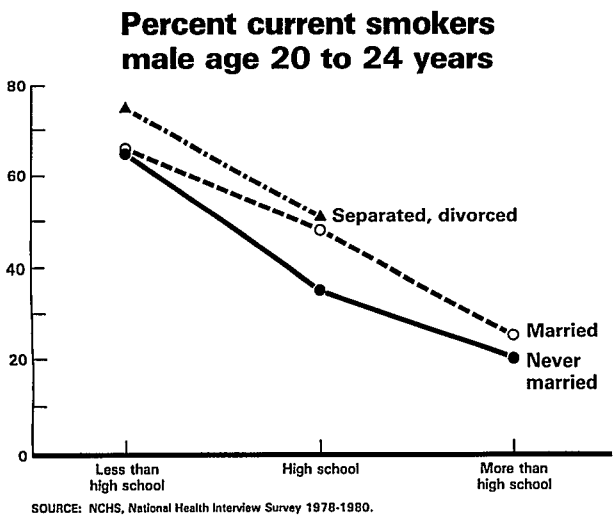
National Center for Health Statistics: Births, marriages, divorces and deaths for 1984. Monthly Vital Statistics Report, Vol. 33, No. 12. DHHS Pub. No. (PHS) 85-1120. Public Health Service. Hyattsville, Md. Mar. 26, 1985.

Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service. U.S. Department of Health, Education, and Welfare. Public Health Service Pub. No. 1103. U.S. Government Printing Office, 1964.

Somers, AR: Marital status, health and use of health services: An old relationship revisited. JAMA. 241(17):1818-1822, April 27, 1979.

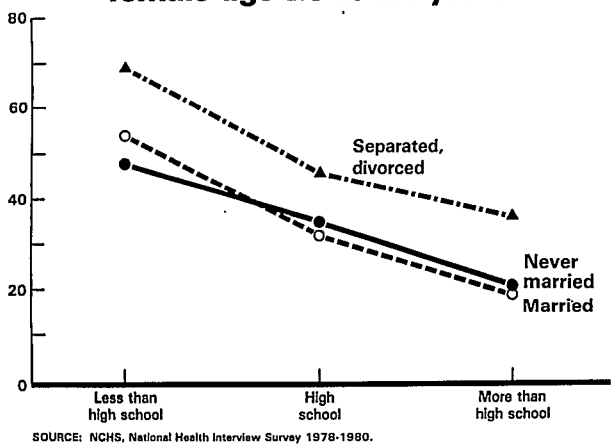
Verbrugge, LM: Marital status and health. J of Marriage and the Family. 41(2):267-285, May 1979.

FIGURE 1



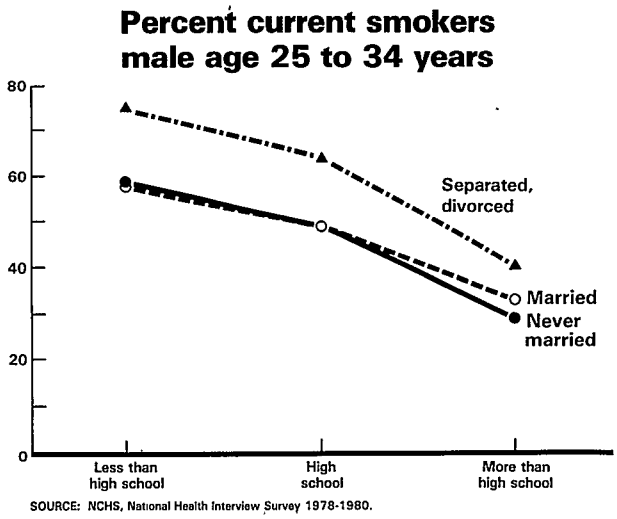
SOURCE: NCHS, National Health Interview Survey 1978-1980.

FIGURE 2



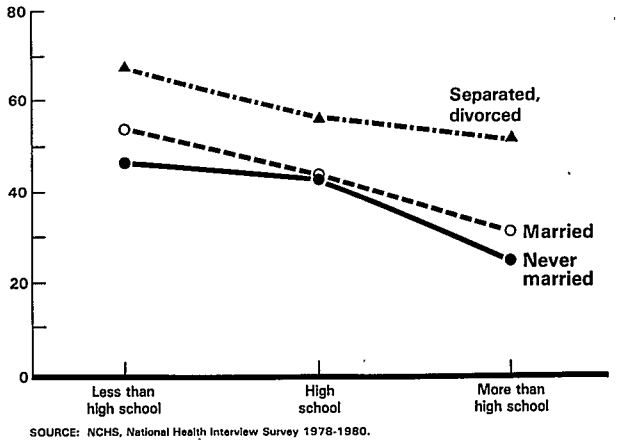
SOURCE: NCHS, National Health Interview Survey 1978-1980.

FIGURE 3



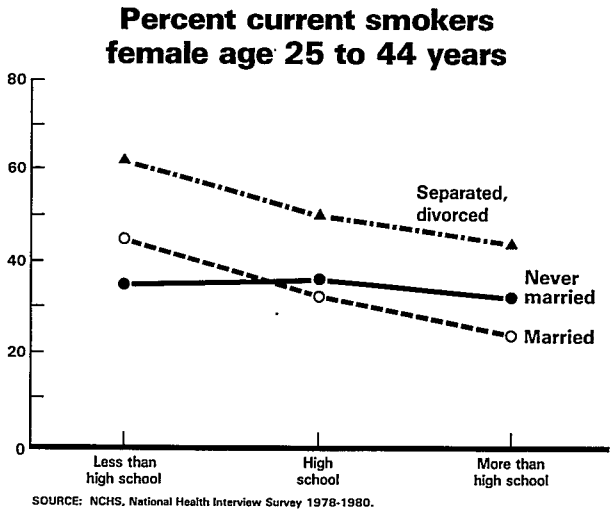
SOURCE: NCHS, National Health Interview Survey 1978-1980.

FIGURE 4



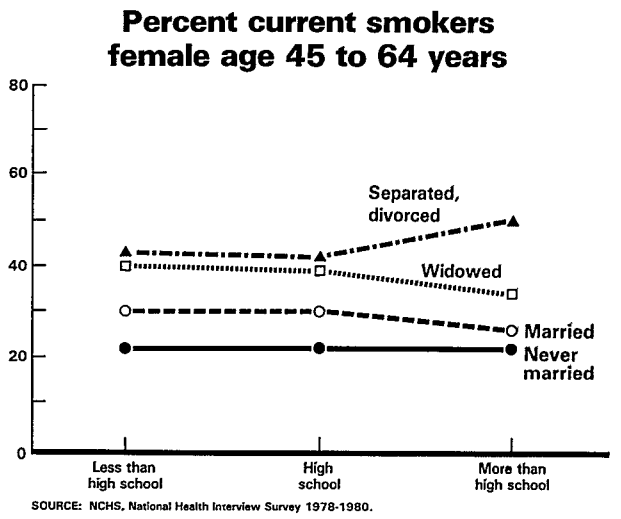
SOURCE: NCHS, National Health Interview Survey 1978-1980.

FIGURE 5



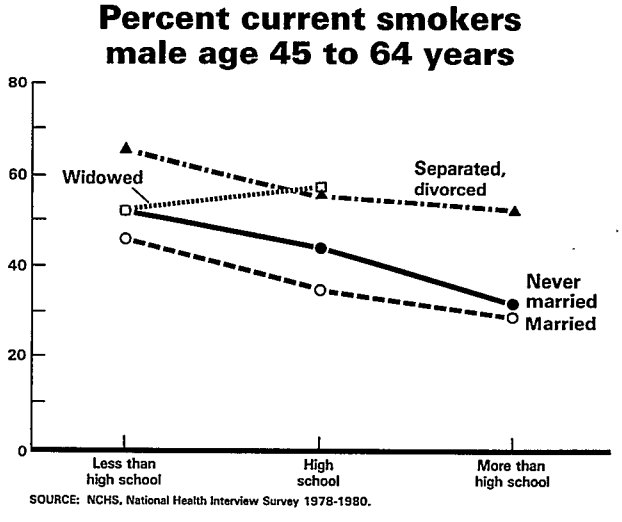
SOURCE: NCHS, National Health Interview Survey 1978-1980.

FIGURE 7



SOURCE: NCHS, National Health Interview Survey 1978-1980.

FIGURE 6



SOURCE: NCHS, National Health Interview Survey 1978-1980.

APPENDIX

Table 1. Percent current smokers according to sex, age, and education: United States, 1978-1980

Sex and age	Education		
	Less than high school	High school	More than high school
	Percent of persons		
Male			
20-24	66	41	22
25-34	60	50	32
35-44	55	45	33
45-64	48	37	31
Female			
20-24	55	34	21
25-34	47	35	26
35-44	46	35	29
45-64	31	31	28

SOURCE: National Center for Health Statistics, National Health Interview Survey, 1978-1980.

Table 2. Percent current smokers according to age, education, and marital status: United States, 1978-1980

Sex, age and marital status	Education		
	Less than high school	High school	More than high school
	Percent of persons		
Male			
20-24 years of age			
Never married	65	35	20
Married	66	48	25
Divorced or separated	75	51	*
25-34 years of age			
Never married	59	49	29
Married	58	49	33
Divorced or separated	75	64	40
35-44 years of age			
Never married	47	43	25
Married	54	44	31
Divorced or separated	68	56	52
45-64 years of age			
Never married	52	44	32
Married	46	35	29
Divorced or separated	66	56	52
Widowed	52	58	*
Female			
20-24 years of age			
Never married	48	35	21
Married	54	32	19
Divorced or separated	69	46	36
25-34 years of age			
Never married	38	35	31
Married	45	32	22
Divorced or separated	62	50	42
35-44 years of age			
Never married	29	38	36
Married	44	32	26
Divorced or separated	61	49	46
45-64 years of age			
Never married	22	22	22
Married	30	30	26
Divorced or separated	43	42	50
Widowed	40	39	34

SOURCE: National Center for Health Statistics, National Health Interview Survey, 1978-1980.

*Less than 25 sample persons in this category.

Table 3. Percent who ever smoked according to age, education, and marital status: United States, 1978-1980

Sex, age and marital status	Education		
	Less than high school	High school	More than high school
	Percent of persons		
Male			
20-24 years of age			
Never married.....	73	42	31
Married.....	75	64	37
Divorced or separated..	83	69	70
25-34 years of age			
Never married.....	69	62	47
Married.....	74	70	54
Divorced or separated..	77	76	56
35-44 years of age			
Never married.....	57	63	53
Married.....	78	71	62
Divorced or separated..	83	71	74
45-64 years of age			
Never married.....	70	63	57
Married.....	80	74	69
Divorced or separated..	79	85	79
Widowed.....	83	86	*
Female			
20-24 years of age			
Never married.....	56	43	28
Married.....	64	47	28
Divorced or separated..	74	58	53
25-34 years of age			
Never married.....	42	44	44
Married.....	54	48	40
Divorced or separated..	69	61	59
35-44 years of age			
Never married.....	31	50	49
Married.....	57	48	47
Divorced or separated..	70	61	59
45-64 years of age			
Never married.....	27	27	44
Married.....	43	45	48
Divorced or separated..	54	56	68
Widowed.....	53	53	54

SOURCE: National Center for Health Statistics, National Health Interview Survey, 1978-1980.

*Less than 25 sample persons in this category.

Table 4. Percent of ever smokers who began smoking at age 20 or older according to age, education, and marital status: United States, 1978-1980

Sex, age and marital status	Education		
	Less than high school	High school	More than high school
	Percent of persons		
Male			
20-24 years of age			
Never married	2	5	13
Married	6	4	15
Divorced or separated	*	5	*
25-34 years of age			
Never married	18	17	22
Married	9	16	24
Divorced or separated	6	15	26
35-44 years of age			
Never married	10	29	37
Married	13	17	27
Divorced or separated	17	24	29
45-64 years of age			
Never married	25	39	48
Married	18	26	35
Divorced or separated	22	27	33
Widowed	30	33	*
Female			
20-24 years of age			
Never married	4	8	15
Married	4	8	18
Divorced or separated	2	14	*
25-34 years of age			
Never married	22	27	33
Married	16	24	33
Divorced or separated	15	40	31
35-44 years of age			
Never married	*	34	50
Married	26	34	39
Divorced or separated	25	33	29
45-64 years of age			
Never married	42	65	54
Married	43	55	48
Divorced or separated	39	46	47
Widowed	48	59	63

SOURCE: National Center for Health Statistics, National Health Interview Survey, 1978-1980.

*Less than 25 sample persons in this category.

Table 5. Former smokers as a percent of all who ever smoked according to age, education, and marital status: United States, 1978-1980

Sex, age and marital status	Education		
	Less than high school	High school	More than high school
	Percent of persons		
Male			
20-24 years of age			
Never married.....	12	16	34
Married.....	13	25	34
Divorced or separated..	*	26	*
25-34 years of age			
Never married.....	13	21	38
Married.....	21	31	40
Divorced or separated..	2	16	28
35-44 years of age			
Never married.....	17	33	54
Married.....	31	37	50
Divorced or separated..	18	20	29
45-64 years of age			
Never married.....	26	30	44
Married.....	42	53	58
Divorced or separated..	17	34	33
Widowed.....	38	33	*
Female			
20-24 years of age			
Never married.....	14	17	27
Married.....	15	31	33
Divorced or separated..	7	21	*
25-34 years of age			
Never married.....	9	21	29
Married.....	16	33	44
Divorced or separated..	9	17	28
35-44 years of age			
Never married.....	*	24	26
Married.....	23	33	45
Divorced or separated..	13	19	21
45-64 years of age			
Never married.....	19	17	50
Married.....	30	33	46
Divorced or separated..	19	26	27
Widowed.....	24	27	38

SOURCE: National Center for Health Statistics, National Health Interview Survey, 1978-1980.

*Less than 25 sample persons in this category.

Table 6. Percent of current smokers smoking 25 or more cigarettes a day according to age, education, and marital status: United States, 1978-1980

Sex, age and marital status	Education		
	Less than high school	High school	More than high school
	Percent of persons		
Male			
20-24 years of age			
Never married.....	17	23	10
Married.....	19	25	32
Divorced or separated..	*	*	*
25-34 years of age			
Never married.....	30	23	22
Married.....	28	33	34
Divorced or separated..	40	29	48
35-44 years of age			
Never married.....	29	38	*
Married.....	43	44	44
Divorced or separated..	47	64	44
45-64 years of age			
Never married.....	36	41	75
Married.....	39	44	50
Divorced or separated..	30	50	47
Widowed.....	17	34	*
Female			
20-24 years of age			
Never married.....	19	16	12
Married.....	18	15	13
Divorced or separated..	44	25	*
25-34 years of age			
Never married.....	14	24	25
Married.....	26	21	24
Divorced or separated..	31	32	26
35-44 years of age			
Never married.....	*	*	25
Married.....	30	29	23
Divorced or separated..	29	34	24
45-64 years of age			
Never married.....	*	39	11
Married.....	22	23	28
Divorced or separated..	26	36	31
Widowed.....	21	20	19

SOURCE: National Center for Health Statistics, National Health Interview Survey, 1978-1980.

*Less than 25 sample persons in this category.

THE USE OF EVENT RELATED DATA TO IMPACT DRINKING DRIVING BEHAVIOR

Lance B. Segars and Barbara E. Ryan
San Diego County Department of Health Services - Alcohol Program

Introduction

Motor vehicle crashes are a significant public health problem in the United States and the leading cause of death among males aged 15 to 25. Significant among the causes of these fatalities is drinking driving. The National Highway Traffic Safety Administration reports that 55% of all fatal accidents between 1979 and 1981 were alcohol related (1).

The emergence of grassroots groups such as Mothers Against Drunk Drivers (MADD) and Remove Intoxicated Drivers (RID) as well as the formation of the Presidential Commission on Drunk Driving has resulted in greater public attention on the drinking driving problem. Much of that attention has focused on the prevention of drinking driving and rehabilitation of the convicted drinking driver. Indeed, DUI education and rehabilitation programs are the fastest growing treatment industry in the State of California.

The State of California responded to the increased public concern regarding drinking driving by instituting major legislative changes in 1982. One provision under new law requires all individuals convicted of a first driving under the influence (DUI) offense who are granted probation by the Court to attend a treatment/education program. In San Diego County alone this resulted in an additional 18,000 individuals a year entering alcohol programs as a result of a DUI conviction.

The literature on drinking driving is immense. However, most of the research is limited to two perspectives. The psycho-social perspective is focused on the diagnosis of alcoholism or the individual characteristics of drinking drivers, e.g., ego identity, locus of control, etc. The traffic safety perspective is focused on the impairing effect of alcohol on the driving task or analysis of crash involvement on the basis of blood alcohol level. Such reductionist approaches fail to consider the environment in which DUI behavior occurs (2). Very little research on the DUI event has been conducted. A better understanding of DUI behavior can be gained by de-emphasizing the view of human behavior as separate from the broader social context in which it occurs (3).

The existing literature on DUI leaves several questions unanswered. Who are alcohol impaired drivers? Under what conditions do individuals drive after drinking? Is the event which results in a DUI conviction typical behavior for these individuals? At what points in the DUI event can interventions occur?

In order to respond to these and other questions the San Diego County Alcohol Program embarked on a series of data collection activities to describe the system in which drinking driving occurs, comes to the attention of law enforcement agencies, is adjudicated in the courts, and, finally, results in attempts at intervention via an alcohol treatment/education program.

The data presented in this paper represent early efforts from a number of studies to describe the DUI event. It is suggestive of the importance of and problems associated with viewing DUI behavior within the environmental context.

Who Are Drinking Drivers?

Descriptions of the drinking driver vary substantially. For example, rates of alcoholism among those convicted of driving under the influence (DUI) have been reported to range from 4% to 87% (4). The magnitude of alcoholism in the DUI population, however, is critical in determining service needs. Therefore, shortly after changes in the DUI laws in California made mandatory education/treatment programs a condition of probation, the San Diego County Alcohol Program began collecting data concerning the characteristics and behavior of individuals convicted of DUI. The initial results demonstrated that, given the structure of law, arrest, and adjudication practices in San Diego, the DUI offender referred to a program is in general not alcoholic.

The largest group of program participants are young (51% under 30) males (81%). While DUI offenders report frequent heavier drinking, their behavior is not atypical for their general age cohort (Tables 1 and 2). In most cases, the DUI conviction represents the first reported problem associated with their consumption of alcohol (5). Thus, programs are serving a population predominantly at risk for both continued alcohol-impaired driving and progression of alcohol problems. However, consistent with general population findings these individuals may often mature out of alcohol problems with age (6). Consequently, the DUI program for first offenders in San Diego County is intended to hasten that maturation process through education, opportunities for self-assessment of alcohol use, and voluntary referrals for additional alcohol services when appropriate.

The population entering a first DUI conviction program is dependent upon specific local arrest and adjudication practices. Consequently, the characteristics of the San Diego County first conviction program participants may differ dramatically from those in other areas due to local variances in the criminal justice system.

How Does the DUI Event Occur?

DUI can be viewed from a perspective of general behavior or as a function of alcoholism. The levels of consumption and alcohol problem indicators reported by the largest group of DUI offen-

TABLE 1
ALCOHOL CONSUMPTION
OF DRINKING DRIVERS IN
COMPARISON TO ALCOHOL PROGRAM
PARTICIPANTS AND THE GENERAL POPULATION

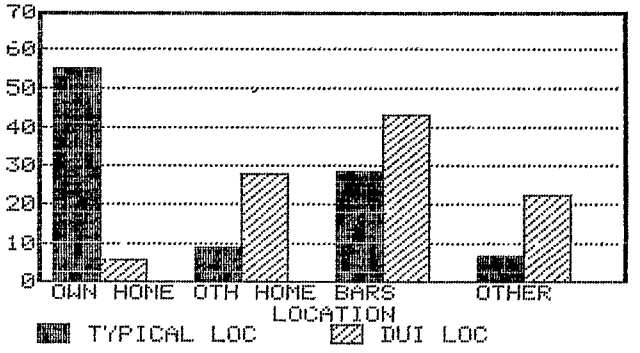
Consumption Category	Daily Avg. Oz. Alcohol	Example	Population Group GP	FCP	CDDP	Recovery
			Percent in Category			
Light Drinker	0.01 - 0.21	Up to 3 drinks per week	56.7	9.1	12.5	0.6
Moderate Drinker	0.22 - 0.99	Up to 13 drinks per week	28.4	50.3	35.9	3.2
Heavier Drinker	1.00 - 2.99	2 - 5 drinks per day	12.2	32.7	38.3	9.3
High Risk	3.00 +	six or more drinks per day	2.7	7.9	13.3	86.9

GP General U.S. population excluding abstainers n=1582(9).
FCP First conviction drinking drivers n=168(5).
CDDP Multiple conviction drinking drivers n=272.
Recovery - Alcohol recovery program participants n=314(10).

TABLE 2
CONSUMPTION OF DRINKERS BY AGE
FOR MALES IN THE GENERAL POPULATION AND
FIRST DUI OFFENDERS

Age	Drinking Category: Light Group GP	Light Group FCP	Moderate Group GP	Moderate Group FCP	Heavier Group GP	Heavier Group FCP
	Percent in Category					
18-20	35	38	30	31	35	31
21-34	30	3	37	51	33	46
35-49	26	5	45	43	29	52

Figure 1
Location of Typical Drinking and DUI Event



ders in San Diego suggests that viewing DUI in the environmental context rather than as an individual alcoholism problem provides useful information both for changing individual behavior and preventing alcohol-impaired driving.

The first environmental factor which is important in understanding DUI is the location in which drinking occurs. The drinking driving behavior which results in conviction in San Diego begins with consumption of relatively large amounts of alcohol away from home. The location of consumption is most often bars or private parties (Figure 1). In comparison, the typical drinking location for these individuals is their own home. Of particular interest from a public health perspective is the identification of specific environments from which almost half of all convicted DUI's originate. Obvious implications for prevention are server training activities. Additionally, the availability of public transportation could be a consideration in the issuance of on-premise liquor licenses.

Information concerning less frequently reported drinking locations provide examples of more unique environmental problems. Two areas found among 'other' drinking locations were sports events and work sites. Recently, attention has been focused on alcohol use during sporting events. Local responses to both drinking driving and other risks associated with alcohol use at sporting events, include limiting carry-in beverages, and stopping alcohol sales prior to the end of the event. Responses in other communities have been to reduce the alcohol content of beverages available in stadiums.

Work site drinking presents an interesting example of how accepted drinking practices relate to DUI problems. Among those convicted, 5% reported drinking at work. Further questioning indicated that, for the most part, these cases represented a common practice among construction and trades workers. At the end of the work shift, employees regularly purchase beer and remain at the work site socializing. While in

general such practices may improve camaraderie, the implications of impaired employees on construction sites should be considered by contractors and managers from both a DUI and a risk management perspective.

In addition to location differences the amount consumed prior to arrest also appears atypical for these individuals. While DUI offenders report typical drinking of four drinks on days when they drink, consumption at the time of arrest averages slightly over eight drinks consumed during a four-hour period. These self-reports of consumption are consistent with the blood alcohol level at arrest (Table 3). Information regarding quantity and duration of consumption are particularly relevent to programs which stress "knowing your limits." Often these programs and public service campaigns provide information concerning acceptable levels of consumption by body weight. While generally accurate, these messages often reflect consumption over a one-hour period. Such information may give unrealistic estimates of acceptable drinking rates over longer periods of time. Thus, rather than suggesting that for a 150-pound person consumption of five drinks in an hour is impairing, a more useful message would describe consumption of two drinks per hour over the course of a four-hour party as impairing for driving.

Beverage preference of DUI offenders provides additional information on DUI behavior that has implications for prevention and program responses. For the most part the beverage of choice is beer (Figure 2). Conversely, persons entering alcoholism recovery services report a very different beverage preference. Alcohol sales data suggests that the DUI offender beverage preference is similar to the general population, which consumes beer at the highest per capita rate. In addition, a recent study suggests that beer drinkers are more likely to participate in high risk behavior such as drinking and driving (7).

The identification of beer as the beverage of choice for DUI offenders has implications for the design of educational campaigns and social policy decisions. This information contradicts the widely-held belief that beer is a relatively harmless beverage of moderation. DUI offenders in San Diego, as elsewhere, view beer as less impairing than distilled spirits (8). This view is consistent with social policies such as lower legal purchase ages in several states for beer than for spirits and a voluntary ban on radio and television advertising for spirits with no similar constraints on beer advertising. Further, much of beer advertising involves active, often hazardous male-dominated behavior. In addition, beer drinking is usually presented as a group activity in bars or taverns. This type of drinking may be more likely to result in buying rounds, staying longer, and drinking more than usual.

Other data on this population include the perceptions of offenders about their behavior (Table 3). For the most part, DUI offenders in San Diego were not arrested as a result of serious driving problems; only 14% reported accident involvement. Again, this is very different than reports from other areas. In addition, while most believed they were under the influence of alcohol, less than half felt they were drunk. Similarly, almost half believed they passed the field sobriety test. These perceptions suggest that reliance on personal judgement after drinking is not sufficient to reduce DUI behavior. Additionally, the availability of others in the car, particularly someone who is not impaired for driving, is rare. These pieces of information indicate that attempts to intervene must occur not at the point of driving, but rather prior to the point of impairment.

Figure 2
Beverage of Choice of DUI and Recovery Program Participants

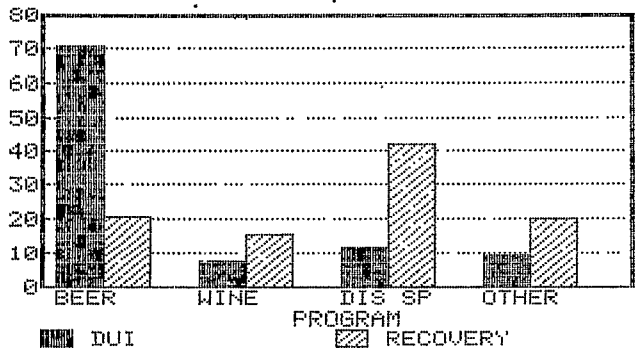


TABLE 3
ALCOHOL IMPAIRED DRIVING EVENT OF CONVICTED DRINKING DRIVERS

	Percent	
Arrest blood alcohol level		
< .10	0.0	
.10 - .14	27.6	
.15 - .19	41.7	
.20 - .24	21.8	
.25 +	7.0	
Refused test	1.9	
Accident involved		
No	86.2	
Yes	13.8	
Do you believe you passed the field sobriety test?		
No	45.3	
Yes	43.7	
No test	10.8	
Were you under the influence?		
No	19.6	
Yes	80.4	
Were you drunk?		
No	53.6	
Yes	46.4	
Were others in car?		
No	67.9	
Yes	32.1	n=168

Summary

The findings of these studies suggest that viewing drinking driving from a behavioral, event focused perspective as well as a traffic safety and/or psycho-social perspective provides a greater understanding of the problem. An increased understanding of the environmental context of this public health and safety problem suggests strategies for intervention and prevention focused on high risk individuals and behaviors.

Future research on drinking and driving behavior would be useful in further describing the event. Greater detail on the context of the event is needed, especially as it differs from locale to locale. Such information includes where they are drinking, why, with whom, where they are coming from, where they are going, etc. In addition, larger samples would allow for greater breakdown by age, sex, or other important independent measures to identify specific high risk groups. The findings of these studies in San Diego County have highlighted some of these issues and provided a focus for future research activities.

REFERENCES

1. Rubin, D., Hoar, R., and Klinger, D. "Report on Traffic Accidents and Injuries, 1981-The National Accident Sampling System," United States Department of Transportation, National Highway Traffic Safety Administration, March 1983.

2. Ross, H.L. Deterring the Drinking Driver. Lexington, Mass. D.C. Heath and Co., 1982.

3. Wallack, L. Drinking and driving: Toward a broader understanding of the role of mass media. Journal of Public Health Policy 5:471-496, 1984.

4. Vingilis, E. Drinking Drivers and Alcoholics, Are They From the Same Population? In: Smart, Glaser, Israel, Kalant, Popham, and Schmidt (Eds.) Research Advances in Alcohol and Drug Problems, Volume 7. Plenum Publishing Corp., 1983.

5. Ryan, B.E. and Segars, L.B. "San Diego County First Conviction Program Population Description," Unpublished Manuscript. San Diego County Department of Health Services - Alcohol Program, January, 1983.

6. Cahalan, D. and Room, R. Problem Drinking Among American Men. Rutgers Center of Alcohol Studies. New Brunswick, New Jersey, 1974.

7. Berger, D.E. and Snortum, J.R. Alcoholic beverage preferences of drinking-driving violators. Journal of Studies on Alcohol 46:232-239, 1985.

8. Ryan, B.E. and Segars, L.B. "San Diego County First Conviction Drinking Driver Program Participant Knowledge and Attitude Change," Prepared for The Ninth Annual California Conference on Alcohol Problems, Burlingame, California. September, 1984.

9. Johnson, P., Armor, D.J., Polich, S., and Stambul, H. U.S. Adult Drinking Practices: Time Trends, Social Correlates and Sex Roles. Rand Corporation, Santa Monica, California, 1977.

10. Segars, L.B. and Wynne, J.D. "The San Diego County Client Outcome Study, Part II: Baseline data on selected characteristics of problem drinkers," Unpublished manuscript, San Diego County Department of Health Services--Alcohol Program, February, 1981.

THE USE OF HEALTH STATISTICS IN PREDICTING THE EFFECTS OF INTERVENTION PROGRAMS TO INCREASE EARLY DETECTION OF BREAST CANCER

N. Urban, E. Avlon, B. Thompson, Fred Hutchinson Cancer Research Center

Background

In 1982, the American Cancer Society (ACS) prepared and published a statement endorsing the use of mammography as a valuable tool in the detection and diagnosis of breast cancer. Pointing out that breast cancer is the number one cancer killer of American women, the ACS called for 1) monthly breast self-examination (BSE) for women aged 20+; 2) annual physical examination for women over 40; 3) a baseline mammogram for women between ages 35-40; and 4) annual mammography for women age 50+ [1].

The first major study of the use of mammography in breast cancer detection was a randomized trial at the Health Insurance Program (HIP) of greater New York during the 1960's [2]. A combination of physical examination and mammography was used annually to screen women aged 40-64. Among Shapiro's findings were that

1) mortality was reduced by 28.9% in the group invited for screening, relative to the controls;
2) mortality reduction was demonstrated only in the over-50 age group;
3) cancer incidence over the five-year period following entry into the study increased negligibly (about 5%) as a result of the screening program; and
4) annual screening improved the rate of early detection within detection modality: a higher proportion of cases detected at follow-up screens, than at initial screens, had no axillary nodal involvement.

More recently, during the 1970's, the Breast Cancer Detection Demonstration Project (BCDDP) employed annual mammography, physical examination and BSE instruction in a nationwide breast cancer screening demonstration program. The BCDDP demonstrated that mammography improved over the intervening decade; mammography detected only 55% of the cases detected at screening in the HIP study, in the BCDDP it detected 89% of screening-detected cases [3].

Also of interest are the results of a recent Swedish study in which investigators found that single-view mammography employed alone and at longer intervals still reduced mortality by 31%. A screening interval of thirty-three months was used in screening the over-age-50 women with resulting mortality reduction of 40% in that age group. Like the HIP, the Swedish trial has been unable to demonstrate significant mortality reduction in the women under age 50 [4].

Approach

In our judgement, these and other studies have established the efficacy of screening by mammography and physical examination in reducing mortality from breast cancer. What has not yet been established is the relative cost-effectiveness of different intervention strategies.

Our objective is to investigate the impact, on detection costs as well as mortality, of alternative intervention strategies involving mammography, physical examination, and breast self-examination instruction, including 1) the use of combinations of detection modalities at varying intervals, and 2) the targeting of high-risk women for intensive intervention, employing estimates from the literature and published health statistics.

Ours is a population-based approach to analyzing the problem, taking into consideration non-participation as well as baseline detection activities. We focus on a hypothetical population of women age 45+. This age group accounts for about 87% of breast cancer cases in Western Washington. A baseline model was constructed, describing breast cancer detection in the absence of any systematic intervention program.

It was assumed that the conditional probabilities (e.g. the predictive value of each screening modality) used in the baseline model to describe the detection of breast cancer are not affected by the introduction of a screening program. It was further assumed that the hypothetical population of women aged 45+ in whom breast cancers are to be detected remains constant, with annual incidence of 200 breast cancer cases, and that

1) among women exposed to mammography, breast cancer cases are detected in accordance with the sensitivity of mammography, assumed to be .8, as estimated from the HIP data by Walter and Day [5];
2) breast cancer cases not detected by mammography, because of either non-exposure or false negativity, remain eligible for detection by BSE;
3) breast cancer cases not detected by mammography or BSE remain eligible for detection by physical examination; and
4) those breast cancer cases destined to be incident within the year which are not detected by physical examination, BSE, or mammography surface as symptomatic.

It is important to clarify our assumption that incidence remains constant, because in the case of cervical cancer, screening raises the steady-state annual incidence relative to incidence prior to implementation of the screening program. This is because screening detects some early lesions that would not otherwise surface clinically. Our model can accommodate this, as long as the new incidence rate remains constant at the higher level once the prevalent cases have been detected.

However, in the estimates presented here, we have assumed that there is no increase in incidence as a result of screening, primarily because the evidence in the literature did not seem to us to support an assumption that breast cancer incidence would increase as a result of screening [2]. Note that we refer to cases detected at the initial screen as prevalent cases, following convention in most of the literature, and include among incident cases all new breast cancer cases occurring after the initial screen.

Starting with the baseline model, a screening intervention strategy's impact is evaluated by 1) changing the proportion of women exposed to each screening detection modality, and 2) adjusting the probability of positivity of each screening detection modality to reflect breast cancer

incidence in the population following screening, in some proportion of the population, by other modalities more likely to detect early cancer.

The Baseline Model

The approach described above yields a simple, decision-analysis-type model describing the services utilized in the detection of breast cancer, and the resulting distribution of cases by detection modality. In the absence of any systematic intervention program, and assuming unit costs for each service as shown in Table I, the total annual cost of services utilized is $1,104,680 to detect 200 cancers in a cohort of 100,000 women age 45+. We refer to this baseline model as Model A, to which we compare alternative intervention strategies.

We simulated the simplest forms of intervention by inviting all women for mammography or physical examination or BSE instruction--models B, C, and D respectively--and generated the services that would be utilized if 90% of women participated. The resulting services and total costs generated are shown in Table I.

We employed estimates of service costs that were consistent with the literature and with our own experience. We assumed that service costs are independent of the implementation of a screening program. It should be noted too that it is the relative costs rather than the actual costs that affect conclusions about the relative cost-effectiveness of alternative intervention strategies.

We then compared the additional costs of the strategy to its mortality reduction, relative to the baseline. Estimates of absolute and incremental mortality reduction and costs generated for Models A, B, C, and D are given in Table II.

It can be seen that, assuming 90% participation, mammography for all women approximately triples system detection costs relative to the baseline. It averts 15.9 deaths, a mortality reduction of 28.3%, for a total of $2,377,645, implying a marginal cost per death averted of about $150,000. Physical examination alone for all women increases costs about a sixth as much as mammography, but saves only one-third as many lives. BSE instruction alone achieves about 38% of the mortality reduction of mammography, at about 30% of its additional cost.

Methods

In order to generate the estimates reported in Table II it was necessary to make several assumptions and to employ estimates from published and unpublished health statistics. First, we estimated five-year survival rates for women age 45+, by stage at diagnosis, from the cancer registry data for the last decade for the Puget Sound area of Washington State. Early-stage cases accounted for 53.7% of the 7,103 cases analyzed. Among these, five-year survival was 85.6%, implying mortality of 14.4%. Among the late-stage cases, which accounted for 46.3% of all cases, five-year survival was 56.0%, implying mortality of 44.0%.

These estimates are based on 7,103 women aged 45+ whose breast cancer was detected since 1973. Five-year survival rates were estimated using life-table survival analysis. Deaths due to all causes that might possibly be attributable to the breast cancer were counted as deaths; only when there was no evidence of cancer at death were deaths due to extraneous causes censored.

In order to estimate the effect of a screening intervention strategy on mortality, we assumed that, other things being equal, the five-year survival rate for breast cancer is determined by clinical stage at diagnosis.

A search of the literature was then conducted for information about breast cancer detection in the absence of a screening program, to obtain the parameters for our baseline model. A 1981 paper based on data on women in Georgia provided the distribution of cases by detection modality, and the stage distribution [6]. The relevant data are shown in Table III.

Cases detected by BSE were earlier-stage, on the average, than cases detected by physical examination. Note that BSE-detected cases do not include all cases detected by BSE-practicers; rather, they are only the cases detected during routine BSE.

We then assumed that, other things being equal, the stage distribution of breast cancers is determined by the mode of first detection of the cancers. However, the stage distribution by detection modality reported by Huguley and Brown reflects an unscreened population. Therefore, based on data reported by Shapiro [2], we have estimated that the overall stage distribution shifts toward earlier detection by a factor of .13 for all incident cases among women invited to participate in an annual screening program. For lack of more specific data, we have assumed that this factor applies when screening is offered annually, regardless of the detection modality.

The proportion of cancer cases that are early stage was estimated as the weighted sum of the rates of early stage within cancer modality, weighted by the proportion of cases detected by each modality. Adjustment was made as necessary to reflect the impact of annual screening.

Expected mortality was calculated as a weighted sum of the mortality rates by stage, where the weights are the proportion of cases in late and early stage.

Results

From comparison of Models A, B, C, and D shown in Table II it is clear that annual mammography for all women is costly, and the use of physical examination or BSE instruction instead of mammography is significantly less costly but results in less than half of the mortality reduction achievable by mammography.

For our first set of intervention strategies, we considered synchronous combinations of screening modalities. In North America and the United Kingdom, a combination of physical examination and mammography is usually employed, with BSE instruction either initially or annually. For example, the Canadian National Study combines initial BSE instruction with a combination of physical examination and mammography applied annually [7]. However, because annual mammography is very costly when applied to an entire population of women, variations are being tried. For example, in the United Kingdom a two-year interval is being used for mammography [8], and at Group Health Cooperative of Puget Sound, the mammography interval is being varied according to the risk-level of the woman. The higher

the risk level of the woman, the more frequently mammography is given [9].

In our first look at intervention strategies we considered three models that combined detection modalities synchronously. The results are shown in Table IV.

In Model G, mammography and physical examination are provided synchronously--i.e. in combination at the same time--once a year. In Model H, mammography, physical examination, and BSE instruction are combined similarly. In Model I, the three modalities are combined at three-year synchronous intervals.

It can be seen from Table IV that combining physical examination with annual mammography adds more to the costs of the program than to its effectiveness. Model G costs 13.4% more than Model B (annual mammography) but it reduces mortality by only 1.9%. Relative to Model B, it reduces mortality by only 0.3 lives at an additional cost of $380,000, implying that it saves an additional life at a marginal cost of over $1 million. Similarly, adding annual BSE instruction along with annual mammography and physical examination increases costs by about 35% while improving survival by less than 5% (Model H relative to Model B).

From Model I it can be seen that increasing the screening interval improves cost-effectiveness somewhat. Using a combination of the three modalities at three-year intervals achieves 52.2% of the mortality reduction achieved by annual mammography for all women (Model B) for 43.6% of the costs.

The reason for the poor performance of synchronous combinations of detection modalities is their redundancy. From the BCDDP experience, we have learned that of cases detected at screening, 88.9% were detected by mammography. Of these, 53% were also picked up by physical examination. Only 8.7% of screening-detected cancers were detected by physical examination alone [3].

BSE instruction can be similarly redundant. National surveys have found that about 24% of women practice BSE. We also know from the literature [10] that about 75% of women receiving interactive BSE instruction continue to practice for at least a year. Thus, BSE practice is higher in a year immediately following BSE instruction. The redundancy in screening is less if women are influenced to increase their BSE practice during a time period when they have not recently received mammography or a physical examination, since more undetected cancers will be prevalent during such a time period.

Next we considered targeting high-risk women for mammography as a means of improving cost-effectiveness. First, we considered the case in which 10% of the population having a relative risk of 5.0 are offered annual mammography, while the remaining women receive normal care. This is Model E, shown in Table V. Second, we considered the probably more realistic case in which 10% of the population having a relative risk of 2.0 are offered annual mammography while the remaining 90% receive normal care. This is Model F, also shown in Table V.

As shown in Table V, based on our assumptions about half the mortality reduction achievable by annual mammography for all women can be achieved for only 6% of the costs, if 10% of the women can be identified who are truly at relative risk of 5.0 (Model E). Unfortunately, selection on risk factors such as family and reproductive histories and prior benign breast disease probably will not yield such a group [11].

When the relative risk of the highest-risk 10% is actually only about 2.0 (Model F), only about 21% of the mortality reduction is achieved, for about 9% of the costs. Since the costs of identifying and targeting the high-risk women have not been taken into account in this analysis, this approach appears less promising than we had originally anticipated.

These results suggest that potentially three approaches improve cost-effectiveness. These are targeting high-risk women, increasing the screening interval, and eliminating redundancy among detection modalities. Therefore, we considered combining targeting high-risk women for annual mammography with staggering the interventions for the remaining women. First, we staggered the three detection modalities, each at a three-year interval, so that every woman who participates receives one intervention every year, but receives each intervention only once every third year. This is Model J.

Next, we combined staggering the three detection modalities as in Model J with targeting high-risk women for annual mammography. In Model K we assumed that the relative risk of the highest risk 10% of the population of women was 2.0, as in Model F. In Model L we optimistically assumed that women with relative risk of 5.0 could be identified and targeted (as in Model E). The results of Models J, K, and L are shown in Table VI.

Staggering the interventions appears to improve cost-effectiveness. Under the assumptions that we have made, 76% of the mortality reduction can be achieved for 43% of the cost of annual mammography by staggering the three interventions over a three-year interval (Model J). Note that an important assumption here is that the cost of performing the physical examination and BSE instruction is not raised as a result of separating them from the mammography visit.

Combining the staggering of the three detection modalities at three-year intervals with the targeting of high risk women for annual mammography appears to be the most cost-effective approach if women with relative risk of 5.0 can be reliably identified and targeted. We estimate that 88% of the mortality reduction can be achieved for 46.4% of the cost of annual mammography using this approach (Model L). As before, we have not accounted for the costs of identifying and targeting the high-risk women.

In Model K we assume that the relative risk of the high-risk women is only 2.0. Mortality reduction is improved only modestly by targeting these women for annual mammography. We estimate that 81.1% of the mortality reduction achievable by annual mammography can be reached this way at 48.2% of the costs. Since we have not accounted here for the costs of identifying and targeting the high-risk women, we therefore conclude that the strategy of targeting high-risk women for annual mammography should be analyzed further before it is adopted widely in the community.

Summary of Conclusions

We have made a number of assumptions, many of which cannot be verified. Sensitivity analysis is being performed but is not described here. Conclusions should therefore be viewed as tentative. In summary, they are as follows:

1) Annual mammography for all women age 45+, assuming 90% participation, results in mortality reduction of almost 30% for a cost of about $150,000 per death averted. In a cohort of 100,000, assuming annual incidence of 200 breast cancer cases, this intervention can be expected to avert about sixteen breast cancer deaths annually, once a steady state has been reached.
2) Including physical examination in an annual mammography screening is probably not cost-effective. Assuming that mammography detects about 90% of the cancers detectable on annual screening, and that the cost of the physical examination is not reduced by combining it with mammography, inclusion of physical examination in the annual mammography screen saves an additional life at a marginal cost of over $1 million.
3) Mammography can be made more cost-effective by using it at longer intervals, such as once every three years, and selecting high-risk women for annual mammography, assuming that sufficiently high-risk women can be identified at reasonable cost.
4) Combining detection modalities in an intervention strategy can be made more cost-effective by staggering their use. Annual intervention consisting of mammography one year, BSE instruction the next, and physical examination the third, so that all women age 45+ receive mammography once every three years, achieves over 75% of the mortality reduction of annual mammography alone at less than half the cost.
5) Combining selection of high-risk women for annual mammography with staggering of three detection modalities for remaining women further improves cost-effectiveness; however, this approach requires that sufficiently high-risk women be reliably identified and targeted.

References

1. American Cancer Society: Mammography 1982: A Statement of the American Cancer Society. Cancer J for Clinicians 32:226-230, 1982.
2. Shapiro S: Evidence on Screening for Breast Cancer from a Randomized Trial. Cancer 39:2772-2782, 1977.
3. Baker L: Breast Cancer Detection Demonstration Project: Five-year summary report. Cancer J for Clinicians 32:195-225, 1982.
4. Tabar L, et al: Reduction in Mortality from Breast Cancer After Mass Screening with Mammography. Lancet 829-832, 1985.
5. Walter SD and Day NE: Estimation of the Duration of a Pre-Clinical Disease State using Screening Data. Am J Epidemiol 118:865-86, 1983.
6. Huguley CM and Brown RL: The Value of Breast Self-Examination. Cancer 47:989-995, 1981.
7. Miller AB, Howe GR, and Wall C: The National Study of Breast Cancer Screening. Clin Invest Med 4:227-258, 1981.
8. Roberts MM, et al: The Edinburgh Randomised Trial of Screening for Breast Cancer: Description of Method. Br J Cancer 50:1-6, 1984.
9. Carter AP, et al: Report on Early Detection of Breast Cancer Implications for Group Health Cooperative of Puget Sound. Sept 1983.
10. Lee C: Community Center Control of Los Angeles (CCC/LA): Breast Examination Training (BET) Centers. Prog in Cancer Control IV pp. 283-292, 1983.
11. Soini I and Hakama M: Failure of Selective Screening for Breast Cancer by Combining Risk Factors. Int J Cancer 22:275-281, 1978.

TABLE I - Breast Cancer Detection by Mammography, Breast Self Exam, and Physical Exam Services utilized and detection costs in a cohort of 100,000 for three intervention strategies: Single Detection Modalities on All Women Age 45+

Service	Unit Cost	A	B	C	D
BSE Minimal Instruction	$ 1	20,000	20,000	20,000	20,000
BSE Interactive Instruction	5	0	0	0	90,000
Screening Mammogram	30	2,720	90,000	2,720	2,720
Diagnostic Mammogram	100	2,716	771	3,769	4,502
Screening Physical Exam	5	30,925	30,825	89,783	30,799
Physical Exam for Symptoms	35	1,973	560	1,609	4,156
Consult for Suspicious Mammogram	50	136	2,430	136	136
Biopsy with Surgical Consult	500	1,002	780	1,002	1,002
Total Cost		$1,104,680	$3,482,325	$1,491,530	$1,809,055

Intervention Strategies

A Baseline: No systematic intervention
B Annual mammography for all women; 90% participation
C Physical Exam annually for all women; 90% participation
D BSE instruction annually all women; 90% participation

TABLE II - Breast Cancer Detection by Mammography, Breast Self Exam, and Physical Exam
Detection costs and mortality reduction in a cohort of 100,000 for three intervention strategies
Single Detection Modalities on All Women Age 45+

	A	B	C	D
Total Cost	$1,104,680	$3,482,325	$1,491,530	$1,809,055
Incremental Cost (relative to A)		$2,377,645	$ 386,850	$ 704,375
Expected Deaths Annually	56.2	40.3	51.2	50.2
Deaths Averted (relative to A)		15.9	5.0	6.0
Marginal Detection Costs per Death Averted (relative to A)		$ 149,537	$ 77,370	$ 117,396
Percent Mortality Reduction		28.3	8.9	10.7
Relative Percent Mortality Reduction (relative to B)		100.0	31.4	37.7
Percent of Incremental Cost (relative to B)		100.0	16.3	29.6

Intervention Strategies
A Baseline: No systematic intervention
B Annual mammography for all women; 90% participation
C Physical Exam annually for all women; 90% participation
D BSE instruction annually for all women; 90% participation

TABLE III - BREAST CANCER CASES BY DETECTION MODALITY. Baseline: No Screening Program

	Cancer		Stage Distribution*	
	Number	Percent	Early	Late
Mammography & Screening	85	4.1	78.3	21.7
BSE	431	20.7	58.1	41.9
Physical Exam	358	17.2	54.1	45.9
Self (accidentally)	1191	57.2	50.4	49.6
Other	18	.8	46.7	53.3
All Methods	2083	100.0	53.7	46.3

* To be consistent with the SEER staging distribution, the percent early was calculated from Huguley and Brown's data as the percent in stage 0 or I plus .54 times the percent in stage II. This approximation yields an overall percent early stage of 53.7%, equivalent to that estimated from local SEER data. (Huguley CM and Brown RL: Cancer 47:989-995, 1981.)

TABLE IV - Breast Cancer Detection by Mammography, Breast Self Exam, and Physical Exam
Detection costs and mortality reduction in a cohort of 100,000 women age 45+ for four intervention strategies: Synchronous Combinations of Detection Modalities at Varying Intervals

	B	G	H	I
Total Cost	$3,482,325	$3,801,960	$4,309,865	$2,140,885
Incremental Cost (relative to A)	$2,377,645	$2,697,280	$3,205,185	$1,036,205
Expected Deaths Annually	40.3	40.0	39.6	47.9
Deaths Averted (relative to A)	15.9	16.2	16.6	8.3
Marginal Detection Costs per Death Averted (relative to A)	$ 149,537	$ 166,499	$ 193,083	$ 124,844
Percent Mortality Reduction	28.3	28.8	29.5	14.8
Relative Percent Mortality Reduction (relative to B)	100.0	101.9	104.4	52.2
Percent of Incremental Cost (relative to B)	100.0	113.4	134.8	43.6

Intervention Strategies
B Annual mammography for all women; 90% participation
G Synchronous mammography and physical exam annually
H Synchronous mammography, physical exam and BSE instruction annually
I Synchronous mammography, physical exam, and BSE instruction at three-year intervals for all women; 90% participation

TABLE V - Breast Cancer Detection by Mammography, Breast Self Exam, and Physical Exam
Detection costs and mortality reduction in a cohort 100,000 women age 45+
for three intervention strategies: Selection of High Risk Women for Annual Mammography

	A	B	E	F
Total Cost	$1,104,680	$3,482,325	$1,246,180	$1,319,435
Incremental Cost (relative to A)		$2,377,645	$ 141,500	$ 214,755
Expected Deaths Annually	56.2	40.3	48.2	52.9
Deaths Averted (relative to A)	15.9	15.9	8.0	3.3
Detection Costs per Death Averted (relative to A)		$ 149,537	$ 17,688	$ 65,077
Percent Mortality Reduction		28.3	14.3	5.9
Relative Percent Mortality Reduction (relative to B)		100.0	50.3	20.8
Percent of Incremental Cost (relative to B)		100.0	6.0	9.0

Intervention Strategies

A Baseline: No systematic intervention
B Annual mammography for all women; 90% participation
E Annual mammography for high-risk women only; 90% participation; relative risk of high-risk women = 5.0
F Annual mammography for high-risk women only; 90% participation; relative risk of high-risk women = 2.0

TABLE VI - Breast Cancer Detection by Mammography, Breast Self Exam, and Physical Exam
Detection costs and mortality reduction in a cohort of 100,000 women age 45+
for four intervention strategies: Staggered Combinations of Screening Modalities

	B	J	K	L
Total Cost	$3,482,325	$2,130,075	$2,250,965	$2,207,045
Incremental Cost (relative to A)	$2,377,645	$1,025,395	$1,146,285	$1,102,365
Expected Deaths Annually	40.3	44.1	43.3	42.2
Deaths Averted (relative to A)	15.9	12.1	12.9	14.0
Marginal Detection Costs per Death Averted (relative to A)	$ 149,537	$ 84,743	$ 88,859	$ 78,740
Percent Mortality Reduction	28.3	21.5	23.0	24.9
Relative Percent Mortality Reduction (relative to B)	100.0	76.1	81.1	88.0
Percent of Incremental Cost (Relative to B)	100.0	43.1	48.2	46.4

Intervention Strategies

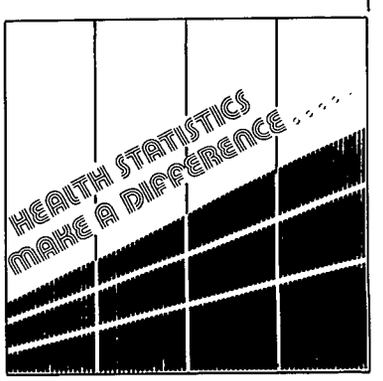
B Annual mammography for all women; 90% participation
J Mammography, physical exam and BSE instruction at staggered three-year intervals for all women; 90% participation
K Combination of annual mammography for high-risk women with rr = 2.0 and staggered mammography, PE, and BSE instruction at three-year intervals for remaining women
L Combination of annual mammography for high-risk women with rr = 5.0 and staggered mammography, PE, and BSE instruction at three-year intervals for remaining women

Acknowledgements

The authors wish to acknowledge the research assistance of Tierney Norsted, who performed the survival analyses, and the secretarial support of Le Anne Symonds, whose excellent skills on the word processor were much appreciated. We also wish to thank Emily White for helpful comments.

Special Session 3 (HRSA)

Some Analytical Perspectives on the Changing Supply of Physicians



PHYSICIAN DISTRIBUTION - URBAN AREAS STILL GET MORE THAN THEIR SHARE.

C. Howard Davis, Health Resources and Services Administration

INTRODUCTION

Since the mid-1970's, the environment for the delivery of health services changed appreciably. The supply of physicians has expanded 27.5 percent, far outpacing the population growth of 8.9 percent. The number of professionally active physicians per 10,000 population in the U.S. rose from 15.5 in 1970 to 17.4 in 1975 and to 19.7 in 1980. This ratio is expected to increase further to 21.9 by 1985. Much of this increase was stimulated by policies implemented by Congress in 1963.

With the expanding physician supply, there is increasing evidence that market forces have been effective, although with some lag, in increasing the number of physicians in formerly less well-served areas. Yet, despite an increase in the overall physician-to-population ratio, concern is still expressed, and indeed corroborated, that certain geographical areas and population groups are not participating in the increasing spread of physicians. Nevertheless, the extent that diffusion has or has not occurred may be obscured by definitional problems.[1/]

This study attempts to add further to our understanding of the geographical spreading of medical doctors in response to market disequilibrium caused by a rapid increase in physicians relative to population between 1970 to 1980. Selected States were focused upon as study areas. Counties within a State were used as units of observation. The change in physicians in the county-aggregated urban areas was compared to the non-urban areas. Also regression analysis was performed, using counties as observations, to determine if causal factors can be identified that influence the location of physicians and if such factors are common among States.

The literature reviewed in a lengthier version of this paper supports the contention that geographical spreading of physicians, including some types of specialists, has indeed been taking place. This spreading, or diffusion, has been occurring in response to an increasing supply of physicians as would be expected in efficiently functioning manpower markets. Yet, there is evidence that some areas do not provide sufficient incentives to participation in this sharing and will persist as medically underserved areas for an indefinite time.

Traditionally, proportionately more physicians specializing in family medicine have selected smaller urban and more rurally oriented communities in which to practice than have physicians trained in other medical specialties. The American Academy of Family Physicians provides further evidence from their annual surveys of graduating family practice residents that this tradition is continuing. These surveys, covering the years 1980 through 1984, record the responses of these graduating residents regarding the size of the community in which they intended to serve.

While numerically more residents planned to establish practice in urban than in rural areas, the annual rate of increase was under that for rural areas. (3.2 percent vs. 8.9 percent) Within the general urban category, the thrust of the growth was in the "In urbanized areas" (4.7 percent per year), and in towns between 2500 to 25,000 in population (3.5 percent per year).

A comparison of the trends of residents with population trends emphasize the differential shift towards the rural areas. In making this comparison, an assumption is made that the momentum of the growth patterns manifested between 1970 to 1980 carried through 1984. The favorable differential is evident toward those areas defined as outside urban but included within urbanized areas. Such areas include small towns and cities. The results also indicate a differential change in favor of the rural area relative to all urban areas.

There are two aspects to the movement of physicians: the absolute number and the number relative to a change in population. Large numbers of physicians can be expected to continue congregating in large metropolitan areas in which there is already a sizable concentration of physicians. Physicians engaged in patient care include the spectrum of specialties. Highly specialized physicians types need access to adequate supporting health services and to be located in the center of a large market area that can sustain their activity. If a large proportion of the movement of physicians is represented by newly graduating medical students, then many will be assuming residencies at large urban hospitals. Also, in the perception of individual physicians specializing in family or internal medicine, establishing new practices in large urban areas would not significantly adversely partition the market. If, however, they were to locate in communities of, say, 10,000 in which there were already three practicing physicians in the same or cognate specialties, then the population-to-physician ratio would be lessened rather dramatically, from 3333 to 2500, a reduction of 25 percent. For the above reasons as well as their putative amenities, larger urban areas can be expected to draw large numbers of physicians.

The differential movement of physicians relative to population change is at least as important as the absolute movement of physicians. Is the increase of physicians keeping pace with population growth or might the number of physicians in a community be diminishing faster than its population is declining? Even though the geographical spreading of physicians may be taking place, its pace may not be proportional to population growth in the nonurban areas or even among smaller urban areas. Consequently, the population-to-physician ratio would increase.

While the literature provides strong evidence of geographical spreading or diffusion, the process is not uniform. Some communities are able to attract more while others attract fewer than their proportional share of physicians. Can we improve our

understanding of the factors which conduce to a proportionally significantly greater increase of physicians in certain communities? The authors of the NCHSR study conclude that "underserved areas with poor prospects for economic or population growth will not attract physicians.[2]

METHODOLOGY

Individual States were examined because the author feels that there is a uniqueness common to individual States imposed by socio-legal factors. Some of these factors are State licensure requirements, a professional and collegiate attachment to the institution in which residency was done, and a predisposition to locate in an area with which one is both familiar and comfortable. Because of legal requirements, movement within a State in response to economic factors may be easier than between States. After the individual States were studied, the States were compared. Indeed, as discussed below, covariance analysis supports the supposition of uniqueness among individual States.

Three States were selected for examination: Tennessee, North Carolina, and Pennsylvania. Tennessee was chosen because the author attended college there and has an interest in the State's development. Both Tennessee and North Carolina had relatively greater growth in physician supply than the rest of the Country between 1970 and 1982. North Carolina and Tennessee had increases of 73.7 and 64.2 percent, respectively, compared to 50.2 for the total U.S. In contrast, Pennsylvania increased only 38 percent. North Carolina was one of the first States to have an Area Health Education Center Program (AHEC) and Pennsylvania is an industrialized State with a relatively small rural population. As in many research endeavors, resource constraints limit sample size and selection.

Physicians engaged in patient care were selected as the measure of physician supply. A Relative Measure was used as an indicator of comparative change in physician supply among areas by relating physician growth to population growth. A value larger than one indicates improvement while the opposite indicates a deterioration in the population-to-physician ratio. An improvement can be produced either by a fall in population greater than the relative reduction in the number of physicians or by an increase in physicians proportionally greater than in population.

Counties, although used as units of observations, do not represent the most conceptually desirable units of observation; their use is a compromise with reality. Counties neither necessarily comprise homogeneous socioeconomic characteristics nor constitute an integrated market. A small county may appear as a medically underserved county. This might not actually be the case for most of its residents when the service plexus is considered. If adjacent to a major urban area, this county may be totally integrated into the latter area with respect to the services and commodities purchased. For instance, the residents of a small county that is essentially a bedroom community to an urban center may obtain a substantial part of their medical needs from that urban center. Despite its having a very high population-to-physician ratio as a county, the residents of such a county may have adequate access to physicians. Also, a county encompassing a larger physical area may have most of its population living immediately contiguous to a major urban area located in another county. Much of the remainder of this county may not be easily accessible because of physically difficult terrain. As in the previous illustration, the socioeconomic characteristics of most of its residents would be influenced by the adjoining urban area. However, the residents in the remainder of the county may be socially, economically, and educationally insulated from the influence of that urban area.

Nevertheless, in spite of these and other drawbacks, counties are useful observational units. They are clearly defined geographic and legal entities, migration across county boundaries is relatively easy, and much of the published socioeconomic data are conterminous with their boundaries.

Since statewide averages obscure important deviations, it is useful to look at changes in counties and in other identifiable areas. Certainly not all counties in a State participate in an overall improvement. Researchers can speculate upon the circumstance fostering an improvement in certain counties but not in others. For instance, is specific area improvement linked to causal socioeconomic factors in addition to the relative improvement in the State's overall population-to-physician ratio? If so, then population in many geographic areas that lack propitious factors may not readily share the benefits of a greater number of physicians.

Regression analysis was used to identify socioeconomic determinants that significantly affect the relative change in medical doctors engaged in patient care (subsequently referred to as physicians). The relative change is the dependent variable and is labeled as RMDPTN.[3]

The independent variables were as follows:

(a) The population to patient care physician ratio [4] in 1975 (SEVMD 75), which may indicate opportunity as much as medical need

(b) Relative change in retail trade income from 1975 to 1980 [5] (RRETRADE)

(c) Per capita income in 1980 (CAPINC80)

(d) Relative change in workers employed in manufacturing from 1975 to 1980 (RMFGWKRS)

Area population-to-physician ratios were found by Mathematica Policy Research to be positively and significantly associated with (1) shortage county designations, (2) mean distance to a doctor, and (3) other measures of access. This variable also effectively controls for population size. Relative change in retail income reflects the development in the capacity of the area to sustain tertiary activity. Per capita income controls for the capability of an area to afford and maintain the facilities and amenities consequential to a sophisticated urban area. Relative change in

manufacturing employment indicates the growth of income generated from an export activity. This variable had been expected to be positively associated with the increase in physicians. This was not the case, however. I am speculating that those counties which showed large relative gains in manufacturing employment started at a low level and had not yet gained a sufficient population, income, and infrastructure to support a tertiary activity.

Attention was focused on the 1975 to 1980 period since the increase of physicians in each State examined was considerably larger than during the prior five year period. The interest was on the process of assimilation of these physicians. These physicians selected their practice sites based on judgments formed regarding the favorableness of an area to a new or additional practitioners. These judgments were presumably based on conditions existing or perceived to be existing at the time of their selection based upon trends apparent in the immediately prior time period. If newly settling physicians are influenced by the contemporary environment of an area, then implicitly a five year interval captures the lagged response of physicians to changes in the influential determinants. Indeed, the impetus of changes occurring in the two to three year period in an area prior to the observed 5 year period would extend the response time to include these additional years. It is significant that the population-to-physician ratio and the degree of economic vitality are important determinants.

TENNESSEE

URBAN AREAS

The years 1975 to 1980 saw the assimilation of about two-thirds (1,542) of the 1970 to 1980 increase of physicians (2,229) of which the Major Urban Areas (MJUAs) absorbed 77 percent. Certainly, if the distribution pattern of physicians were to be affected by a large increase relative to the population, it should have been evident over this period 1975 to 1980.

Indeed, the distribution was affected, both between the Rest of State (R of St.) and the MJUAs as well as within the MJUAs. The Relative Measure for the minor urban areas increased moderately compared to the relatively large gain in MDs in patient practice. This resulted from a relatively stronger population growth compared to the other two areas. Even then, much of the improvement in the Relative Measure occurred in only three of the Minor Urban Areas. Nevertheless, the Relative Measure for the R of St. posted a better gain relative to the MJUAs than during the 1970 to 1975 period. The uniformity exhibited among the MJUAs between 1970 to 1975 no longer held.

There was a substantial improvement in the ratio in the Tri-Cities. Chattanooga and Nashville both displayed a relatively greater improvement than did either Knoxville or Memphis. The greater comparative improvement in the Tri-Cities might have been in response to the larger population-to-physician ratio. There was, however, no correspondence between this ratio and the degree of improvement among the other MJUAs. Nor did the improvement in the ratio appear to be causally related merely to population growth, although the Tri-Cities area did post the largest gain between 1970 to 1975.

REST OF STATE

The population residing outside of the MJUAs amounted to 47 percent (2,158,864) of the State in 1980. This part of the population was divided into three parts, those residing in: (1) minor urban areas; (2) the remainder of the state; and (3) the medically underserved areas. Counties were assigned into the minor urban areas and labeled according to the town or city which seemed to be their central influence.

In aggregate, the population growth of the counties comprising the minor urban areas was nearly twice that of the MJUAs and slightly greater than that of the remainder of the state. The gain in physicians was substantially higher. As a consequence, the Relative Measure showed considerable improvement compared to either the MJUAs or the R of St. There was considerable spread in the Relative Measure, with Peripheral Nashville and McMinnville showing the largest improvement. Yet, a comparison of the Relative Measure among the minor urban areas with their respective changes in population again demonstrate that the relative increase in physicians could not be ascribed merely to population growth.

Further differentiating the remainder of the State into adequately medically served and medically underserved areas casts some additional light on the diffusion of physicians. In the context of this paper, the ratio of population-to-physicians engaged in patient care in excess of 3500 is used to differentiate between adequately and inadequately medically served counties. While the medically adequately served counties gained physicians, although certainly much fewer in numbers than did the major urban areas, the tempo of population increase (19 percent) was well above that of either the major or minor urban areas, which caused the Relative Measure of improvement to increase less than in either of the two types of urban areas.

The counties that were medically underserved in 1975 actually lost three physicians (57 in 1975 compared to 54 in 1980). But since population paced an even larger decline, there was a moderate improvement in the Relative Measure. The condition in these counties would have been much less improved were it not for a rather considerable gain of physicians in Hawkins County. Hawkins, although still classified as medically underserved in 1980 with a population-to-physician ratio of 3,645.9, is adjacent to the Tri-Cities area and began to share in its population gain, especially between 1975 to 1980. Its population (43,751 in 1980) rose nearly 7 percent over this period while the number of physicians increased from 8 to 12, a gain of 50 percent.

REGRESSION ANALYSIS

If population change is not a major determinant of the movement of physicians, are their other identifiable factors that serve to influence their location? This section

presents the results of regression analysis in which an attempt was made to identify certain socioeconomic characteristics that have had a significant influence on the relative change in the number of physicians.

In Tennessee, only three of the 95 counties were assigned a weight of 1/2 of a physician, a proxy for no physicians present 3/. The R-square was .735; that is, nearly 3/4 of the variation was explained. The coefficient of each of the independent variables was statistically significant. Multicollinearity appears not to be a problem; nor is the significance of any of these variables affected by any outlier observations. SEVMD__75, RRETRADE, and CAPINC80 each had the expected signs for their coefficients. The sign of RMFGWKRS was negative, which was not as anticipated. The reason for the negative value is considered below.

By far the most important explanatory variable was SEVMD__75. Physicians were attracted to counties with high population-to-physician ratios. By, itself, this variable was responsible for 84 percent of the total sum of squares accounted for by the model. Per capita income was important and indicated that physicians located in counties in accordance with their respective degree of affluence. More affluent counties afforded better income potentials to physicians.

The increase of physicians also was positively related to the change in retail trade income (RRETRADE). The RRETRADE variable reflects the economic activity of an area. Change in this variable was positively correlated with both changes in population and changes in per capita income. The importance of the change in retail trade may be related not only to the vitality of the tertiary sector but also may be a proxy for all the trappings of a sophisticated urban life style correlative with an economy having a level of income necessary to sustain tertiary activity.

One might infer that the growth of retail trade is linked to the establishment of new shopping centers with concomitant office space and a generally good location with respect to population density. These factors might be perceived by physicians to constitute a favorable environment. The relative change in the employment of manufacturing workers was negatively correlated with the dependent variable, an unexpected result. Yet the degree of urbanization during this time period was negatively related to both the growth of manufacturing employment and of manufacturing income and the growth of manufacturing income was negatively correlated with the change in retail trade. What then seems to be the case is that the lesser urbanized counties were expanding their manufacturing activity but had not yet attained a population or income level that would produce a viable tertiary level of economic activity. In many of these counties with a very small manufacturing base, a comparatively small numerical increase can have a substantial impact on relative growth.

One county (Morgen) presented itself as an extremely influential outlier. The number of physicians grew from zero to 4 in 1980. The county posted a slightly greater than statewide population growth, but its relative growth in manufacturing employment and income was each comparatively strong. Its border is adjacent to Anderson County (Oak Ridge) and to Roane County on the South. Roane is a growing county with 23 physicians serving a population of 48,425 in 1980. Undoubtedly, that part of Morgen County is experiencing a spill-over. (Excluding Morgen did lower the R-square but did not materially affect the statistical significance of the coefficients of the independent variables.)

PENNSYLVANIA

Pennsylvania had a relatively favorable population-to-physician ratio in 1970. Over the decade, population increased by 0.6 percent while medical doctors engaged in patient care (MDPTNs) rose 27 percent. The larger increase in MDPTNs was between 1975 to 1980, during which the gain in MDPTNs was more than twice that of the 1970 to 1975 period. As a consequence, the population-to-physician ratio dropped appreciably more than between 1970 to 1975.

Pennsylvania is predominantly an urban state. Although both the number and the share of persons living in non-urban areas have increased over the decade, those areas contained less the 19 percent of the State's population. There were only three counties (not all the same) in each reference year which were medically underserved. Because of the change in composition of the counties, the largest population in these counties was no more the 0.7 percent of the total State population and less the 4 percent of the non-urban population. By 1980, these percentages were 0.5 percent and 2.5 percent, respectively; perhaps of rather trivial concern except for the persons residing in these counties. Perry County, which contained nearly 36 thousand of the 55,000 persons residing in those three counties had a population-to-physician ratio of 3,572, just barely above the accepted cutoff level. Even most of Fulton County was within close proximity to Franklin County, which is quite adequately served.

Population-to-physician ratios among the 14 urban areas varied from a low of 538 to a high of 1,550 in 1970, and gains over the decade in their Relative Measures were not particularly uniform. These 14 urban areas comprised 28 of the 67 counties in the State. The counties constituting the remaining part of the State faired less well than did 12 of the 14 urban areas. Can we detect circumstances promoting an improvement in certain areas or counties and not in others?

About 71 percent of the total 1970 - 1980 increase of physicians was absorbed into the total stock of physicians during the second part of the decade. While increasing in both the total urban areas and the Rest of State, concentration increased comparatively more in the total urban areas because population gained slightly in the Rest of State but declined very slightly for all urban areas.

Among the urban areas, the relative increase in MDPTNs was lower in the areas that experienced population losses. Also, there was

somewhat more uniformity in the 1975 - 1980 Relative Measures, with the measures for Allentown - Bethlehem - Easton, Northeast, and Williamsport areas predominating. The Allentown - Bethlehem - Easton had the highest population-to-physician ratios in 1975 while that of the Northeast was moderately high. Philadelphia and Pittsburg together gained 57 percent of the total 1975 to 1980 increase in physicians. This was so even though the two areas each had a population loss and had comparatively low population-to-physician ratios.

REGRESSION ANALYSIS

The R-square was significant but not very strong, accounting for approximately 15 percent of the total variation. Only two of the four variables were statistically significant. The condition indices obtained in the collinearity diagnostics showed a rather high value. This high value, however, is related to the strength of the association between the change in retail trade income and the slope of the intercept, not with the other independent variables. The variance inflation factors were not unduly influenced.

RRETRADE and CAPINC80 were both positively correlated with the RMDPTN variable with the first variable being only somewhat more influential. Thus, physicians appeared to be attracted to economically dynamic areas as evidenced by increased retail trade and per capita income. The variables SEVMD__75 and RMFGWKRS were not significant.

NORTH CAROLINA

Population in North Carolina rose just under 16 percent between 1970 and 1980. The increase was 7.3 percent during the initial five years and 7.9 percent during the second five years. The number of physicians increased 21.4 percent (992) during the first half of the decade and 35.9 percent (2,023) during the latter half of the decade. The increase amounted to 64.9 percent over the entire decade.

The population-to-physician ratio fell from 1093.3 in 1970 to 966.7 in 1975 and then plunged to 767.7 in 1980. The statewide improvement in the population-to-physician ratio accrued only moderately to the medically underserved counties. In 1970, there were 17 of these counties with a population of 340,400, about 6.7 percent of the total State population, with a total of 67 physicians. In 1975, there were still 17 medically underserved counties with 51 physicians serving 269,500 persons (about 4.9 percent of the population). By 1980, these counties numbered 14 with 52 physicians serving 281,621 persons (about 4.8 percent of the State's population).

The average population-to-physician ratio in these counties rose successively from 5080.6 in 1970, to 5284.3 in 1975, and 5415.8 in 1980. Furthermore, there remained a set of the same seven medically underserved counties in each of the years. These counties contained 1.8 percent of the total State population. In 1975, three additional counties fell into the classification and their presence persisted into 1980.

Virtually all these medically underserved counties are in the eastern coastal plain of North Carolina and nearly all are within the proximity of an urban area with a relatively favorable population-to-physician ratio. Thus, medical service to residents of these counties can be presumed to be available although with perhaps some inconvenience. There was no county in the medically underserved category in the mountainous Western part on North Carolina.

Over the period 1975 to 1980, population increased proportionately slightly less in the minor urban areas than in either the MJUAs or the Rest of the State. Physicians increased slightly more in the minor than for the major urban areas but considerable less than for the Rest of the State. Thus, the Rest of the State experienced a much larger gain in the Relative Measure than did either the major or minor urban areas. Among the minor urban areas, Cherry Point, Plymouth (which still had only 6 physicians in 1980), Rocky Mount, and Henderson all experienced comparatively large gains in their Relative Measures.

REGRESSION ANALYSIS

The R-square is significant but rather weak, accounting for only 10 percent of the total variation. Only one of the four variables (SEVMD__75) was statistically significant. The condition indices obtained in the collinearity diagnostics were within acceptable bounds and the variance inflation factors were not especially large. Thus, all that can be inferred from this regression is that physicians do move to areas with high population-to-physician ratios.

SOME GENERALIZATIONS

The geographic spreading depends on the location decisions of physicians. Physicians locate their practices in accordance with their preferences for an agreeable living environment and with their perception of the income potential. These preferences and perceptions are influenced by a number of attributes which include population growth, population density, per capita income, economic vitality, libraries, hospitals and other health facilities, and professional, social, and economic amenities. Population growth reflects the economic vitality of an area and is less important as an explanatory variable than are the other measures of economic vitality. Although tested as an explanatory variable, population growth was not significant.

In the context of this study, geographic spreading appears to have taken place since 1975. The nature of process differs among the three States examined in this study. The Rest-of-State category (excluding both minor and major urban areas) in Tennessee and Pennsylvania experienced a lesser increase in their respective Relative Measures than occurred in the major and minor urban areas. In contrast, the major and minor urban areas in North Carolina realized a slower pace in the Relative Measure than in the Rest-of-State category. This contrast underscores the lack of uniformity among States.

Regression analysis indicated that the ratio of population-to-physicians in 1975 was positive and significant as a predictor of physician increases in Tennessee and North Carolina but not in Pennsylvania. Other

variables were also influential, although to a lesser extent and to different degrees among the three States. Changes in income generated in retail trade and in per capita income were significant in Tennessee and in Pennsylvania. What is suggested is that an area (county) must exhibit some degree of economic vitality, in addition to other requisite factors such as adequate population density and the availability of hospitals, to attract physicians. Many medically-underserved areas with a significant total population are not likely to experience such economic vitality and thus are unlikely to attract physicians. Consequently, these areas will remain medically underserved despite an adequate overall number of physicians.

Merging the counties for all three States and contrasting these results with those obtained using a covariance analysis is instructive. Performing a regression analysis on all 262 observations with RMDPTN as the dependent variable produces an R-square of 48 percent, which is highly significant. Each of the four variables is significant. CAPINC80 was significant overall despite its lack of significance for North Carolina. RRETRADE, significant in the overall regression as well as for Tennessee and Pennsylvania.

Covariance analysis indicated that common slopes for SEVMD_75 and CAPINC80, despite their significance, could not legitimately be computed across States. There is a difference between the State-specific slopes of each of these two variables attributable to the FIPS variable (which is a class variable identifying the state). Only in the case of RRETRADE was the slope significant and a possible commonality of the slope among the States could not be rejected. The variable RMFGWKRS was not significant. This result is not surprising since this variable was not significant for either North Carolina or Pennsylvania. The dependent variable (RMDPTN) was not State specific.

Policy decisions derived from the analysis of the merged observations may not be applicable to individual States. The response of the spreading process can not be considered uniform among States with respect to the dependent variables tested with the merged observations. Instead, the process is differentially determined by State-specific parameters. Inferentially, this finding is relevant to models based upon observations (county or otherwise) aggregated to a national level. If some of the more important variables are uniquely affected by the environment peculiar to individual States, the nationally determined parameters may not be appropriately applied to regional, State, or other disaggregated levels.

CONCLUSION

In the three states examined, the follow conclusions seem warranted:

1. Geographic spreading is taking place, i.e., an increasing number of physicians are spreading out to serve areas outside of the major urban areas.
2. As between minor urban areas and the rest-of-state groupings, the change in physicians relative to the population was not uniform either within or among the three States.
3. Also, the population-to-physician ratios have been compressed over the ten year period, from 1970 to 1980 and especially since 1975.
4. Although in two of the three States, the ratio of population-to-physician ratio was a significant determinant, the county also needed to exhibit economic viability to attract physicians.
5. The spreading process may be unique among States, depending upon the environment of the individual State.

REFERENCES AND FOOTNOTES

1/ Hickson, Jesse, L. Jackson Brown, and Stephen J. Boehlert. "The Geographic Diffusion of Physicians and Dentists," unpublished paper, Modeling and Research Branch, Bureau of Health Professions, Health and Resources and Services Administration, Room 8-55, Parklawn Building, Rockville, MD, 20857. These authors distinguish between two definitions of diffusion: (a) the process of spreading freely over a large area; and (b) the process of equilization among areas of differing degrees of concentration. They note that the two concepts involve distinctly different processes. They point out that the distribution of primary care physicians and dentists towards a greater equilization is conditioned upon the distribution of economic demands for their services. See pages 4 to 5.

2/ U.S. Department of Health and Human Service, The National Center for Health Services Research. Physicians Supply and Distribution: Issues and Options for State Policy Makers. Lewin and Associates (September 15, 1983).

3/ A technical problem arose in employing the dependent variable. A few counties had no physicians in 1975; thus, the gain of one physician represented an infinite increase. To circumvent this problem, a county with zero MDPTN's in 1975 was arbitrarily assigned 1/2 of a physician.

4/ Area population-to-physician ratios were found by Mathematica Policy Research to be positively and significantly associated with (1) shortage county designations, (2) mean distance to a doctor, and (3) other measures of access. See "Physician Capacity Utilization Surveys: Project Summary (October, 1978), U.S. Department of Health, Education, and Welfare, Public Health Service, Human Resources Administration, Bureau of Health Manpower, DHEW Publication No. (HRA) 79-17.

5/ Lawlor and Reid cite a study by J. W. Leyes which found "a close relationship existed between the level of retail services and the level of basic health care services in several hundred Wyoming communities." See Lawlor, Ann C. and Jack T. Reid. "Hierarchical Patterns in the Location of Physician Specialists Among Counties." Inquiry (Spring, 1981) Vol. 18, page 79.

IMPACT OF PRODUCTIVITY CHANGES ON 1990 PHYSICIAN REQUIREMENTS AND ANTICIPATED SURPLUS

Karen Rudzinski and Jerald Katzoff, Health Resources and Services Administration

INTRODUCTION

Since the publication of the final report of the Graduate Medical Education National Advisory Committee (GMENAC), there has been widespread acceptance of the impending "glut" of physicians in the marketplace (Graduate Medical Education National Advisory Committee, 1981). According to revised estimates of that work, a surplus of nearly 63,000 physicians is projected (Bowman, Katzoff, Garrison, et al., 1983). In an effort to curb the undesirable consequences of this perceived surplus, some have advocated decreasing medical school class sizes and the availability of residency positions as well as further restricting the entry of foreign medical graduates into the country.

Recently the size of this projected surplus has been a subject of debate, stimulated by the declining productivity of physicians (Freiman and Marder, 1984). In the earlier work, physician requirements were calculated based upon the division of 1990 total population "norms of care" by the expected annual productivity of physicians. Productivity for each specialty was defined in terms of one of the following measures: total patient visits, total ambulatory visits, total work hours, or total patient hours per week, multiplied by the expected average number of weeks worked per year. Hence, if productivity were lower than originally estimated, the requirements would be greater and the overall resulting surplus would be substantially reduced, if not eliminated.

This paper presents a methodological sensitivity analysis of productivity on physician normative requirements. The intent is to isolate changes in the original specialty-specific physician surplus/shortage estimates under differing productivity assumptions in order to demonstrate the degree to which these estimates are sensitive to productivity changes. As such, the study does not represent an "update" of the entire original work. Furthermore, the study does not indicate the results ensuing from it will occur. The study will extrapolate aggregate specialty-specific trend data on productivity using four differing methods in order to discern the degree to which potential changes in productivity alter normative requirements.

BACKGROUND

Between 1970 and 1980, physician hours worked per week declined 3 percent or 1.5 hours. In their critique of the original work, Freiman and Marder translated this decline in productivity into 8,000 non-Federal physicians leaving the labor force. The authors further stated that a continual productivity decline of only a few hours per week would virtually eliminate the projected surplus. Hence, implicit in these comments is that the original work overestimated 1990 productivity and, consequently, overstated the extent or even the presence of a future physician surplus. Thus, while the original work implicitly considered such factors as gender, lifestyle, practice setting and nonphysician providers in the derivation of its productivity estimates, the full impact of these factors, which are briefly discussed below, may not have been taken into account.

Women generally work fewer hours than men. In 1983, women MDs worked 7.9 percent fewer hours and saw 18.9 percent fewer patients than men (American Medical Association, March 1984). Further, women MDs saw 37 percent fewer patients per hour than men (Langwell, 1982). While the productivity of women increased over the 1970-80 decade, the lower levels of productivity of women physicians, coupled with their greater relative growth in supply (27.9 percent increase compared to an 11.1 percent total supply increase between 1980-83) (American Medical Association, 1984), are generally expected to produce a continuing decline in future physician productivity.

Lifestyle, in combination with practice setting changes, is also expected to result in lower productivity levels, especially given the increase in the supply of young physicians, which grew by 51.6 percent over the past decade (American Medical Association, November 1984). Young physicians are more likely to practice in group settings. Physicians in these settings, who prefer more predictable work schedules, work fewer hours and see fewer patients (Goodman and Swartwout, 1984).

Substitution of ancillary personnel can also offset productivity levels. The nonphysician/physician substitutability ratio approached 0.5 to 0.75:1, when visits were used as an indicator of productivity (Record et al., 1979). It should be noted that this substitutability only holds for visits which are delegable. However, substitution may result in both increases and decreases in the productivity of physicians. An increase in productivity may stem from their decreased time involvement per patient, but a decrease in productivity may result from the increased time involved in supervision.

The size of the physician supply can also be expected to affect productivity. Cross-sectional analyses have pointed to a decline in productivity in areas with larger physician supplies (Davis, 1982; Mathematica Policy Research, 1980). Data from 1975 to 1983 also show a leveling of total patient visits beginning in 1982, and a slight decline observed between 1982-83. The same data also indicate that total visits per physician decreased by 5.2 percent from 1982 to 1983, whereas prior to that period, the decrease observed approximated only 0.8 percent annually (American Medical Association, October 1984).

Despite the observed trends, the empirical relationship between supply and productivity has never been firmly established. For example, physician contact declined more between 1959 and 1964 than was observed through part of the 1970s, although physician supply grew more in the latter period

(Wilson and Begun, 1977). Others reported a decline in visits to a Health Maintenance Organization over time, although the MD/enrollee ratio remained constant (Luft and Trauner 1981). Similarly, at the national level, one researcher found that a doubling in physician supply would be needed to increase physician-initiated visits by 1 percent (Willensky, 1982).

In the study presented here, no attempt will be made to separate the above types of influences on productivity or explicitly adjust the productivity estimates. However, assuming that each factor plays a role in modifying physicians' productivity and that the original work implicitly incorporated these factors in the determination of productivity estimates for each of the specialties, this analysis will compare physician requirements with supply after incorporating four productivity adjustments to the original estimates.

METHODS

Data

Part of the data used in this analysis is obtained from the 1990 specialty-specific supply, requirements and productivity estimates utilized in the original and revised work (Graduate Medical Education National Advisory Committee, 1981; Bowman, Katzoff, Garrison, et al., 1983). These data have been used in conjunction with productivity trend adjustments based on 1970, 1978 and 1980 data (American Medical Association, 1984). The AMA data are used since the original Delphi Panels heavily relied on 1978 AMA productivity estimates in the determination of future productivity changes, and the critique by Freiman and Marder cited changes in productivity as reported by the AMA. Whereas the Freiman and Marder paper used "time" as a measure of productivity, this analysis measures annual productivity in terms of one of the following: total patients visits, total ambulatory visits, total work hours and total patient care hours, in accordance with the specific measure utilized for each specialty in the original work.

Analysis

The basic mode of analysis used in this study parallels the original approach. This analysis accepts the 1990 norms of care component and adjusts only the productivity denominator, based on four different assumptions.

The basic formula used in the analysis is as follows:

Physician Requirements = 1990 Norms of Care/Productivity

In this study, changes are made only to the denominator. If productivity declines, based on any of the methodologies, total requirements will increase, and the total projected surplus will decrease. Conversely, if productivity increases, based on any of the four methodologies, requirements will decrease, and the projected surplus will increase.

Since some specialties were deemed in a shortage in the original study, productivity increases for these specialties would result in a decrease of the shortage. Conversely, productivity decreases in these specialties may point to an exacerbation of the impending projected shortage.

Specialties were aggregated according to the level at which the AMA publishes productivity estimates. Hence, subspecialties were grouped respectively within the overall specialties of internal medicine, pediatrics or surgery. Consequently, the resulting requirements and corresponding 1990 shortages or surpluses projected in the original work and in the revisions, as displayed in this paper, are aggregates; the pediatric and internal medicine subspecialties are subsumed within the larger specialties of general internal medicine and general pediatrics. In particular, the projected original surplus of nearly 21,000 internists in 1990 is primarily due to the subspecialists who constitute over 17,000 of the surplus.

Also, the 1990 subspecialty productivity measures used in the original work were weighted and aggregated based on the projected 1990 supply and productivity of each specific subspecialty within the larger aggregate specialty. The average AMA productivity changes and the estimated average 1990 productivity levels in the original work were used in the absence of published AMA specialty-specific productivity estimates. As a product of these calculations, revised productivity estimates were developed for nine specialties: internal medicine, family/general practice, pediatrics, obstetrics/gynecology, psychiatry, surgery, anesthesiology, radiology, and all others.

The different 1990 productivity adjustments are as follows:

1. Application of the 1980 productivity levels as published by the AMA. This alternative is considered to be the most conservative since no future decline is incorporated.

2. Application of the 1970 to 1980 relative (percentage) change in productivity, as reported by the AMA. This alternative is anticipated to produce a large decline in overall physician surplus.

3. Application of the 1970 to 1980 absolute change in productivity, as reported by the AMA. This alternative is expected to have the most significant impact on the originally projected surplus. As the baseline levels of productivity continually shrink, large reductions in the projected surplus can be expected.

4. Application of an average absolute change in productivity estimated as the midpoint between the observed 1970-80 AMA reported change and the originally projected 1978-90 change (the latter was adjusted by a factor of five-sixths in order to equalize time spans). This change assigns equal weights to the recent observed changes and to the normative changes incorporated in the original work.

RESULTS

Total Requirements and Surplus Changes

The revised original requirements estimates (Bowman, Katzoff, Garrison, et al., 1983) produced a 1990 physician requirements of 473,000 and a projected surplus of nearly 63,000 physicians. Utilizing the 1980

productivity estimates yields a lower 1990 requirements estimate of 466,600. Thus, the original productivity estimates were lower than those observed in 1980. In fact, if productivity levels were to stabilize at the 1980 level, the projected surplus of physicians would increase from 62,750 to 69,150, a growth of nearly 10 percent, other factors in the model held constant.

On the other hand, a projection of a continuation of the large relative decrease in productivity observed between 1970 and 1980 produces a decline in the surplus to 22,950. A continuation of the absolute declines in productivity observed between 1970 and 1980 would more significantly decrease the projected surplus, by nearly 82 percent to 11,100.

A decline in productivity based on the weighted midpoint between the 1970-80 AMA reported productivity change and the one incorporated in the original time-adjusted expected estimate, results in a 1990 surplus of 39,500 physicians, nearly a 37 percent decline from the revised projected surplus. This result indicates that the overall productivity decline from 1978 to 1990 originally projected is less than that observed between 1970 to 1980. If future productivity were to continue to decline at this rate, the projected overall surplus of physicians would decline but not disappear.

Thus, a projected surplus remains even assuming a continuation of significant declines in productivity. The requirements/supply ratio varies from a low of 0.87:1 to a high of 0.98:1. Nevertheless, the aggregate surplus is sensitive to future changes in productivity. The sensitivity is particularly demonstrated in the specialty-specific analysis.

Specialty Specific Requirements and Surpluses (Shortages)

The productivity of physicians is not uniform for all specialties. The actual productivity estimates used in the original work are lower than those derived using the four modifications for several specialties. In particular, in surgery, all four modifications result in an increase in the surplus. In surgery, the projected surplus may actually be 14 to 22 percent greater than what was originally projected. Consequently, one-third of all the 1990 projected supply will be in excess of requirements.

In radiology, the revised surplus estimates are also higher, but not to the same degree. In fact, anesthesiology and radiology generally exhibit similar levels of shortage (anesthesiology) and surplus (radiology) as that projected earlier, regardless of the productivity level used. The only exception is if 1980 productivity levels were to continue in anesthesiology. A continuation of 1980 patterns in anesthesiology would result in more than a 40 percent reduction in or a virtual elimination of the shortage for this specialty.

In psychiatry, the original productivity estimate is also lower than that used with any of the four modifications. However, since psychiatry was projected to be in shortage, the higher productivity levels derived here result in a lesser shortage. Thus, the 12,900 shortage of psychiatrists in 1990 may decrease by 24 to 40 percent.

For three specialties, general/family practice, internal medicine and obstetrics/gynecology, the original productivity estimates were actually higher than those used in the four modifications. Since all three specialties were previously forecasted to be in surplus or near balance, lowering the productivity estimates substantially decreases the projected surplus. In fact, if other than a continuation of the 1980 productivity levels is assumed for internal medicine and general/family practice, shortages in both specialties may occur, as great as 16,850 and 13,000 respectively under the most liberal productivity assumption. These shortages translate into respective requirements/supply ratios of 113.1:1 and 120.2:1.

If productivity in internal medicine and general/family practice is forecasted to decline more modestly (i.e., midway between the observed 1970-80 change and the originally projected 1978-90 change), a shortage of approximately 9,000 would still exist for the latter specialty. The ratio of requirements to supply in general/family practice may be as high as 114:1.

Particular note should be paid to internal medicine. Supply projections used in the original study indicate that subspecialists will comprise approximately 43 percent of all internists and their ambulatory productivity levels are projected to be lower - 53.4 visits - than that of general internists - 80.0 visits. Subspecialists were also previously seen to comprise the bulk of the original projected oversupply in internal medicine. As a result of the above, the finding of a surplus of 5,000 internists using the midpoint of 1970-1980 and adjusted 1978-1990 productivity changes (method 4), may actually indicate a potential shortage of general internists. If subspecialization rates in internal medicine training stabilize at their most recent reported levels - 60 percent (Schleiter and Tarlov, 1985), an exacerbation of this undersupply of general internists may ensue.

The specialty of obstetrics/gynecology was originally projected to be facing a surplus of over 10,000 physicians by 1990. Current productivity trends in this specialty are declining, which would result in possible productivity levels lower that that originally projected. Consequently, lowering the productivity levels for this specialty may decrease the projected surplus from 10,450 to between 2,700 and 8,200.

The alternative productivity assumptions used in this paper utilize 1980 data which indicate an unusually low level of productivity for obstetrician/gynecologists. In 1978, the average total visits for this specialty was 126.2, and in 1980 it was 113.3. Indications that this 1980 estimate may be atypical are supported by the 1981 and 1982 AMA productivity estimates for obstetrics/gynecology which approach the level found in 1979 (American Medical Association, 1984).

Another interesting specialty comparison

can be made by observing the relative parity between requirements and supply. Regardless of the productivity estimate employed, near parity between requirements and supply is observed for all specialties not disaggregated. Again, large disparities in requirements and supply are found for psychiatry, anesthesiology (although the absolute numbers for anesthesiology are fairly small), obstetrics/gynecology, general/family practice, surgery, and, to a degree, internal medicine.

CONCLUSION

A key factor that determines the future magnitude of a surplus of physicians is the set of 1990 productivity levels utilized. Recently, the productivity levels of physicians have been declining. A continuation of declines was hypothesized by some observers to result in an elimination of the projected surplus.

In this paper, the utilization of various productivity levels in the recalculation of aggregate specialty-specific requirements has shown that the requirements are highly sensitive to productivity changes. Results also indicate that the projected surplus of physicians into the next decade will not be completely eliminated under any of the four scenarios employed in this paper. However, under the assumption of a continuation of the 1970 to 1980 absolute productivity decline, the originally projected surplus of physicians will decline substantially.

Specialty-specific comparisons indicate that surpluses may be more likely to occur in obstetrics/gynecology, radiology, and surgery. Specialty-specific shortages may remain in psychiatry, anesthesiology, and may occur in general/family practice, and in general internal medicine.

A caveat to these results should be noted. In the GMENAC effort, as well as in the sensitivity analysis discussed in this paper, productivity estimates are derived independently of either the (economic) demand for or supply of physician manpower. In technical parlance, productivity is determined "exogenously" or outside, of the needs-based requirements model. Yet, in actuality, physician productivity (as measured by office visits per physician) may be affected by the interplay of the supply and demand for physicians If physician supply increased relatively faster than the total number of office visits generated by the population, the average number of office visits per physicians will diminish. If this happens, productivity (as defined above) will decrease. A lessening of the average number of office visits would not, however, necessarily indicate specific states of shortage, surpluses, or balance.

What we can say based on this study is that positing different levels of projected productivity of physicians due to some of the factors described previously (i.e., changing gender composition, changing lifestyles, etc.) will change the requirements and resulting requirements - supply assessments, all other factors held constant.

References

American Academy of Family Practice: Report on Survey of 1982 Graduating Family Practice Residents. AAFP Reprint No. 155H, July 1982.

American Medical Association: Changing Medical Practice Arrangements. SMS Report November 1983, 3:7:1-4.

American Medical Association: Characteristics of Young Physicians - 1982. Excerpts from the American Medical Association Masterfile November 1984.

American Medical Association: Differences in Practice Characteristics Between Female and Male Physicians. SMS Report March 1984;3:2:1-4.

American Medical Association: Socio-Economic Characteristics of Medical Practice - 1984. Chicago, IL: American Medical Association, 1984.

American Medical Association: Trends in Patient Visits with Physicians: 1975-1983. SMS Report October 1984; 3:7:1-4.

Black R, Schmittling G, Stern TL: Characteristics of Practice Patterns of Family Practice Residency Graduates in the United States. Journal of Family Practice 1980; 11:767-778.

Bowman MA, Katzoff, JM, Garrison LP, et al.: Estimates of Physician Requirements for 1990 for the Specialties of Neurology, Anesthesiology, Nuclear Medicine, Pathology, Physical Medicine and Rehabilitation, and Radiology. JAMA 1983; 250:2623-2627.

Davis K: The Impact of Physician Supply on Health Care Costs. Ballman, MD: Johns Hopkins University, 1982 (mimeographed).

Freiman MP, Marder WD: Changes in the Hours Worked by Physicians, 1970-80. AJPH 1984; 74:1348-1352.

Geyman JP: The Emerging Role of the Residency-Trained Family Physician. The Journal of Family Practice 1980; 11:717-719.

Geyman JP, Ciriacy EW, Mayo F et al.: Geographic Distribution of Family Practice Residency Graduates: the Experience of Three Statewide Networks. Journal of Family Practice 1980; 11:761-766.

Glenn JK, Hofmeister RW: Rural Training Settings and Practice Location Decisions. Journal of Family Practice 1981; 13:377-382.

Goodman LJ, Swartwout JE: Comparative Aspects of Medical Practice: Organizational Setting and Financial Arrangements in Four Delivery Settings. Med Care 1984: 22:255-267.

Graduate Medical Education National Advisory Committee: Report of the Graduate Medical Education National Advisory Committee to the Secretary, Department of Health and Human Services VOL II: Modeling, Research and Data Technical Panel. Washington, DC: Government Printing Office, 1981.

Hecht RC, Farrell JG: Graduate Follow-Up in the University of Wisconsin Family Practice Residency Programs. Journal of Family Practice 1982; 14:549-555.

Langwell K: Factors Affecting the Increases of Men and Women Physicians: Further Explorations. Journal of Human Resources 1982; 17:261-275.

Luft H, Trauner J.: The Operations and Performance of Health Maintenance Organizations: A Hypothesis of Findings from

Health Services Research. National Center for Health Services Research, 1981.

Mathematica Policy Research: The Physician Capacity Utilization Surveys: A Report on the Telephone Resurvey. DHEW Publication No. (HRA) 80-18, 1980.

Record JC et al.: Final Report: Provider Requirements, Cost Savings and the New Health Practitioner in Primary Care: National Estimates for 1990. Health Resources Administration, Bureau of Health Manpower, Contract No. (HRA) 231-77-0077, 1979.

Schleiter, MK and Tarlov, AR: National Study of Internal Medicine Manpower: IX. Internal Medicine Residency and Fellowship Training: 1984 Update. Annals of Internal Medicine 1985, 102:681-685.

Willensky G: New Information about Physician-Initiated Demand. Washington, D.C.: The National Health Policy Forum, June 29, 1982.

Williams, DC: Surgeons and Surgery in Rhode Island, 1970 and 1977. New England Journal of Medicine 1981; 305:13139-1323.

Wilson G, Begun JW: Trends in Physician's Patient Volume. Inquiry 1977; 14:171-175.

PROJECTIONS OF PHYSICIAN SUPPLY IN THE U.S.*

John Drabek
Bureau of Health Professions

Summary and Highlights of Projections

The Nation's supply of active physicians (MDs and DOs) grew substantially over the past decade rising from 326,200 in 1970 to 467,000 in 1981. Although much of this increase was a result of increased capacity of U.S. schools of allopathic and osteopathic medicine, foreign medical schools also contributed significantly to this growth.

In order to predict the impacts of this increased supply of physicians in future years the Bureau of Health Professions has developed a computer model which generates projections of several aspects of the physician supply. The model indicates that:

- The number of active physicians will continue to increase over the next two decades but at a slower rate than in the past decade.
- Although the growth in physician supply is expected to slow down during the next two decades, this growth is still expected to substantially exceed population growth.
- Foreign medical graduates are projected to contribute less to the growth in active physician supply during the next two decades than in the previous decade.
- Women physicians are projected to continue to increase substantially both numerically and as a percentage of the total physician supply.
- Although the number of physicians in the primary care specialties is projected to increase substantially, the percentage of primary care physicians among the total active physician supply is projected to change by only a small amount between 1981 and 2000.

National Projections of Active Physician Supply in 1990 and 2000

The supply of active physicians (MDs and DOs) in 1990 is projected to range from 592,600 to 608,200 with a "most likely" estimate of 594,600 (see Table 1). By the year 2000 the Nation can expect to have from 695,800 to 749,900 physicians with a "most likely" estimate of 706,500. In the most likely series of estimates this amounts to a growth in physician supply of 27 percent between 1981 and 1990 and 51 percent between 1981 and 2000. While the supply of physicians is expected to grow at a slower rate during the next two decades than in the last decade, this growth is expected to be substantially greater than the growth in the population which is expected to amount to 16 percent between 1981 and 2000. Consequently, the 1981 ratio of 199 physicians per 100,000 population is projected to increase to 235 by 1990 and 260 by 2000.

U.S.-trained physicians (MDs and DOs) are projected to account for a larger percentage of growth in active physician supply over the next two decades than in the previous decade. Approximately 86 percent of the increase in physician supply between 1981 and 2000 is expected to result from the increase in U.S.-trained MDs and DOs. Although foreign medical graduates are projected to contribute substantially less to the growth in physician supply in the coming decades than in the last, they will continue to comprise a substantial percentage of the total active physician supply. Approximately one in five active physicians in 1990 and 2000 is expected to have received their training in medical schools outside the U.S. and Canada. As is further shown on this table, the supply of U.S. born FMGs are projected to increase at a substantially higher rate than the total supply of foreign-trained physicians (81 percent compared with 35 percent by the year 2000). Consequently, U.S. born foreign medical graduates are projected to comprise an increasing percentage of the total supply of foreign medical graduates. In 1981, 12 percent of all foreign medical graduates were U.S. born. By the year 2000 this proportion is expected to increase to 16 percent.

The number of Doctors of Osteopathy is projected to more than double between 1981 and 2000. Yet their numbers are expected to remain relatively small and they are expected to comprise only 6 percent of the total active physician supply in 2000 compared with 4 percent in 1981.

Female physicians are projected to number approximately 141,000 by the year 2000, an increase of more than 150 percent from the number in 1981. As a result of this continued rapid increase in the number of female physicians, approximately one out of five physicians in the year 2000 is projected to be a woman compared with about one out of eight in 1981. Thirty-six percent of the projected growth in physician supply from 1981 to 2000 is expected to result from the increase in the supply of women physicians.

Specialty - The numbers of physicians are projected to increase in all but a few specialty categories during the next two decades. In terms of broad specialty groups, the largest increases are projected to continue to occur in the medical specialties other than the primary care specialties (see Table 2). On average, specialties in this category are projected to increase 79 percent from 1981-2000. Increases in the number of primary care specialists during this period are projected to average 55 percent. The surgical specialties are projected to grow the least, as a category with a projected average increase of only 36 percent.

Among the primary care specialties, the retirement of older general practitioners results in a below average growth rate for general and family practice. Although internal medicine and pediatrics are likely to grow at a higher than average rate, pediatrics is projected to grow at a faster rate than internal medicine.

Although no attempt has been made to forecast the demand for care, in those areas where it appears that the demand for care will increase substantially in future years the model predicts significant increases in specialists (see Table 3). Cardiovascular disease, gastroenterology, and pulmonary disease specialists are likely to be in demand as the diseases treated by these specialists are expected to increase. Plastic surgery is projected to increase in response to increased demand for such operations. Diagnostic radiology is expected to experience further substantial growth since these physicians are heavily involved with the recent technological developments.

On the other hand, certain specialties which face particularly intense competition are projected to grow at slower rates. For example, general surgeons may provide a mix of routine patient care and surgical care. They may experience substantial competition from the increased supply of primary care providers and from the surgical subspecialties.

The percentage distribution of physicians by specialty is projected to remain relatively constant over the next two decades as it did in the last decade. However, because the growth in the number of physicians in most specialties is projected to exceed population growth the ratios of practitioners to the population in most specialties are expected to increase. For example, while the percentage of MDs in the primary care specialties (excluding Ob/Gyn) is expected to remain stable around 40 percent during the interval from 1981 to 2000 the ratio of primary care MDs to the population is projected to increase from about 77 per 100,000 population in 1981 to 103 per 100,000 by 2000.

The projected slowdown in growth of the foreign medical graduate supply is expected to be accompanied by a stabilization of the distribution by specialty. During the period from 1970 to 1980 there was a notable increase in the percentage of FMG physicians in the primary care specialties. However, during the next two decades the percentage of FMGs in the primary care specialties is projected to remain constant at 38 percent (excluding Ob/Gyn) or 45-46 percent if Ob/Gyns are counted in the primary care totals. Similarly, the percentages in other specialties are projected to show very little change from 1981 to 2000.

The significant increase in the number of women physicians projected for the next two decades is expected to be reflected in substantial increases in the number of women in most specialty categories. On the whole, the surgical specialties and medical specialties other than primary care are projected to show the largest gains in female practitioners. The number of female physicians in these broad categories is projected to triple in the interval from 1981 to 2000. Surgical specialists are projected to comprise a slightly larger proportion of female physician supply in the year 2000 than in 1981, increasing from 13 percent to 16 percent. Concomitantly, the percentage in "other" specialties is projected to decline from 34 to 30 percent in the period between 1981 and 2000. Like total physician supply, the percentage in primary care specialties is projected to stabilize at about 50 percent for the next two decades, or at about 60 percent if obstetrician-gynecologists are included in the primary care totals.

State Level MD Projections

Projections of the number of active MDs by State are given in Table 4. These projections are generated by allocating the expected national growth in the supply of MDs across the States based on the location patterns exhibited by MDs as of December 31, 1981. Since there are relatively few restrictions on the ability of young physicians to locate their practices as economic conditions may have considerable influence on their location choices. The figures in Table 4 should be interpreted as projections of the trends that existed in 1981, which could be substantially affected by changes in the attractiveness of the States.

Over the period 1981-1990 the U.S will experience a 26 percent increase in the supply of MDs. However, the individual States are expected to demonstrate a wide range in their growth rates--from 19 percent in Michigan to nearly 66 percent in Alaska. Although State population growth is not explicitly considered in the model, this pattern parallels the growth rate of the general population with States in the South and West usually growing at greater rates than those in the North Central and Northeast.

Outlook for the Future

It is evident that physicians in all specialties will experience much greater competition than they do at the present time because the physician supply will continue to increase in future years at a more rapid rate than the demand for care. It is unlikely that this conclusion will change because of supply side developments. All of the physicians who will practice in 1990 are either enrolled in or have already graduated from medical school. Even if there were a substantial decrease in applicants to medical schools late in the 1980s, there will still be many more applicants than spaces. The substantial decrease in applicants to medical schools late in the 1980s, there will still

be many more applicants than spaces. The basic projection assumes that some schools will reduce enrollments. But, this does not yet appear to be happening to any significant extent. There were 16,395 first-year medical students enrolled in the 1984-1985 academic year, which is only three percent below the peak of 16,910 students in 1981. The physician supply projections in the year 2000 are not particularly sensitive to changes in first-year enrollments. The basic projection already assumes a 5 percent decrease in enrollment and the low alternative with a 10 percent decrease in enrollment only reduces the projected supply of U.S. graduates in 2000 from 527,900 to 518,200--a difference of 9,700 physicians or 1.9 percent.

*For a more detailed discussion of the BHPr physician supply model and projections see: Projection of Physician Supply in the U.S., March 1985, ODAM Report No. 3-85

Table 1 - Projected Numbers of Active Physicians by Country of Medical Education, Estimated 1981, Projected 1985-2000, Basic Series

	1981	1985	1990	1995	2000	Percent Change in Supply 1981-1990	Percent Change in Supply 1981-2000
All Active Physicians	467,000	527,900	594,600	653,800	706,500	27.3	51.3
MDs	449,000	506,000	566,900	620,500	667,900	26.3	48.8
U.S. Trained	343,300	387,100	439,300	485,400	527,900	28.0	53.8
Canadian Trained	7,000	7,000	7,000	7,100	7,100	--	--
Foreign Trained	98,700	111,900	120,500	128,100	133,000	22.1	34.8
U.S. Born	11,600	16,900	18,200	19,800	21,000	56.9	81.0
DOs	18,000	21,900	27,800	33,300	38,600	54.4	114.4
Total U.S. Trained	361,300	409,000	467,100	518,700	566,500	28.7	55.7
			Rate Per 100,000 Population				
All Active Physicians	198.8	217.8	234.5	248.0	259.6	18.0	30.6
MDs	191.0	208.8	223.6	235.4	245.4	17.1	28.5
U.S. Trained	146.2	159.7	173.3	184.1	194.0	18.5	32.7
Canadian Trained	3.0	2.9	2.8	2.7	2.6	-6.7	-13.3
Foreign Trained	42.0	46.2.	47.5	48.6	48.9	13.1	16.4
U.S. Born	4.9	7.0	7.2	7.5	7.8	46.9	59.2
DOs	7.8	9.0	11.0	12.6	14.2	41.0	82.1
Total U.S. Trained	153.8	168.8	184.2	196.8	208.2	19.8	33.4
			Percent Distribution				
All Active Physicians	100.0	100.0	100.0	100.0	100.0		
MDs	96.2	95.9	95.3	94.9	94.5		
U.S. Trained	73.6	73.3	73.9	74.2	74.7		
Canadian Trained	1.5	1.3	1.2	1.1	1.0		
Foreign Trained	21.2	21.2	20.3	19.6	18.8		
U.S. Born	2.5	3.2	3.1	3.0	3.0		
DOs	3.9	4.1	4.7	5.1	5.5		
Total U.S. Trained	77.4	77.5	78.6	79.3	80.2		

Source: Health Resources and Services Administration, Bureau of Health Professions
Rates based on population estimates from Current Population Reports Series P-25 No. 900 and 925 and unpublished estimates of the civilian population in the outlying areas.

Table 2 - Number of Active Physicians (MDs)[1/] by Specialty and Percent Change, Estimated 1981 and Projected 1990 and 2000

Specialty	1981	1990	2000	Percent Change 1981-1990	Percent Change 1981-2000
Total	448,800	566,660	667,790	26.3	48.8
Primary Care	180,210	233,030	279,630	29.3	55.2
General and Family Practice	65,600	77,910	89,270	18.8	36.1
Internal Medicine	82,020	109,420	132,470	33.4	61.5
Pediatrics	32,590	45,710	57,890	40.3	77.6
Primary Care with Ob/Gyn	209,390	270,260	323,510	29.1	54.5
Other Medical Specialties	28,340	39,660	50,620	40.0	78.6
Allergy	1,640	1,670	1,700	1.6	3.8
Cardiovascular Disease	10,730	14,850	18,860	38.4	75.8
Dermatology	6,000	7,880	9,630	31.2	60.4
Gastroenterology	4,600	7,330	9,960	59.6	116.8
Pediatric Allergy	450	420	370	-7.3	-18.0
Pediatric Cardiology	750	1,110	1,430	48.6	91.8
Pulmonary Diseases	4,180	6,420	8,670	53.8	107.6
Surgical Specialties	121,210	145,480	164,370	20.0	35.6
Colon and Rectal Surgery	740	790	860	6.5	15.8
General Surgery	37,990	42,710	45,920	12.4	20.9
Neurological Surgery	3,600	4,370	4,930	21.3	36.8
Obstetrics and Gynecology	29,180	37,230	43,880	27.6	50.4
Ophthalmology	13,680	16,590	19,090	21.3	39.5
Orthopedic Surgery	15,200	19,090	22,170	25.6	45.9
Otorhinolaryngology	6,870	7,890	8,620	14.9	25.5
Plastic Surgery	3,370	4,810	6,110	42.7	81.5
Thoracic Surgery	2,280	2,580	2,780	12.8	21.8
Urology	8,310	9,430	10,010	13.5	20.5
Other Specialties	119,048	148,490	173,180	24.8	45.5
Aerospace Medicine	740	860	920	16.3	24.8
Anesthesiology	18,400	22,550	25,710	22.6	39.7
Child Psychiatry	3,540	4,620	5,610	30.5	58.6
Diagnostic Radiology	8,820	13,760	18,470	56.1	109.5
Forensic Pathology	260	290	320	10.0	21.2
General Preventive Medicine	890	970	1,030	9.4	15.8
Neurology	6,510	9,590	12,460	47.5	91.5
Occupational Medicine	2,500	2,260	2,020	-9.7	-19.3
Psychiatry	30,250	35,300	38,980	16.7	28.9
Public Health	2,520	1,900	1,290	-24.6	-48.8
Phys. Medicine and Rehabilitation	2,570	2,950	3,220	14.5	25.1
Pathology	15,050	18,620	21,330	23.7	41.7
Radiology	12,040	12,370	11,880	2.7	-1.4
Therapeutic Radiology	1,830	2,390	2,860	30.8	56.4
Other Specialties	13,130	20,070	27,080	52.9	106.3

1/ These figures differ from those published by the AMA since they reflect adjustments to include approximately 90 percent of the physicians who are not classified according to activity status by the American Medical Association and whose addresses are unknown.

NOTE: Figures may not add to totals due to independent rounding.

Source: Health Resources and Services Administration, Bureau of Health Professions

Table 3 - Growth Rates by Specialty Ranked According to Expected Percent Change 1981-1990

Rank	Specialty	Percent Change 1977-1981	1981-1990	1981-2000
1	Gastroenterology	70.4	59.6	116.8
2	Diagnostic Radiology	102.5	56.1	109.5
3	Pulmonary Diseases	61.5	53.8	107.6
4	Other Specialties	14.8	52.9	106.3
5	Pediatric Cardiology	22.4	48.6	91.8
6	Neurology	34.5	47.5	91.5
7	Plastic Surgery	28.7	42.7	81.5
8	Pediatrics	23.5	40.3	77.6
9	Cardiovascular Disease	44.3	38.4	75.8
10	Internal Medicine	21.6	33.4	61.5
11	Dermatology	18.7	31.2	60.4
12	Therapeutic Radiology	33.0	30.8	56.4
13	Child Psychiatry	13.5	30.5	58.6
14	Obstetrics and Gynecology	16.4	27.6	50.4
	ALL SPECIALTIES	18.5	26.3	48.8
15	Orthopedic Surgery	18.3	25.6	45.9
16	Pathology	12.1	23.7	41.7
17	Anesthesiology	21.0	22.6	39.7
18	Neurological Surgery	13.9	21.3	36.8
19	Ophthalmology	14.4	21.3	39.5
20	General and Family Practice	9.9	18.8	36.1
21	Psychiatry	14.6	16.7	28.9
22	Aerospace Medicine	-2.8	16.3	24.8
23	Otorhinolaryngology	9.2	14.9	25.5
24	Physical Medicine and Rehabilitation	31.4	14.5	25.1
25	Urology	10.6	13.5	20.5
26	Thoracic Surgery	-2.9	12.8	21.8
27	General Surgery	7.1	12.4	20.9
28	Forensic Pathology	30.6	10.0	21.2
29	General Preventive Medicine	7.4	9.4	15.8
30	Colon and Rectal Surgery	14.8	6.5	15.8
31	Radiology	-10.3	2.7	-1.4
32	Allergy	-0.2	1.6	3.8
33	Pediatric Allergy	-19.1	-7.3	-18.0
34	Occupational Medicine	22.1	-9.7	-19.3
35	Public Health	-7.4	-24.6	-48.8

Table 4 - Active Physicians (MDs) by Geographic Region, Division, and State and Percent Change Estimated 1981 and Projected 1990 and 2000

	Number of Physicians[1]			Percent Change	
	1981	1990	2000	1981-1990	1981-2000
UNITED STATES[2]	448,800	566,600	667,630	26.2	48.8
NORTHEAST	117,650	144,190	166,120	22.6	41.2
NEW ENGLAND	30,420	39,400	47,450	29.5	56.0
Connecticut	7,940	9,990	11,750	25.8	48.0
Maine	1,670	2,180	2,660	30.5	59.3
Massachusetts	16,070	21,180	25,820	31.8	60.7
New Hampshire	1,560	1,980	2,320	26.9	48.7
Rhode Island	2,050	2,480	2,890	21.0	41.0
Vermont	1,130	1,580	2,020	39.8	78.8
MIDDLE ATLANTIC	87,230	104,800	118,670	20.1	36.0
New Jersey	15,060	17,900	19,970	18.9	32.6
New York	48,480	57,220	63,670	18.0	31.3
Pennsylvania	23,690	29,680	35,030	25.3	47.9
NORTH CENTRAL	100,620	124,080	143,680	23.3	42.8
EAST NORTH CENTRAL	71,670	87,230	99,900	21.7	39.4
Illinois	22,440	27,580	31,850	22.9	41.9
Indiana	7,340	8,930	10,310	21.7	40.5
Michigan	15,330	18,290	20,480	19.3	33.6
Ohio	18,580	22,450	25,680	20.8	38.2
Wisconsin	8,000	9,970	11,580	24.6	44.8
WEST NORTH CENTRAL	28,950	36,850	43,780	27.3	51.2
Iowa	3,940	4,820	5,610	22.3	42.4
Kansas	3,960	4,980	5,800	25.8	46.5
Minnesota	8,160	11,010	13,680	34.9	67.6
Missouri	8,540	10,610	12.410	24.2	45.3
Nebraska	2,520	3,150	3,670	25.0	45.6
North Dakota	960	1,210	1,410	26.0	46.9
South Dakota	870	1,060	1,200	21.8	37.9
SOUTH	131,640	167,350	198,030	27.1	50.4
SOUTH ATLANTIC	72,680	91,870	108,260	26.4	49.0
Delaware	1,050	1,250	1,420	19.1	35.2
District of Columbia	4,010	5,030	5,920	25.4	47.6
Florida	18,540	22,490	25,480	21.3	37.4
Georgia	8,604	10,860	12,760	26.2	48.3
Maryland	13,320	17,590	21.260	32.1	59.6
North Carolina	9,470	12,120	14,480	28.0	52.9
South Carolina	4,460	5,480	6,260	22.9	40.4
Virginia	10,330	13,510	16,510	30.8	59.8
West Virginia	2,890	3,550	4,170	22.8	44.3
EAST SOUTH CENTRAL	21,100	26,960	32,080	27.8	52.0
Alabama	5,230	6,500	7,510	24.3	43.6
Kentucky	5,250	6,940	8,460	32.2	61.1
Mississippi	2,960	3,810	4,590	28.7	55.1
Tennessee	7,640	9.720	11,530	27.2	50.9
WEST SOUTH CENTRAL	37,880	48,520	57,690	28.1	52.3
Arkansas	2,970	3,850	4,650	29.6	56.6
Louisiana	6,870	8,650	10,170	25.9	48.0
Oklahoma	4,220	5,560	6,920	31.8	64.0
Texas	23,820	30,460	35,950	27.9	50.9
WEST	93,490	123,310	150,610	31.9	61.1
MOUNTAIN	20,240	27,640	34,500	36.6	70.5
Arizona	5,220	6,840	8,190	31.0	56.9
Colorado	6,220	8,520	10,570	37.0	69.9
Idaho	1,100	1,550	2,020	40.9	83.6
Montana	1,130	1,490	1,830	31.9	61.9
Nevada	1,200	1,730	2,210	44.2	84.2
New Mexico	2,200	3,180	4,250	44.5	93.2
Utah	2,560	3,500	4,370	36.7	70.7
Wyoming	620	840	1,080	35.5	74.2
PACIFIC	73,250	95,670	116,110	30.6	58.2
Alaska	640	1,060	1,550	65.6	142.2
California	57,430	74,090	89,010	29.0	55.0
Hawaii	2,140	2,800	3,390	30.8	58.4
Oregon	4,980	6,820	8,570	36.9	72.1
Washington	8,070	10,900	13,590	35.1	68.4

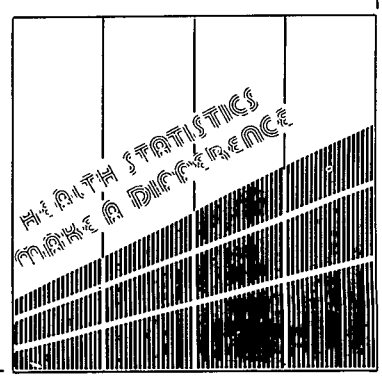
1/ These figures include about 90 percent of those MDs not classified according to activity status by the American Medical Association.

2/ Includes physicians in the U.S. Possessions.

Source: Health Resources and Services Administration, Bureau of Health Professions

Session V

Health Surveillance Systems in Action



The Development of a Model Disease Surveillance System for State Health Departments

Denton Peterson, Karen White, and Craig Hedberg
Minnesota Department of Health

Recent developments in the computer industry have made it possible to consider the application of computer technology to areas that have previously had limited access to this technology. An example of such an application is the operation of reportable disease surveillance systems by state health departments. Disease surveillance is basically an information processing application, and similar systems in other industries have been shown to increase dramatically in efficiency and effectiveness when computerized systems are applied. However, a recent survey (1983) found that less than one-third of the states had any access to computers and that only 2 states actually had computerized surveillance systems.

The Centers for Disease Control (CDC), Division of Surveillance and Epidemiologic Studies, have been involved in developing computer applications for disease surveillance for the past few years. They developed the Epidemiologic Surveillance Project (ESP), which had as one function the automated collection of weekly disease surveillance data from selected states. From these projects CDC recognized the need for states to have access to computer technology for state disease surveillance systems. They developed the Model Disease Surveillance System Program, which would fund the development of computerized surveillance systems in four states. Each state was to develop its own system according to general CDC guidelines. CDC planned to make the resulting systems and their components available to other states that were interested in developing automated surveillance systems.

The Minnesota Department of Health (MDH) was one of the four state health departments selected to develop a model surveillance system. The Division of Disease Prevention and Health Promotion is responsible for disease surveillance at MDH. Although a computerized surveillance system did not exist at the time of the award, computers had begun to be used quite extensively for data analysis and office support in recent years. The following paper will describe the planning and development work that has been completed on that system during the first year.

General Requirements of Surveillance Systems.

General CDC guidelines pertaining to model surveillance systems emphasized the comprehensive nature of disease surveillance at state health departments. They recognized that a computerized disease surveillance system should go beyond the record storage and analysis functions of surveillance to include all activities associated with operating a disease surveillance system. This included stressing that the model surveillance system should reinforce the existing mechanisms of disease surveillance information communication between state and local public health agencies and between the state and the CDC. In addition, states were encouraged to develop a communication interface with the general medical community.

Development of such a system requires several different types of computer applications technology. The most obvious pertain to data storage, data retrieval, and data analysis. Disease surveillance systems have large data storage and retrieval requirements. Database Management Systems (DBMS) are the application technology that can address these needs. DBMS allow for records and record structuring that is sufficiently detailed for surveillance applications, and DBMS are flexible enough for continued development of the surveillance systems records. DBMS data retrieval functions are ideal for analyzing and investigating subgroups of selected cases, a common disease surveillance activity. DBMS also contain the analytical tools to perform simple analysis and have the capability to interact directly with more advanced analysis software such as statistical or graphics packages. Advanced statistical analysis is a requirement of many disease surveillances and is impossible without computerized statistical packages. A number of statistical and graphics packages already exist which can be adapted to the surveillance application.

The communication components of a surveillance system require that an application include computer communications technology. Computer communications technology allows users to send data files and messages between computers or computer terminals. These systems use telephone lines to connect one location with another and communication software to provide the specific functions that make up computerized communications (electronic mail, separate user accounts, bulletin boards, etc.). Applying computerized technology to disease surveillance systems will facilitate the collection of disease surveillance data by making possible the transfer of data files between those collecting data and those

storing data. Computer communications also enhances communications between those involved in disease surveillance. Finally, computerized communication technology can disperse surveillance findings to the general medical community in a manner that is faster and more flexible then current methods. A requirement of computer communication technology is that anyone included in a communications network have access to a computer or a computer terminal that can connect to the communication system.

In addition to requiring that systems or programs perform the function of an automated surveillance system, and that equipment be available to run these programs, users who can operate and use this technology are also required. The users of this automated surveillance system are the existing epidemiologists and clerks who currently operate the manual system and those who would be included in the communications component of the system. Both groups are composed primarily of individuals not experienced in computer operations. A general requirement of an automated surveillance system is that these individuals be trained to use the system in an effective way. Related to training the users in systems operation is the need to develop automated systems that are easy to use and learn.

The Minnesota Model Surveillance System

The basic design strategy for developing the Minnesota model disease surveillance system was to base the system on the IBM PC family of microcomputers and to use existing, well supported, commercial software packages for developing the systems. The reasons for this decision included: 1. the size and power requirements of an automated surveillance system are within the performance range of this configuration, 2. the standardization of this configuration in the computer industry insures continued technical support and development, 3. commercial software products are developed for flexibility which makes them easy to modify for the surveillance application, 4. successful commercial programs have interfaces that are designed for easy use by inexperienced users, and 5. the components are low priced compared to most computer hardware and software.

The surveillance system itself is divided into four parts. These parts represent a division of the system into separate functions. The four parts are:

1. Automated Disease Surveillance System (ADSS)
   The centralized computer system which performs the data storage, data retrieval, and data analysis functions.
2. Local Epidemiologist Work Station (LEWS)
   A computer work station for epidemiologists working outside of the central office. Its function is to provide computer support to these individuals for their daily activity.
3. Minnesota Public Health Communication Network (MPHnet)
   Performs the electronic communication function for communication between public health professionals.
4. Surveillance Reporting and Communication System (SRCS)
   Public health dissemination and disease report collection sys-tem for all health professionals.

The Automated Disease Surveillance System (ADSS) functions as the data collection system for all reportable diseases. The collected data is stored in a database system developed with dBASEIII. The data structure is such that all diseases included in the system have a common formatted record of basic information with additional records used for specific diseases where more information is collected. Currently 82 diseases are included in the general portion of the surveillance and 1Ø of these diseases include additional information records. The ADSS system includes the analysis routines that are used to produce the reports generated by the system. Statistical software is being evaluated for eventual inclusion in the system.

ADSS currently interfaces directly to MPHnet and CDC. via an internal modem and the MINET network. It sends to the ESP at CDC, a weekly update on 32 diseases. This report is produced internally by the system and requires no manual intervention. MPHnet can be used to send surveillance information to local health departments or to MDH's outstate district offices. In the future ADSS will be able to collect surveillance data gathered at local health departments via transfer over the MPHnet network.

ADSS was developed on IBM PC and IBM XT computers that are connected together in an ETHERnet network. The IBM XT's hard disk stores the collected data. A streaming tape drive provides backup for the system. The system could be expanded by adding more systems to the network. It has been collecting all of the state's acute disease surveillance data since January 1, 1985. Program components that have not been completed include incorporating an advanced statistical analysis and graphics program in the system.

The model surveillance system's comprehensive goal requires that the system include all epidemiologists that contribute to collecting disease surveillance data.

The local epidemiologist workstation (LEWS) was conceived as a system to provide computer support for these individuals for outbreak investigations, general office support and communication of surveillance information.

The system's hardware was chosen so that the system would be usable in both a field or an office setting. This was done by basing the system on a portable microcomputer and including an extra full size monitor for long term viewing. An integral part of LEWS is a modem and communication software. It allows the LEWS user to use the work station for computer communications. Examples of communication functions which are used include contacting other epidemiologists using MPHnet, listing surveillance findings from ADSS, and sending surveillance data to ADSS.

LEWS can run the dBASE III programs which were developed for ADSS and give a local health department the option of the same type of surveillance system as the state system. Currently, software development on the system has been limited to only outbreak investigation software. A LOTUS 1,2,3 program has been developed to aid the epidemiologist foodborne outbreaks. It allows the investigator to develop a foodborne questionnaire which is automatically translated into a line listing when the data from the questionnaire is entered. This listing can be evaluated for any combination of symptoms (case definition) and significance tests generated on the associations.

The Minnesota Public Health Network (MPHnet) was developed to provide electronic communication between public health professionals and public health institutions in the state. Part of this communication relates to collecting surveillance data and communicating surveillance related information among those conducting disease surveillance in Minnesota. Its current status is that each MDH outstate district office has an account, 2 local health agencies and 8 separate offices within the Department. Plans call for all local health agencies to have access to this network eventually.

MPHnet was implemented on the Medical Information Network (MINET) developed by TELENET. MINET is a national medical network and has a large number of users among public health agencies. Using an established network, such as MINET, is the easiest way to implement a communication network. User accounts are assigned to each user and communication between two users can begin immediately. MINET can be accessed by any computer terminal or microcomputer with a modem. The flexibility of the user interface with the system is an advantage when a limited amount of equipment is available to users.

The function of the surveillance communication and reporting system (SRCS) is the disposition and collection of public health information including surveillance information. Unlike MPHnet, SRCS's audience is the entire health community, including both the public health and the non-public health sectors.

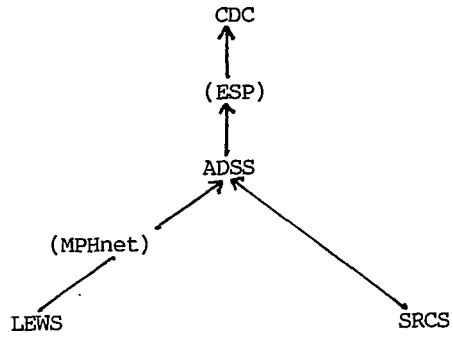
SRCS is conceived as a system that all health professionals could use to access certain types of public health information pertaining to surveillance or emergency public health recommendations such as rabies prophylaxis. The system uses menus to guide the users to the appropriate information that they are requesting. In addition, the system is designed so disease reports can be entered directly into the system by the user. The advantage to a system such as SRCS is that it provides a common known resource for public health information and has the capability to be immediately upgradeable for types of information that are rapidly evolving. An application in public health where these characteristics would be very important is AIDS, where there is a very high level of interest about a rapidly changing situation.

SRCS is also implemented on MINET. MINET was chosen because its Inform Script facility was designed for these types of applications. The drawback to developing SRCS on MINET is that only those who have access to a MINET account will have access to SRCS. Membership is growing in Minnesota, but it is still a small proportion of the intended audience. The current state of the SRCS system development is that the system code is being written and the system is expected to begin operation in late 1985.

The following two figures demonstrate the integration of the four components of the Minnesota Model Surveillance System. The first figure shows how surveillance data can be collected using the different systems that have been described. Local agencies that choose to have a locally based surveillance system could operate the system on LEWS and upload it via MPHnet to ADSS. Physicians would have the alternative of directly transmitting case reports to MDH via SRCS. Not shown on this diagram is the manual entry of disease reports collected in the current manner. This diagram also shows weekly transfer of surveillance data from ADSS to CDC via ESP.

Figure 1.

MODEL SURVEILLANCE INFORMATION COLLECTION

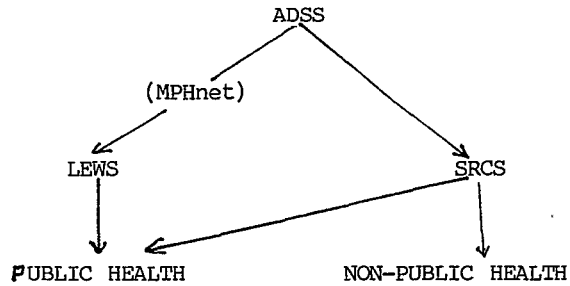


The second figure shows the dispersion of collected information using the model surveillance system. Information is shown to originate at ADSS following the analysis and evaluation of the collected data. Results that are pertinent to the SRCS system would be entered into the SRCS format and made available to the entire health community. Requests for surveillance information for local agencies and epidemiologists would be sent using the MPHnet communication network.

The four components that make up the model surveillance system address the functions of a comprehensive automated disease surveillance system. Partitioning of these functions into specific components is somewhat arbitrary and depends on the perspective of those developing the system and existing systems that may be incorporated into the system. The components represent a conceptual structure that is important for relating the different aspects of a comprehensive system to each other. After the system is finished and functioning this structure can aid the system users in its continued development.

Figure 2.

MODEL SURVEILLANCE INFORMATION DISPERSION



Discussion

Integrating the automated surveillance system into the health departments general work flow is a two part process. The first part includes the designing and developing an automated system. The second is system implementation. Three of the four components that make up the automated surveillance system have been finished or are nearly finished with the first part of this process. These components are in different stages of the implementation process and illustrate the different factors that are important in that process. These factors are developing systems that meet real user needs, having the equipment available to run the system, and trained users that can operate the system.

The ADSS has been collecting disease surveillance data since January 1, 1985, and is already being used extensively for analysis and development of disease surveillance summaries. The system is based within MDH and is operated by epidemiologists and their support staff. It meets a critical need for automated data handling and analysis. The staff recognize this and have been actively involved in learning to use the system. It is already functioning in the general work flow of disease surveillance.

The LEWS is very important to the overall function of a model disease surveillance system because it provides the access of computerization to those who do not currently have access and who are necessary for a comprehensive automated disease surveillance system. Two LEWS systems have been distributed and six more have been ordered for epidemiologists outside of the central MDH office. The immediate goal for LEWS is that all full time epidemiologists that are working on disease surveillance would have access to a system such as LEWS. Each user needs to be trained in the system's basic operation. Such training is currently in progress. LEWS usage by epidemiologists depends on their perceiving that LEWS automated tools can make their job more efficient and effective. Outbreak investigation and communication are felt to be two applications that will be of immediate benefit, and these were the first to be incorporated into the system.

MPHnet is implemented within the MINET communication system and so almost no technical development was necessary. However, before integrating the network into the disease surveillance process two tasks need to be completed. They include 1) locating a sufficient number of network users to make the network's function practical and 2) identifying communication

activities to be included on the network. The difficulty in locating network users is that most potential users do not have access to computer equipment they can use to contact MPHnet. LEWS development and distribution helps to alleviate this problem. Identification of applications to use on MPHnet can be difficult because of the newness of computer communication technology and the lack of understanding of where it can be efficiently applied.

The most important implementation problem to be faced when SRCS is completed is to communicate to the system's audience (the general medical community) its availability and the information that it can provide. Since SRCS represents a completely new application in disease surveillance information dispersion, this task will be more involved. SRCS's success depends on it being identified by its audience as a good information source that they need. This will require that it be well supported by its developers.

The model surveillance project by CDC is a very good method of making computer technology available to states. It helps eliminate duplication of effort and makes available systems that have worked for other states to states lacking computer expertise or experience. By supporting four different projects a larger variation of systems can be developed which will result in systems that may be more appropriate for a given environment. Other benefits include increased standardization of reporting records and the possibility of future evolvement of the systems being shared by groups of states. This user group model has been shown to be a very effective method in other industries of insuring continued development of systems and sharing development costs.

UTILIZATION OF THE VETERANS ADMINISTRATION PATIENT TREATMENT FILE TO ASSESS READMISSIONS PATTERNS

N. Martin MacDowell, University of Iowa, Iowa City, IA
Robert L. Ludke, S. Ann Hunter, Brenda M. Booth, Iowa City VAMC, Iowa City, IA

INTRODUCTION

As a consequence of budgetary constraints, the appropriateness and effectiveness of inpatient treatment within the Veterans Administration is being closely examined. An area of particular concern is readmissons that occur within a short period of time following a recent hospitalization. When a patient is readmitted within two weeks of a previous discharge questions regarding the appropriateness of the readmission, previous discharge and initial admission must be addressed.

With continued development of computer-based health information systems, such as the Veterans Administration (VA) Patient Treatment File (PTF), patient utilization and quality assurance issues like the appropriateness of readmissions can be more easily investigated. However, to fully understand these issues, the output from the health information system must often be supplemented with other existing data sources.

This paper discusses a study which merged PTF and medical record abstracted data to examine medical/surgical readmissions at the Iowa City Veterans Administration Medical Center (VAMC) in FY 1984. Two research questions were addressed: (1) are readmissions within two weeks of a previous discharge necessary; and, (2) are they the consequence of an inappropriate initial admission, initial discharge, or readmission? The study had four major objectives:

1) Identify the incidence of readmissions within two weeks of previous discharge among medical/surgical patients admitted to the Iowa City VAMC in FY 1984.

2) Document the reason for each readmission within two weeks of a previous discharge.

3) Assess, using admission and discharge criteria developed by Inter-Qual, the appropriateness of the initial admission, initial discharge, and readmission.

4) Document the nursing acuity level of each readmitted patient using the VA classification instrument.

METHODOLOGY

One of the two major sources of data used in this research was the VA PTF. For each inpatient discharge at a VA medical center a discharge abstract form is completed. The data are keypunched and verified locally and then submitted to the VA Data Processing Center in Austin, Texas for inclusion in the PTF. A variety of socio-demographic, medical, and resource utilization data on all VA inpatient discharges from October 1, 1983 through September 31, 1984 are in the PTF.

Identification of Patients

Patients readmitted within two weeks of a previous hospitalization were identified through a computer program written to access the Iowa City VAMC PTF data base stored online at the National Institutes of Health (NIH) Computer Center. Without use of the PTF a cumbersome manual (and error prone) process using the medical center's Gain and Loss Sheets would have been needed to identify these patients. A total of 689 medical/surgical readmissions to the Iowa City VAMC in FY 1984 were identified. In addition, six occurrences of medical/surgical readmissions not identified using the PTF file were found during medical record abstracting and added to the data base. This total of 695 readmissions comprised about 10% of the 7,038 medical/surgical admissions to the Iowa City VAMC (excluding one day hemodialysis admissions).

The patient's social security number and dates of admission and discharge for each occurrence were obtained and used to locate medical records for abstraction. Other data items retrieved from the PTF in this study included:

Sociodemographic data: birth date, marital status, race, sex, payment source, and residence location.

Medical data: admission source, facility, primary diagnosis, secondary diagnoses, surgical procedures, type of anesthesia, whether or not outpatient treatment was recommended after discharge, and discharge disposition (e.g., home, nursing home).

Resource utilization information: number of surgical procedures performed, inpatient length of stay (leave and pass days during the stay are also recorded), and bed section transfers within the facility.

These variables were stored in a password protected SAS library sorted by patient social security number and date of readmission. Using these two data items, PTF data were linked to readmission data abstracted from the patient's medical record.

Abstraction of Medical Records

For the purposes of the study, additional information about the circumstances of the previous hospitalization and readmissionwas needed from the patient's medical record. Data were abstracted for 90% (627/695) of the identified medical/surgical readmissions. Less than 100% of records were abstracted because patient charts were not available in the Iowa City VAMC. In the VA system, a single patient medical record is used across the system and goes with the patient wherever he/she seeks care.

Four trained abstractors (three registered nurses and a medical records specialist) obtained the necessary data from the medical charts. For each readmission and preceding hospitalization, the primary and secondary diagnoses were obtained. The classification of the major reason for the readmission; appropriateness of the previous admission and discharge; characteristics of the previous hospitalization; the appropriateness of the readmission; and the nursing acuity level at the previous admission, previous discharge and readmission were deduced from the medical chart.

To ascertain the intra- and inter-rater reliablity of the medical record abstraction, reliability assessments were performed weekly. Two random samples of 81 readmissions each were selected without replacement. One of these samples was used to measure intra-rater reliability and the other inter-rater reliability. As shown in Table 1, the level of reliability was acceptable for all data items involving abstractor judgment. The reliability of recording the diagnosis data will be assessed by comparing abstractor-recorded primary diagnosis and secondary diagnoses with the diagnosis data obtained from the PTF.

Classification of Reason for Readmission

Abstractors used the flow chart shown in Figure 1 to ask a series of yes/no questions to determine one of eight reasons for the readmission. A similar flow chart was developed and tested during a pilot study of readmissions at the Iowa City VAMC[1]. All decisions made by the abstractor about the reason for a given readmission are based on written documentation in the medical record. The validity of this classification scheme will be assessed in the data analysis by comparing the reason for readmission assigned by an abstractor to the reason for readmission assigned by a physician for a sample of about 80 readmissions.

Appropriateness of Admission and Discharge

Appropriateness of the previous admission, discharge, and readmission were determined using Inter-Qual standards[2]. These standards provided a series of screens (generic and specific to the body system indicated in the primary diagnosis) that the medical record abstractors could use to determine the appropriateness of care. The patient must have had one of the symptoms mentioned in the generic or body system screens for the admission or discharge to be considered appropriate.

Preliminary review of the data indicates that about 42% (261/627) of the previous admissions were inappropriate by Inter-Qual standards. The major reason for inappropriate previous admissions, accounting for 44% (116/261), was no procedure was scheduled within 24 hours. Likewise, about 46% (291/627) of readmissions were judged inappropriate. Of the these, 42% (123/291) were inappropriate because a procedure was not scheduled in 24 hours.

Among previous discharges, 48% (298/627) were classified as inappropriate. Of these, 59% (177/298) were inappropriate because Inter-Qual standards were not applicable due to lack of sufficient detail.

Nursing Acuity Level

The nursing acuity level was assessed for each patient on the first day of the previous admission, day of previous discharge, and first day of readmission using the VA nursing care classification instrument shown in Figure 2. Abstractors reviewed the admission or discharge notes, nursing notes, and other relevant information in the medical record to determine the documented nursing acuity level for the patient. A score was assigned by totaling the points for each of the four possible categories. The patient was assigned to the category with the highest point total.

This patient classification instrument is typically used in combination with visual observation of the patient. The abstractors encountered problems using this instrument because documentation for each item on the instrument was not available in the medical record. When data were collected from the chart, a documentation level was noted on the data collection form. A "firm" nursing acuity level was recorded if documentation for each aspect of nursing care could be found in the chart. If a "firm" level could not be established, the abstractor sought to assess an "inferred" nursing acuity level from the information present in the chart. The abstractors deduced missing information based on their professional judgement and classified the assessment as "inferred". A "firm" nursing acuity level level was established for about 80% of the readmissions and was "inferred" for other readmissions. These steps were necessary to obtain as much information as possible while still maintaining valid methods of data collection.

Patients with the highest nursing acuity level at previous admission or discharge were hypothesized to be more likely to have appropriate readmissions than patients with the lowest nursing acuity level. In the preliminary analyses this pattern was supported. The data analysis will examine the relationships between the appropriateness of the readmission and nursing acuity level at the previous admission and discharge using both a "firm" measure and an "inferred" measure.

CONCLUSIONS

There are several advantages associated with the creation of merged data files like the one described in this study. First, it is possible to detect data entry errors and assess the reliability of important data items because the diagnoses and operative codes may be entered into both the computerized information system, such as the PTF, and the abstracted record. Second, availability of the computerized information system may reduce the number of items that need to be abstracted from medical records, thus minimizing the time and cost of medical record abstraction. Third, the computerized information system may contain a number of items either not contained in the medical record (e.g., DRG) or time consuming to obtain from chart review (e.g., discharge status).

Although not addressed in this study, it is possible to use the computerized data bases for longitudinal studies of patients' inter-institutional utilization patterns. For example, about 20% of medical records in the PTF (records with a SSN ending with a 1 or 5) are selected annually for a study of outpatient visits. For many patients, the VA is their sole provider of health care, and thus, total health care utilization, both inpatient and outpatient, may be available.

The PTF file contains data about all VA inpatient hospitalizations and can be used to provide data efficiently for studies about patient utilization and care within the VA. As illustrated in this study, the PTF can be combined with data abstracted from the medical record to study characteristics of VA inpatient care that can be used to improve the efficacy and efficiency of medical care.

REFERENCES

1 MacDowell NM, Hunter SA, Ludke RL. Readmissions to a Veterans Administration Medical Center. Journal of Quality Assurance. Spring: 20-23, 1985.

2 Jacobs C and Lamprey J: A Guide to Systematic Utilization Monitoring. Chicago: Inter-Qual Inc., 1979

Figure 1 -- Method for Classification of Reason for VA Readmission

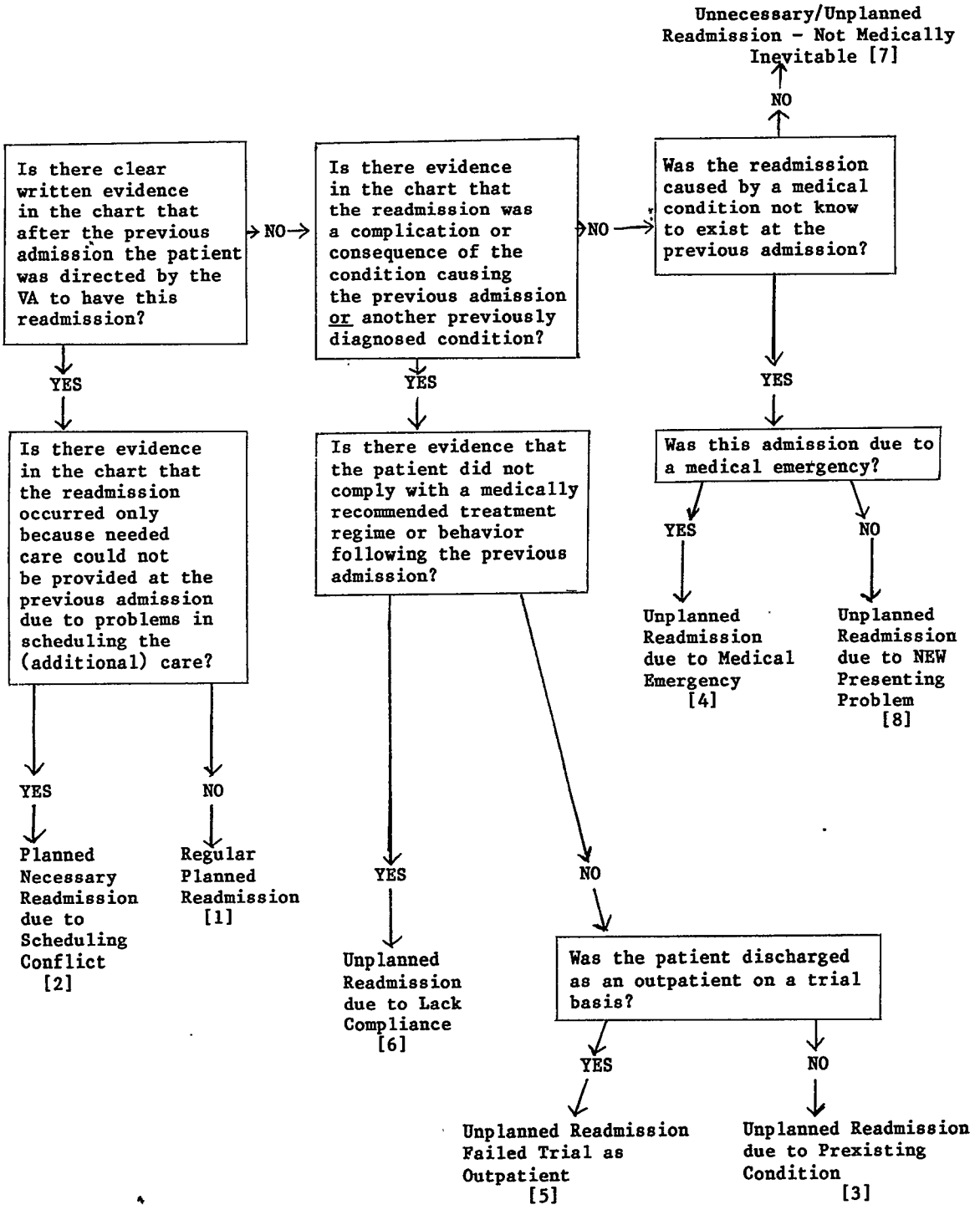
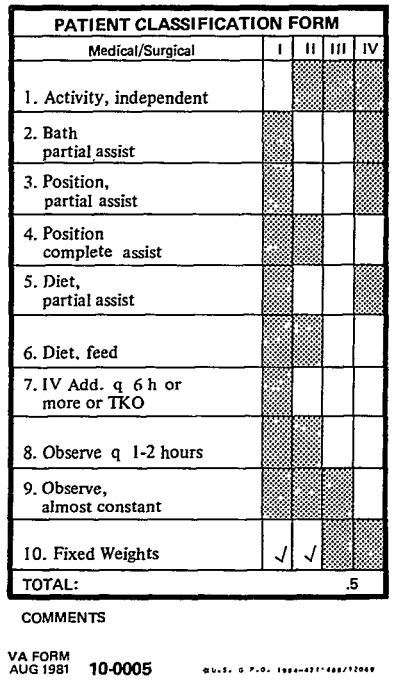


Figure 2 - Classification of Nursing Acuity Level



PATIENT CLASSIFICATION FORM

Medical/Surgical	I	II	III	IV
1. Activity, independent				
2. Bath partial assist				
3. Position, partial assist				
4. Position complete assist				
5. Diet, partial assist				
6. Diet, feed				
7. IV Add. q 6 h or more or TKO				
8. Observe q 1-2 hours				
9. Observe, almost constant				
10. Fixed Weights	√	√		
TOTAL:				.5

COMMENTS

VA FORM AUG 1981 10-0005

Table 1 - Intra and Inter-Rater Reliability of Medical Record Abstraction, Iowa City VAMC Readmissions Study, June/July, 1985

Variable	Intra-rater Reliability (n=81)	Inter-rater Reliability (n=81)
Classification of Reason for Readmission	88.9%	75.3%
Appropriateness of Previous Admission	91.4	87.7
Appropriateness of Previous Discharge	84.0	70.4
Appropriateness of Readmission[a]	90.1	90.1
Nursing Acuity Level at Previous Admisssion	82.7	67.9
Nursing Acuity Level at Previous Discharge	96.3	80.3
Nursing Acuity Level at Readmission	85.2	67.9

[a] By coincidence intra- and inter-rater reliability has same value.

THE DEVELOPMENT OF A CONTINUALLY UPDATED NATIONAL DATA BASE ON PREVENTIVE BEHAVIORS ADOPTED BY THE AMERICAN PUBLIC

R. Barker Bausell, University of Maryland and the Prevention Research Center
Peggy L. Parks, University of Maryland

Beginning in 1983 the Prevention Research Center, created by Robert Rodale, began a data collection program designed to track the state of preventive behaviors among the American public on an annual basis. In order to accomplish this relatively formidable task, a number of critical decisions regarding the selection of preventive behaviors had to be made early on in the project's planning phase. Some of the more important of these were:

(1) Only behaviors subject to personal modification would be studied. This meant that we would exclude such crucial concerns as air and water quality in favor of preventive acts (e.g., not smoking) more directly under the individual's control. (We made this decision for two reasons: individual behavior is easier to measure and it probably precedes major societal efforts.)

(2) Only behaviors were selected for which a clear consensus existed with respect to a documented relationship between compliance and the prevention of disease or injury. (This meant that we would wind up with a relatively conservative group of behaviors, but hopefully one that would stand the test of time.)

(3) Each selected behavior had to be applicable to the entire adult population. [Actually the Preventive Research Center also collects annual data on preventive behaviors relevant only to women (e.g., breast self-examination) and to children (e.g., DPT inoculations), but these are separate issues not related to the subject at hand.]

After reviewing the literature and consulting with dozens of content specialists, the project director (Suzanne Irvine) contracted Louis Harris and Associates to conduct a telephone interview with one hundred public health professionals as a final validation check of the behaviors selected. After receiving a written version of the final questionnaire, these respondents were asked to rate the behaviors on a scale from one (of low importance) to ten (of utmost importance) with respect to protecting the overall health of the general population. All of the behaviors that constitute the basis for this paper received a mean rating of at least 6.90 out of a possible 10-points. (The grand rating mean for the final set of behaviors was 7.92.)

The final set of 21 behaviors along with their compliance definitions are listed in the Appendix. Basically we wound up with six safety related behaviors [three driving related acts (wearing seatbelts, obeying the speed limit, avoiding drinking after driving) and three centered in the home (owning a smoke detector, avoiding smoking in bed, and avoiding home accidents in general), six dietary behaviors (avoiding excessive fat, cholesterol, sodium, and sugar; consuming adequate fiber and vitamins/mineral), two health monitoring acts (regular blood pressure screening and dental exams), and seven general lifestyle variables (moderate or no alcohol consumption, avoiding smoking, exercising regularly, maintaining recommended bodyweight, stress reduction, and proper sleep)].

Once our behavioral core was developed, the next step was to devise a data collection scheme that could be replicated from year-to-year. For this we again contracted Louis Harris and Associates, who conducted a telephone interview of 1250 adults using a random digit dialing procedure stratified by geographic region and metropolitan versus non-metropolitan residence within those regions.[1]

The first survey was conducted in the fall of 1983, a second (employing an identical methodology) in the fall of 1984, and final plans are being made for the third survey (which will be conducted in the fall of 1985) at the time of this conference. Because each survey contains the same core of 21, identically worded behaviors, year-to-year tracking is possible with respect to (1) each individual behavior, (2) a composite index made up of the sum of the 21 behaviors, and (3) demographic breakdowns of both individual behaviors and the composite measure.

These data are collected for two basic purposes: (1) health promotion and (2) basic research into the dynamics of preventive behavior. The first objective is made possible by an annual press conference conducted several months following the survey itself. The second objective is made possible by the inclusion of a certain amount of unique information each year that is designed specifically for exploring the determinants and dynamics of preventive behavior. Although the primary focus of this paper is on the second objective, we will briefly discuss the data's health promotional uses as well.

The Prevention Index

Originally conceived by Robert Rodale, editor-in-chief of Prevention Magazine, the Prevention Index is designed as a sort of report card on the nation's prevention efforts. Like a report card, its primary purpose is to serve as a mechanism for both feedback and corrective action. Toward this end we spend a substantial part of our budget on a large press conference each spring in which the results from the previous survey are announced and at which the Prevention Index itself [i.e., a composite of the 21 preventive behaviors weighted by importance scores obtained via our professional survey) is unveiled. So far many millions of television viewers, radio listeners, and newspaper/magazine readers have been reached with the messages that prevention is an important enterprise and that a great deal of room for improvement exists with respect to compliance with each of our targeted behaviors.

Basic Research Uses

As important as the health promotional aspects of these data are, their research potential may in the long run prove to have the most lasting impact. To illustrate this possibility we will first discuss some of the information contained in this annually evolving data bank, followed by a summary of some of (1) our findings to date, (2) the uses to which we would like to see these data put in the future, and (3) opportunities for the collaborative use of the data and the instrument used to generate it.

Information Contained in the 1983 and 1984 Surveys

As previously mentioned, each survey contains (and will always contain) both a basic core of common content and some unique elements specifically designed to further our knowledge concerning preventive behavior. The 1983 survey, for example, contained the following items in addition to the 21 core behaviors:

(1) self-rated health status, (also present in the 1984 survey),

(2) perceived control over future health (also present in 1984),

(3) opinions regarding whether more emphasis should be placed on prevention or on treatment modalities (repeated in 1984),

(4) whether the respondent had received specific advice regarding improving his or her health habits (e.g., nutrition, exercise, smoking) from a doctor during the past five years (almost 2/3 had not),

(5) The importance the respondent placed upon a subset of the 21 core behaviors with respect to "helping people in general to live long and healthy lives",

(6) several open-ended questions [i.e., specific steps taken to reduce stress and events that lead people to improve their habits (if indeed they had done so)],

(7) three behaviors specific to women's health (i.e., calcium intake, frequency of breast self-examinations and pap smears - also present in the 1984 survey),

(8) a number of preventive behaviors relevant to children [administered to households in which one or more children under the age of 18 were present (n=428) - also present in the 1984 survey], and

(9) several miscellaneous preventive behaviors included on a one-time basis that did not meet our above mentioned criteria for various reasons (e.g., ownership of a fire extinguisher, eating breakfast, not being exposed to industrial accidents or toxins).

Unique information contained in the 1984 survey included:

(1) a number of additional preventive behaviors: cholesterol blood tests, avoiding caffeine, avoiding food additives, consuming additional vitamin A and C, the use of dental floss, not living in a household in which another person smokes, the existence of a family fire escape plan, and the avoidance of recreational drug usage,

(2) health and medical utilization information: number of days missed from work, number of days worked below peak efficiency, number of sick visits to a health provider, and number of days the respondent was forced to stay in bed for at least half a day,

(3) a brief cholesterol knowledge test,

(4) sources of information most influential (i.e., magazines, TV/radio, books, classes) in changing the respondent's health habits, and

(5) miscellaneous questions related to preventive behavior (e.g., opinions regarding the prohibition of smoking in public places).

This relatively eclectic assortment of information was generated to serve the surveys' two basic purposes: (1) to enhance its newsworthiness, hence health promotional potential, and (2) to identify potential determinants and correlates of preventive behavior. We hope that this paper will further help to generate professional input with respect to additions to future surveys. As will be discussed later in this paper, Prevention Research Center staff are truly eager to collaborate with other public health professionals interested in the study of preventive behavior.

Completed Research

Obviously there is no way that a single paper can detail all of the interrelationships found among two such large sets of variables. What we will attempt to do, therefore, is summarize some of the more global questions we have been able to address. Thus, while each individual behavior can be (and has been) studied separately, we will focus on the overall picture of preventive behavior presented by our respondents' self-reported compliance with the 21 core behaviors (i.e., the Prevention Index) as a whole.

(1) Do recognizable patterns of preventive behavior exist? Said another way, are people who engage in one form of preventive behavior more (or less) likely to engage in another? The answer is yes, although the magnitude of the intercorrelations among our 21 variables were quite low and the resulting factor pattern was quite sparse. Viewed as a 21 item composite (i.e., a score ranging in value from a low of 0 to a high of 21), the Prevention Index possessed an adequate (but moderate) internal consistency of .56. Compliance scores for our two samples are normally distributed and, interesting, out of 2500 respondents: none reported complying with fewer than four behaviors and only five people claimed perfect compliance.

(2) Can correlates of this generalized measure of preventive behavior be identified? Yes. The following are a sample of the types of people who are more likely to report compliance with the set of behaviors as a whole.

a) women,

b) older persons,

c) persons with higher educational attainments,

d) persons with higher income (excluding persons at the extreme upper end of this continuum),

e) people reporting themselves to be in excellent health,

f) people reporting themselves to be most satisfied with life in general,

g) fulltime workers who missed fewer than three days due to illness,

h) people who believe they have a great deal of control over their future health status,

i) people who believe that preventive behaviors are important in promoting long and healthy lives,

j) people possessing the most knowledge about the cholesterol content of food,

k) people who report that the cholesterol content of food influences their purchasing decisions, and

l) people who have received preventive information from magazines, books, TV/radio, or health related classes.

(3) How do public opinions regarding the importance of various individual preventive behaviors compare to professional opinions? There is very little congruence. In general the public rates most behaviors slightly higher than our public health sample, but there is little consistency in the rank orders assigned by the two groups.

(4) Is there a relationship between the care we take of our children and the care we take of ourselves? Yes. Although mothers generally report greater compliance with respect to pediatric preventive behaviors than they do with their personal preventive behavior, a positive correlation was obtained between overall pediatric preventive behavior (as reported for a randomly selected child under 18 living in the household) and scores on the Prevention Index.

Proposed Research

Although the above questions are only a sample of the types of analyses we have completed so far at the Prevention Research Center, there is a great deal more research that should be done. Here are a few examples:

(1) Combining individual preventive behaviors into risk factors for specific diseases. Instead of identifying determinants and effects of the Prevention Index as a whole, a priori subscores could be studied instead. For example, exercise, smoking, weight control, stress reduction, fat and cholesterol consumption, blood pressure screening, and alcohol consumption could be studied as composite risk factors for cardiovascular heart disease.

(2) Comparative Studies. Given the representativeness of the sampling method employed, the Prevention Index as well as each

of the 21 individual behaviors can be viewed as norm referenced measures. Thus at risk subgroups within the population as well as groups of individuals who have already contracted a preventable disease (or experienced a preventable accident) can be compared with our continually updated national norms. To facilitate this use of the data, age and sex specific norms will be published in the March, 1986 (Vol. 9, No. 1) issue of Evaluation and the Health Professions.

Examples of the types of subgroups that might be profitably studied in this way are:

a) trauma victims,

b) CHD or cancer patients,

c) welfare mothers,

d) indigent elderly, or even

e) college (including medical) students.

(3) Predictive Studies. Although we have identified quite a few correlates of preventive behavior, we've probably only scratched the surface. At this point we just do not know a great deal about what preventive behavior is related to. Possible correlates include:

a) compliance with medical (e.g., diabetic) regimens,

b) personality,

c) political orientation, or simply

d) intelligence.

Collaborative Opportunities

The Prevention Research Center possesses a relatively unique mix of characteristics. It is entirely corporately funded by Rodale Press, yet its data is freely available to the professional community (e.g., the forthcoming norms to be published in Evaluation and the Health Professions this spring). It carries out its own research program, yet it is not competitive in any traditional academic sense, hence its staff is more than willing to collaborate with other professionals with respect to the use of the Prevention Index itself, its norms, or even the analysis of the collected data. Toward that end we[2] would be delighted to supply any additional information on the data sets described in this paper.

---

[1] A more detailed description of this methodology (and the overall survey results) may be obtained from the Prevention Research Center, 33 East Minor Street, Emmaus, PA 18049.

[2] For additional information please contact the senior author: Evaluation and the Health Professions, University of Maryland, 655 West Lombard Street, Baltimore, MD 21201.

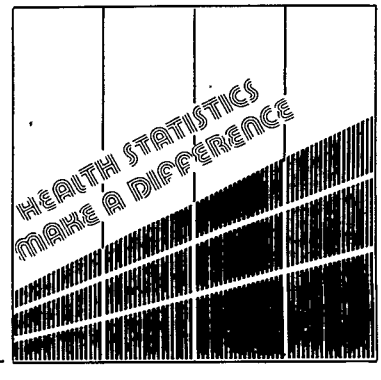
APPENDIX

Compliance Definitions for the 21 Behavioral Items

	Item	Definition
1.	How often do you have a blood pressure reading?	At least once a year.
2.	How often do you go to the dentist for treatment or a checkup?	At least once a year.
	Thinking about your personal diet and nutrition, do you try a lot, try a little, or don't you try at all to:	
3.	Avoid eating too much salt or sodium.	Try a lot.
4.	Avoid eating too much fat.	Try a lot.
5.	Eat enough fiber from whole grains, cereals, fruits, and vegetables.	Try a lot.
6.	Avoid eating too many high-cholesterol foods, such as eggs, dairy products, and fatty meats.	Try a lot.
7.	Get enough vitamins and minerals in foods or in supplements.	Try a lot.
8.	Avoid eating too much sugar and sweet food.	Try a lot.
9.	In feet and inches, what is your height without shoes on? What is your present weight without clothes? What kind of body frame or bone structure would you say you have - small, medium, or large?	In range based upon Metropolitan Life Insurance tables.
10.	How often do you exercise strenuously - that is, so you breathe heavily and your heart and pulse rate are accelerated for a period lasting at least twenty minutes?	At least 3 times/week.
11.	Do you smoke cigarettes now or not?	Do not smoke.
12.	Do you consciously take steps to control or reduce the stress in your life?	Take steps.
13.	How many hours do you usually sleep each 24-hour day in total?	7-8 Hours.
14.	In general how often do you consume alcoholic beverages? On a day when you do drink alcoholic beverages, on average, how many drinks do you have? (By a "drink" we mean a drink with a shot of hard liquor, a can or bottle of beer, or a glass of wine.)	No more than 4 drinks per day for a total of no more than 15 per week.
15.	How often do you wear a seatbelt when you are in the front seat of a car - all the time, sometimes, or never?	Always.
16.	How often do you drive above the speed limit?	Never does so.
17.	How often do you drive after drinking alcoholic beverages?	Never does so.
18.	Do you have a smoke detector in your home.	Yes, owns one.
19.	Does anyone in your household ever smoke in bed?	No one does.
20.	Do you take any special steps or precautions to avoid accidents in and around your home?	Yes, takes steps.
21.	About how often do you socialize with close friends, relatives, or neighbors?	At least once a week.

Session W

Developments in Telephone Survey Methodology



RESULTS FROM THE NATIONAL HEALTH INTERVIEW SURVEY RANDOM DIGIT DIALED FEASIBILITY STUDY

Anthony M. Roman, Bureau of the Census
Joseph E. Fitti, National Center for Health Statistics

I. Background Information

Late in 1982, the Bureau of the Census and the National Center for Health Statistics formed a Joint Agency Telephone Task Force to plan a three-year program of research and development leading to the implementation of random-digit-dialing (RDD) sampling techniques in the National Health Interview Survey (NHIS). In their final report on the three-year plan, the Task Force recommended that a study be conducted to test the feasibility of conducting the NHIS by telephone and to investigate a number of major issues in the use of RDD in this survey.

A Joint Agency Work Group was convened and research questions and methodology for a feasibility test were formulated by September 1983. The 1984 NHIS/RDD Feasibility Study with over 3,000 households, was conducted in the spring of 1984.

Although many objectives were identified and investigated in the Feasibility Study, three particular areas will be discussed in this paper due to their general applicability to RDD surveys. They are the use of an automated case management system, response rate issues, and selection of a household respondent. Initially though, a short discussion of the sample design of the Feasibility Study is required.

II. Sample Size and Design

The sample for the Study was an RDD sample of telephone households in the 48 contiguous United States. The telephone households in the sample were selected using the method described by Waksberg [1]. The sample was selected in 12 independent replicates. One replicate was introduced each week for 12 consecutive weeks. Each replicate was interviewed for three weeks. Hence there was some overlap in the data collection phases of adjacent replicates. A schema of the study schedule is shown in Figure 1.

The total sample size for the study was 3,024 telephone households with a sample size per replicate of 252. Each replicate consisted of 21 primary sampling units (PSU's) with 12 telephone households selected from each. A primary sampling unit was composed of a block of 100 telephone numbers, each having the first eight of ten digits the same.

The procedure for sampling was according to the two-stage operation for the Waksberg design. The PSU's generated from the AT&T tape file were sorted on the basis of geography, population density, and proximity to urbanized areas. For each replicate, a systematic random sample of 135 PSU's was selected from the sorted list and one telephone number from each PSU was called in a random order until 21 residential PSU's were obtained for the sample (primary screening). A PSU was declared residential if the one randomly selected telephone number from the PSU was classified as belonging to a residential unit.

For the selection within each retained PSU and interviewing of telephone households for the sample (secondary screening), twelve telephone numbers were selected at random from the PSU. Each number was dialed to determine whether or not it was residential. If a number was determined not to be residential, it was replaced by another randomly selected number from within the same PSU. In this manner, twelve eligible telephone households were selected from each PSU. A schema of the sample design is displayed in Figure 2.

III. The Automated Case Management System

With an understanding of the sample design and interviewing schedule, it is now possible to address the automated portions of the Feasibility Study. The computer hardware used in the study consisted of a DEC VAX 11/750 minicomputer as host with 30 HAZELTINE and WYSE video display terminals as interviewer work stations. Residing on this system was customized software entitled the automated case management system (ACMS). This software was designed to handle three vital survey functions: 1) sample selection, 2) call scheduling, and 3) record keeping.

Regarding sample selection, the ACMS performed all steps required by the Waksberg procedure. This included 1) selection of the primary telephone numbers from the AT&T tape to be screened, 2) selection of the final 21 PSU's for each replicate, 3) selection of the twelve telephone numbers from within each PSU, and 4) selection of replacement numbers for ineligible survey units. The ACMS has additional sample selection capabilities including the ability to interpenetrate balanced experimental designs within the study and generating substitutes for nonrespondent units.

In call scheduling, the ACMS performed several activities. Perhaps the most important was determining when each case should be called. Since within the Feasibility Study, an upper limit of 20 attempts to a telephone number was implemented, the spacing of the call attempts could dramatically affect results. For instance, if it was discovered that all 20 attempts were made during afternoon hours and the telephone number always had a ring-no-answer result, one could only wonder if the case would have easily been contacted if attempted in the evening. Therefore, the major goal of the ACMS was to assure a random dispersion of call attempts across hours of the day and days of the week. In order to accomplish this, it used an algorithm for assigning priorities for calling each telephone number and attempted to use past call

information to maximize the probability of a contact. In assigning priorities, the ACMS considered such factors as 1) scheduled callbacks, 2) the number of attempts made to the telephone number within designated time slots, 3) the number of days remaining in the interviewing period, 4) the amount of time since the number was last attempted, and 5) the outcome of the previous call attempts. As part of this overall scheduling process, the ACMS needed to determine the next action to take based upon the outcome of each call. For example, if a call resulted in a completed interview, the ACMS would close out the case from any future attempts. If the call resulted in a ring-no-answer, the ACMS would place the case back into the scheduling queue. If the call resulted in the fifth attempt with no contact to a case, then the ACMS would assign the case for a call to the telephone business office. Finally, if a call resulted in a busy signal, the ACMS would hold the case aside to be attempted again in a short period of time.

In examining the effectiveness of the call scheduling operation, it was felt that the ACMS performed quite well. The proper dispersion of call attempts was obtained and past call information was used quite effectively. Still, it is felt that certain modifications can be made to the scheduling algorithm to obtain even better results. This may imply raising the limit of 20 call attempts per telephone number and perhaps scheduling more than one attempt to get information from the telephone business office. It is felt that especially since the divestiture of AT&T, cooperation from telephone business offices varies greatly from one telephone company to another and often even within the same company. Failure to gain information in one business office call does not imply that success is unlikely in another.

The third function of the ACMS was to record all pertinent information about the call attempts and to create a database for analysis purposes. This database contains such items as: 1) the date and time of each call attempt, 2) the interviewer who made the attempt, 3) the outcome of the attempt, 4) time marks for progress of the interview, 5) last item answered in the interview, and 6) any notes made by the interviewer.

It must be noted that although the Feasibility Study used an automated case management system and a computer assisted introduction to the interview, once actual interviewing commenced, a paper and pencil approach was used. This was necessitated by the fact that a computer assisted telephone interviewing (CATI) questionnaire form was not available at the time of the study. Such a form has since been developed and may be used in future research.

IV. Response Rates

A. General Information

Response rates for an RDD survey are not as readily calculated as those from a traditional field survey. There are several reasons for this but one in particular stands out. In a field survey, since the interviewer physically visits the sample unit, there are numerous methods available for answering questions such as: is the unit vacant, is the unit a vacation residence only occupied at specific periods of time, is the unit residential, does the unit need to be subsampled, and who resides at the unit. The interviewer may inspect the unit and the neighborhood and also talk to persons who reside near the unit in order to answer such questions. Far less is available to an interviewer in an RDD survey where the sample unit is a randomly selected telephone number. This section describes problems encountered and results obtained from the Feasibility Study regarding these matters.

The sample design of the Feasibility Study called for 3,024 eligible units to be selected, but only 2,957 were used in computing response rates. Of the 67 missing units, 36 were lost when three PSU's were discovered to be ineligible for the study. These discoveries were made too late to generate replacement PSU's. The problem encountered here appears to be related to identifying special places over the telephone. For example, college dormitories were considered special places and ineligible for interview within the Feasibility Study. When a PSU containing only college dorm rooms was misclassified during primary screening as residential, the error was not uncovered until it was too late to correct. In addition to these 36 units, three more were lost when one PSU was found to have only nine eligible units within its 100 telephone numbers. The remaining 28 cases that were lost were units which were contacted and determined to be ineligible too late to generate a replacement unit. It may be possible to reduce this problem with modifications to the automated call scheduler but this is a topic that needs more investigation.

The 2,957 units used in computing response rates are displayed in Table 1

Table 1: Catagorization of Interview Outcomes

Outcome	Number of Units
Complete Interviews	2,251
Partial Interviews	42
Refusals	370
Other Noninterviews	36
Unresolved	258

As in the continuing NHIS, partial interviews were considered a form of noninterview. Of the 36 units classified as "Other Noninterviews", 35 were described as language barriers which could not be converted. This indicates a potential problem for centralized RDD interviewing. Calling from a far away central location into an area does not allow the flexibility of field interviewing where a local interviewer who speaks the language can be hired.

B. Methods of Computation

The 258 unresolved units pose a problem for response rate computation. Assuming that each of these telephone numbers belongs to an ineligible survey unit is not realistic and could artifically inflate the response rate. Likewise, assuming that each of these telephone numbers belongs to an eligible survey unit is equally unrealistic. It is probable that a proportion of the unresolved units are eligible for the survey but the value of that proportion is unknown.

For the Feasibility Study, three response rates were computed using the following notation:

C = number of completed interviews
E = number of units determined upon contact to be eligible
I = number of units determined to be ineligible
U = number of unresolved units

The first response rate computed (R_L) assumed that all unresolved units were eligible for the survey. As such, it served as a lower bound on the true response rate obtainable if the eligibility status of all sample units could be determined. It was computed as:

$$R_L = \frac{C}{E+U}$$

The second rate computed (R_U) assumed that all unresolved units were ineligible for the survey. This rate serves as an upper bound on the true response rate. It was computed as:

$$R_U = \frac{C}{E}$$

The final response rate computed (R_C) was a compromise that assumed that a proportion, p, of the unresolved units were eligible. This proportion was estimated from the sample using only those units whose eligibility status had been determined. It basically assumes that the unresolved sample units are eligible in the same proportion as resolved sample units. This rate was computed as:

$$R_C = \frac{C}{E+pU}$$

where:

$$p = \frac{E}{E+I}$$

C. Results

The computed response rates are displayed in Table 2. Three important results are evident from this study. The first is the effect that unresolved units can have on reported response rates. The value of R_L was 76.12%. The corresponding value of R_U was 83.40%. This demonstrates that unresolved units can affect the response rates by over seven percentage points. This is not a desirable characteristic of RDD surveys, and more research is needed on methods to reduce the number of unresolved units. Of course, the true response rate lies somewhere between R_L and R_U as evidenced by the estimator R_C which has a value of 78.91%. The second result is an obvious improvement over time that was made within this study. The value of R_L was 68.48% in replicates 1 through 3 and steadily increased to a value of 83.13% in replicates 10 through 12. This gain was accomplished by modifying the survey operations to make better utilization of available resources but more importantly by the interviewers gaining experience and sharpening their skills in administering the forms over the telephone. It is not believed that any additional major gains would have occurred had the study been extended, but certainly future gains are not impossible after an appropriate time to study and digest the experiences of the Feasibility Study.

The third result is an extension of the second. By concentrating on replicates 10 through 12, it is seen that the value of R_L was 83.13%, R_C was 84.92% and R_U was 88.09%. This indicates that response rates of 85% or higher are possible for the NHIS using RDD procedures. These rates could only be obtained provided that a well trained and experienced staff of interviewers was maintained.

Table 2: Response Rates

	R_L	R_C	R_U
Replicates 1-3	.6848	.7121	.7581
Replicates 4-6	.7371	.7697	.8242
Replicates 7-9	.7936	.8269	.8727
Replicates 10-12	.8313	.8492	.8809
Replicates 1-12	.7612	.7891	.8340

Some interesting results were uncovered when 223 of the unresolved units encountered earlier in the study were included in an extended follow-up. These telephone numbers were attempted again in order to: 1) determine if they could be contacted in another time period, 2) determine their eligibility status at the time of the follow-up, 3) reconcile their eligibility status to the time of the original interviewing period, and 4) determine why they could not be contacted originally. These calls were made from 2 to 13 weeks after the units were originally attempted.

The results of the extended follow-up are presented in Table 3. It is interesting to observe that nearly 90% of the original unresolved cases were resolved during the follow-up. This indicates that the largest portion of unresolved units are not of some chronic form that can never be resolved. To the contrary, resolution appears to be more a matter of timing and effort. This is displayed

even more dramatically when it is revealed that 66% of those units resolved in the follow-up, were resolved by contacting the unit directly (i.e., as opposed to classifying the unit from information gained from a telephone business office). In a related issue, it was found that 78% of the resolved cases in the follow-up required fewer than 10 call attempts to attain resolution. This indicates even further that units that go unresolved may at another time be quite easy to resolve. A somewhat disturbing result is that over 20% of the units were resolved by the telephone business office (TBO). Since the original interviewing procedures required the TBOs to be called after five attempts to a unit with no contact, the question arises as to why the TBOs did not resolve these units originally. It is possible that the TBOs had obtained more information about these units between the time of initial interviewing and follow-up, but this seems unlikely for such a large percentage of units. It is more likely that cooperation from the TBOs is sporadic with individual operators, time of day and how busy the office is at a particular moment all playing a role in the ability to obtain information.

Table 3: Percentage Comparisons from the NHIS Extended Follow-Up

	Percentages
Of total cases in follow-up:	
% resolved during follow-up	89.69
Of resolved cases:	
% resolved by contact	66.00
% resolved through TBO	20.50
% resolved through other source	13.50
Of resolved cases:	
% residential	60.50
% nonresidential	14.50
% other out-of-scope	25.00
Of resolved cases:	
% reconciled as residential at time of initial interview	56.50
% reconciled as nonresidential at time of initial interview	5.50
% not reconciled	38.00
Of reconciled cases:	
% absent entire interviewing period	18.55
% with no valid reason for noncontact	66.13
% seasonal residence/business	15.32

Another interesting result is that 60.5% of the resolved units in the follow-up were found to be residential at the time of the follow-up. Also, 56.5% of the resolved units were reconciled as residential at the time of initial interviewing. These numbers compare very favorably with the value of p (59.54%) computed from the original resolved sample units. This may indicate that perhaps the original unresolved units are approximately the same percentage residential as the original resolved units. If this is the case, then R_C is a very good estimator of the true response rate. Of course, more research is needed in this area.

The final result of note from Table 3 is that of those units which could be reconciled as to their residential status during the initial interviewing period, 66.13% stated they had no valid reason as to why they could not be contacted. If this result is to be believed, then perhaps the modifications to the call scheduling algorithm cited earlier may help reduce the problem of unresolved units significantly.

V. Respondent Selection Within Households

The NCHS/Census Joint Task Force on Telephone Surveys devoted considerable attention to the selection of a respondent rule for the Feasibility Study. A number of respondent rules were considered with respect to cost, sampling error, and non-sampling error. The recommendation made was that a Most Knowledgeable Respondent (MKR) rule be used. An important factor in the determination of a rule for this study was the desire to approximate the respondent rule used in the face-to-face NHIS.

Under the rule developed, the interviewer asked the telephone answerer to identify the MKR for the household, that is, the person most knowledgeable about the health of the household members. It is the MKR who then becomes the household respondent that should be interviewed. Those who favor an MKR rule list as their reasons: 1) the better quality of data that may be obtained from a respondent who is in some sense the most knowledgeable household member regarding the questions of interest and, 2) the fact that asking to speak to an MKR impresses upon the respondent the importance of the survey. Those who do not favor an MKR rule generally list response rate concerns (i.e., A potential respondent is on the telephone and may not be interviewed because of the need to pursue the MKR. If the MKR is never reached, an interview may have been lost). It was the intent of the Feasibility Study to address these concerns.

Two important results came from this study of respondent rules. The first is that approximately 78% of the time, the person answering the telephone identifies themselves as the MKR, another 2.5% of the time, the phone answerer identifies another household member as MKR and that person comes to the phone to be interviewed, while about 19.5% of the time, either an MKR cannot be identified or is not present at the time of the call. The second major finding involves the cases in which callbacks are required to contact an MKR. The ratio of completed interviews to noninterviews in this instance was 0.21. This compared very unfavorably to the ratio of 4.18 when callbacks were not required.

The two results considered together perhaps suggest the best approach. It is fine to ask for an MKR and interview them as long as they are present during the call attempt. If the MKR is not available then interviewing some other household member should be preferable to scheduling a callback.

VI. Additional Issues

The following is a list of additional issues that were investigated in the Feasibility Study: 1) comparison of two NHIS questionnaire versions, 2) examination of the suitability of a three-week interviewing period, 3) investigation of substitution as a method of nonresponse adjustment, 4) examination of the ability to identify "special" types of living quarters, 5) a cost analysis, 6) methods of monitoring telephone interviewers, and 7) calculation of intracluster correlations in order to optimize future sample designs.

Information on these issues as well as more detailed discussions of the topics in this paper are available in a joint report prepared by the Census Bureau and NCHS [2].

VII. Further Research

Although the Feasibility Study has provided a wealth of information concerning the use of RDD with telephone health surveys, there are still many issues that need further research. Possibly the most immediate need is that of accurate cost data on the components of an RDD survey. As many private and public survey groups are contemplating the move to increased use of RDD techniques, the question constantly arises as to the amount of savings that can be expected. It is difficult to estimate these savings from a research vehicle such as the Feasibility Study. The large development costs associated with research should not be included in making comparisons, but it gets very difficult to decide just where development ends and production begins in a research project which undergoes constant changing and growing. What is needed is an accurate model based approach which compares the expected costs of the various components of a field and RDD survey.

A second need is for a more in depth study of where, inside the telephone interview, a breakoff or refusal is most likely to occur. This was planned for the Feasibility Study but problems in creating an appropriate data base limited the utility of any results. The Census Bureau does plan to study this problem in future research endeavors in order to locate the critical points for getting and keeping a respondents cooperation. Then survey questionnaire design experts may help alleviate some of the refusal rate problems with cold contact telephone surveys.

Yet another area that needs general research is that of effectively monitoring and rating interviewers. Unlike a field survey, in which each interviewer gets an assignment of cases specifically for him/her and on which he/she exclusively works, for cost and convenience purposes, a centralized telephone facility produces an environment in which all interviewers share a common workload. Therefore, the computation of such interviewer performance measures as individual response rates have no meaning. Research is needed on how to effectively monitor an interviewer's performance in such a way that problems can be detected and a consistent measure of interviewer performance can be created. Alternatively, research is needed on how to effectively assign designated subsamples to each specific interviewer. This would allow for measures of interviewer performance and effect to be computed. Research would be required on how to best blend the automated call scheduling procedures with randomizing the shifts each interviewer would work. Potential problems exist with scheduled callbacks and interviewer morale in this changing work schedule concept.

Finally, there is a need for comparing the telephone data from the Feasibility Study to field data from the continuing NHIS from the same time period. It is only in this manner that questions regarding data quality can be addressed. This analysis becomes particularly important if one considers the possibility of a dual frame approach to the NHIS in which data from both a field and an RDD component must be combined. The potential biases that exist from mixing sample frames must be carefully studied and understood before such an approach can be implemented.

References

[1] Waksberg, Joseph (1978) "Sampling Methods for Random Digit Dialing", JASA, 73, pp. 40-46.

[2] The Bureau of the Census and the National Center for Health Statistics (1985). "The Results of the 1984 NHIS/RDD Feasibility Study: Final Report", submitted to the NCHS-Census Joint Steering Committee on Telephone Surveys.

FIGURE 1

SCHEMA OF INTERVIEWING SCHEDULE

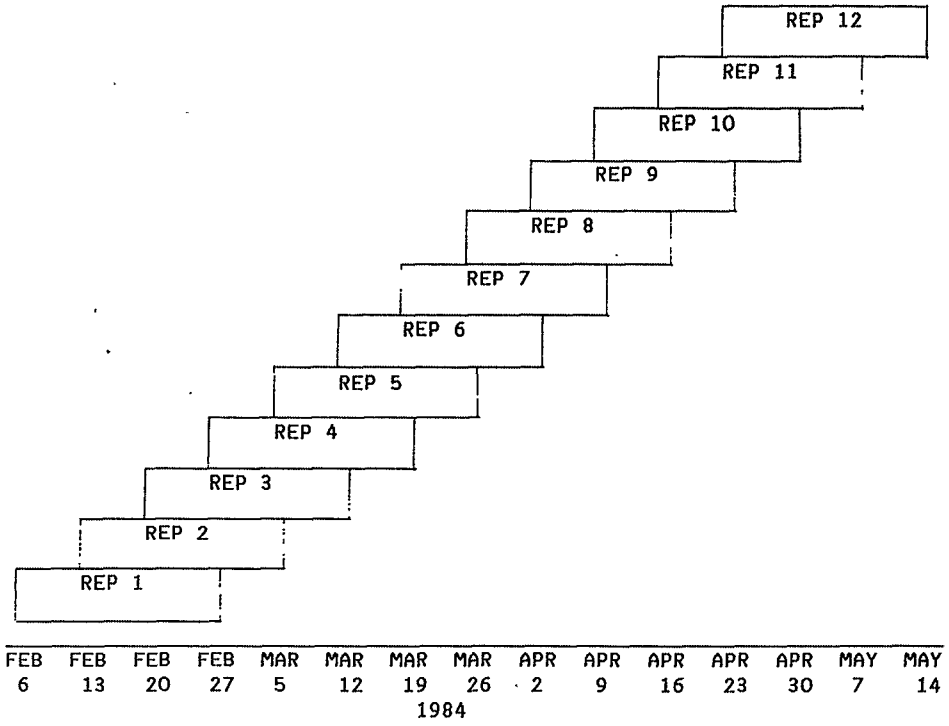
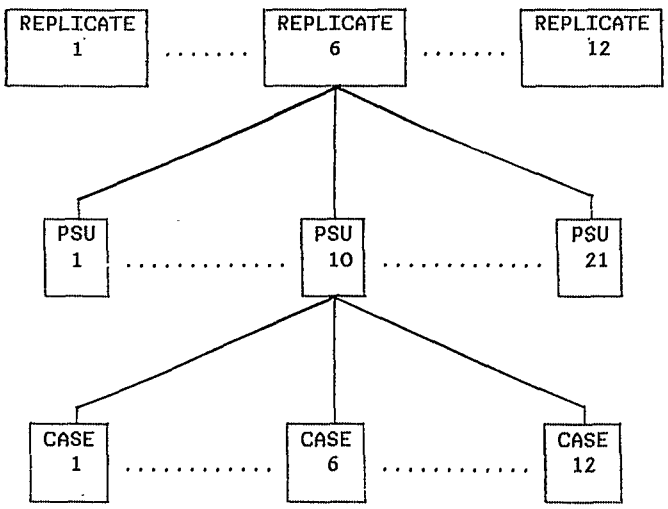


FIGURE 2

SCHEMA OF SAMPLE DESIGN



THE BEHAVIORAL RISK FACTOR SURVEILLANCE SYSTEM

Gary C. Hogelin
Centers for Disease Control

Two years ago at this conference, we reported the first results of our efforts at the Centers for Disease Control to assist State health departments to collect behavioral risk factor information on their adult populations. This report elaborates on our recent efforts, particularly regarding the development of a surveillance system for behavioral risk factors. But, in order for you to better understand our present efforts, I will provide you with some brief background comments.

We began assisting State health departments to collect behavioral risk factor information because State-level health education/risk reduction programs have a basic need for data upon which to develop Statewide objectives and priorities for health promotion efforts. The basic elements of our assistance have been a training program, a standard questionnaire, a standard sampling plan, and data processing.

The program of assistance began in 1981 and continues today, but the nature of our assistance has changed. The data are collected by telephone survey methods. The method was chosen because other means of data collection, such as sales or tax receipts, could not provide the required information. And household interviews were too costly and technically difficult for most State health departments to do regularly. The telephone mode of interviewing was chosen for its relative simplicity and cost features. To date, we have assisted 33 States and the District of Columbia to collect this information. Another 12 States have collected this information but without CDC assistance.

These data collection efforts have been in the form of point-in-time surveys. By that I mean the entire data collection process for each State was conducted during a 1 to 2 week period, and then the data were processed and tabulated. We have now moved our efforts to assist the States into a surveillance system. We use the term "surveillance" to describe data collection which is of a continued, systematic nature. The most distinguishing aspect of surveillance versus a point-in-time survey is that the sample is divided into twelve parts, and interviewing is conducted monthly throughout the year.

This regular State data collection program is called the Behavioral Risk Factor Surveillance System. There are several reasons for moving to surveillance versus point-in-time surveys. Primary among the reasons is to control in some way for the seasonality of the behaviors and to gain the additional analytic flexibility of monthly data collection. Some of the behaviors (i.e., drunk driving) are influenced by certain holidays, and monthly data collection will allow better detection of this seasonality. The data base for each State will increase month by month over time. Future analyses which are dependent on time of data collection will be able to take advantage of this additive aspect.

The staff providing the assistance for the Behavioral Risk Factor Surveillance System are organizationally located in the Division of Nutrition, Center for Health Promotion and Education within CDC. The objectives of the Behavioral Risk Factor Surveillance System are: 1) to monitor the State-specific prevalences of personal health behaviors related to the leading causes of premature death, 2) to determine the seasonal patterns of personal health behaviors, 3) to respond to a health crisis as warranted by providing a rapid means of acquiring population based information, and 4) to incorporate the health-related behaviors that are critical indices of the health status of Americans into State disease control planning and intervention efforts.

The first objective, State-specific estimates, is to provide States with the necessary data upon which to monitor trends in the behaviors and to permit continuing evaluation of priorities for health promotion efforts. The estimates the States use are annual estimates. The second objective, seasonal patterns, is a short term research effort designed to assist the States in the timing of their health promotion efforts. The third objective, responding to a health crisis, takes advantage of the system's routine nature and flexibility. A module of questions about a given health problem can be added to the questionnaire for one or more months to acquire the desired information. This is the only system we are aware of which can respond within 3-4 months of a crisis with data important to addressing the crisis. The final objective, displaying health behavior data, is to bring more emphasis to health behavior as a primary determinant of health. There is an extensive body of literature now available which supports the notion that future advances in health and longevity for Americans will arise more out of changes in behavior versus improvements in medical care. And attention to this new public health perspective by decision makers is critical to future funding priorities.

The Behavioral Risk Factor Surveillance System is standardized whenever possible. The data collection is conducted according to a prescribed schedule. All the States participating in the system conduct interviews at the same time each month and even at the same approximate hours. The sampling plan, a three-stage cluster design, is the same for all participants. The sampling will vary where stratification is desired. The procedures, such as number of attempts for each household and refusal conversion, are prescribed so each participant will not introduce variations affecting the comparability of the data. To provide for greater uniformity in methods, the supervisory staff have been trained according to a set lesson plan and questions which arise are called in to a CDC staff person. And, of course, each participant uses a standard questionnaire developed by CDC with input from the State participants.

During the coming months, the system will implement computer-assisted-telephone

interviewing or CATI. This CATI feature uses microcomputers for direct data entry of respondents' answers. Because the responses are edited at the time of the interview, the CATI feature reduces errors in the data. But, for our purposes, the big advantage in CATI is the rapid turn around time in producing the data.

The system has the ability to adapt and respond to issues of growing importance through the use of new modules of questions. In 1986 we plan to offer the participants a new module of questions on smokeless tobacco, a product which is coming under increasing scrutiny in the public health sector and for which there is a paucity of information available.

A final feature of the system (and, in my judgment, the most important one) is the State-specific nature of the data. Public health programs are traditionally applied at the State and local level. But our best data on health issues is generally gathered at the national level, leaving the program users of the data at a disadvantage. The data are designed to serve the State program users. It is their data base from which to determine their directions and priorities.

The Behavioral Risk Factor Surveillance System is not without its disadvantages. Telephone coverage is not complete in any of the States, and, of course, non-coverage is a source of bias in the estimates. Response rates are another potential source of bias. The States' response rates are comparable to other telephone surveys, about 75%, but this rate is considerably less than that which could be obtained by in-person interviewing. All the data are self-reported and subject to effects of societal attitudes. Alcohol consumption, for example, is uniformly underreported in surveys. One final disadvantage is the lack of participation. Presently, 23 States and the District of Columbia participate in the system. These participants represent 55 percent of the adult U. S. population. Since not every State participates, national estimates cannot be determined from the data. At least nine additional States have indicated interest in participating if sources of funding can be identified.

The subject areas in the system include seat belt usage, hypertension, physical activity, overweight, dieting, cigarette smoking, alcohol misuse, and a host of demographic items. These subject areas were selected for a number of reasons. The first criterion for selection was that the subject item had to be linked in the scientific literature to one or more of the 10 leading causes of premature death in the U. S. Next the data items in each subject area had to be directed to current, personal behavior. The general notion of surveillance implies the current status of a condition so only present behaviors are monitored. The survey deals with personal behavior in order to restrict the items to a reasonable number and because people respond best to questions of an individual nature. Not every behavioral risk associated with premature death is amenable to health promotion strategies, so only those which are amenable were selected. Some behaviors (e.g., family violence and drug abuse) were excluded because they are such sensitive topics that we felt the results would be rendered useless by underreporting. Our last criterion was questionnaire length. The questionnaire developed at CDC is designed to collect the standard core of information in 10 minutes. The participating States may wish to add questions of their choice and by virtue of the short standard questionnaire, can do so without making the interview excessively long.

The States participating in the system are responsible for providing all the personnel, facilities, and data entry. The Centers for Disease Control provides a standard questionnaire, a sampling plan, training, and some portion of the funds necessary to conduct the project. Funds are provided to the participants through the cooperative agreement mechanism. The awards average approximately $17,000 per State.

Surveillance systems are generally evaluated by such qualities as timeliness, representativeness, flexibility, sensitivity and specificity. With regard to timeliness, immediate availability is not nearly so critical for these data as it is for infectious disease data. We set our objectives for final data to be ready within 2 months of their receipt at CDC. Thus far it has taken approximately 4 months to process and return the data to the submitting State -- which is still rapid enough for the intended purposes of the system. However, as we move into CATI we expect the turn around time to diminish dramatically to the point where unweighted tabulations can be produced immediately following interviewing.

Because not every household has a telephone, some households are systematically excluded. Nationally about 7 percent of households do not have telephones. And as pointed out earlier only 55 percent of the adult U.S. population is represented by participating States. This coverage problem leaves us with estimates just for participating States and just representing telephone households. In the future we plan to address the coverage issue by adding more States, developing dual-mode sampling strategies, and conducting a survey of non-participating States to aggregate with participating States.

The system is sufficiently flexible to adapt to another questionnaire within 3-4 months. I mentioned earlier that a module of questions on smokeless tobacco would be added for calendar year 1986. And four States will be attempting new methods in sampling and questionnaire design to obtain more information on the behavioral risks of pregnant women. Both of these are evidence that the system is adequately flexible for its purposes.

The evaluation issues of sensitivity and specificity relate primarily to the design of the questionnaire and the subsequent analyses. The risk factor definitions are continually evaluated for accuracy so sub-populations at risk are correctly identified. And new questionnaire designs which increase self reporting of undesirable behavior may be included if the increase in data quality outweighs the loss of comparability with previous data. As an example, the alcohol misuse questions currently used were added because they increased the reporting of total alcohol consumed.

The Behavioral Risk Factor Surveillance System has been in operation since January 1984. The nature of surveillance versus point-in-time surveys is such that our organization has had to adapt to new procedures. With 24 participants conducting monthly interviews, the data processing workload is several times greater than before. Monthly interviewing has also proven to be less efficient for the States, and is thus somewhat more expensive.

Perhaps the biggest adjustment, though, has been with the evolution into CATI. While CATI produces cleaner data quicker, it requires a large investment in equipment and training. The state-of-the-art in CATI is not such that software can be obtained "off the shelf." Rather the software has to be designed to the users procedures unless the user is willing to adapt to a given CATI software package. In our case we have chosen to adapt the software to our procedures and to accept the resulting need for continuing software design and debugging. We feel our CATI system will suit all our needs in a few more months.

When examining the data for trends, we can compare some of the State results for 1984 with comparable 1982 data. For the majority of the behavioral risk factors, no evidence of any trends can be determined. Each State shows some change but the changes are usually small and occur in both directions. One risk factor, seat belt use, clearly describes a pattern. The proportion of adults reporting that they always or nearly always use seat belts has increased, in some cases quite dramatically, in each State where comparable data exist.

In closing, I feel there is one last organizational adjustment that is worthy of discussion, and this has to do with the use of the data. While there are many applications for which the information is designed and others for which it could be used, reality has not yet matched our expectations for the application of the data. We find, just as others who have been associated with data collection for years have found, that data collection is easier for some than is data application. For some, data collection has become the program itself rather than a mechanism for supporting the program.

In the coming months we will continue to adapt our organization to fulfill the needs of the system. As the data collection and processing become more routine we will direct our attention to a series of workshops designed to provide the participants in the system with the skills to both analyze and apply the data appropriately. Although the system is new, we envision that it will grow and eventually become the most useful source of State-specific information for those in the health promotion and discuss prevention arena.

SURVEY MANAGEMENT AND COST ANALYSIS IN A TELEPHONE SURVEY DESIGN

J. Michael Bowling

The North Carolina State Center for Health Statistics conducted a simple random/random digit dialed telephone survey of NC households during November of 1984. The survey population consisted of parents of children either residing at home or away at college. The Childhood Injury Survey was designed to collect information on parental knowledge, attitudes and practices that impinge upon their children's safety at home or in an automobile. Further, injury rate calculations were to be made of the age group 19 and under for the preceeding year of childhood experience.

This being the first completely in-house telephone survey of its kind conducted by the State Center, efforts were made in the design process to allow for a study of factors under the control of the survey coordinator that may improve the efficiency and quality of subsequent telephone surveys.

Interviewer training for the survey was conducted on November 2 with the survey commencing the next day. A pool of interviewers was drawn from Division of Health Service volunteers who were given time off for time worked in the case of professional staff and overtime compensation in the case of clerical volunteers. Fifty-three volunteers worked on the project with afternoon screening conducted by the more inexperienced interviewers and night dialing conducted by a core of 15 interviewers that averaged over 25 completions during the month. Paid interviewers were also used for day and night dialing. Hours for the survey were 1:00 - 5:30 and 6:30 - 9:30 on weekdays and 1-5 on weekends.

Each number chosen was dialed either until a terminal result code was obtained or until 5 contacts had been made without a callback request. For those numbers for which a callback was requested, dialing continued until the end of the survey.

Post stratification and adjustment for undercoverage were employed after the survey was completed. Variables chosen for adjustment and post-stratification included race, single-parent status, county level of urbanization, and region of the state.

Initially, we analyzed the respondent contact sheet data to determine the importance of interviewer characteristics on completion rates. As stated earlier, we had a large pool of volunteer and paid interviewers with 15 core interviewers to conduct a majority of the interviews. Of our total interviewer pool 81 percent were female, and the racial breakdown was 56 per cent white and 44 per cent nonwhite. No interviewers had prior experience in telephone survey interviewing.

Table 1 displays the results of a logistic regression of the likelihood of a completion in a call attempt on the interviewer characteristics of race, with white interviewers given a value of 1 and nonwhite a value of 0, sex of interviewer with males having a value 1 and females 0 and experience of the interviewer which is a continuous variable created as the sum of all previous interviews completed by each interviewer beginning with 0. Two other measures of experience were also created but are not shown on Table 1. These include the total number of contacts attempted by the interviewer prior to a call attempt and the total number of times the interviewer talked with someone in the call attempt irrespective of a completion. We have calculated the effect of experience in Table 1 after 17 completed surveys which was an average level of experience for interviewers.

We have presented both additive effects on the log odds of completion and the multiplicative effect on the odds of completion. We will interpret the multiplicative parameters as they are somewhat more understandable.

The intercept (.027) indicates the very small likelihood of a completion in a call attempt for the interviewers in the omitted categories i.e. nonwhite, female interviewers with no experience on a day telephone contact. Values of the effect parameters greater than 1 indicate an increased likelihood of a completion by that proportion to the right of the 1. As an example the odds of a completion increase for night dialing, other things being equal, from .027 to .05 an increase of 90 per cent. Parameters less than 1 indicate a decline in the odds of completion when he value of the variable is 1.

As you can see from the table white interviewers were 32 per cent more likely to complete an interview on a call attempt than nonwhite interviewers, male interviewers were slightly less likely to complete an interview than females and interviewers with 17 previous completions were 24 per cent more likely to complete an interview than interviewers with no experience. Both experience with contacts and through talking with someone in a contact increase the odds of a completion with contacts much important than talking with a contact and experience gained from completing the interview most important of all.

Night dialing, here used as a control, is much more productive than day dialing with almost double the chances of a completion in a call attempt. The likelihood of a completion increases to .05 for night dialing with experienced, female, white interviewers having odds of .084 of completing an interview.

For every refusal, 22.5 additional calls were made to attain a completion. That translated into an average of 19 minutes of dialing, 17 minutes for the additional calls and 2 for the refusal. If the actual ratio of completion to calls had improved to the extent that it increased far experienced over nonexperienced interviewers over 4,000 calls could have been prevented. While this is perhaps overly optimistic given inefficiences of simple random sampling telephone sampling,

considerable savings in fewer refusals and call back requests for an eligible respondent could have been made.

The likelihood of a completion in a call attempt is one factor that contributes to overall interviewer proficiency. Another factor that is important is the duration of time required by interviewers to complete an interview. Table 2 presents an OLS regression of the duration of time to complete an interview on interviewer characteristics. Again, interviewer experience is an important predictor of interviewer proficiency. An interviewer who has completed 17 interviews averaged 1 minute less than inexperienced interviewers. Males tend to complete interviews is less time than female interviewers by over 1.5 minutes. This perhaps is the result of female respondents answering questions with little elaboration when talking with male interviewers. Seventy per cent of the respondents to the CIS were female.

After the survey was completed, call sheets containing ID, time of call, date of call, and result code (HH result codes or respondent codes) were key punched for each number contacted. We also obtained call charge information for each charged number (6,868). Each call sheet was disaggregated into each call attempt on each selected telephone number and charge information with the duration of the call was matched by number and day to the call sheet data. This information allows us the opportunity to determine precisely the number of calls needed to complete the survey and illuminate factors that may be manipulated in future surveys to reduce costs through possibly an improvement in the response rate and a decline in the number of calls needed to complete the survey.

In conducting a cost analysis of the CIS, we will concentrate upon the number of calls made and the duration of calls without applying a dollar value to each.

Analysis of call sheet information indicates that 13,485 randomly generated numbers were selected for contact by our interviewers. A total of 23,935 calls were made to these numbers with a likelihood of completion on each call attempt of 4 per 100 calls. A response rate of 85 per cent was attained after refusal conversion procedures were used to increase the response rate 3%.

The likelihood of completion in a call attempt is a useful construct in the consideration of factors that may improve survey efficiency and more importantly reduce the most visible flaw in the survey operation - unit nonresponse in which a unit of importance to the survey fails to participate.

Through the use of logistic regression for binary dependent variables the importance of interviewer characteristics and timing of call influences can be assessed on the odds of a completion in a call attempt (Harrell, 1980).

The use of the odds of completion as a basis for an evaluation of the data collection process is consistent with the stochastic view of nonresponse in which factors such as survey topic, the level of interviewer training, and demographic characteristics of each population member affect whether a selected respondent will complete a questionnaire.

Through monitering completions per hour during the survey, we were aware that completion rates varied by both day and hour of the call. This prompted us to determine with the use again of logistic regression the best days and hours of dialing in terms of odds of a completion. Figure 1 is a graph of day of the week by the multiplicative effect parameter calculated with Friday as the omitted category. Friday also happens to be the worst day of dialing with Mondays and Sundays the most productive.

The next graph displays the effects on the odds of a completion by hour of the day in which the call was dialed. Here the 1:00 - 2:00 calls have been omitted. Afternoon dialing declines from that hour with 5:00 - 5:59, the least productive hour. The 6:00 - 6:59 hour is the most productive period to obtain completions from parents. The 9:00 to 9:59 hour (most calls 9:00 - 9:30) was not sufficiently late to increase the nonresponse. In summing up these analyses, interviewer characteristics did make a difference in the likelihood of a completion. Frey (1983) argues for the development of a trained interviewer pool, probably more so to limit training time and have interviewers familiar with the data collection format. On the basis of this analysis, higher response rates and shorter data collection periods may also be the product of using experienced interviewers and scheduling personnel to maximize the dialing during the most effective periods.

The last question we wanted to answer with this analysis is how would we have fared in terms of total calls expended using a Waksberg cluster approach (Waksberg, 1978). If numbers had been chosen and screened prior to conducting interviews we would have made 5874 additional calls to attain 340 clusters with 3 completions anticipated from each cluster. This, all things being equal, would have been 21.5 per cent less efficient than the SRS procedure. We are now attempting to secure call sheet data from another survey which will enable us to determine the average proportion of numbers in NC clusters that are active. This will allow the determination of increased efficiencies of the Waksberg design over a simple random survey design.

REFERENCES

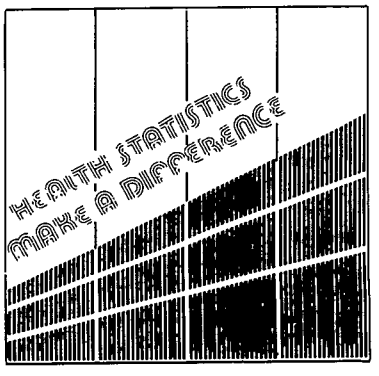
Frey, James H. (1983) Survey Research by Telephone. Sage Library of Social Research, Vol. 150: Beverly Hills.

Harrell, Frank (1980) "Logist Procedure." Supplemental Library User's Guide. SAS Institute: North Carolina.

Waksberg, Joseph (1978) "Sampling Methods for Random Digit Dialing." Journal of the American Statistical Association. 73: 40-46.

Session X

Vital Statistics Related to Infant Mortality and Low Birth Weight



NATIONAL BIRTHWEIGHT-SPECIFIC INFANT MORTALITY SURVEILLANCE

Jack C. Smith
Carol J. R. Hogue
James W. Buehler
Lilo T. Strauss
William M. Sappenfield
Centers for Disease Control

I. INTRODUCTION

The 1990 objectives for the nation state that the greatest single problem associated with infant mortality is low birthweight (1). Nearly two-thirds of the infants in this country who die are low birthweight. Two specific 1990 objectives relate to lowering the risk of low birthweight, but no national data exist that provide birthweight-specific mortality rates. I will report today on an interim project to develop national birthweight-specific infant mortality rates for the 1980 live-birth cohort.

II. BACKGROUND

The last national data on birthweight-specific infant mortality are now over 20 years old. These data were collected in a sample follow-back survey of infant deaths conducted by the National Center for Health Statistics (NCHS) for the 1960 live-birth cohort. Statistics from this study were published by NCHS in their Rainbow Series, Vital and Health Statistics (2).

In 1981 NCHS proposed to develop a national computerized file of infant death certificates, linked to birth certificates plus fetal death certificates. However, after considerable development of this plan, implementation was postponed because of resource constraints.

In 1982 the Centers for Disease Control (CDC) began to discuss an interim approach to collecting national infant mortality statistics. We refer to the project that has evolved as "NIMS," the National Infant Mortality Surveillance project. The purpose of the interim project was not to supplant a national computerized system proposed by NCHS, but rather to fill an existing data gap as expeditiously as possible. The Division of Reproductive Health at CDC sought and received support for the NIMS project from the National Institute of Child Health and Human Development (NICHD). Specifically, the purpose of NIMS is twofold:

1. To produce a national report describing the maternal and infant factors related to birthweight that are associated with infant mortality.
2. To provide expertise to guide and assist the development of ongoing State and national surveillance and research on infant mortality.

III. METHODS

Methodologically, the primary issue of the project centered around the variable of birthweight. Non-birthweight-specific infant mortality rates at both the State and national level are routinely produced to be used for a variety of epidemiologic and programmatic purposes. Producing birthweight-specific infant mortality rates, however, presents a much more complex methodologic and logistic problem. Briefly, the problem centers around the fact that the variable of interest, birthweight, is not available from an infant's death certificate. Rather, for a deceased infant, that infant's birth certificate containing birthweight must be located and linked with the death certificate. This linkage responsibility currently occurs at the State level and is usually done by the State Office of Vital Statistics. While linkage of records of any kind is seldom easy because of difficulties with identifiers such as name and address, the problem of linkage for infant mortality is compounded by the fact that the State of birth and the State of death may not be the same, thus requiring exchange of certificates between States to effect linkage.

In addition to the methodologic problems presented by linkage, birthweight-specific infant mortality statistics that have been produced have lacked uniformity and thus have not lent themselves to either comparison or aggregation. For example, some States produce birthweight-specific infant mortality statistics based on a death cohort (i.e., infants dying in a given calendar year), while other States use a birth cohort (i.e., infants born in a given calendar year). NIMS uses the resident birth cohort for 1980 which includes all infants who were born in 1980, but died within a year of birth in either 1980 or 1981.

Another important lack of uniformity is the birthweight variable itself. Some States produce birthweight-specific statistics by pounds and ounces, some by grams. Also, States vary in the categories of birthweight intervals they use in tabulation of their data. Even the cutoff point to dichotomize into low birthweight and non-low birthweight is not uniform. Some States use less than or equal to 2500 grams, some use less than 2500 grams. NIMS uses the World Health Organization (WHO) definition of less than 2500 grams as low birthweight.

To discuss the issues above--linkage, birth cohort versus death cohort, categories of birthweight--and other methodologic and logistic issues, in May 1983 CDC convened a planning session in Atlanta. In attendance were seven members of the Executive Committee of the Association of Vital Records and Health Statistics (AVRHS), and representatives from the American Academy of Pediatrics, American College of Obstetrics and Gynecology, State Directors of Maternal and Child Health (MCH), the National Center for Health Statistics, the National Institute of Child Health and Human Development, and the Centers for Disease Control. As a result of the planning session,

CDC was encouraged to proceed with the NIMS project. The decision was also made to ask each State to provide CDC with a set of birthweight-specific tabulations of their infant deaths based on the 1980 resident birth cohort. States, insofar as possible, would provide the birthweight-specific tabulations for neonatal deaths (deaths at less than 28 days of age) and postneonatal deaths (28 days to one year) separately by race (white, black, and total). It was also agreed that the birthweight-specific denominator of births necessary to calculate birthweight-specific infant mortality rates would be produced at CDC from the 1980 national natality data file provided by NCHS.

At the annual meeting of the AVRHS in Portland, Oregon, July 1983, CDC and NCHS both made presentations to the Association's membership urging their States' participation in NIMS as an interim project and their States' cooperation as plans were to be made for an eventual ongoing national linked infant mortality data base.

The NIMS project also received support from the membership of other organizations such as the Association of State and Territorial Health Officers (ASTHO) and the Association of State MCH Directors.

IV. IMPLEMENTATION

In the summer of 1984, letters went to the 50 State health departments and the health departments of New York City, the District of Columbia, and Puerto Rico requesting their participation in NIMS. All 53 reporting areas agreed to participate.

The amount of effort to participate in the NIMS project was not equal for all States. Some States already had linked record files and had the capacity to readily produce birthweight-specific infant mortality statistics. Some States, however, had neither linked record files nor the capacity to readily produce birthweight-specific infant mortality statistics, even if linked files were available.

CDC staff worked closely with State health department staff to resolve a myriad of definitional and operational problems, but all States responded with data. For States with comparatively poor or nonexistent systems for creating a linked birth-death data base, NIMS provided the impetus for the health department to invest the resources necessary to move forward with defining their infant mortality data needs and implementing or improving their data system. For States with comparatively good systems for creating a linked birth-death data base, NIMS provided an opportunity to re-examine definitions, linkage procedures, and data quality.

Once the tabular data of birthweight-specific infant deaths requested from the States were received at CDC, the data were entered into a computer file and edited. A similar file of birthweight-specific live-birth data was developed by CDC from the national natality file. These two files of data will eventually yield rates of birthweight-specific infant mortality for a national report.

Because all States could not respond uniformly to our request for data, CDC has had to do a certain amount of manipulation of the data to make it as definitionally uniform as possible. For example, one State could not provide the 250 gram birthweight intervals requested; thus CDC elected to redistribute that State's data based on the birthweight distribution of States that provided 250 gram intervals. Cases of nonuniformity of data that cannot be rectified through statistically acceptable data manipulation will be footnoted in the report. For example, two States' racial definition of infants differed from the NIMS definition.

V. DISSEMINATION OF FINDINGS

The NIMS data set contains infant mortality data from over 43,000 linked records. This represents approximately 95% of all infant deaths occurring in the United States to the 1980 birth cohort.

Findings from NIMS will be disseminated in three ways: (1) at a national conference, (2) in a national surveillance report, and (3) through scientific articles presented at professional meetings and published in professional journals.

CDC is planning to host a national conference in the spring of 1986 in Atlanta. The conference will bring together both programmatic people involved in developing and implementing programs aimed at lowering infant mortality and statistical people involved in producing better linked infant mortality data for research and program evaluation. At the conference, the preliminary findings from the NIMS project will be presented and conference participants will provide feedback regarding the suggested final content of a national surveillance report.

A national surveillance report will be published by CDC and disseminated to national, State, and local health agencies. The report will provide data on maternal and infant characterisitics by birthweight. The report will feature analysis for neonatal and postneonatal mortality rates by race.

Scientific articles will explore certain issues related to infant mortality, such as cause of death, that can be addressed specifically by analysis of information in the NIMS data set. These articles will also help to formulate strategies for reducing infant mortality, based on findings from NIMS.

Preliminary findings from NIMS indicate that 1980 infant mortality rates for both white and black and other races are about one-half the rates for 1960. For whites the decrease was from 22 to 10 infant deaths per 1,000 live births. For black and other races the decrease was from 41 to 19. Analysis also indicates that both for whites and for black and other races, approximately 90% of the decrease in infant mortality can be attributed to a reduction in birthweight-specific mortality, while approximately 10% can be attributed to a more favorable birthweight distribution in 1980 than in 1960.

VI. CONCLUSION AND SUMMARY

Based on the enormous interest and concern about the nation's infant mortality rate, it

is fitting and appropriate that the State health departments have joined together with CDC to produce the NIMS data set. Since more than two decades have passed since the last such data set was developed and since no other similar national data set is currently forthcoming, the NIMS data set provides a unique source of information for the health community to use in developing strategies for lowering infant mortality. And, importantly, in the process of developing the NIMS data set, much has been learned methodologically and statistically to guide the development and improvement of ongoing infant mortality surveillance at both the State and national level.

REFERENCES

1. Promoting Health/Preventing Disease, Objectives for the Nation, Public Health Service, Dept. of Health and Human Services, U.S. Government Printing Office, Washington, D.C., Fall 1980.
2. Armstrong RJ. A Study of Infant Mortality from Linked Records by Birth Weight, Period of Gestation, and Other Variables, United States, Vital and Health Statistics, Series 20, Number 12, DHEW Publication No. (HSM) 72-1055, May 1972.

CLASSIFICATION OF 1982 PERINATAL DEATHS IN MASSACHUSETTS

Edward J. Lammer, Massachusetts Department of Public Health and Massachusetts General Hospital; Lisa E. Brown; Marlene T. Anderka, and Bernard Guyer, Massachusetts Department of Public Health

Infants are healthier now than ever. But after more than a decade of decline in infant and neonatal mortality rates, there is evidence of a recent increase. In 1982, the Massachusetts Infant Mortality Rate (IMR) increased for the first time in nine years and by the largest amount in 17 years (from 9.6 to 10.1 deaths per 1000 live births between 1981 and 1982). We developed this analytic scheme to better understand the causes of perinatal deaths and to identify specific groups of infants who may benefit most from intervention strategies.

Background

Our classification scheme and others [1,2,3,4,5] share in the attempt to better define causes of perinatal deaths. Previous classification schemes depended on the biases of the classifier [6] and hence, interpretation of their findings may have limited generalizability. For example, classification by pathologic findings, as suggested by Bound,[1] included only infants examined by autopsy. The cause-of-death assignment required of the pathologist may lead to loss of information and depends on uniform decision as to which of a number of findings is the major one. Classification according to clinical cause, as postulated by McNay and Baird,[2,3] relies heavily on clinical interpretation rather than objective data, making it difficult to assess potentially preventable factors. The roots or our scheme derive from the pathophysiological classification developed by Wigglesworth[4] and a cause-specific classification recently developed by McCarthy.[5]

Wigglesworth grouped causes of perinatal death into five groups that were not dependent on pathological information but rather had common intervention strategies. This was the unique feature of his scheme. He, however, did not provide adequate quidelines for assigning abstracted death reports into his five cause-of-death categories. Both Wigglesworth and McCarthy incorporated birthweight into their schemes, the most powerful determinant of infant mortality, into the analysis. [5] McCarthy expanded and operationalized Wigglesworth's classification by including postneonatal deaths in the analysis and outlining the International Classification of Diseases (ICD) codes appropriate for each cause-of-death group. He also extented Wigglesworth's concept of linking the cause-of-death classification scheme to specific intervention strategies.

Wigglesworth's and McCarthy's models include similar groups into which perinatal deaths may be classified. We adopted these cause-of-death groups for our analysis. Our classification scheme is unique since we created an algorithm to assign underlying causes of death. This was done rather than using ACME (Automated Classification of Medical Entities), the algorithm used for most national mortality analyses. The ACME program assigns the single underlying cause of death to each record depending on the relative position held by items entered in Parts I and II on the completed death certificate. Thus, the ACME algorithm could select a different single underlying cause of death if the order of items entered on a death record was changed. Our algorithm was developed to help compensate for the inconsistencies and inaccuracies associated with the order in which causes are completed on the perinatal death records.

Methods

We used the vital records filed at the Massachusetts Department of Public Health, Registry of Vital Records and Statistics. All Massachusetts birth, death, and fetal death records conform to the U.S. and W.H.O. standard format (National Center for Health Statistics, 1978 revision). Selection criterion for inclusion in the analysis scheme were as follows: (1) all death certificates of live-born infants (neonates aged less than 28 days at the time of death and infants less than one year old at the time of death) who were delivered to Massachusetts residents in 1982, and (2) all 1982 fetal death certificates for Massachusetts residents. Massachusetts law requires that fetal deaths of 20 or more weeks' gestation or of a weight of at least 350 grams be reported. Only these were included in the analysis. A fetal death per chapter 111 section 202 of the Massachusetts General Law is a death prior to the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of the pregnancy; the death is indicated when, after such separation, the fetus does not breathe or show any other evidence of life, such as the beating of a heart, pulsation of the umbilical cord, or definite movement of the voluntary muscles. A fetal death does not include an induced abortion as defined in section 12K of chapter 112 of M.G.L. Spontaneous abortions are included.

The key variables used in our classification of both fetal and infant deaths were birthweight, age at the time of death, and cause of death. Causes of infant and fetal deaths were abstracted from the death records using the 9th revision of the ICD.[7] The first three entries of Parts Ia, Ib, Ic, and Part II were abstracted for the infant death records and keypunched onto computer tape. The first entry for the Congenital anomaly section on the birth certificate was also abstracted. Infant death certificates were

hand-linked to their corresponding birth certificates and the abstracted versions of the two records were merged on one computer tape. For fetal death records, only the first entries in Parts Ia, Ib, Ic, Part II, and the Congenital Anomaly section were abstracted and keypunched. Altogether, there were thirteen fields possible from which to select the underlying cause of death for each infant death, and five from which to select the underlying cause of fetal death (see Table 1).

We assigned each perinatal death to one of six underlying cause-of death categories appropriate for either the neonatal or fetal deaths. These groups were, as much as possible, mutually exclusive and had distinguishable intervention strategies (see Table 2).

Each record was scanned and hierarchically assigned to a cause-of-death category using a SAS array procedure.[8] Initially, the records were scanned for ICD codes allocated to the Congenital Anomalies category. If a record was identified with one of these ICD codes, it was classified as a death due to Congenital Anomalies and outputted from the file. After all records were scanned and the appropriate assignments made to Congenital Anomalies, the remaining records were scanned for ICD codes corresponding to the next of the six categories (for either fetal or infant deaths) listed in hierarchical order, Infection (see Table 2). This process continued until each record was assigned to one of the cause-of-death groups.

TABLE 1

Data Fields for Assigning Underlying Cause of Death

Time of Death	Source	Number of Fields *
Neonatal	Death Certificate:	
	Part Ia	3
	b	3
	c	3
	Part II	3
	Birth Certificate:	
	Congenital Anomaly Box	1
	Total number of fields =	13
Fetal	Fetal Death Form:	
	Part Ia	1
	b	1
	c	1
	Part II	1
	Congenital Anomaly Box	1
	Total number of fields =	5

* One ICD cause-of-death code per field

TABLE 2

Hierarchical Cause-of-Death Categories and Corresponding ICD Codes

Fetal Deaths	Infant Deaths
Congenital Anomalies: 2594, 5531, 7400-7469 7471-7478, 7480-7599	Congenital Anomalies: 2594, 5531, 7400-7469 7471-7478, 7480-7599
Infection: 7712, 7718, 7602, 7627, 7700, 7711	Infection: 0010-1398, 3200-3249, 3910-3919 4200-4229, 4610-4619, 4640-4661, 4800-4809 4820-4829, 4840-4848, 485, 486, 4870-4878 5070, 5100-5109, 7700, 7710-7714, 7718
Asphyxia: 7670, 7680, 7681	Prematurity: 4310, 7640-7651, 7690, 7704, 7707 7721, 7722, 7775
Maternal Conditions Affecting Pregnancy: 7600 7601, 7603, 7604, 7610-7614, 7620-7626, 7628 7629, 7638, 7902	Trauma During Delivery: 4160, 7479, 7670-7679 7680-7689, 7701, 7990
Other Conditions: 7605-7609, 7616, 7618, 7619 7730, 7733, 7765, 7780	Sudden Infant Death Syndrome (SIDS): 7980
Unknown: 7615, 7636, 7640-7642, 7649-7651, 7660 7704, 7762, 7798, 7799, 7999	Others: 1940, 2040, 2280, 2396, 2500, 2521 2710, 2758, 2762, 3350, 3358, 3456, 3578 4275, 4279, 4289, 4298, 430, 4349, 5119 514, 5570, 5602, 5734, 586, 5938, 7339, 7611 7627, 7703, 7705, 7708, 7725, 7732, 7757 7780, 7785, 7798, 7823, 7991, 7999, 803 851, 854, 861, 8798, 9010, 912, 933, 9590 9679, 9682, 986, 9878, 9916, 9941, 9947

The hierarchical arrangements for fetal and infant causes of death were determined by first, ordering causes from the least to the most preventable. Not all causes of death can be distinguished in this way since, more often than not, the underlying cause is a subjective determination from several causes, all of which may have contributed to the perinatal death. When the distinction between preventable versus non-preventable was unclear, we ordered the categories by trial and error to reduce the potential for misclassification since some records had codes corresponding to more than one of the six categories. The potential for misclassification was greatest for infant death records with ICD codes corresponding to both Prematurity and Trauma categories. We found that misclassification was much lower if Prematurity preceded Trauma in the hierarchy (see Matrix 2 comments), even if trauma-related infant deaths may have been more amenable to interventions. SIDS was the only cause listed on records allocated to that category. There was thus no chance for misclassification and it made no difference where SIDS was placed in the hierarchy. The Other categories were formed since the occurrence of the same conditions was not high enough to warrant separate categories. Records assigned to the Other categories were those remaining after all other listed causes were excluded.

Following the cause-of-death assignments, we created a 3x4 matrix that further separated our perinatal population into birthweight- and age-specific groups. First, the 1982 perinatal death records were stratified into four birthweight groups: $\leq$ 1499 grams, 1500-2499 grams, $\geq$ 2500 grams, and unknown. The birthweight categories were then separated into three age-at-time-of-death groups: fetal, neonatal, and postneonatal (see Figure 1).

FIGURE 1: Perinatal Mortality Matrix Scheme[5]

AGE AT DEATH	BIRTHWEIGHT ≤1499	1500-2499	≥2500 gms
Fetal			
Neonatal		Causes of Death	
Post-Neonatal			

* Causes are assigned among all cells

Analysis of 1982 Perinatal Deaths: One Application of our cause-specific classification scheme

There were 574 fetal death records, 561, neonatal death records, and 193 postneonatal death records included in our analysis. The records were assigned to cause-of-death groups and organized into fetal, neonatal, and postneonatal matrices. The cells of each matrix contain the actual number of deaths from the different causes followed by the corresponding rate in parenthesis (see Table 3). The matrices are presented below with brief comments for each.

TABLE 3

Definition of Mortality Rates

Fetal Mortality Rate = $\frac{\text{Number of 1982 resident fetal deaths in a birthweight- and cause-specific group}}{\text{Total 1982 live births* and fetal deaths in a specific birthweight group}}$ x 1000

Neonatal Mortality Rate = $\frac{\text{Number of 1982 resident deaths at < 28 days in a birthweight- and cause-specific group}}{\text{Total 1982 live births in a specific birthweight group*}}$ x 1000

Postneonatal Mortality Rate = $\frac{\text{Number of 1982 resident deaths at } \geq \text{ 28 days and < 1 year in a birthweight- and cause-specific group}}{\text{Total 1982 live births* in a specific birthweight group*}}$ x 1000

* Total 1982 live births in Massachusetts = 75,749:

843 weighed	≤ 1499 gms
3,613	1500-2499 gms
71,244	≥ 2500 gms
58	Unknown

MATRIX 1: 1982 Massachusetts Fetal Deaths *

CAUSES	BIRTHWEIGHT ≤ 1499	1500-2499	≥ 2500 gms	Unknown
Congenital Anomalies	35 (40.3)	16 (4.4)	17 (0.2)	6 (93.7)
Infection	15 (17.7)	2 (0.5)	5 (0.1)	3 (49.2)
Asphyxia	62 (69.2)	37 (10.1)	55 (0.8)	4 (64.5)
Maternal Conditions	92 (99.3)	25 (6.9)	51 (0.7)	3 (49.2)
Other	9 (10.7)	4 (1.1)	3 (<0.1)	4 (64.5)
Unknown	76 (83.5)	16 (4.4)	34 (0.5)	0

* Each matrix cell contains the number of deaths followed by the Fetal Mortality Rate in parenthesis.

Matrix 1: Fetal Deaths

Asphyxia accounts for approximately one sixth of the mortality among stillbirths since there are close to 600 fetal deaths. The fetuses dying at ≥ 2500 gms from asphyxia are of interest to us because they are among the most preventable deaths. Further study of deaths from asphyxia can be accomplished by looking at time of death, either during or before labor. The intervention strategies most important to consider for stillbirths in the ≥ 2500 gm category would perhaps be better obstetrical services and training of medical personnel. There is also a large proportion of deaths from maternal conditions related to the pregnancy. These deaths are likely to be preventable and warrant further attention and study.

Many deaths among the ≤ 1400 gm category are from unknown causes. Autopsy reports and perinatal audits will be informative and useful to better distinguish the underlying causes of death. Perhaps asphyxia was the assigned cause of death when the cause was really unknown in the ≥ 2500 gm fetuses.

MATRIX 2: 1982 Massachusetts Neonatal Deaths*

CAUSES	BIRTHWEIGHT ≤ 1499	1500-2499	≥ 2500 gms	Unknown
Congenital Anomalies	52 (62.3)	40 (11.1)	65 (0.9)	8 (137.9)
Infection	14 (16.8)	8 (2.2)	11 (0.1)	0
Prematurity	277 (332.1)	6 (1.7)	6 (0.1)	3 (51.7)
Trauma	3 (3.6)	6 (1.7)	26 (0.4)	3 (51.7)
SIDS	0	0	8 (0.1)	0
Other	7 (8.4)	6 (1.7)	11 (0.1)	1 (17.2)

* Each matrix cell contains the number of deaths followed by the Neonatal Mortality Rate in parenthesis.

Matrix 2: Neonatal Deaths

Congenital anomalies are the leading cause of death for all except the lowest birthweight category. As we would expect, prematurity is the leading cause for the ≤ 1499 gram neonates. Intervention strategies to help reduce the number of deaths from prematurity, might be identifying women at risk for delivering premature infants and providing intensive care transport and ICU services.

The greatest potential for misclassification is between the Prematurity and Trauma categories. Only about 10% of the 277 infants assigned to the Prematurity category have ICD codes corresponding to Trauma, in addition to codes belonging to Prematurity. If the deaths were separated in the reverse order, i.e., Trauma before Prematurity, 43% of the records assigned to Trauma would have codes corresponding to the Prematurity category. More attention should be given to these two groups of infants to better assess the need for, and applications of intervention strategies. Interventions appropriate to consider for infants dying from trauma or prematurity in the ≥ 2500 gram category, may be improved obstetrical and community services.

Matrix 3: Postneonatal Deaths

There are no unusual findings in this matrix. SIDS is the leading cause of death and the postneonatal mortality rates are inversely related to birthweight. Most would agree that SIDS is difficult to prevent, but interventions considered appropriate for infants dying from causes such as congenital anomalies and infections may be to encourage access to, and promotting the quality of community health services and prenatal screening. Some premature infants who previously would have died during the neonatal period, are surviving until the postneonatal period because of improved medical care. A resulting shift in mortality trends may lead to the interpretation that mortality from prematurity is decreasing when, in fact, it may not be. This shift did not have much influence on the postneonatal mortality rates in our analysis since there were only nine prematurity-related deaths during the postneonatal period.

MATRIX 3: 1982 Massachusetts Postneonatal Deaths*

CAUSES	BIRTHWEIGHT ≤ 1499	1500-2499	≥ 2500 gms	Unknown
Congenital Anomalies	3 (3.6)	11 (3.0)	22 (0.3)	2 (34.5)
Infection	5 (6.0)	4 (1.1)	11 (0.2)	0
Prematurity	7 (8.4)	2 (0.5)	0	0
Trauma	0	0	2 (<0.1)	0
SIDS	5 (6.0)	12 (3.3)	68 (1.0)	2 (34.5)
Other	2 (2.4)	4 (1.1)	29 (0.4)	2 (34.5)

* Each matrix cell contains the number of deaths followed by the Postneonatal Mortality Rate in parenthesis.

Discussion and Conclusion

This classification scheme is one way to begin using and understanding the causes of fetal and infant deaths as they relate to the intervention strategies which might most effectively help reduce the Massachusetts IMR. Continued improvement in infant mortality depends on the commitment of resources for those who will benefit most from such interventions. We demonstrated with our analysis of 1982 perinatal deaths, how this model can be used to identify subgroups of deaths for which perinatal audits may be justified. The outcome of a perinatal audit would be to develop appropriate methods for learning how to reduce perinatal mortality. Eventually, perhaps our scheme could be used by policy makers, as a tool to evaluate the necessity for, and impact of intervention programs. Before this can happen, we need to understand and decrease the limitations of our model.

A scheme using multiple cause-of-death categories, as ours does, provides useful information about the potential for misclassification. For example, we are aware of the potential misclassification between the Trauma and Prematurity cause-of-death categories and special attention should be used in assessing intervention needs for either group. Our algorithm assigns underlying cause of death both for the fetal and infant deaths while the ACME program cannot assign causes for fetal deaths. We attempted to compensate for inconsistensies in death reporting; however, the assumption made when we devised our algorithm was that all cause-of-death information has equal significance, independent of the relative position held on the record. This may or may not be true. Our scheme is limited by the accuracy with which infants and fetuses are assigned to the underlying cause-of-death categories. Further refinement and tests of accuracy, such as comparing our analysis with autopsy and perinatal audit findings among geographically distinct populations or groups of deaths from a specific cause over time will enable us to better assess temporal or geographic trends in perinatal mortality.

Acknowledgments

The authors extent their thanks to Elaine Trudeau, Phyllis Zeuli, Michael Healey, Marion Chapman, Ellen Cussler, Linda Juster, and Dan Torosian for their assistance in obtaining this data.

References

1 Bound JP, et al: Classification and Causes of Perinatal Mortality. *BMJ* 1956;2:1191-96.

2 McNay MB, et al: Perinatal Deaths: Analysis by Clinical Cause to Assess Value of Induction of Labor. *BMJ* 1977;1:347-350.

3 Baird D, Walker J, Thomson, AM: The Causes and Prevention of Stillbirths and First Week Deaths Part III: A Classification of Deaths by Clinical Cause: The Effect of Age, Parity and Length of Gestation on Death Rates by Cause. *J Obstet Gyn Brit Empire* 1954;61:433-488.

4 Wigglesworth, JS: Monitoring Perinatal Mortality: A Pathophysiological Approach. *Lancet* 1980;2:684-686.

5 McCarthy B: The use of Birthweight in Perinatal Surveillance and Evaluation. Presented November 1, 1983, Beijing, Peoples Republic of China.

6 Walker J: Classification, in Chalmers I, McIlwaine G (eds): *Perinatal Audit and Surveillance.* Proceedings of the 8th Study Groups of RCOG. 1980, pp 211-220.

7 *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death.* Based on the recommendations of the Ninth Revision Conference, 1975, and adopted by the 29th World Health Assembly WHO Geneva, 1978.

8 *Statistical Analysis System Users Guide: Basics,*1982 edition. SAS Institute, Inc. Cary, NC.

PRENATAL CARE AND ITS RELATIONSHIP TO MEDICAID COSTS

Wayne Schramm and Garland Land
Missouri Center for Health Statistics

Numerous studies have examined whether prenatal care reduces the risk of a bad pregnancy outcome. The results of these studies have been somewhat contradictory due to the use of different populations and methodologies, but most have shown that the receipt of adequate prenatal care generally improves pregnancy outcome, in particular reducing prematurity by birth weight.

If adequate prenatal care does indeed reduce the risk of prematurity, then adequate prenatal care should also reduce the costs for the mother and child at and immediately following birth. Among Missouri Medicaid births in 1980, it was found that Medicaid costs within 30 days after birth were three times higher for low birth weight (under 2,500 grams) infants than for normal weight babies. For very low birth weight infants (under 1,500 grams), the cost differential was 8 to 1 (Missouri Monthly Vital Statistics, 1982). An extremely premature delivery for which the baby is transferred to a neonatal intensive care unit can cost tens of thousands of dollars.

There has been a scarcity of studies that have examined the cost benefits of prenatal care. Among the few studies that have been done, most have primarily involved the calculation of synthetic estimates based on various assumptions. Among these are Blackwell, 1983, which calculated a 2 to 1 benefit cost ratio for prenatal care in California and Behrman, et al., the Institute of Medicine, 1985, which estimated a 3 to 1 ratio if the low birth weight rate were reduced from 11.5 to 9.0 percent among an indigent population.

The study most similar to the current study was done by Malitz, 1983, on a sample of Texas Medicaid births in 1981. Malitz examined birth-related costs by trimester prenatal care began. He found that the greatest costs occurred for mothers with no prenatal care. However mothers beginning care in the first trimester had net costs greater than those beginning care in the second or third trimester.

Therefore, the results of the Malitz study do not demonstrate the large cost benefits of prenatal care shown by Blackwell and Behrman. It is clear that paying for prenatal care will probably increase the mother's medical costs. It is also reasonable to assume that adequate prenatal care will probably reduce the newborn's costs since it appears to be related to a reduction in prematurity. However, whether the reduction in newborn costs is greater than the increase in mother's costs is not clear, and the present study will attempt to further clarify the answer to this question.

This study will use a sample of 1981 and 1982 Missouri Medicaid births to test the following hypotheses:

1. Adequate prenatal care reduces Medicaid costs within 45 days after birth and reduces them more than it increases mother's costs before birth; and
2. The Missouri Unborn Child Program, which provides prenatal care to Medicaid-eligible women in their first pregnancy, is cost beneficial.

Medicaid mothers will be divided into those receiving adequate prenatal care (defined by beginning care by the fourth month of pregnancy with at least five visits for preterm deliveries and at least eight visits for full-term births) and those having inadequate prenatal care.

The Missouri Unborn Child Program was implemented in October, 1980. Before this implementation many mothers in their first pregnancy did not become eligible for Medicaid until their baby was born, and therefore their prenatal care was not paid for. Since the primary purpose of this program is to provide prenatal care for previously ineligible pregnant women, one would expect the level of prenatal care for these women to be higher than that for other Medicaid mothers in their first pregnancy who are not on this program. If the level of prenatal care is higher, then applying the same logic as in the original hypotheses about prenatal care, one would also expect the newborn costs to be reduced for these births. A cost benefit analysis will be performed for these births similar to the one testing the primary hypothesis of this study.

METHODS

The basic design of the study involves the linking of three separate data files. These files are: (1) Medicaid, (2) birth certificates, and (3) death certificates. The Medicaid file was needed to obtain Medicaid cost data. The birth certificate file provided data on prenatal care, maternal characteristics, and birth weight. The death file was linked in order to obtain a neonatal (under 28 days) death rate for babies of mothers with adequate prenatal care compared with babies of mothers with inadequate prenatal care.

Initially, computer tapes of 9,827 newborn Medicaid records for babies born in 1981 and 10,196 records for babies born in 1982 were created from Missouri Medicaid claim tapes. All claims with a date of service within 45 days of birth were included. Claims were included if they were submitted by the hospital or physician to the Medicaid Program within eleven months after the end of the calendar year of birth.

These Medicaid newborn records were then matched with their mother's claim records using the Medicaid identification number as the matching criteria. All Medicaid claim records for the mother with a date of service within 45 days after or nine months before the birth of the child were included. A matching mother's record was found for 91 percent of the 1981 Medicaid births and 86 percent of the 1982 births.

The newborn Medicaid records were then matched against their corresponding birth records using name and date of birth as matching criteria. Matching birth records were found for 99 percent of the 1981 births and 98 percent of 1982 births.

The accuracy of the prenatal care variable is of vital importance in this study. Records with incomplete information on prenatal care need

to be excluded. Adequate prenatal care, as defined in this study, is a conservative estimate allowing for the inaccuracy of information recorded on the birth certificate. ACOG recommends that pregnant women begin care in the first trimester and obtain at least nine visits. The study definition allows for care to begin in the first four months of pregnancy with at least eight visits. Therefore, many of the women falling into the adequate category may not meet the ACOG standard, but nearly all the women falling into the inadequate category should meet the inadequate standard.

Records excluded because prenatal care was unknown on the birth record totalled about 290 in 1981 and 390 records in 1982. Prenatal care was considered unknown if month prenatal care began or if number of prenatal visits were unknown. If length of pregnancy was less than 26 weeks or if length of pregnancy was unknown and prenatal care could not be determined from month care began or number of visits, then prenatal care was also considered unknown.

Additional exclusions were made for third-party liabilities and lack of hospital claims in order that the final study file contain Medicaid cost data as complete as possible.

A summary of the entire matching process is illustrated in Figure 1. After all exclusions, 7,046 records in 1981 and 7,245 records in 1982 remained in the study sample. These figures represent slightly over 70 percent of the original newborn Medicaid files and over 82 percent of the Medicaid paid claim amounts. Approximately 40 percent of the mothers in each of these Medicaid files had inadequate prenatal care which was over double the Missouri general population state percentage of 18 percent.

Before proceeding with the cost benefit analyses, tests will be made to study the relationship between the prenatal care categories and two outcome variables, low birth weight (under 2,500 grams) and neonatal death rates. These tests will be done with the file before the exclusions for the incomplete Medicaid information. These rates will also be stratified by race.

In order to determine the best covariates, a stepwise regression analysis was performed between a number of possible confounding variables and the primary dependent and independent variables, Medicaid costs and prenatal care. Variables significantly correlated with either prenatal care or Medicaid costs in either year, 1981 or 1982, after adjustment for the other variables in the equations, were given a rank based on order of entrance into the stepwise regression. Ranks for each year were then averaged.

Variables selected as covariates in testing the null hypothesis that there is no difference in Medicaid costs between adequate and inadequate prenatal care births are per diem hospital reimbursement, age of mother, number born, birth spacing, and metropolitan residence. Analysis of covariance will be used to control for these variables. Least square estimates of mother, newborn, and total Medicaid paid claim amounts will be calculated and compared between adequate and inadequate prenatal care categories for each year, 1981 and 1982. Selected as covariates in testing the Unborn Program cost hypothesis were per diem hospital reimbursement, race, education and age of mother and number born. (Only first births were used in this analysis.)

Because of a change in Missouri Medicaid newborn reimbursement procedures in late 1981, it was not appropriate to combine 1981 and 1982 data together. During most of 1981, Medicaid reimbursed hospitals at 100 percent of newborn submitted charges if the newborn was Medicaid-eligible. In 1982, Medicaid reimbursed hospitals for newborn claims on a per diem formula which was calculated on a hospital's entire population of patients. Since nursery costs are generally lower than adult hospital costs, Medicaid frequently reimbursed hospitals at a rate higher than submitted charges. Because of this change in procedure, newborn Medicaid paid claims averaged about 50 percent more in 1982 than in 1981. To see if this change had any effect on the results of this study, submitted charges to Medicaid as well as actual Medicaid paid claim amounts will also be analyzed.

RESULTS

Low Birth Weight Rates

For each year 1981 and 1982, babies of mothers with inadequate prenatal care had low birth weight (LBW) rates 19 percent higher than babies of mothers with adequate prenatal care, a statistically significant difference. (See Table 1.) In 1981 the inadequate prenatal care LBW rate was 12.6 percent compared to 10.6 percent for adequate prenatal care while the comparable rates in 1982 were 13.1 and 11.0, respectively. The low birth weight differential between prenatal care categories was nearly the same for both race groups and both years studied.

If one assumes a causal relationship, it is estimated that approximately 115 LBW Medicaid births per year were prevented by adequate prenatal care.

The pattern of very low birth weight (VLBW) births by level of prenatal care is very similar to the LBW rate pattern with an average difference of 0.4 percent.

Neonatal Death Rates

Overall, the inadequate prenatal care neonatal death rate for 1981-1982 was 29 percent higher than the adequate rate (11.2 vs. 8.7 deaths per 1,000 live births, respectively), but this differential was not statistically significant. As Table 2 shows, this differential primarily occurred for white Medicaid infants. The neonatal death rate for white babies of mothers with inadequate prenatal care was 13.4 per 1,000 live births compared to 7.9 for those with adequate prenatal care. This differential of 70 percent was statistically significant. For infants in the all other racial group, there was virtually no difference in the neonatal death rates by level of prenatal care.

Assuming a causal relationship again, it is estimated that approximately 28 Medicaid infants' lives were saved during 1981 and 1982 by the receipt of adequate prenatal care.

Medicaid Costs by Level of Prenatal Care

The net Medicaid costs for the births involving mothers with adequate prenatal care were

generally greater than the net costs for the inadequately cared for births, as Table 3 shows. In 1981 net costs for the adequate prenatal care category were $94 greater than the inadequate category, a difference which was not quite statistically significant (p = .06). In 1982, despite the change in newborn reimbursement procedures, this differential was very similar to the previous year, $110, and it was statistically significant.

The pattern of Medicaid claim amounts by level of prenatal care by recipient was very similar for both years studied. Mothers' costs were greater for adequately cared for mothers while newborn costs were greater for babies in the inadequate category. The increases in costs for mothers with adequate prenatal care ($143 in 1981 and $125 in 1982) were greater than the decreases ($50 in 1981 and $15 in 1982) in costs for their babies. For both years, 1981 and 1982, the increases in the mothers' costs for the adequate prenatal care category were statistically significant while the decreases in the newborn costs were not statistically significant. Since the increases in mothers' costs were generally greater than the decreases in the newborn costs, overall, the net costs were greater for mothers and their children with adequate prenatal care than those without this care.

The total submitted charges to Medicaid were also examined by level of prenatal care to provide a more complete cost file. These charges include all costs submitted to Medicaid, regardless of whether Medicaid paid for them or not. The pattern of these charges by level of prenatal care by recipient was very similar to that for Medicaid paid claim amounts. In 1981 the increase in total submitted charges for adequate births compared with inadequate births was $138 and in 1982 it was $188. Both of these increases were statistically significant at the .05 level.

Unborn Child Program

Table 4 shows that generally, the mothers on the Unborn Child Program received more prenatal care than other first-time mothers on Medicaid. In 1981, 34.5 percent of mothers on the Unborn Program received inadequate prenatal care compared to 39.7 percent of other first-time Medicaid mothers. In 1982, this differential expanded as 34.5 percent of Unborn Program mothers had inadequate prenatal care contrasted to 42.1 percent for the comparison group. The percentage of first-time mothers on the Unborn Program also greatly expanded in 1982, increasing to 56 percent from the 38 percent participation rate in 1981.

Unborn Program mothers usually did not get in for prenatal care any sooner than other mothers, but they generally stayed in contact with their physicians and obtained more prenatal visits. As Table 4 shows, Unborn Program mothers obtained an average of 0.7 more visits than other first-time Medicaid mothers in 1981 and 1.0 more visits in 1982. But the month prenatal care began was virtually the same for both groups in both years studied.

Low birth weight rate comparisons by Unborn Program participation were completely different between 1981 and 1982 (see Table 5). In 1981 there was virtually no difference between Unborn and other first birth low birth weight rates (11.0 percent for Unborn and 10.2 for other). In 1982 Unborn rates were significantly lower than other births for the all other race and total race categories. The 1982 low birth weight rate for Unborn Program babies was 8.5 compared to 12.2 for those not on the program.

The overall costs of the Unborn Program outweighed any savings resulting from having more prenatal care. As Table 6 shows, in both 1981 and 1982, the Medicaid mean paid claim amounts were $157 higher for first-time mothers on the Unborn Program than it was for other first-time Medicaid mothers.

While the costs are greater than the savings for the Unborn Child Program, they are not as great as those that appear in the annual budget for the program. Multiplying the $157 difference between Unborn and other first-time Medicaid mothers by the number of participants in the program results in total costs of $165,000 in 1981 and $270,000 in 1982 for the Unborn Child Program. Actual expenditures charged to the program were $1.1 million in 1981 and $2.2 million in 1982. The reason for this apparent discrepancy is that many of the charges to the Unborn Program would be picked up by Medicaid anyway if the mother applied for eligibility after birth. The mother does not go off the Unborn Program until the end of a month. So if she delivers early in a month, all of her delivery and hospital costs would be charged to the Unborn Program. If there was no Unborn Program, they would have been picked up by AFDC eligibility.

DISCUSSION

The pregnancy outcome results of this study are comparable to the findings of other studies of prenatal care. The adequate-inadequate prenatal care differentials were generally not quite as strong as those found in the other studies but they were statistically significant. This may be due to the high risk Medicaid population used in the current study.

Cost benefits of the present study were much lower than those found by Blackwell and Behrman but were comparable to the Malitz Texas Medicaid study. The Blackwell and Behrman cost benefit analyses of prenatal care were based on synthetic estimates. Assumptions which may have been faulty included a possibly overestimated very low birth weight reduction in the Blackwell study and an apparent overestimate of the percentage of low birth weight infants that require neonatal intensive care in the Behrman study.

A primary source of potential error in the present Missouri study was incomplete Medicaid cost data. Deleting records with third-party liabilities and records without hospital claims improves the data, but may not completely eliminate the problem. All eligible costs may not have been claimed. For example, claims for 1982 newborns were still being paid as late as November, 1983. Billing problems with many rural hospitals also may have reduced claims. It is possible, although not probable, that the adequate and inadequate prenatal care populations varied with respect to these complicating factors.

The exclusion of nearly 30 percent of Medicaid records may possibly have biased the results

of this study. The Missouri file is, however, more complete than the Texas Medicaid file used by Malitz. The Texas match rate was only 58 percent.

The low birth weight relationship between prenatal care categories was nearly identical for inclusions and exclusions. This suggests that the large number of excluded records did not appreciably affect the results.

Many physicians combine prenatal care and delivery into one package billing. Women obtaining less than adequate prenatal care may be billed for the entire package even though they don't use it all. This factor may tend to obscure differences in mothers' Medicaid paid claim amounts between prenatal care categories.

Another possible source of error is inaccuracy on the birth certificate. Month prenatal care began and number of prenatal visits are entered on the birth certificate from a variety of sources, depending on hospital procedures (Land and Vaughan, 1984). Some hospitals obtain the information from the mother and others obtain it from physician's prenatal records, and still others use a combination of these two sources. Mothers may tend to overreport number of visits and the prenatal care record at the hospital may miss the last visit or two. The prenatal record at the hospital also may not reflect visits at a public health clinic. This is why a broad definition of prenatal care was used in this study rather than the more exacting definitions used in some other studies. In this study mothers with adequate prenatal care began care an average of two months earlier and averaged over five visits more than mothers with inadequate prenatal care. So as a group it is clear that the adequately cared for mothers began care earlier and had more care than the inadequately cared for mothers. Nevertheless, misclassification of level of prenatal care still may have occurred. This misclassification may tend to obscure differences in both outcome and claim amounts between prenatal care categories.

It should also be emphasized that only quantity of care was studied, not quality of care. It is possible that quality would be more cost beneficial.

In addition, the long-term costs possibly averted because of healthier, normal-weight babies was not assessed. But other studies involving synthetic estimates (Blackwell and Behrman) have shown that most newborn savings from prenatal care occur in the initial hospitalization including neonatal intensive care. Most of these costs should have been covered in this study.

Errors on the birth record for other variables may also have affected the results of this study. The variables used as covariates for testing the principal hypothesis of this study, age of mother, number born, birth spacing, metropolitan residence, and race, however, are generally considered to be fairly accurate. Moreover, Missouri has extensive editing and querying programs to improve completeness and consistency of reporting, although some errors still might occur.

Other factors not available from the birth or Medicaid records may have influenced the results. Women obtaining prenatal care are self-selected in that they were motivated to obtain physician care regularly before pregnancy. These women therefore may be more concerned for the health of themselves and their babies. These factors could have influenced the outcome results. Mothers who did not obtain adequate prenatal care may not trust doctors. Therefore they may be less willing to seek medical help when they are sick. This in turn could reduce their Medicaid paid claims.

Most of the results of this study were consistent for both years studied, 1981 and 1982, which helps to validate these results. One exception was the low birth weight rate comparisons between Unborn and other Medicaid first-born births. The Unborn Child Program was much more active in 1982 and many more first-time mothers participated in the Unborn Program in 1982 than in 1981. This may have improved the low birth weight outcomes somewhat. It is more likely that small numbers and random chance caused the complete reversal of results between 1981 and 1982.

The results of this study probably cannot be generalized to the total population for two reasons. First, the Medicaid population is much different from the general population of the state. More Medicaid mothers are black, urban, unmarried, and under 20. There is generally less difference in prematurity rates by level of prenatal care for women in these high risk categories. Therefore, one would also expect less difference in newborn costs for these women and therefore less favorable benefit/cost ratios. A second factor arguing against generalizability is that Medicaid paid claims are not equivalent to medical costs. Not all medical expenses are covered by Medicaid and eligibility is not always constant throughout pregnancy and post-partum period.

CONCLUSION

In this study of 1981 and 1982 Missouri Medicaid births, adequate prenatal care was associated with an apparently improved pregnancy outcome as measured by low birth weight and neonatal mortality rates. Similarly, participation in the Unborn Child Program was associated with a reduced low birth weight rate, at least in one of the two years studied.

However, despite these improvements in pregnancy outcome, the costs to the Medicaid program of providing adequate prenatal care or prenatal care in the Unborn Child Program were greater than any short-term savings in newborn costs. The overall increased costs of providing adequate prenatal care represent about two percent of the $40 million in total Medicaid costs for providing maternal and infant care.

While providing adequate prenatal care did not prove to be cost efficient, this does not negate the positive aspect of the program. Cost-effectiveness is only one way of measuring the worth of a program. Results of this study show that the program may have averted approximately 115 LBW births per year and 14 neonatal deaths per year.

As Vladeck, 1984, stated, arguing cost-effectiveness alone and ignoring service quality and simple humanity may be self-defeating. The most economic course in most cases would simply be to do nothing. The argument against doing

nothing is ultimately a moral one. The fact that adequate prenatal care was associated with improved pregnancy outcome makes prenatal care a beneficial program in and of itself. Although costs are greater than savings, providing adequate prenatal care to Medicaid mothers appears to be a reasonable investment.

ACKNOWLEDGEMENTS

This study was financed through a Prevention Grant from the Health Care Financing Administration.

The authors thank the Division of Family Services, Medical Services for providing the use of their Medicaid data files, assisting us in understanding these files, and providing financial assistance to this project.

We thank Debbie Heckman, Division of Data Processing, for writing the computer programs to access the Medicaid files and for merging the mother and newborn Medicaid records together.

The authors also thank Carolyn McMillin for typing this report and providing other needed clerical assistance to this project.

REFERENCES

American College of Obstetrics and Gynecologists: Standards of Obstetric-Gynecologic Services. Chicago, 1974.

Behrman, R.E., Chairman, Committee to Study the Prevention of Low Birthweight. Institute of Medicine. Preventing Low Birthweight, National Academy Press, pp. 212-240, 1985.

Blackwell, A., An Administrative Petition to the United States Department of Health and Human Services, June 29, 1983.

Drillien, C.M., The Social and Economic Factors Affecting the Incidence of Premature Birth Part I. Premature Births without Complications of Pregnancy. J. Obstet. Gynaecol. Br. Emp. 64:161-184, 1957.

Jonsen, A.R., and Garland (eds.), Ethics of Newborn Intensive Care. A Joint Publication of Health Policy Program, University of California at San Francisco and the Institute of Governmental Studies, University of California at Santa Barbara, p. 82, 1976.

Kane, S.H., Significance of Prenatal Care. Obstet. Gynecol. 24:66-72, July, 1964.

Kessner, D.M., Singer, J., Kalk, C.E., and Schlesinger, E.R., Contrasts in Health Status. Vol. I Infant Death: An Analysis by Maternal Risk and Health Care. Washington. Institute of Medicine, National Academy of Sciences, 1973.

Land, G., and Vaughan, W., Birth Certificate Completion Procedures and the Accuracy of Missouri Birth Certificate Data. Journal of AMRA, August, 1984.

Malitz, D., A Cost-Benefit Analysis of Extending Medicaid to Provide Prenatal Care to Pregnant Women, Texas Department of Human Resources, May, 1983.

Missouri Monthly Vital Statistics. Provisional Statistics from the Missouri Center for Health Statistics, Vol. 16, No. 9, November, 1982.

Schramm, W.F., Prenatal Care and Prematurity in Missouri. Proceedings of the Fourth Annual NCHS Data Use Conference in Salt Lake City, Utah, pp. 118-134, October, 1979.

Showstack, J., Budetti, P.P., Minkler, D., Factors Associated with Birthweight: An Exploration of the Roles of Prenatal Care and Length of Pregnancy. Am. J. Public Health, 74:1003-1008, 1984.

Shwartz, S., and Vinyard, J.H., Prenatal Care and Prematurity. Public Health Rep. 80:237-248, 1965.

Terris, M., and Glasser, M., A Life Table Analysis of the Relation of Prenatal Care to Prematurity. Am. J. Public Health, 64:869-875, 1974.

Terris, M., and Gold, E.M., An Epidemiologic Study of Prematurity. II. Relation to Prenatal Care, Birth Interval, Residential History and Outcome of Previous Pregnancies. Am. J. Obstet. Gynecol. 103:371-379, 1969.

Tokuhata, C.K., Digon, E., and Mann, L., Prenatal Care and Obstetric Abnormalities: Experiences of 185,000 Pennsylvania Births, J. Chron. Dis., 26:163-185, 1973.

Vladeck, B.C., The Limits of Cost-Effectiveness. Am. J. Public Health, 74:652-653, 1984.

Weiner, G., and Milton, T., Demographic Correlates of Low Birth Weight. Am. J. of Epidemiology. 91:260-272, 1970.

Figure 1
Selection of Study Samples
1981, 1982

	1981	1982
• Initial newborn Medicaid population.....	9,827	10,196
Exclusions		
• No matching Medicaid record for mother.....	791	1,331
• No matching birth record...	96	231
• Unknown prenatal care.....	229	349
• Third-party liability.....	267	266
• No hospital claims.....	1,398	774
• Total exclusions.....	2,781	2,951
• Final study sample.....	7,046	7,245
• Adequate prenatal care.....	4,261	4,321
• Inadequate prenatal care...	2,785	2,924
• Unborn 1st births.....	1,052	1,781
• Other Medicaid 1st births..	1,719	1,420

Table 1
Low Birth Weight (Under 2,500 Grams) and Very Low Birth Weight (Under 1,500 Grams) Rates per 100 Live Births by Race by Level of Prenatal Care: Missouri Medicaid Births 1981 and 1982

	Total		White		All Other	
	Adeq	Inadeq	Adeq	Inadeq	Adeq	Inadeq
1981						
VLBW..	1.2	1.7	0.8	1.2	1.6	2.0
LBW...	10.6	12.6*	8.7	10.9*	12.1	14.2*
N.....	5,684	3,793	2,542	1,821	3,142	1,972
1982						
VLBW..	1.2	1.6	1.1	1.5	1.3	1.7
LBW...	11.0	13.1*	9.0	11.0*	12.9	15.1*
N.....	5,577	3,954	2,641	1,998	2,936	1,956

VLBW = Very low birth weight (under 1,500 grams) rate.
LBW = Low birth weight (under 2,500 grams) rate.
*Statistically significantly elevated at .05 level.

Table 2
Neonatal Death Rates per 1,000 Live Births by Race by Level of Prenatal Care: Missouri Medicaid Births 1981-1982

	Total		White		All Other	
	Adeq	Inadeq	Adeq	Inadeq	Adeq	Inadeq
No.....	98	87	41	51	57	36
Rate...	8.7	11.2	7.9	13.4*	9.3	9.1

*Statistically significantly elevated at .05 level.

Table 3
Medicaid Mean Paid Claim Amounts (Dollars) by Recipient by Level of Prenatal Care: Missouri Medicaid Births 1981 and 1982

	Recipients		
Prenatal Care	Total	Mother	Newborn
1981			
Adequate.....	2,384	1,548	836
Inadequate.....	2,290	1,405	885
Difference.....	94	143*	-50
95% conf. int.	(-2)	87	(-120)
of diff.	190	199	20
1982			
Adequate.....	2,829	1,580	1,249
Inadequate.....	2,719	1,455	1,264
Difference.....	110*	125*	-15
95% conf. int.	22	77	(-49)
of diff.	198	173	34

*Statistically significant at .05 level.

Table 4
Inadequate Prenatal Care Percentage by Type Mean Month Prenatal Care Began and Mean Number of Prenatal Visits by Unborn Program Participation: Missouri Medicaid 1st Births 1981 and 1982

	1981		1982	
	Unborn	Other	Unborn	Other
Inadequate %.....	34.5	39.7*	34.5	42.1*
Late care %.....	24.0	25.8	24.1	27.7*
Too few visits %.	22.8	29.8*	23.9	33.3*
Month care began...	3.4	3.4	3.4	3.3
Visits.....	9.5	8.8*	9.3	8.3*
N.....	1,052	1,719	1,781	1,420

*Statistically significantly different at .05 level.

Table 5
Low Birth Weight Rates by Race by Unborn Program Participation: Missouri Medicaid 1st Births 1981 and 1982

	Total		White		All Other	
	Unborn	Other	Unborn	Other	Unborn	Other
1981						
No.....	115	179	46	56	69	123
Rate...	11.0	10.2	8.8	8.0	13.0	12.1
1982						
No.....	149	177	73	55	76	122
Rate...	8.5	12.2*	7.5	10.1	9.4	14.0*

*Statistically significantly different at .05 level.
NOTE: Total rates are race adjusted.

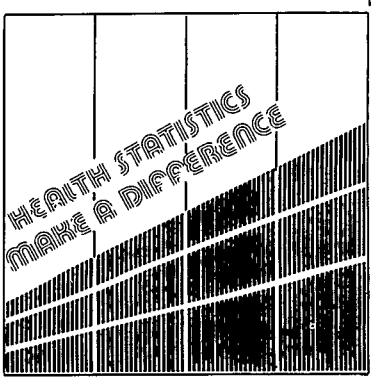
Table 6
Medicaid Mean Paid Claim Amounts (Dollars) by Recipient by Unborn Program Participation: Missouri Medicaid 1st Births 1981 and 1982

	Recipients		
	Total	Mother	Newborn
1981			
Unborn.....	2,254	1,408	846
Other.....	2,097	1,260	837
Difference.....	157*	148*	9
95% conf. int.	20	76	(-96)
of diff.	294	220	114
1982			
Unborn.....	2,776	1,524	1,252
Other.....	2,619	1,305	1,314
Difference.....	157*	219*	-62
95% conf. int.	31	155	(-159)
of diff.	283	273	35

*Statistically significantly different at .05 level.

Special Session 4 (NCVHS)

Does the National Committee on Vital and Health Statistics Make a Difference?



AN OVERVIEW OF THE NATIONAL COMMITTEE ON VITAL AND HEALTH STATISTICS: WHO - WHAT - HOW

Robert H. Barnes

Introduction

The National Committeè on Vital and Health Statistics is an official advisory body on health statistics to the Secretary of Health and Human Services. In my presentation, I wish to tell you what I perceive as the purpose or mission of the Committee, to describe how the Committee functions, its membership, and its authority. Further, I want you to learn about its relationship to federal and state governments, to the private sector, and to the World Health Organization. By telling you about its activities in the recent past and its goals for the future, I hope that you will able to affirmatively answer the question, "Does the National Committee on Vital and Health Statistics make a difference?"

Background

The National Committee on Vital and Health Statistics is one of the oldest advisory committees in the federal government. It was created in 1949 by the Surgeon General of the U.S. Public Health Service, and by an act of Congress in 1974 (Public Law - 93-353) it was established legally. Its charter is renewed every two years, having last been renewed by the Secretary of Health and Human Services in 1984. Over time, nothing stays the same. This is certainly true of the National Committee. It has evolved from an original membership of 12 prior to 1974, to the present membership of 15. The members originally were selected for their technical expertise in health statistics. Now the members are selected by the Secretary for having distinguished themselves in fields such as health statistics, epidemiology, health planning, and the provision of health services. Prior to 1974 the Committee accomplished its purpose through Technical Consultant Panels, whose members applied their expertise to highly technical statistical questions. Now the committee works through subcommittees which study statistical issues closely linked to both national and international health policy. The charter states that the Committee is advisory to the Secretary of Health and Human Services, and that it reports through the office of the Assistant Secretary for Health. However, as you will see when I describe its study of the revision process of the 9th edition of the International Classification of Diseases, the Committee has become a national forum for diverse interests to voice their concerns, not only about the revision process of ICD-9, but many other topics. Instead of being a group of highly technical and often academic leaders in health statistics and epidemiology, the background of the membership is diversified, with some being the above, but others having experience as practicing physicians, teachers in medicine, leaders in the health insurance industry, or managers of state vital and health statistics programs. This has broadened the functional style of the Committee, being not only technically but also socially and economically sensitive to issues. There has been some criticism of this new orientation, but others have felt that the broadened background of Committee members has increased its effectiveness.

Purpose

The function of the Committee is advisory, but to do this, it studies the statistical aspects of various high priority topics. The members are constantly aware of the need for high quality health information and statistics to develop appropriate health policy. Its strength is based on the expertise and commitment of its membership and the remarkable support of the staff of the National Center for Health Statistics. It is limited in its work only by its part-time volunteer status and budget restrictions. During the last three years, dramatic changes have taken place in the entire health field, particularly in health care delivery, financing, and technological breakthroughs. Economic power has been shifting from organized medicine to mega-corporations. The federal government, in its mandate from Congress to control costs, has implemented the DRG Prospective Payment Plan for hospitals, frozen Medicare payments to physicians, and is now looking at new ways of paying them. The Health Care Finance Administration has contracted with a vast network of Professional Review Organizations to monitor the health care system to not only control costs, but to "guarantee quality of care." Paralleling these "control activities," the technological breakthroughs, and the increase in number of elderly requiring long-term care have created the possibility that all cost control hopes will be thwarted. What

might seem at first glance to be a minor issue compared to the above dramatic ones, is the use of the World Health Organization's International Classifications of Diseases (ICD-9) to code the financially-driven DRG Prospective Payment System. The concern of epidemiologists and statisticians is that health data will be skewed by using a code for a purpose for which it was never intended. In the meantime, the World Health Organization grinds along its 15 year process of revising the ICD, a code for international mortality and morbidity data. As you might readily agree, there is no question about the dramatic changes that are taking place in the health care field, as well as in the field of health statistics.

It has been in this atmosphere that the National Committee on Vital and Health Statistics has been working to implement its mandate from Congress to be an advisory body on health statistics to the Secretary for Health and Human Services. It has been active in a variety of areas that have touched on some of the issues mentioned. The charter from the Congress has given the Committee continuity and stability. Even though the Committee has no direct power, it has ready access to government at the highest level in the health care field. Over the last two years it has increased its close relationship to the Health Care Finance Administration, particularly with the Office of Data Management and Strategy.

You next will hear a report from Dr. William Felts, Chairman of a Subcommittee on Statistical Aspects of Physician Payment Systems. The staff of the H.C.F.A. has been working closely with the subcommittee on this topic. As mentioned previously, the National Committee provides a forum where representatives from the private sector, state and federal government meet to discuss the often conflicting opinions about the statistical aspects of complex health issues.

After Dr. Felts speaks you will hear Dr. Ronald Blankenbaker, acting Subcommittee Chairman on Data Gaps in Health Promotion and Disease Prevention. Dr. Blankenbaker's subcommittee's work illustrates the National Committee's desire to link its efforts to important national health policy.

In the spirit of augmenting communications and reducing conflict, the National Center for Health Statistics and the Committee this past year cosponsored three conferences on the revision of ICD-9. The information obtained will go to the World Health Organization to reflect the concerns of the users of ICD in this country.

The National Committee, through its subcommittee structure, focuses its attention on a few prioritized topics. These topics are agreed upon not only by the Assistant Secretary for Health, but also by the Committee itself and the National Center for Health Statistics. The Committee is aware of the fragmented and uncoordinated nature of health data and health statistics programs in this country. It is aware of the hundreds of sources of data flowing upwards from state and local governments, insurance companies, research institutions, and downwards from many federal sources. It has not been a goal of the National Committee to develop or even attempt to visualize a national health data statistics system. It is a goal of the Committee to coordinate some of the efforts in this direction. Development of the minimum health data sets -- the uniform hospital discharge data set, the long-term data set, and eventually the ambulatory data sets are a step in that direction. Playing a supportive role in the Cooperative Health Statistics System of the federal and state governments, is another example of the coordinating role of the National Committee. I wish to emphasize again the desire of the National Committee to lend its efforts as closely as possible to health policy. The National Committee is not a political organization, nor does it participate in budget development of any health statistics system. However, if budget restrictions make the Committee fear that health statistics systems would be jeopardized, it might very well take a position against such a restriction. The Committee feels strongly its obligation to support timely and high quality data.

In the immediate future, the National Committee is committed to its work on long-term care minimum data sets, the monitoring of the revision of the International Classification for Diseases, and fostering the idea of one procedure code. It has a work force evaluating the role of the National Committee on Black and Minority Health Data.

Through its Subcomittee activities and its ability to be a forum, and through the support of the National Center for Health Statistics, the Committee has been making a "difference" and will continue to do so. As more dramatic changes continue to take place in the health care field, the National Committee will focus its attention on the

needs of the federal and state governments, as well as the private sector, to help assure a health statistics system that is geared to producing a healthier America.

RELATIONSHIP OF THE NATIONAL COMMITTEE ON VITAL AND HEALTH STATISTICS TO THE STATISTICAL PROGRAMS OF THE HEALTH CARE FINANCING ADMINISTRATION

William R. Felts, George Washington University Medical Center

Introduction: Administrative records increasingly are becoming a significant source of data for policy formulation and research in the health care field. These records, many of which are generated as part of the reimbursement process, can provide valuable information on patient characteristics, types of services performed, and provider practice patterns.

Recognizing this development, the National Committee on Vital and Health Statistics (NCVHS) identified the statistical aspects of reimbursement systems for federal health care programs as an important area for Committee attention. Specifically, the NCVHS noted concerns about potential adverse impacts upon existing statistical data bases resulting from changes in reimbursement methodology. These concerns were accentuated by the passage of the Tax Equity and Fiscal Responsibility act (TEFRA) of 1982 and the subsequent enactment and implementation of prospective pricing for hospitals using the Diagnosis Related Group (DRG) methodology. They were further heightened by the consideration of extending the method to encompass physician reimbursement, and the passage of the Omnibus Deficit Reduction Act of 1984 and its freeze on physician reimbursement.

To evaluate these issues, the NCVHS appointed a work group in May and a subcommittee in December 1984. Their initiatives serve as a focal point for inquiry among various federal agencies and the private sector to identify the major forces in the dynamic arena of health care data, sources of data, the principal users and their needs, items of possible redundancy and/or adverse impact upon data quality, and various assumptions utilized by data analysts that could result in distortion and inequities.

The chain of Medicare Part B data flow from a patient/physician encounter in institutional and ambulatory settings to the Health Care Financing Administration (HCFA) was selected for initial review and illustrates the magnitude of problems of data quality. The influence of Medicare law and regulation is pervasive upon all insurers. The subcommittee is attempting to review and summarize these effects, with special emphasis upon ambulatory settings in view of their changing character.

The three major sources of data entry and manipulation are the physician's office, the carrier, and HCFA with respect to the flow of Part B claims, although other participants exist.

Physician's Office Function: The physician usually generates the initial description of the reason for the encounter depicted either as symptom or diagnosis, and the service(s) provided. This important level is of critical importance for accuracy. The taxonomies and codes utilized to describe these attributes exemplify the complexities. The International Classification of Diseases, 9th edition, Clinical Modification (ICD-9-CM) is the basis for diagnostic communication in the U.S., and the Health Care Procedure Coding System (HCPCS) promulgated by HCFA is dominant for services and procedures. The latter has as its level I core the Current Procedural Terminology (CPT-4), developed and maintained by the American Medical Association. Two additional levels include that added by HCFA to accommodate providers other than doctors of medicine or osteopathy for whose services benefits are paid (level II), and a means for carriers to add local codes for special circumstances (level III). The development and implementation of HCPCS is acknowledged to represent a giant step forward toward uniformity in claims processing in the last five years.

Within the hospital prospective pricing system and its DRG methodology, procedures and codes from volume 3 of ICD-9-CM are utilized rather than HCPCS in determining the institution's reimbursement. Research projects sponsored by HCFA are evaluating substitution of HCPCS for volume 3 codes. If successful, additional means would be provided by which to "link" part B claims with those from Part A. Thereby a better coordination of information about beneficiary utilization would be possible to assist in the elimination of duplicative services, and to improve data by which to evaluate episodes of illness.

Following the encounter and its description the doctor's office usually generates a statement to the patient or completes an insurance (claim) form for submission to the insurer, in this instance the Part B Medicare carrier. A second major improvement in claims processing in the last five years identified by HCFA personnel has been the adoption of the "universal" claim form 1500. Again, a detailed review of the attributes and inadequacies of this form, especially as viewed by private sector insurers, is beyond the scope of this presentation. Suffice it to say that some difficulties persist. Changes in law dictate requisites for such instruments to capture additional detail, as do employer desires for additional information, and changing research objectives.

The Carrier Function: Form 1500 is transmitted by the physician (when assignment is accepted), or by the patient (when assignment is not accepted) to the Part B carrier. If the form is not provided by the physician (at his expense) the patient must complete another form and attach the physician's statement (bill). This often is accompanied only by a lay-term statement from the patient as to the reason for seeking services; a point that illustrates HCFA's dilemma in seeking to capture and analyze Part B diagnoses. The problem is compounded by the volume of such claims (230 million in 1984, averaging 2 to 3 itemized services per claim). As a result, HCFA has not required carriers to cap-

ture and retain diagnoses under Part B other than in certain specific instances. The latter include electronically submitted claims (currently representing 10 to 20% of claims received); aphakia to which payment for optometry services is limited; subluxation for which chiropractors can be paid; and Alzheimer's disease, an excluded benefit for non-inpatient psychiatric services.

This major absence of diagnoses from ambulatory care settings can be construed as a "weak link" in the system. It restricts the ability to make judgments regarding medical necessity for the majority of such claims. Over 50% of claims received under Part B are for office visits, almost all of which are assumed to be medically necessary. The progressive expansion of services provided outside hospitals is a factor which may bring increased pressure to transmit and retain diagnoses under Part B, at least for selected items.

Claims processing entails numerous carrier functions. The form 1500 contains three major sections of information: descriptions and identifiers of the patient, general information from the physician, and specific information concerning services provided. The eligibility of the patient must be verified, means sought for coordination of benefits (for in some instances Medicare has become the secondary rather than primary insurer), and the services provided and charges.

When codes for services have not been submitted from the physician's office they must be added by the claims clerk. These individuals have received relatively minimal training in coding, and if guidelines fail to provide an easy "match" with the uncoded descriptor the claim may be assigned to the lowest remunerative level designated for such a service. The appropriateness of the service is judged, and the usual, customary and reasonable (UCR) fee schedule consulted to determine the allowable charge. The latter is adjusted against the beneficiary's deductible status. A payment authorization and the payment are generated and mailed, either to the physician under assignment or the patient when unassigned. The claim data is then stored in the carrier archives where it is maintained for varying timespans.

The claims file is used to expand the specific history of the beneficiary. It permits screening for duplicate claims and review of deductibles; updating reasonable charge screens; review of materials for aberrant practice patterns (post-payment); monitoring fees of non-participating physicians; and matching selected Part B claims with those from Part A which are now being received regularly from the intermediaries. While the carriers have enjoyed a considerable autonomy in their data processing methods, the adoption of the form 1500 and HCPCS, along with HCFA requirements for transmittal of greater detail during annual reporting cycles have increased the administrative demands upon them. The "translation" of prior coding taxonomies and databases into HCPCS represents one example, and is an exercise in which the medical profession may have legitimate concerns. Multiple assumptions have been made by the carriers in the translations that materially affect reimbursement levels, and may seriously skew procedure profiles and UCR levels.

Some of these claims processing details also affect the private sector insurers. The problems of coordination of benefits is significant. Sufficient data to allow determination of primary insurer and allocation of liability is not always included on the form 1500 (it has not been needed by Medicare where each individual is insured), but may be necessary to allow determination of the primary insurer when both husband and wife are employed and covered for health benefits. Private sector policies covering entire families (dependents) necessitate access to the names of all members.

Coordination of benefits also must frequently be determined on a state by state basis where requirements vary. Some payment schedules are employer specific. "Place of service" is used under some plans to determine levels of benefits provided, and the National Association of Blue Shield Plans (NABSP) identified 33 options for such sites. Differences in reporting requirements and usage of principal, primary, secondary and comorbid diagnoses also can cause data variations.

These examples indicate why some economists and administrators estimate that "packaging" claims into "episodes of care" or "spell of illness" rather than the detailed itemization under the UCR system would reduce the claims volume from about 260 million to 150 million per year, and significantly lessen costs of processing.

Other Recipients of Data: Although the next step to be reviewed herein addresses data utilization at the HCFA level, it should be noted that selected data is routed to a number of additional agencies and files, including Peer Review Organizations (PRO's), state agencies for the Medicaid Management Information System (MMIS), rate setters in those states performing such functions, commercial abstractors, the Hospital Associations, the Joint Commission for Accreditation of Hospitals (JCAH), DHHS Regional Offices which maintain a role in administrative and statistical evaluation of carrier performance, and independent researchers - to name but a few.

HCFA Functions: HCFA recognized the inadequacy of its databases for addressing its expanded assignments prior to the enactment of TEFRA and the Omnibus Deficit Reduction ACt of 1984. It had a beneficiary file with an individual identification number providing general information about the person, and an Health Insurance Master (HIM) file containing demographic information used to monitor utilization data for deductibles. Several additional files were maintained which had emerged in parallel with the evolution of computer systems and state-of-the-art HCFA statistics. Among them were a Part B Bill Summary Record, representing a 5% sample of beneficiaries with aggregate charges by type and place of services but without diagnostic or procedure codes and with date of service noted only by the month in which it was provided; a Prevailing Charge Directory containing data on

the 110 highest volume procedures by locality and carrier; and a Payment Record file reflecting 100% of payments made, used to administer Part B. A HCFA Task Force was appointed to evaluate these datafiles against the added statutory demands.

The Task Force recommended modifications and data additions which culminated in the development of four datafiles known collectively by the acronym "BMAD" (Part B Medicare Data). Effective in July 1984, carriers were required to provide HCFA with the necessary data for updating these files on an annual basis, concurrent with the carrier's reasonable charge update cycle. HCFA, in turn, is developing BMAD into a modern online system with software capability for access by individual carrier and in the aggregate. Heretofore, approximately 50 different files had to be individually accessed to compile aggregate figures. More timely and sophisticated information will result.

BMAD has the potential to become one of the most powerful of HCFA's data systems. Its four core components are:

* Beneficiary file: This will contain 5% of beneficiaries and data items as the bill summary record, claims detail about all End Stage Renal Disease (ESRD) beneficiaries, and HCPCS procedure codes. It allows HCFA to link a beneficiary's Part B and Part B services utilization data.

* Provider file: This is composed of a 1% representative sample of providers (physicians and non-physicians) entitled to Part B payments with ALL of their services charged to Medicare patients. It accumulates data on each sampled physician over several years, allowing longitudinal analysis of impact of actual and projected program changes upon physicians and suppliers. Currently, this data is maintained on approximately 6,000 individuals.

* Procedure file: This accumulates the information on each procedure code in HCPCS by each carrier, with its frequency, the charge, and the amount paid.

* Prevailing Charge file: This is designed to ultimately replace the prevailing charge directory and contains the prevailing charge limits for every procedure by each carrier. It allows HCFA to study and accurately project payment levels.

The antecedent files to BMAD were utilized for HCFA's administrative purposes, and have been a principal source for many articles published in such journals as Health Care Financing Review. The datafile changes will enhance such studies. Recognition that commercial health insurers often defer to governmental policy to "lead the way" further emphasizes their importance. This data also is cited by other agencies dedicated to statistical compilations and study of health trends in the United States. While the steps taken toward data standardization should upgrade the quality of these files, the numerous sources for error identified previously are worrisome.

The Immediate Future: What changes may be anticipated if the methods for physician reimbursement under the federal programs are "reformed"? The UCR method has become distorted almost beyond recognition from its original composition by a continuing series of "adjustments".

The subcommittee has explored possible changes in reimbursement with HCFA representatives and identified a number of related studies under way or in various stages of development intramurally and/or by contractors. Three categorical thrusts of possible change from the current UCR system can be identified. One is that of physician DRG's; a second consists of various forms of capitation payment, including the possibility of placing the carriers at risk; and a third is represented by relative value scales (RVS). Data on Medicare patients enrolled in Health Maintenance Organizations (HMO's) already represents a significant gap of detailed knowledge. The HMO's are reimbursed for Medicare enrollees under the formula known as the Adjusted Average Per Capita Cost (AAPCC), defined by TEFRA as the estimated average per capita amount that would be payable if Medicare services for HMO members were furnished in the local fee-for-service sector. Capitation and RVS approaches have implications for ambulatory as well as inpatient care. A randomized list of some of these initiatives may be summarized as follows:

1. Hospital Specific Calibration of Physician DRG's: In response to TEFRA and the Omnibus Deficit Reduction Act, HCFA has prepared data for the DRG's (in-house) based upon physician charges from the most recent years available. It contains all inequities inherent in the present system that have accumulated in past years. It also correlates poorly for non-surgical services. However, it could be implemented on October 1, 1985 should precipitous legislative action so dictate.

2. DRG Adapters: Several experimental software packages have been designed for use with the current DRG Grouper. They utilize CPT codes (HCPCS) rather than those of volume 3 of ICD-9-CM. One such version, currently being tested within HCFA on 1983 data, handles the 202 most common surgical DRG's.

3. Case Mix: A number of projects are funded to achieve greater fiscal equity by including severity of illness parameters under the hospital DRG system. Others focus upon comparisons between hospital outpatient facilities and physician offices. Another thrust of case mix is being tested for emergency rooms.

4. Part A and Part B Linkages: Contractural and intramural activities which seek to link Part A and Parb B data are being tested in several states. Success could provide another administrative tool by which the impact of selected changes in physician reimbursement may be assessed and devised.

5. Episodes of Illness: Several approaches to "lump" payments for services into episodes of illness, capitation or similar clusters are being subjected to research development and evaluation.

6. Survey: 5,000 physicians are being surveyed on practice costs and incomes, percentages of reimbursement from various payers for selected services,

participation in Medicare, and "productivity". In addition to monitoring the data for statistical purposes, various proposals for changes in reimbursement methods can assist in evaluating the potential impact upon individual physicians.

7. HCFA Relative Value Scale: An in-house proposal has been calculated based upon HCPCS procedures utilizing 1983 actual charge data. HCFA will update it to the 1984 calendar year data as quickly as possible. The statements regarding inequities under #1 also apply.

8. Other Relative Value Scales: In addition to the HCFA RVS, responses from a number of independent contractors to a request for proposal to develop new relative value scales are under review. One or several may be funded. Methodologies to be evaluated include such attributes as resource cost, time, training of provider, skill required, etc.

9. Laboratory Schedules: By statute, laboratory fees must attain the 75th percentile of the national average by 1987, with Medicare paying 60 to 62% of the scheduled amount. Legislative action will be required if regional adjustments are to be considered by HCFA.

10. Durable Medical Equipment: Items of durable medical equipment are confronted with problems of authority similar to those designated under Laboratory Schedules.

11. HCFA Publication Requirements: Statutory requirements for HCFA to publish names of participating physicians, and to monitor the performance of individual physicians under the "fee freeze" imposed in 1984 necessitate data systems to permit:

a. determination of all participating physicians and suppliers;
b. monitoring all physician's actual charges to Medicare enrollees;
c. monitoring all physicians services "to determine any changes in the per capita volume and mix of physician services provided to beneficiaries", classified by participating physicians, assigned and unassigned claims, specialty and geographic area.

12. Other Groups: The Subcommittee has communicated with various other public and private sector groups in attempts to identify additional activities that relate to data flow needs resulting from contemplated changes in methods by which physicians are reimbursed. Most appear to incorporate ideas similar to those pursued by HCFA. Employers and the private insurance industry are especially interested in various preferred provider options which allow comparison of providers.

The Office of Technology Assessment has embarked upon a study of physician reimbursement under Medicare to examine relative prices, regional differences, assignment rates, and utilization of services (employing HCFA's BMAD) which will include information by procedure on:

a. utilization
b. assignment rate of physicians and suppliers
c. actual, customary and prevailing charges
d. the difference in charges by physician specialty and locality

Discussion: The dynamic state of the current environment makes it impossible to render a final judgment of the degree of data distortions resulting from recent and contemplated changes in reimbursement under the Medicare system.

The present thrust of cost-containment of health care appears to be dominant within the federal system, and the impact upon data files, statistics, and longitudinal comparability of data is afforded a lesser priority. The changes required in diagnostic emphasis (e.g., major medical problems of the patient, primary diagnosis, or principal diagnosis as defined by HCFA) may have a significant negative impact, especially relating to longitudinal comparability. It is possible, however, that improved accuracy in the medical record documentation of diagnoses and procedures will result in gains that exceed the negative effects upon statistical applications. Great care should be taken to identify and stipulate changes in data collection requirements, and in the analytic methods applied to them.

Inaccuracies in data entry remain of concern; especially those distal to the site of the actual patient/physician encounter. Unless they are eliminated or minimized, distorted interpretations and decisions can contribute to injustices to individual patients and providers, and to inappropriate national policy decisions.

HCFA does not anticipate a need for additional items of data from encounter sources under the various system changes under study.

Conclusion: The NCVHS is in a unique position to provide the extensive liaison with both the public and private sectors which is needed to insure full input from the relevant policy makers and data users. The Subcommittee which has been charged with addressing these issues will continue to seek information in the following areas:

* Determine more clearly the specific needs of users of data from patient/physician encounters in ambulatory care settings.

* Develop a schematic overview of the flow of data from multiple ambulatory settings into the various data bases.

* Define better the different sites of care in the ambulatory settings and the types of services delivered so that understanding of data requirements can be improved.

The result of these tasks will be considered in a review of the minimum Data Set for Ambulatory Medical Care.

The ultimate goals of the Subcommittee inquiry will be to encourage comparability and standardization where feasible; to enhance the multiple utility of data bases; to assure that data requirements are justified; and to prevent unnecessary duplications.

NCVHS AND DATA GAPS IN THE 1990 HEALTH OBJECTIVES

Ronald G. Blankenbaker, NCVHS

As long as we have had a recorded history, there is evidence that man as a social being has been concerned about the health of all of his fellow humans. One of the most significant developments to date must be the discovery of microbes as a cause of disease. This led to the so-called first public health revolution in the United States in the late 19th and early 20th centuries.

During this period the death rate decreased rapidly and we brought under control or eradicated most of the infectious diseases (major killers) of the past. According to Dr. William Foege, former Director of the Centers for Disease Control, "For every week experienced since the year 1900, two days of life expectancy have been gained" (1). This is a rather dramatic statistic, however most of this success can be attributed to measures that could "protect" the public through the use of vaccines, sanitary measures, etc. and the rapid advances of this first revolution have now slowed dramatically.

The key modern-day diseases and disabling conditions are much different than the infectious problems of the past. Heart disease, cancer, stroke, accidents, homicide, suicide, cirrhosis and bronchitis/emphysema are now the primary causes of death and disability between the ages of 20-65. We do not have the luxury of being able to provide programs that protect the public from these diseases, however essentially all of them are preventable through individual behavior modification or intervention.

This development prompted some investigators to study the natural history of modern diseases so that intervention measures could be taken to prevent these diseases prior to the development of signs or symptoms. Prospective Medicine was coined by Drs. Robbins and Hall in 1970 (2) to describe a comprehensive system of health care which evaluates the risk or probability of developing a particular disease and provides measures for the individual to intervene and prevent that disease (3).

Through the work of Robbins and Hall, along with many others, it is now estimated that the most significant factor affecting modern disease is life-style. This led to a report by the Surgeon General in 1979, Healthy People (4), which refers to the promotion of healthy life-styles and prevention of disease as the second public health revolution. This publication summarized the various risks to health and presented major goals for the health care delivery system to minimize the effects of these risks

Subsequently, these goals were translated into 227 objectives for the decade of the 1980's (5). Organized into 15 target areas, these objectives addressed personal prevention services, health protection mechanisms and health promotion activities. Implementation plans and strategies were then prepared for each of the various federal agencies responsible for the attainment of these objectives by 1990 (6,7).

In the overall development and evaluation of these objectives, along with the progress in attaining them, the National Center for Health Statistics (NCHS) has a vital role. NCHS has been charged to identify the available data, judge the quality of that data, assist in the development of objectives which are measurable, and to determine the data base needed to track the progress toward meeting the objectives. The first statement about movement toward attainment was made in Health United States, 1983 (8). Simultaneously, Dr. Lawrence Green and his colleagues expressed a serious concern regarding the adequacy of available data for many of the objectives and pointed out the need for much additional data to assess the current and future status of health in this country (9).

In the Fall of 1984, the National Committee on Vital and Health Statistics (NCVHS) addressed the data gap situation and the potential negative impact of budgetary constraints at NCHS. Consequently, the NCVHS appointed a work group to evaluate the role that NCVHS might assume in this matter. Doctors Lawrence Green, Fernando Trevino, Grayson Miller and the author were appointed to the work group with Dr. Green as chairman.

The first goal of the work group was to try to determine what agencies or organizations were also working on this problem. We learned that NCHS has evaluated the baseline data and did define significant data gaps in approximately thirty objectives. To partially resolve this dilemma, a health promotion and disease prevention supplement was added to the 1985 Health Interview Study, the results of which will be available in the Fall of 1985. This survey will be repeated in 1990. Additionally, a physical fitness conference was held by NCHS in June, 1985 to determine what information was needed in this area and what questions should be asked on future health surveys. Other areas which need similar evaluation are alcohol and drug abuse.

The Centers for Disease Control (CDC) have developed a survey (telephone interview) to track the objectives at the state level. NCHS is working with CDC to standardize these surveys and to assist the states wherever possible. To further evaluate progress in the states, the Intergovernmental Health Policy Project was created to determine what the priorites were and what was being measured. NCHS is working with this project.

The Office of Disease Prevention and Health Promotion has developed a tracking system (computerized) on the status of the objectives. This system assigns a priority for the objectives and assess information on data availability and problems with this data. An annotated critique (midcourse review) will be published in latter 1985 or early 1986 which will provide: the rate of progress in attaining the objectives; the level of achievement likely to be reached by 1990; issues which are enhancing or impeding progress; measures which might be adopted to overcome problems; appropriateness of the objectives; recommended changes in wording, focus or target of the objectives; and a preliminary draft of the recommended objectives for the year 2000.

In addition to the above developments, the work group learned that the Association of

Schools of Public Health was interested in developing an ad hoc task force of public health/prevention organizations to develop strategies to collect the data which is currently unavailable. The initial meeting of this group was held in Atlanta in conjunction with Prevention '85 and we met with them. There we heard that the following activities were taking place: the New England states were assessing progress toward the objectives via a regional approach with funding from the Office of Disease Prevention and Health Promotion; the Foundation of the Association for State and Territorial Health Officials (ASTHO) is developing data on the objectives; the Interagency Public Health System Committee (ASTHO, NCHS, CDC and the National Association of City Health Officers) is addressing data needs in disease prevention/health promotion; the Institute of Medicine has a major study underway to evaluate the effects of health promotion; the National Model Standards for Community Public Health Services are being revised, will be published the Fall of 1985 and will include a complete cross-indexing with the objectives; the Department of Health Administration at John Hopkins University has several projects of interest to these endeavors; a two-day meeting "Prospects for a Healthier America: Achieving the Nation's Health Promotion Objectives" was held in February of 1984 to discuss the objectives and formulate recommendations on how to achieve these objectives (10); the U.S. Prevention Services Task Force is adressing many of these issues; many states have incorporated the objectives with their state health plans, i.e., Indiana, Texas, Utah, Tennessee, South Carolina, New Jersey, etc.; health risk surveys are being carried out annually in many states; Health United States, 1986 will have a prevention section; and many other organizations are working on projects that will provide information on progress toward achieving the objectives, such as the American Public Health Association (APHA), the American College of Preventive Medicine, and the Society of Prospective Medicine.

The Atlanta meeting concluded by recognizing a serious need for funding at the national, state, local and private levels. Further recommendations included: a need for a consortium of interested agencies, a need for local/regional models for data selection that can be applied to the development of national models, and a need to coordinate activities with the Interagency Public Health Systems Committee.

The potential members of this consortium met again in June 1985 and further discussed the need for collection and coordination of national, state and regional data. Concern was raised about the need to postpone data collections in some areas, such as hypertension where national standards are soon to be developed. Additionally, it was announced that CDC would present the results of the Behavior Risk Factor Survey in August (11). The next meeting of the consortium group will be in conjunction with the 1985 annual meeting of APHA.

At this point the work group of NCVHS has determined that: much is being done or planned by many organizations and agencies at the national, regional, state and local levels; there is good reason to believe that these efforts are often duplicative and fragmented; there is general consensus that significant data gaps do exist; and there is a need for coordination and/or central organization of all of these various activities, such as a "consortium" might provide. An alternative would be the clear designation of one agency as the coordinating agency, i.e., NCHS or the Office for Disease Prevention and Health Promotion.

The work group believes that the options for the role of NCVHS at this time are to: further evaluate the issue; work with the office of the Assistant Secretary for Health in an advisory role; consider the calling of a national conference to include all interested parties; provide assistance and advice where needed to the staff of the NCHS; and emphasize the need to concentrate our efforts on the development of objectives for the year 2000. NCVHS may take all, part or none of these options but hopefully will maintain a strong interest in the Objectives for the Nation as they more and more become the driving force for public health programs in the U.S.

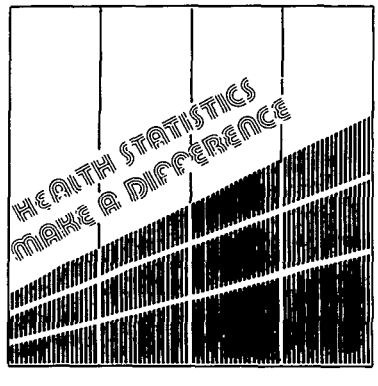
References

1. Foege, William: Personal Communication, May 20, 1983.
2. Robbins, Lewis C. and Jack H. Hall: How to Practice Prospective Medicine, Indianapolis, Indiana, Methodist Hospital of Indiana, Inc., 1970.
3. Robbins, Lewis C. and Ronald G. Blankenbaker: Prospective Medicine and the Health Hazard Appraisal, in Health Promotion: Principles and Clinical Applications, Robert Taylor, Editor, pp 67-121, Appleton-Century-Crofts, Norwalk, Connecticut, 1982.
4. Healthy People, The Surgeon General's Report on Health Promotion and Disease Prevention, U.S. Dept. of Health, Education and Welfare, Public Health Service, Publication No. 79-55071, 1979.
5. Promoting Health/Preventing Disease - Objectives for the Nation, U.S. Dept. of Health and Human Services, Public Health Service, 1980.
6. Public Health Service Implementation Plans for Attaining the Objectives for the Nation, U.S. Dept. of Health and Human Services, Public Health Service, in Public Health Reports, Supplement to the September-October 1983 issue.
7. Strategies for Promoting Health for Specific Populations, U.S. Dept. of Health and Human Services, Public Health Service, Publication No. 81-50169, 1981.
8. Health. United States, 1983, U.S. Department of Health and Human Services, Public Health Service, 1983.
9. Green, Lawrence, Ronald Wilson, and Katherine Bauer: AJPH, Vol. 73, No. 1, January, 1983, pp 18-24.
10. Proceedings of Prospects for a Healthier America: Achieving the Nations Health Promotion Objectives, U.S. Dept. of Health and Human Services, Public Health Service, November, 1984.

11. Kalsbeek, William, G.C. Hogelin, K.L. Eooyrn, M.T. Gotsman and J.D. Marks: The Behavioral Risk Factor Surveys: Design, Methods, and Estimates from Combined State Data, presented at the 145th Annual Meeting of the American Statistical Association, MGM Grand Hotel, Las Vegas, Nevada, August 5, 1985.

Special Session 5 (ADAMHA)

The Use of Health Statistics to Profile Alcohol Abuse and its Consequences



Differential Alcohol-Involved Proportionate Mortality Among Oklahoma Indians: A Tribal Comparison

Mary C. Dufour, Darryl Bertolucci, Henry Malin, National Institute on Alcohol Abuse and Alcoholism
and
Charles Christian, University of Maryland

It has been documented that death rates from alcohol-related causes are disproportionately higher among Native Americans than among other racial/ethnic groups in the U.S. today. Unfortunately, studies which differentiate mortality rates on the basis of tribal affiliation are rare. Death certificates do not list tribal affiliation. The purpose of this study is to examine differential mortality patterns within the Native American population in the State of Oklahoma.

At the present time within Oklahoma there are:

37 Federally Recognized Tribes
4 Federally Recognized Tribes which operate within a larger Tribe
2 Not Federally Recognized (E. Delaware & Cherokee Shawnee)

43 Major Tribes represented in Oklahoma

Additional Tribes are represented by tiny clusters of individuals.

The NCHS multiple cause mortality data files from the years 1968-1978 provided the foundation for the determination of alcohol-related mortality. In the past, only the underlying cause of death listed on the death certificate was tabulated for any given individual. Contributing causes appearing on the death certificate were not evaluated. Since alcohol-related conditions much more frequently contribute to rather than actually cause death, many alcohol-related deaths are not counted if only underlying cause is examined. The multiple cause data files list every entry (up to 20 items) recorded in the "cause of death" portion of the certificate, that is, all contributing and other significant conditions as well as the underlying cause of death. Deaths from 1968-1978 were selected for analysis for two reasons: (1) a ten year sample was necessary to provide an adequate number of Native American deaths for sound statistical analysis, and (2) these years encompass only one version of the International Classification of Diseases (ICDA-8), thus avoiding classification difficulties which sometimes arise when attempting to aggregate data coded in differing ICD codes.

Four causes of death related to chronic alcohol use comprised the measure of "alcohol-relatedness:"

Cirrhosis of the Liver (ICDA-8 code:571)
Alcoholic Psychosis (291)
Alcoholism (303)
Alcohol Poisoning (E860)

It must be kept in mind that this is an extremely conservative measure of the number of alcohol-related deaths. Physicians still hesitate to designate conditions as alcohol-related on the death certificate. In addition, this measure excludes:

- Other Medical Conditions
- External Causes of Death
  - Motor Vehicle Deaths
  - Homicides
  - Suicides
  - Fires
  - Falls

A significant fraction of deaths due to these external causes, particularly motor vehicle crashes are known to be alcohol-related. Since reliable population data is not available on a tribal basis, proportionate mortalities, rather than mortality rates were examined.

$$\text{Proportionate Mortality} = \frac{\text{Number of Alcohol-Involved Deaths in Tribe A}}{\text{Total Number of Deaths in Tribe A}}$$

Summary statistics for the State of Oklahoma include:

Total Deaths by Race-Oklahoma (1968-78)

Native American	7,764
Black	19,653
White	239,696

Percent Alcohol-Involved Deaths by Race

Oklahoma (1968-1978)	
Native American	9.3%
Black	3.2%
White	2.4%

Mean Age (in years) at Death by Race Oklahoma (1968-78)

Native American	57.7 years
Black	61.1
White	68.3

By using information from the Bureau of the Census and the Bureau of Indian Affairs, we attempted to determine in which counties a specific tribe predominates; that is, for a given county--at least 75% of the Native Americans resident in that county are affiliated with that given tribe. These estimates were carefully validated by representatives of the various tribes.

The ideal unit of analysis for this type of study would be the tribe. Due to the format of the death certificate data, deaths can only be disaggregated by county, so the COUNTY was chosen as the unit of analysis. Displayed on the county map of Oklahoma (Figure A) are the predominant tribes. Tulsa County and Oklahoma County were included in this analysis because a large number of Native Americans live in these counties where the cities of Tulsa and Oklahoma City are located. Nearly every tribe is represented in these areas, and, although it was not possible to disaggregate Native American deaths by tribal affilation, the alcohol-related proportionate mortality (APM) was sufficiently high to warrant inclusion. Figure B shows the total number of deaths as well as the percent alcohol-involved deaths for these tribes.

Topping each of these lists is the Cheyenne -Arapaho Tribe with:

Total Deaths	586
% Alcohol-Involved Deaths	29.7
Mean Age at Death (years)	46.5

Observation of Figures C and D, respectively a map and a list of the percent alcohol-involved deaths among Native Americans in Cheyenne-Arapaho predominant counties, reveals that the distribution of deaths is far from homogenous. Eight of the counties had no alcohol-related deaths, while in another county, half of the deaths were alcohol-related.

In contrast, the Seminole Tribe is the lowest on the list (Figure B) of percent alcohol-related deaths.

Summary Statistics-Seminole
Oklahoma (1968-1978)

Total Deaths	322
% Alcohol-Involved Deaths	1.9
Mean Age at Death (years)	57.3

Since the Seminole Tribe predominates in only one county, it is difficult to further discuss the possible heterogeneity of alcohol-related mortality among the members of this tribe.

Summary Statistics-Cherokee
Oklahoma (1968-1978)

Total Deaths	1778
% Alcohol-Involved Deaths	3.5
Mean Age at Death (years)	62.3

A glance at the summary statistics for the Cherokee Tribe discloses that the APM is similar to the 2.4% for Oklahoma Whites and 3.2% for Oklahoma Blacks. A look at the map in Figure E again reveals the heterogeneity of the distribution of APM within the Cherokee counties with no alcohol-related (0.0%) deaths in Nowata County and an APM of 10.8% in Craig County. Figure F provides a tabular display of percent alcohol-involved deaths by race in Cherokee predominant counties which reveals the large within the Cherokee Nation, but among the Blacks and Whites in those counties as well.

A number of important caveats must be kept in mind in interpreting the data from this study. Information on tribal affiliation is inexact. Although a 10 year sample was used, for some tribes, the number of deaths was still extremely small. And, most importantly, the assumption that alcohol-involved mortality among Native Americans in a given county is attributable to the predominant tribe may not be valid.

In conclusion, alcohol-involved mortality among Native Americans in Oklahoma is not homogeneous among tribes, within any given tribe, or among counties. In order to have maximum impact, scarce resources need to be directed toward those most in need of services.

Predominant Oklahoma Tribes Distributed by County

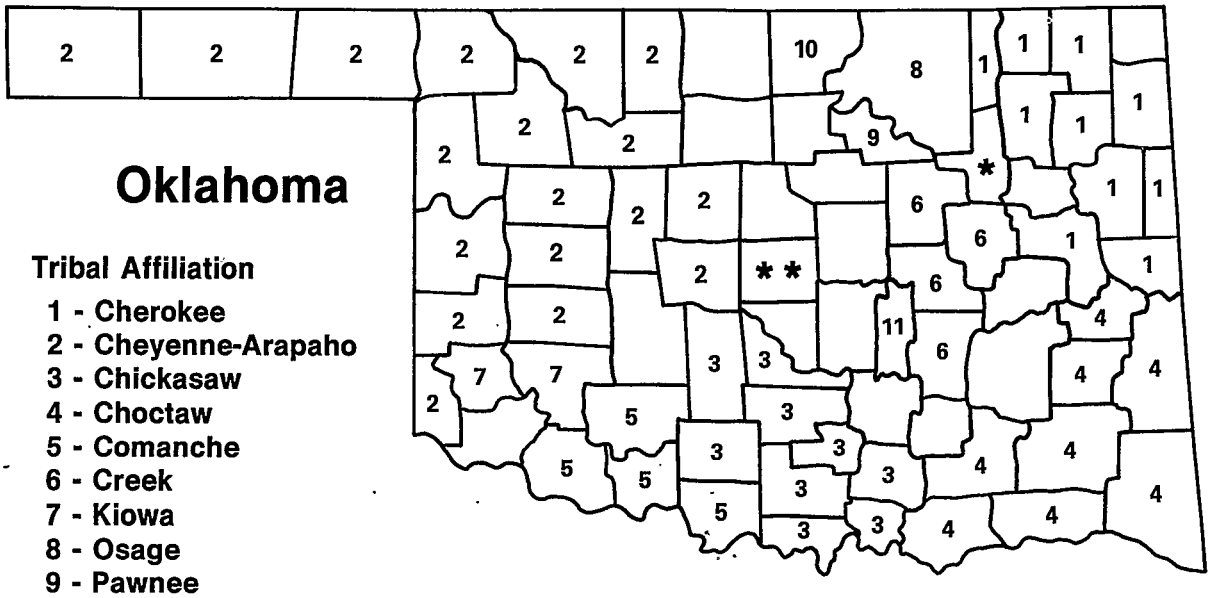


FIGURE A

Alcohol-Involved Deaths Among Native American Oklahomans (1968-78)

Tribe	Total Deaths	% Alcohol-Involved Deaths
Cheyenne-Arapaho	586	29.7
Kiowa	53	15.1
* Oklahoma County	486	14.2
Ponca	183	13.1
Comanche	273	11.7
Osage	258	11.6
Chickasaw	419	11.5
"Other"	1361	10.4
Pawnee	129	9.3
* Tulsa County	391	8.2
Choctaw	864	5.7
Creek	658	4.7
Cherokee	1778	3.5
Seminole	322	1.9

FIGURE B

Percent Alcohol-Involved Deaths Among Native Americans in Cheyenne-Arapaho Predominant Counties Oklahoma (1968-78)

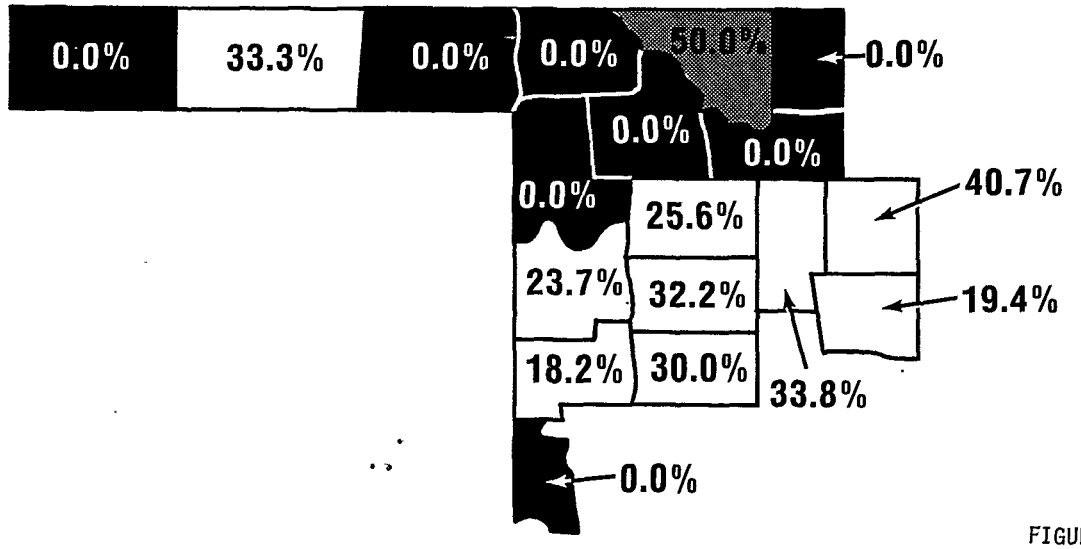


FIGURE C

Percent Alcohol-Involved Deaths Among Native Americans in Cheyenne - Arapaho Predominant Counties—Oklahoma (1968-78)

County	% Alcohol-Involved Deaths
Alfalfa	0.0
Beaver	0.0
Cimarron	0.0
Ellis	0.0
Harmon	0.0
Harper	0.0
Major	0.0
Woodward	0.0
Beckham	18.2
Canadian	19.4
Roger Mills	23.7
Dewey	25.6
Washita	30.0
Texas	33.3
Blaine	33.8
Kingfisher	40.7
Woods	50.0

FIGURE D

Percent Alcohol-Involved Deaths Among Native Americans in Cherokee-Predominant Counties Oklahoma (1968-78)

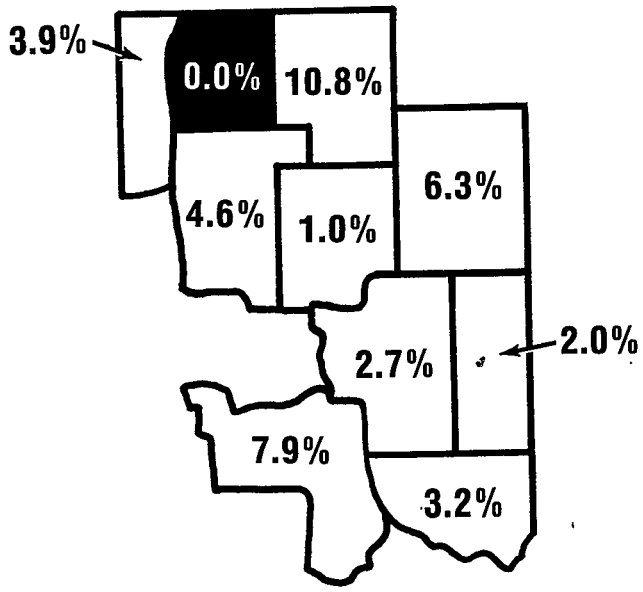


FIGURE E

Percent Alcohol-Involved Deaths by Race— Cherokee Predominant Counties— Oklahoma (1968-78)

	Percent Alcohol-Involved Deaths		
County	Cherokee	Black	White
Nowata	0.0	1.0	1.7
Mayes	1.0	0.0	2.2
Adair	2.0	0.0	1.1
Cherokee	2.7	8.1	1.5
Sequoyah	3.2	3.0	1.3
Washington	3.9	2.5	1.8
Rogers	4.6	5.7	1.8
Delaware	6.3	0.0	1.4
Muskogee	7.9	3.0	2.1
Craig	10.8	1.5	1.2

FIGURE F

1983 NHIS ALCOHOL/HEALTH PRACTICES SUPPLEMENT: PRELIMINARY FINDINGS

Henry Malin, National Institute on Alcohol Abuse and Alcoholism
Ronald W. Wilson, National Center for Health Statistics
Gerald D. Williams, Alcohol Epidemiologic Data System

INTRODUCTION

The Alcohol/Health Practices Supplement of the 1983 National Health Interview Survey (NHIS) was administered throughout the U.S. in 1983 to 22,418 persons using a national household probability sample. The Alcohol Supplement contains questions on drinking practices, health practices, health conditions, problems associated with drinking, and detailed items on the consumption of beer, wine and liquor. The Alcohol Supplement, one of the largest surveys ever conducted on alcohol consumption in the U.S., is a cooperative project between the National Center for Health Statistics (NCHS) and the National Institute on Alcohol Abuse and Alcoholism (NIAAA).

Alcohol abuse and alcoholism are pervasive health problems in the U.S. today (DHHS,1984). The wealth of data in the Alcohol Supplement will allow NIAAA to address a variety of key policy issues regarding the measured prevalence of alcohol use and abuse and its social and health consequences in the U.S. population.

Purpose of the Paper

The purpose of this paper is to present some preliminary findings on several current research projects being conducted by NIAAA's Alcohol Epidemiologic Data System. The preliminary findings give some indication of the content and structure of the extensive and detailed data in the Alcohol Supplement. Also, the preliminary results provide some new information on (a) the prevalence of alcohol consumption and (b) the complex issues of measuring and assessing the alcohol consumption of individuals in the population. Research projects currently underway concern:

1. The prevalence of drinking and differences in constructed levels of alcohol consumption compared to self-reported levels of alcohol consumption;

2. The relationship of different levels of alcohol consumption to both selected health conditions and to various problems related to drinking. This includes an examination of the U-shaped phenomenon in the relationship of drinking levels to overall and selected health conditions; and,

3. Different levels of drinking among various racial/ethnic and other demographic subgroups in the U.S. population.

Limitations

The findings in this paper are preliminary and represent the initial stages of more thorough and complete investigations being conducted by NIAAA. Methodological problems and issues for further research are noted, as appropriate.

METHODS

Constructed Levels of Alcohol Consumption

In order to establish drinking levels for persons in the sample, various items regarding the quantity and frequency (QF) of drinking of each type of alcoholic beverage (beer, wine and liquor) were used to develop levels of alcohol consumption in terms of an individual's average daily consumption of absolute alcohol.

Both a 2-item and a 3-item QF measure were developed using items on (a) the number of days drank, (b) the number of drinks consumed on days that the respondent drank, (c) the number of ounces in each drink and (d) the total number of drinks during the reporting period. The reporting period is a 2-week period preceding the week of the interview or, if the respondent had not had a drink in the 2-week period prior to the interview, a 2-week period prior to, and including, the date of the respondent's last drink.

The 2-item QF estimate consisted of the total number of drinks of (beer, wine, liquor) over the days of the reporting period multiplied times the number of ounces (beer, wine, liquor) in a typical drink. The 3-item QF estimate consisted of the number of days drank (beer, wine, liquor) in the reporting period multiplied times the number of drinks typically consumed on drinking days (beer, wine, liquor) and multiplied times the number of ounces (beer, wine, liquor) in a typical drink. If differences occured with the 2-item and 3-item QF estimations, the higher category of alcohol consumption was used.

Conversions to Absolute Alcohol

The total ounces of beer, wine and liquor consumed were converted to ounces of absolute alcohol using the conversion factors of .04 for beer, .15 for wine and .45 for liquor. The ounces of absolute alcohol from each beverage type were summed and divided by 14 to arrive at average daily consumption expressed in ounces of ethanol.

To establish the constructed levels of drinking, ranges of average daily consumption were used to classify repondents into abstainer, lighter drinking, moderate drinking and heavier drinking categories. The ranges for the constructed drinking levels were zero ounces of average daily ethanol intake for abstainers, .01 to .21 ounces of average daily ethanol intake for light drinkers, .22 to .99 ounces of average daily ethanol intake for moderate drinkers and 1.00 or more ounces of average daily ethanol intake for heavier drinkers.

The conversion factors to derive ethanol levels and the categories of drinking levels, i.e., abstainers, lighter, moderate and heavier drinkers, were developed by Johnson, et al. (1977) when analyzing drinking trends with national surveys conducted from 1971 through 1975. The same classification scheme has been used with more recent drinking surveys by other researchers (Clark and Midanik,1982; Wilsnack, et al., 1984). Also, Williams, et al. (1985) suggest that the QF reliabilities are quite high with either the 2- or 3-item average daily consumption estimates.

Drinking Categories

It is important to note that heavier drinking does not mean excessive nor necessarily problem drinking. The drinking categories have been used frequently in the literature and represent a systematic method for classifying individuals into drinking categories based upon their average daily consumption of ethanol. Also, abstainers are not non-drinkers only.

Abstainers are defined as respondents who reported that (a) they had never had 12 drinks of any alcoholic beverage in their entire life, (b) they had never had 12 drinks of any alcoholic beverage in any one year, or (c) for drinkers, they had not had a drink of any alchoholic beverage in the past year. Drinkers are defined as respondents who reported that they have had 12 or more drinks in any one year and that they had had at least 1 drink within the year preceding the interview. These are the respondents for whom detailed QF measures are available. While most of the analyses in this paper concern both the abstainer and drinking categories, those analyses pertaining to drinkers only are noted, as appropriate.

Weighted Data

Since national prevalence estimates were of most interest for the research questions, weighted data are used almost exclusively

in this paper. Data have been weighted according to sex, age and racial categories to represent the U.S. population. Results in this paper are presented for males and females separately, since these data and previous survey results document substantial differences in the patterns of drinking and alcohol consumption between men and women. While the differences between drinking levels of men and women are important, patterns of relationships to other variables for men and women separately are particularly interesting.

RESULTS

Overall Population Estimates

Table 1 presents the constructed drinking levels of the U.S. population by sex compared to other national surveys conducted since 1971. Results from the Alcohol Supplement suggest that the percentage of abstainers in the U.S. population 18 years age and older has increased in recent years, especially among women. However, the definition of the abstainer category is not identical to the previous surveys and the reporting period in the Alcohol Supplement is 14-days compared to a 30-day reporting period for the other surveys. However, it should be noted that no differences were found on average daily ethanol consumption between 14- and 28-day periods examined in a pre-test of the QF items in the Alcohol Supplement (Williams, et al., 1985).

Table 1: Percentages of Constructed Drinking Levels by Sex, 1971-1983

Constructed Drinking Levels	1971-76 Average	SRG* 1979	NHIS 1983
Males			
Abstainers	27	25	28
Lighter drinkers	28	29	28
Moderate drinkers	28	31	27
Heavier drinkers	18	14	17
(N)	(847)	(755)	(9343)
Females			
Abstainers	43	40	50
Lighter drinkers	36	38	30
Moderate drinkers	17	18	15
Heavier drinkers	4	4	5
(N)	(889)	(1003)	(12479)

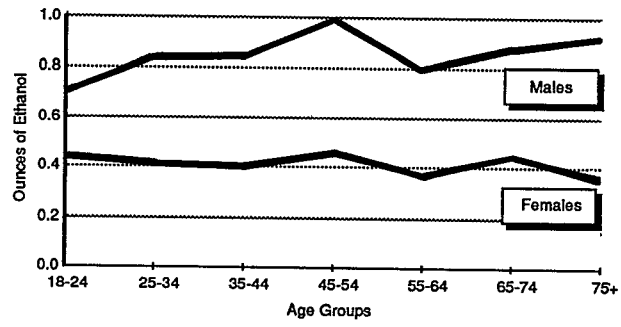
Note.--Adapted from Johnson, et al., 1977; Clark and Midanik, 1982. Percentages based upon weighted frequencies. Sample N's are in parentheses. Column percentages may not add to 100 percent because of rounding. *Social Research Group.

Apparent differences in the national prevalence rates are being examined to determine whether the differences might be attributable to different reporting periods, different definitions of the abstainer group, or to different methods of estimating levels of alcohol consumption. An increase in the number of abstainers, however, would be consistent with the drop in recent years of the total per capita consumption of absolute alcohol estimated from sales data (AEDS, 1984).

Figure 1 presents a graph of the measure of average daily consumption of ethanol by age and sex groups. These data, of course, are for the drinking population only. The figure emphasizes the differences in average daily alcohol consumption between men and women. In terms of average daily consumption, males between the ages of 45 and 54 appear to be the heaviest drinking group. Also of interest is the apparent increase in the average daily consumption of men and women after the age of 64, but the decrease in the average daily consumption for women after age 74 compared to an increase for men after age 74. Such patterns need further examination to understand whether they may be an anomoly in the data or whether they might be attributable to factors such as aging.

Figure 1

Average Daily Consumption of Ethanol for Drinkers by Age and Sex



It is important to note that the variability around the mean levels of average daily consumption is quite large. Standard deviations tend to be around 2 to 4 times the mean values. The distributions of average daily alcohol consumption tend to cluster toward the lower end of the distribution with a large tail representing very high levels of average daily consumption, i.e., the distribution of average daily alcohol consumption is highly skewed overall and within the sex and age groups.

Constructed and Self-Reported Levels of Drinking

Data are available in the Alcohol Supplement on self-reported levels of drinking which are labeled as abstainer, infrequent, light, moderate and heavy drinker. This permits an examination of the constructed levels of drinking compared to the self-perceived levels of drinking by the respondents.

Table 2 presents a comparison of the constructed drinking levels with the respondents' self-perceived levels of current drinking. Even with the difference in time periods, i.e., between the date of last drink and the date of the interview, the differences between what constitutes heavier drinking (the wording for self-reported drinking in the interview form is actually heavy) and the constructed level of heavier drinking is quite revealing.

Table 2: Comparison of Constructed Drinking Levels to Self-Reported Current Levels of Drinking by Sex

Constructed Drinking Levels	Percent M	Percent F	Drinking Levels Now Self Reported	Percent M	Percent F
Abstainer	*	*	Abstainer	5	5
Infrequent	*	*	Infrequent	26	39
Lighter	39	60	Light	42	42
Moderate	38	31	Moderate	26	14
Heavier	23	9	Heavy	2	1
(N)	(6690)	(6251)		(6592)	(6171)

Note.--Percentages based upon weighted frequencies. Sample N's are in parentheses. Weighted product-moment correlations between self-reported and constructed drinking levels: .518 for males and .455 for females.

Table 3 shows the average daily consumption levels of drinkers for the constructed and self-reported classifications. Moderate drinking as reported by the respondents is higher than the constructed heavier drinking category. The higher self-reported mean values for abstainers compared to the infrequent and lighter drinking groups represents a change in drinking by those who drank more in the past year than at the time of the interview.

Table 3: Average Daily Consumption of Ethanol for Constructed and Self-Reported Current Drinking Levels by Sex

Constructed Drinking Levels	Mean M	Mean F	Drinking Levels Now Self Reported	Mean M	Mean F
Abstainer	*	*	Abstainer	1.14	.38
Infrequent	*	*	Infrequent	.25	.16
Lighter	.10	.09	Light	.60	.36
Moderate	.52	.47	Moderate	1.51	1.16
Heavier	2.53	2.26	Heavy	4.21	2.42
(N)	(6690)	(6251)		(6592)	(6171)

Note.--Weighted means. Sample N's are in parentheses.

Further research is underway to re-evaluate the drinking categories as constructed for this paper. While the categories may have value for trend analyses to examine national prevalence rates over time, the discrepancy between the constructed and self-reported drinking levels needs further examination. Finer distinctions can be made, for example, between moderate, heavier and very heavy drinking based upon average daily consumption. It is possible, of course, that the use of the term, "heavy drinker" in the Alcohol Supplement inclined respondents to associate heavy drinking with problem drinking.

Another aspect of drinking patterns suggests that drinking levels are dynamic. That is, drinking levels tend to change over time, e.g., increasing or decreasing one's alcohol consumption, for short time periods or permanently. While most males and females report that they presently are at their highest level of drinking (as measured by lifetime levels), approximately one-third are at an increased or decreased level of alcohol consumption. Further analyses will be conducted to determine potential causal factors for the changes in drinking levels, e.g., health conditions and other factors. Previous results in this paper have already suggested that current abstainers may come primarily from the heavier drinking category.

Health Conditions and Drinking Levels

Table 4 shows the association of self-reported health status by drinking levels for males and females. The relationship indicates that increases in the levels of self-reported health status are related positively to increases in the constructed levels of drinking. In other words, either men and women who tend to drink more are healthier, or drinkers have a tendency to report positively on their health. This relationship is explored further in the relationships of the constructed drinking levels with various health conditions.

Table 4: Percentages of Self-Reported Health Status by Constructed Drinking Levels and Sex

Constructed Drinking Levels	Health Status Poor/Fair	Good	Very Good/Excl
Males (9311)			
Abstainers	20	26	54
Lighter drinkers	11	22	68
Moderate drinkers	7	21	73
Heavier drinkers	10	23	67
Females (12432)			
Abstainers	22	30	49
Lighter drinkers	9	26	65
Moderate drinkers	8	22	71
Heavier drinkers	10	23	68

Note.--Percentages based upon weighted frequencies. Sample N's are in parentheses. Row totals may not add to 100 percent because of rounding. Weighted product-moment correlations between drinking levels and self-reported health status: .136 for males and .191 for females.

The Alcohol Supplement provides a list of 25 health conditions to which survey respondents were asked to report whether or not they ever had each of the conditions. Table 5 presents the lifetime prevalence estimates for the selected health conditions listed (males only). Prevalence rates are expressed in cases per 100 (percentages). The pattern of prevalence rates demonstrates the U-shaped phenomenon in the association of drinking levels to selected health conditions.

Table 5: Lifetime Prevalence of Selected Health Conditions by Constructed Drinking Levels (Males only)

Health Condition	Lifetime Prevalence (%) A	L	M	H
Hypertension or high blood pressure	24.2	20.7	19.3	23.5
Hardening of the arteries	5.8	3.3	2.1	2.8
Tachycardia, arrhythmia or rapid heart	7.9	5.1	5.1	5.9
Arthritis or rheumatism	24.4	16.2	12.1	17.3
Convulsions or seizures	2.3	1.6	1.4	*
Blackouts	8.3	4.6	3.7	7.1
Shortness of breath	20.1	12.1	10.2	17.5
Insomnia or sleeplessness	16.4	12.7	10.2	16.6
Hepatitis	2.5	2.6	2.9	3.6
Any disease of the pancreas	1.6	1.2	*	*
An ulcer, other than skin ulcer	10.8	9.3	7.3	8.5
Any gastrointestinal bleeding	4.3	3.8	2.8	3.4
Diabetes	6.4	3.3	2.7	*
Heart attack or heart failure	7.0	4.7	3.1	3.3
Coronary heart disease	3.9	2.7	*	*
Stroke or hemorrhage of the brain	2.1	1.2	*	*
Angina pectoris	2.9	2.3	1.4	1.9
Cancer	3.3	2.6	1.8	2.6
Yellow jaundice	2.8	2.8	2.4	*
Fatty liver	*	*	*	*
Enlarged liver	*	*	*	2.1
Cirrhosis of the liver	*	*	*	*
Any other liver trouble	*	*	*	*
DT's or delirium tremens	*	*	*	*
Alcoholism	3.1	1.2	1.3	5.2
(N, average)	(2633)	(2591)	(2507)	(1549)

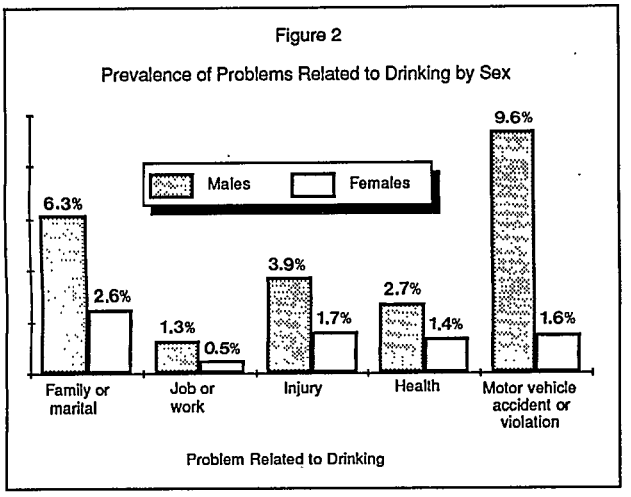
Note.--Percentages based upon weighted frequencies. Sample N's are in parentheses. A=Abstainers, L=Lighter drinkers, M=Moderate drinkers, H=Heavier drinkers. *Less than 30 cases.

The associations of the health conditions to the constructed drinking levels shows a pattern of negative association to the drinking levels. In other words, the lower one's drinking level, the more likely one is to have the particular health condition. Moderate drinkers generally fare better in terms of the health conditions listed. The relationships may reflect the selection of the particular health conditions in the list, but certain relationships, e.g., the negative relationship of drinking to certain heart conditions, are supported in recent literature about the potential benefits of moderate drinking to reduce the risks of cardiovascular disease.

Drinkers in the survey also reported on problems they had

encountered which were related to drinking. These included family or marital problems, job or work problems, injuries, health problems, and motor vehicle accidents or violations. Respondents made the causal link between their drinking and the particular problems and reported both lifetime occurances related to drinking and occurances within the past 12 months.

Figure 2 presents the lifetime prevalence rates (in percentages) of self-reported drinking problems for males and females. As might be expected, the prevalence rates generally were higher for men than for women. Initital examination of the 12-month prevalence rates suggested that the problems, on an annual basis at least, were relatively rare events in the U.S. population as a whole. Correlations between the drinking problems and the constructed drinking levels tended to be significant, but low for both males and females (between .05 and 15 for males and between .04 and .10 for females).



One or more problems due to drinking were reported by 16 percent of the males and 6 percent of the females. Twenty-seven percent of the male former drinkers and 17 percent of the female former drinkers, i.e., those men and women who had not had a drink for over a year, reported one or more problems related to drinking. The correlations between one or more reported problems due to drinking and the constructed drinking levels were .18 and .14 for males and females, respectively.

As noted earlier in the findings regarding the relationship of self-reported health status and the relationship of the constructed drinking levels to health conditions, the relationships are not in the direction that one might expect. In other words, drinking seems to be related positively to good (or better) health. To further examine the relationship of drinking levels to health, a scale score of each of the listed health conditions was developed which was simply the sum of all the conditions. This total score was then related to the constructed drinking levels to examine in more detail the relationship of drinking levels to health.

Table 6 presents the average number of health conditions for each of the drinking levels. It can be noted that abstainers and heavier drinkers tend to have the highest average number of health conditions. Conversely, lighter drinkers and moderate drinkers tend to have a lower average number of health conditions. This is the U-shaped phenomenon regarding the relationship of drinking levels to health.

Some observers believe that the U-shaped phenomenon occurs because former heavy drinkers are included in the abstainer groups. To test this hypothesis, those former drinkers who had not had a drink for over a year were removed from the abstainer group. The third and fourth columns of Table 6 show that the U-shaped phenomenon is still present, even though the

Table 6: Average Total Health Conditions Computed With and Without Former Drinkers

Constructed Drinking Levels	Average Number of Health Conditions: With All Abstainers M	With All Abstainers F	Excluding Former Drinkers M	Excluding Former Drinkers F
Abstainers	1.63	1.79	1.20	1.67
Lighter drinkers	1.14	1.21	1.14	1.21
Moderate drinkers	0.98	1.08	0.98	1.08
Heavier drinkers	1.31	1.54	1.31	1.54
(N)	(9284)	(12434)	(8291)	(11593)

Note.--Weighted means. Sample N's are in parentheses.

average number of health conditions is reduced considerably by the removal of the former drinkers.

Further investigation of the U-shaped phenomenon was attemped by controlling for age and sex of the respondents and, again examining the mean number of total health conditions with and without the former drinkers described above. This reduced the means for the male and female abstainer group, but the general U-shaped relationship was still present: Abstainers and heavier drinkers tended to have a larger average number of health conditions than lighter and, especially, moderate drinkers. More intensive analyses are planned to examine the U-shaped phenomenon with regard to drinking and health.

Demographic Characteristics and Drinking Levels

Differences in the drinking levels of individuals by different demographic subgroups in the U.S. population are well known in the alcohol field. In addition to the differences already mentioned with regard to sex and age, differences in the constructed drinking levels also are present with subgroups according to (a) geographic regions of the U.S. (Census regions), (b) racial and ethnic groups, (c)marital status, (d)years of education and (e) employment status.

Table 7 presents the distributions of the constructed drinking levels in the four U.S. Census regions. Except for the South, there are not large differences in the distributions of the drinking categories. A larger percentage of both men and women in the South claim to be abstainers, women in particular. If one

Table 7: Percentages of Constructed Drinking Levels by U.S Census Regions and Sex

Constructed Drinking Levels	U.S. Census Regions: North East	North Central	South	West
Males				
Abstainers	23	25	36	25
Lighter drinkers	28	30	25	29
Moderate drinkers	30	29	24	28
Heavier drinkers	19	16	16	18
	(2035)	(2544)	(3006)	(1758)
Females				
Abstainers	43	46	61	44
Lighter drinkers	34	32	25	31
Moderate drinkers	17	18	11	19
Heavier drinkers	5	5	3	6
	(2796)	(3174)	(4138)	(2371)

Note.--Percentages based upon weighted frequencies. Sampe N's are in parentheses. Column percentages may not add to 100 percent because of rounding.

examines only drinkers, in the South only women tend to drink less. Southern men appear to drink as much as men in other regions of the country.

Figure 3 presents percentages of the constructed drinking levels by racial/ethnic categories and sex. White males and white females tend to drink more than either Black or Hispanic males and females. The Hispanic subgroup in is not mutually exclusive from the white and Black subgroups. Both Black and Hispanic females tend to be abstainers. One finding in this analysis varies from fairly recent findings regarding the drinking practices of Hispanic men. Hispanic men generally do not tend to drink any more heavily than white men. Some research suggests that young Hispanic males, in particular, tend to be fairly heavy drinkers. Further research will be conducted with the Alcohol Supplement to examine drinking levels of Hispanics by age group.

Figure 3

Percentages of Drinking Levels by Race and Sex

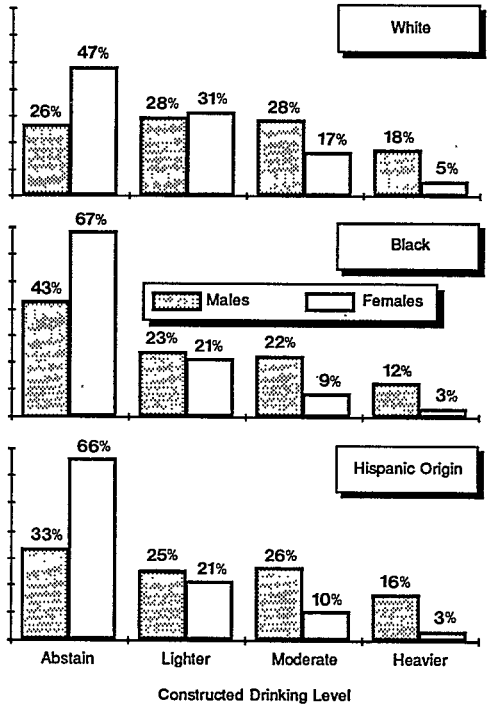


Table 8 presents the percentages of constructed drinking levels by marital status and sex. These results are consistent with studies over the years which have shown that divorced and separated men and women tend to drink more heavily than men and women who are married or widowed.

Table 9 presents the percentages of constructed drinking levels by years of education completed by the survey respondents. In general, the higher the level of education as measured by years of schooling completed, the higher the drinking levels. Family income levels show the same positive relationship to drinking levels, but these data are not presented.

Table 10 presents the constructed drinking levels by the employment status of males and females aged 18 to 64. Results suggest that unemployed males tend to drink more than both employed and not-in-the labor force males. The same relationship, however, does not appear to hold for females. Further research is under way to examine educational levels, employment status and marital status jointly.

Table 8: Percentages of Constructed Drinking Levels by Marital Status and Sex

Constructed Drinking Levels	Marital Status				
	Marr'd	Widow'd	Divor'd	Separ'd	Single*
Males					
Abstainers	29	49	20	17	27
Lighter	30	18	22	22	25
Moderate	26	19	32	33	30
Heavier	15	14	26	28	18
(N)	(6432)	(252)	(505)	(168)	(1952)
Females					
Abstainers	49	72	40	46	42
Lighter	32	18	33	28	30
Moderate	15	7	20	18	21
Heavier	4	3	7	8	7
(N)	(7867)	(1463)	(917)	(337)	(1859)

Note.--Percentages based upon weighted frequencies. Sample N's are in parentheses. Column percentages may not add to 100 percent because of rounding. *Specifically, single is never married.

Table 9: Percentages of Constructed Drinking Levels by Years of Education and Sex

Constructed Drinking Levels	Years of Education		
	Less than H.S.	H.S. Grad	More than H.S.
Males			
Abstainers	43	26	20
Lighter	23	29	30
Moderate	19	27	33
Heavier	15	18	18
(N)	(2429)	(3353)	(3501)
Females			
Abstainers	70	48	35
Lighter	19	32	37
Moderate	9	15	21
Heavier	3	4	7
(N)	(3396)	(5116)	(3902)

Note.--Percentages based upon weighted frequencies. Sample N's are in parentheses. Column percentages may not add to 100 percent because of rounding.

CONCLUSIONS

Results in this paper have concentrated on preliminary findings related to several research projects underway at NIAAA's Alcohol Epidemiologic Data System. Initial findings in the categories of drinkers suggest that problems may be inherent in the different definition of abstainers and the different reporting periods compared to surveys conducted previously.

The constructed drinking categories have been explained and their problems related to self-reported drinking examined. Specifically, there seems to be little correspondence between what experts denote as light, moderate and heavy drinking compared to the perceptions of the drinking respondents. Considerably more work needs to be done in the refinement of the constructed drinking levels.

The relationship of constructed drinking levels to various health conditions suggests, on first blush at least, that moderate drinking may have positive health benefits. The U-shaped

Table 10: Percentages of Constructed Drinking Levels by Employment Status and Sex (Ages 18 to 64)

Constructed Drinking Levels	Employment Status		
	Unemployed	Employed	Not in LF
Males			
Abstainers	26	23	35
Lighter	24	30	26
Moderate	28	30	23
Heavier	23	17	16
(N)	(591)	(6316)	(1043)
Females			
Abstainers	44	40	55
Lighter	35	36	27
Moderate	15	19	14
Heavier	6	6	4
(N)	(578)	(5898)	(3859)

Note.--Percentages based upon weighted frequencies. Sample N's are in parentheses. Column percentages may not add to 100 percent because of rounding.

phenomenon wherein some abstainers and heavier drinkers tend to have more health problems was addressed briefly without any firm conclusions. Even when controlling for age, sex and for former drinkers, light and moderate drinkers still appear to have significantly fewer health problems (from the list of conditions in the Alcohol Supplement at least).

The drinking levels among various demographic subgroups support generally the relationships of different demographic characteristics and drinking levels found previously in the literature. There were, however, some exceptions. Notably, the initial findings from this survey do not support the findings of relatively heavy drinking among male Hispanics. Also, more detailed research is necessery to focus upon different social characteristics of women drinkers in relationship to their levels of alcohol consumption.

The breadth and depth of data available in the Alcohol Supplement will continue to be a source of detailed study on NIAAA policy issues related to the prevalence of alcohol abuse and its consequences for the health status of the U.S. population. Results presented here have only scratched the surface of potential research and policy issues. Hopefully, however, this brief overview has provided an orientation to this new, exciting data base and indicated some of the potential research agendas that can be built from the Alcohol Supplement.

REFERENCES

Alcohol Epidemiologic Data System. Alcohol Data Reference Manual, Vol. 1: U.S. Apparent Consumption. Rockville, MD: National Institute on Alcohol Abuse and Alcoholism, 1985. In press.

Clark, W.B. and Midanik, L. Alcohol use and alcohol problems among U. S. adults: Results of the 1979 National survey. In: Alcohol and Health Monograph 1: Alcohol Consumption and Related Problems. Washington, D.C.: U.S. Government Printing Office, 1982.

Department of Health and Human Services. Alcohol and Health. Fifth Special Report to the U. S. Congress from the Secretary of Health and Human Services. Washington, D.C.: U.S. Government Printing Office (December 1983).

Johnson, P., Armor, D., Polich, S. and Stambul, H. U.S. adult drinking practices: Time trends, social correlates and sex roles. Working note prepared for NIAAA. Contract No. ADM-281-76-0020. Santa Monica, CA: Rand Corporation, 1977.

Williams, G., Malin, H. and Aitken, S. Reliability of self reported drinking in a general population survey. J. of Studies on Alcohol, 1985, 46, 223-227.

Wilsnack, S.C., Wilsnack, R.W. and Klassen, A.D. Epidemiological research on women's drinking, 1978-1984. In: Proceedings of the National Research Conference on Women and Alcohol. University of Washington, Seattle (May 1984). In press.

SELF-REPORTED DRINKING BEHAVIOR AMONG MEXICAN AMERICANS: SOME PRELIMINARY FINDINGS

Charles M. Christian, University of Maryland, College Park, Md.
Terry S. Zobeck, Alcohol Epidemiology Data System, Washington, D.C.
Henry J. Malin, National Institute on Alcohol Abuse and Alcoholism, Rockville, Md.

During the 1970s the National Institute of Alcohol Abuse and Alcoholism (NIAAA) conducted a series of national surveys that characterize the self-reported drinking patterns of the U.S. population. Results across these surveys are fairly similar. Approximately one-third of the adult population described themselves as abstainers, one-third as light drinkers and the remaining third as moderate or heavier drinkers.[1]

While most national surveys have elicited information on the patterns and levels of alcohol use and abuse among the general population, such surveys of specific minority or special populations only recently have been undertaken. Many smaller and more localized surveys have provided evidence of disproportionately high levels of alcohol consumption and higher prevalence and incidence of alcohol-related problems among these populations.

The descriptive survey studies now emerging are enriching our knowledge of alcohol use and alcohol-related problems among minority populations. Clearly, minority groups are culturally different, and within these groups many subcultures exist. Such groups differentially perceive drinking opportunities, limitations, and functions within a culturally-defined environment.[2] It has been reported that drinking behavior is related to the technology and norms of particular cultures, to the organizational, economic, and political characteristics of societies, and to the roles assumed by individuals within their social network.[3] These factors determine the extent to which individuals are exposed to the risk of developing drinking problems. In essence, our further understanding of alcohol use and abuse and alcohol-related problems of minorities requires an assessment of the interplay of these characteristics.

This paper focuses on one specific minority subgroup--the Mexican Americans. Although few studies have focused on Hispanic American subgroups, research on alcohol use patterns among Hispanic Americans in general provides useful information. For example, several studies note higher rates of heavy drinking among Hispanic males and higher rates of abstinence among females[4], perhaps suggesting that their culture places positive sanctions on male drinking and negative sanctions on female drinking.[5] Others have found that Hispanic Americans, particularly males, tend to under-report their drinking behavior.[6] A study of Mexican Americans in California further notes that less acculturated individuals self-report a lower number of heavy drinkers.[7] Still other research tends to show that while Hispanic women are more likely to be abstainers, their abstinence decreases with increasing acculturation.[8] Current drinkers among Hispanic women tend to be young or middle-aged with more than the mean level of education.[9]

While the literature points out a number of socio-demographic attributes which may influence Hispanic American drinking behavior and alcohol-related problems[10], the evidence is particularly inconsistent in both time and space. Research findings suggest that Hispanic Americans tend to have higher alcohol use and alcohol-related problems than Anglos, but offer little information on the recent drinking practices among Hispanic American subgroups.

In this assessment of the drinking patterns of Mexican Americans, specific reference is made to the proportion of abstainers, current drinkers, occasional and former drinkers, their consumption levels, beverage preferences, and self-described drinker categories (abstainer, light, moderate, and heavier). These consumption descriptors are then assessed in relation to several socio-demographic variables: age, sex, income, education, language, and marital status. At this point, it is emphasized that none of the analyses in this paper are age-adjusted, and further, the findings and results of this research are provisional.

DATA AND SELECTED CHARACTERISTICS OF SAMPLE GROUP

Due to the lack of specificity with regard to alcohol use among Hispanic Americans, and particularly among subgroups of this population, the Alcohol, Drug Abuse, and Mental Health Administration developed and sponsored the "Adult Sample Person Supplement" (ASPS) of the Hispanic Health and Nutrition Examination Survey (Hispanic-HANES). This survey allows for a comprehensive assessment of the alcohol use among subgroups of Hispanics (Cuban Americans, Puerto Ricans, and Mexican Americans) linked with a wide array of socio-economic, health, bio-chemical, and nutritional data. NIAAA's section of ASPS provides the alcohol research community, for the first time, with a large population based data source on the drinking patterns of Hispanic Americans.

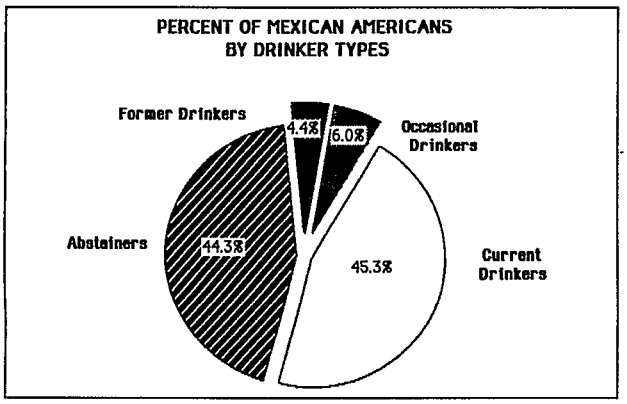
The ASPS was administered to 4,912 Mexican Americans between the ages of 12 and 74. Survey sampling was concentrated in Texas, California, Arizona, New Mexico, and Colorado since about 83 percent of the 8.74 million Mexican Americans enumerated by the 1980 census live in these states. Adding appropriate weights to the sampled population presumes coverage of approximately 7.0 million Mexican Americans.

It is noted here that no tests of significance have been made on these data. Because the sample is drawn via a non-random design and because standard descriptive statistical packages provide variance estimates based on random samples, it is emphasized that the true estimates of variance associated with these preliminary data analyses are considered to be three times greater than those calculated. Future analyses of these data will be adjusted accordingly and will consider non-response bias.

DRINKING BEHAVIOR

Mexican Americans are initially classified according to two types of drinking behavior--drinkers and abstainers. Drinkers comprise 51 percent of all Mexican Americans (Figure 1). Forty-five percent are Current drinkers (defined as those who had consumed an alcoholic beverage during the 28-day reference period prior to being interviewed) and 6 percent are Occasional drinkers (those who had consumed an alcoholic beverage during the reference period, but whose last drink was less than one year from the reference period). Overall, about 88 percent of all drinkers are current drinkers. Abstainers are 49 percent of Mexican Americans, of which 4.4 percent are Former drinkers. Abstainers are defined according to the following criteria: 1) consumed less than 12 alcoholic drinks during their lifetime; 2) consumed less than 12 drinks in any one year; 3) consumed their last drink more

Figure 1



than one year from the four-week reference period--these are classified as Former drinkers; and/or 4) consumed less than 0.01 ounces of absolute ethanol on an average daily basis.

ABSTAINERS: THEIR REASONS FOR NOT DRINKING

The leading reason abstainers (excluding former drinkers) offered for not drinking is that they don't care for and/or dislike alcohol (Table 1). More than one-half selected this choice. The

Table 1

REASONS FOR NOT DRINKING AMONG
ABSTAINERS AND FORMER DRINKERS

REASONS FOR NOT DRINKING	ABSTAINERS Total	Male	Female	FORMER DRINKERS
No Need/ Not Necessary	7.5	11.3	6 0	16 1
Don't Care For/Dislike It	58.1	45.0	63.4	12 4
Medical/Health Reasons	4.3	7.9	2 9	25.5
Religious/Moral Reasons	3.5	1.9	4.1	20 8
Brought Up Not To Drink	2.4	1.4	2.8	N/A
Costs Too Much	0.2	0.5	0.1	1.1
Family Member Alcoholic	1.1	0.8	1.2	0.8
Infrequent Drinker	10.3	8 4	11.0	3.6
Alcoholic/Problem Drinker (Self)	N/A	N/A	N/A	4 5
Other	11.9	21.3	8.0	15.0

NOTE: Totals may not equal 100 percent because of rounding

N/A-- not applicable or respondents were not presented this choice

next two leading reasons are infrequent drinker and no need for alcohol. Only about 4 percent do not drink because of medical or health reasons, and smaller percentages do not drink because of religious, moralistic reasons or an alcoholic family member.

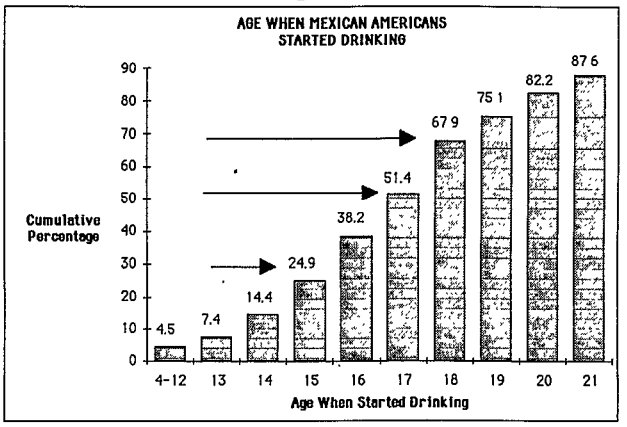
While both males and females selected "don't care for/dislike it" as their leading reason for not drinking, a greater proportion of females selected this reason (63 percent) than males (45 percent). The next two leading responses for males are "no need/not necessary" and "infrequent drinker." For females, the order of these two responses is reversed.

Former drinkers selected different reasons for not drinking. In contrast to abstainers discussed above, one-fourth of former drinkers selected "medical/health reasons" as the most important reason for abstaining, followed by 21 percent who do not drink because of "religious/moral reasons," and 16 percent who indicated "no need/not necessary."

ALCOHOL CONSUMPTION

Of all Mexican American drinkers between the ages of 12 and 74, about 80 percent started drinking between the ages of 14 and 21 (Figure 2). If those who started drinking prior to age 14 are

Figure 2



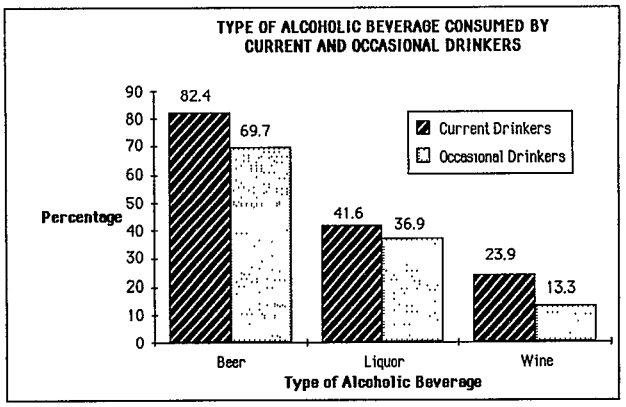
included, then about one-half started drinking by their 18th birthday and more than two-thirds by their 19th. Survey data further reveal that the largest percentage of Mexican Americans started drinking at age 18.

The Quantity, Frequency, and Variability of Beverage Consumption

The following analyses are based on individuals' self-reported drinking practices.

Beer. Overall, beer is the most favored alcoholic beverage among current drinkers. During the reference period, 82 percent consumed beer, 41 percent consumed liquor, and 23 percent drank wine (Figure 3).

Figure 3



Total beer consumption varies significantly among current drinkers. During the reference period, about one-half consumed 15 or fewer beers; almost one-fourth drank between 16 and 36 beers; and 10 percent drank between 37 and 60 beers. About one of every eight current drinkers drank more than 60 beers (in excess of two daily). Ninety-one percent of all beers consumed weighed 12 ounces.

The data further reveal information as to frequency of consumption. Slightly more than 54 percent drank three or fewer beers on the days they drank beer; 17 percent more than a 6-pack; and 5 percent drank in excess of 12 beers (Figure 4).

Liquor. Liquor or spirits (such as whiskey, rum, gin, vodka, and tequila) was the second most frequently consumed alcoholic beverage among Mexican American current drinkers (chosen by 42 percent). Slightly more than one-half of these drinkers consumed less than five glasses of liquor during the reference period, about 20 percent drank between five and nine glasses, and another 10 percent between 10 and 20 glasses. The remaining 13 percent consumed in excess of 20 glasses of liquor, slightly less than one glass per day. Most of these liquor drinks (76 percent) contained one ounce; 13 percent were two-ounce drinks; and 3 percent were

three-ounces.

One-quarter of liquor drinkers consumed only one drink on the days they drank liquor, another one-quarter had two drinks, about 37 percent consumed three to six drinks, and the remainder (about 10 percent) drank more than seven drinks (Figure 5).

Wine. Wine was drank least frequently. Slightly less than one-quarter of the current drinkers (24 percent) consumed wine during the reference period. Cumulatively, about one-quarter of all wine drinkers consumed only one glass of wine, slightly more than one-half drank from one to three glasses, and almost 75 percent consumed less than eight glasses. On the higher end of the scale, about 14 percent consumed eight to 18 glasses, and about 10 percent had more than 20 glasses. Approximately 91 percent of these drinks contained eight or less ounces per drink.

On the days that wine drinkers drank wine, about three-fourths consumed no more than two glasses (Figure 6). Fully 91 percent of wine drinkers drank between one and four glasses on the days they drank wine. About 9 percent drank in excess of five glasses daily.

Figure 4

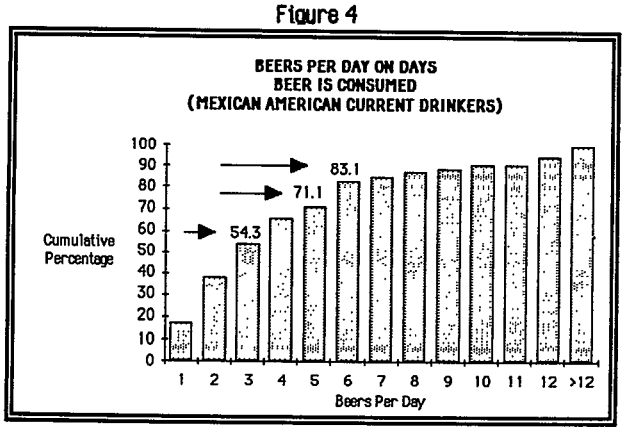


Figure 5

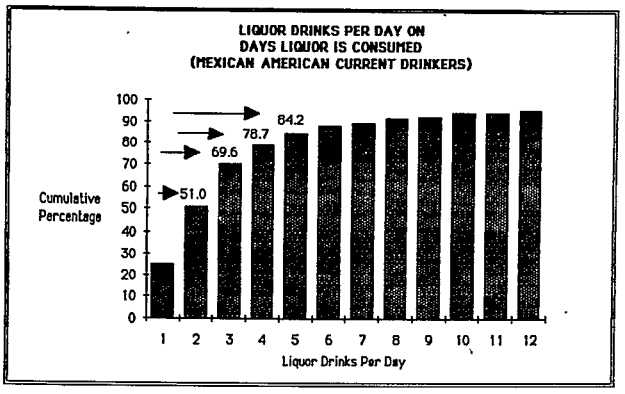
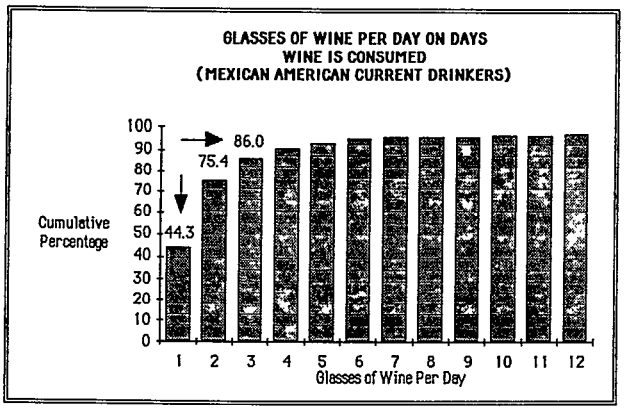


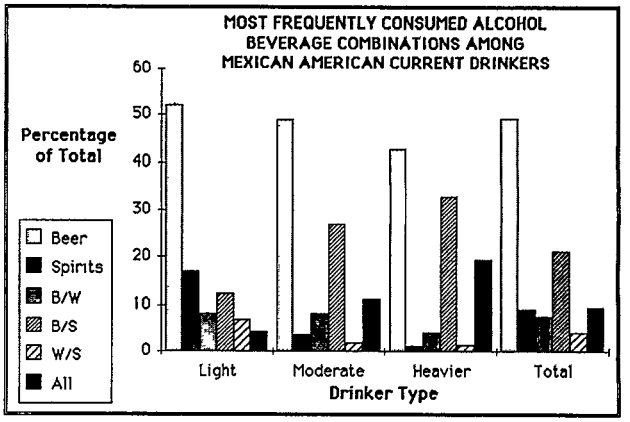
Figure 6



Alcohol Beverage Consumption By Drinker Type

Based on the mean daily absolute ethanol consumed by individuals (hereafter indicated by MDAE), drinkers can be further classified as Light, Moderate, and Heavier. Light drinkers consume between 0.01 to 0.21 ounces of ethanol daily; Moderate, 0.22 to 0.99 ounces; and Heavier, in excess of 1.00 ounce daily. Using this classification, it is found that about one-half (52 percent) of light drinkers consume beer only, 17 percent consume spirits only, and 12.7 percent, beer and spirits only (Figure 7).

Figure 7



Moderate drinkers also are predominantly "beer only" drinkers as slightly under one-half (49 percent) are so classified. About one-fourth (26 percent) drink beer and spirits only; and 11.2 percent drink all combinations of alcoholic beverages.

Heavier drinkers have the lowest percentage (43 percent) of "beer only" drinkers. Almost one-third (31 percent) drink beer and spirits only, and another 19 percent drink all combinations of alcoholic beverages.

In general, these data show that light drinkers are the most likely to be beer only drinkers and the least likely to be drinkers of all alcoholic beverage combinations; moderate drinkers are the least likely to consume spirits only, but are the most likely to consume beer and wine combinations; and heavier drinkers are the least likely to consume spirits only, beer and wine, or wine and spirits combinations only. They are, however, most likely to be consumers of all alcoholic beverage combinations.

SOCIO-DEMOGRAPHIC CORRELATES OF MEXICAN-AMERICAN DRINKING BEHAVIOR

Our attention is now directed toward understanding the extent to which alcohol consumption varies with socio-demographic characteristics of Mexican Americans. Specific attention is given to the following attributes--age, sex, language, education, income, and marital status--and their relationship with type of drinker--abstainer, current, occasional and former. Because occasional and former drinkers form relatively smaller groups, the discussion will focus primarily on abstainers and current drinkers.

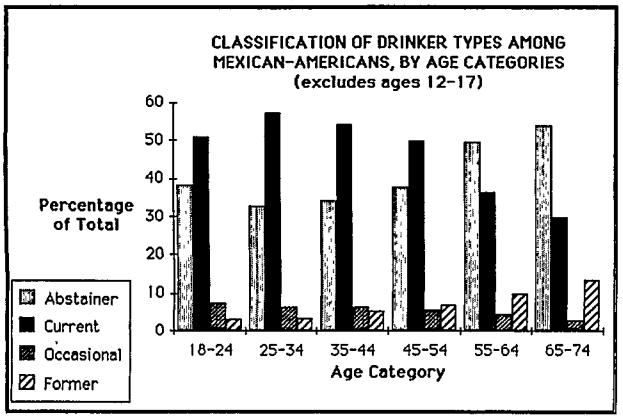
Age. Although the entire sample population ranges in age from 12 to 74 years, for this analysis the 12-17 year olds (17 percent of the population) are excluded for two reasons. First, more than three-fourths of this group are abstainers, thus skewing the overall distribution significantly; and second, youth in this age group are legally denied access to liquor because state laws prohibit sales to minors.

Fully 20 percent of the total population are between 18 and 24 years; 26 percent between 25 and 34; 15 percent between 35 and 44; 10 percent between 45 and 54; 7 percent between 55 and 64; and 4 percent between 65 and 74 years. It is noted that the data are

divided into 10-year intervals, with the exception of the youngest (18-24 years) which has been truncated.

The four age groups that make up the population between 18 and 54 have similar patterns of abstainers and current drinkers (Figure 8). In each group, current drinkers represent more than

Figure 8

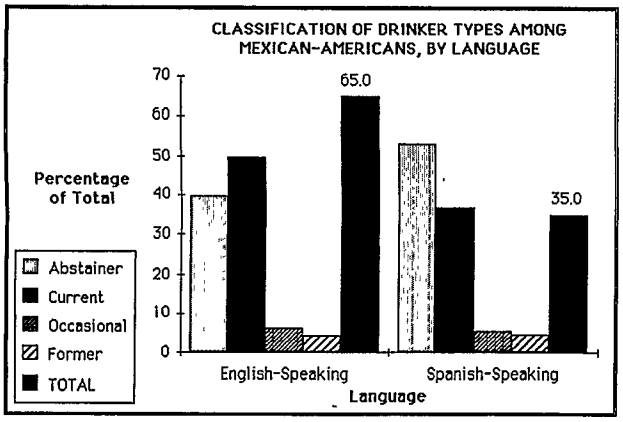


one-half of the total, and abstainers vary between 30 and 40 percent. For those over 55 years of age, however, the pattern is reversed, with abstainers in the majority. The older age groups also have higher proportions of former drinkers than the younger age groups.

Sex. Mexican Americans are almost equally proportioned by sex--50.5 percent are male and 49.5 percent female. Males are more than twice as likely to be current drinkers (63 percent) than females (31 percent), while females are more than twice as likely to be abstainers than males (64 percent versus 25 percent, respectively). Males have slightly higher proportions of occasional and former drinkers.

Language. As many as 65 percent of the Mexican Americans spoke English during the survey interview while 35 percent spoke Spanish (Figure 9). It is not known that those who spoke Spanish

Figure 9

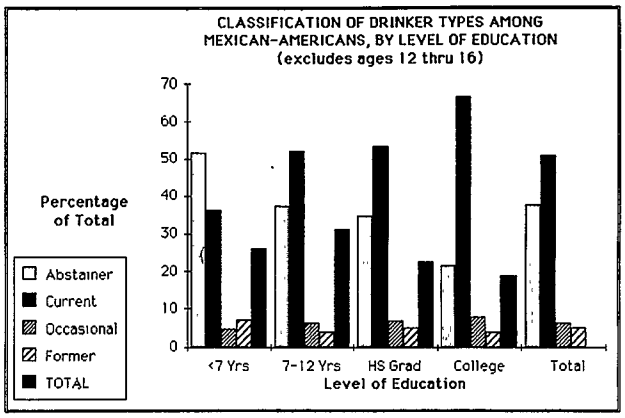


could not speak English, but simply that they felt more comfortable conversing in Spanish. It can be presumed, however, that those who chose Spanish are perhaps less acculturated than those who spoke English.

Among English-speaking persons, one-half are current drinkers, and 40 percent are abstainers. Spanish-speaking respondents have a much higher percentage of abstainers (43 percent) and a lower percentage of current drinkers (37 percent). Both language groups have relatively similar proportions of occasional and former drinkers.

Education. For this particular analysis, youth between the ages of 12 and 16 are excluded since it is assumed that the great majority are still attending school. Of the remaining population, 28 percent have less than seven years of education; 32 percent have 7-12 years; 23 percent graduated from high school; and 19 percent have some college (Figure 10). Overall, slightly more

Figure 10

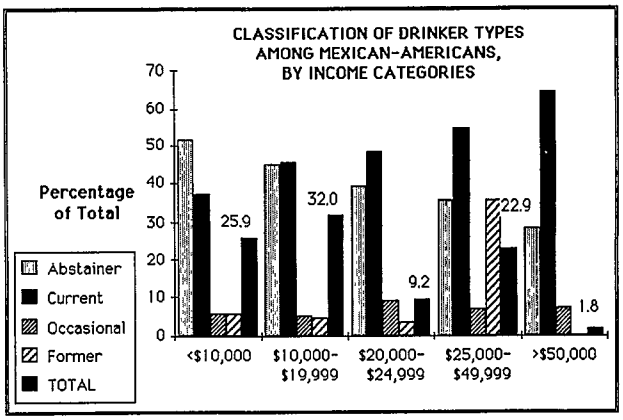


than one-half of these respondents are classified as current drinkers and about 38 percent are abstainers. Approximately seven percent are occasional and five percent are former drinkers.

Among persons with the lowest education level, slightly more than one-half are abstainers and about 36 percent are current drinkers. As the education level increases, the proportion of abstainers decreases to a low of 22 percent for respondents with some college education. Conversely, the proportion of current drinkers increases with education level to a high of 67 percent among those with some college.

Income. Survey data show that as many as one-quarter of all Mexican Americans live in families with annual incomes of less than $10,000, and more than one-half (58 percent) are in families with incomes less than $20,000 (Figure 11). On the

Figure 11



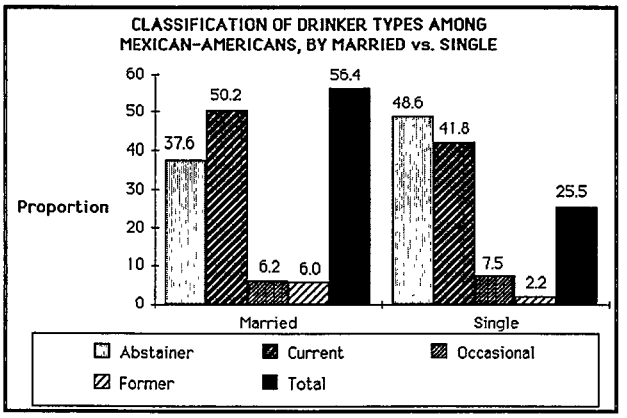
upper income scale, about 23 percent live in families with between $25,000 to $49,999, but less than two percent are found in families with greater than $50,000 annual income. Using the Census income classification (with some minor modifications), several important findings are revealed about income and drinker type.

Slightly more than one-half (52 percent) of persons with less than $10,000 annual income are abstainers--the largest percentage of any income group. The proportion of abstainers decreases with increasing income. The income group with the lowest percentage of abstainers (28 percent) is found among those at the highest income level, i.e., greater than $50,000. Current drinkers, on the other hand, are proportionately more concentrated in the highest income categories. Almost two-thirds of those with annual family incomes exceeding $50,000 are current drinkers

while only about 37 percent of the lowest income group are so classified.

Marital Status. Because of limited data across the various marital status categories, two sets of analyses are performed here. The first focuses on married--spouse in household, and single--never married respondents by all drinker types; the second highlights all marital status categories by current drinkers and abstainers only. In the first analysis, among those who are married (56 percent of Mexican Americans), about one-half are current drinkers, approximately 38 percent are abstainers, and 6 percent are occasional and former drinkers (Figure 12). The

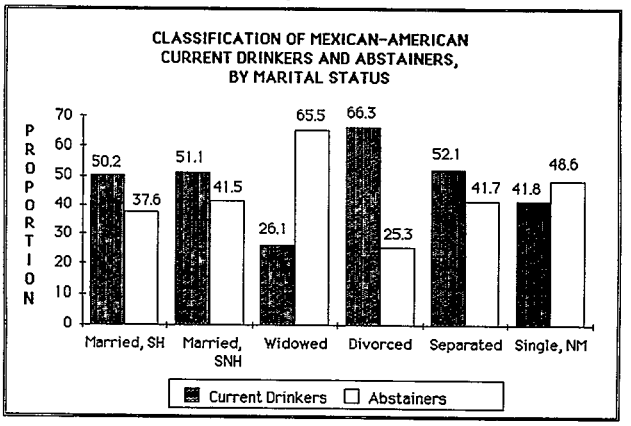
Figure 12



pattern for single respondents (26 percent) is reversed--abstainers make up the largest drinker type (49 percent) with the proportion of current drinkers somewhat smaller (42 percent). Occasional drinkers make up about 8 percent of single persons and former drinkers about 2 percent.

The analysis of all marital status types by current drinkers and abstainers shows that married--spouse not in household, divorced, and separated individuals tend to have higher proportions of current drinkers than abstainers (Figure 13). This is

Figure 13



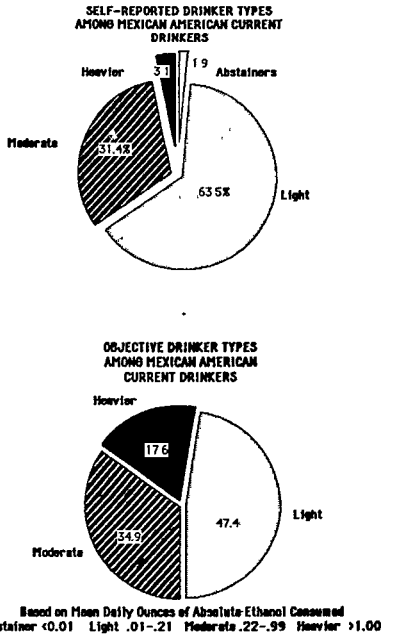
particularly true among the group of divorcees in which current drinkers outnumber abstainers almost 3 to 1. The pattern is reversed among widowed persons with abstainers outnumbering current drinkers by a similar ratio.

MEAN DAILY ABSOLUTE ETHANOL (MDAE) CONSUMPTION AMONG CURRENT DRINKERS

The ASPS survey further collected data to compare one's perception of drinking with self-reported alcohol consumption. Respondents were asked to classify their drinking as light, moderate, or heavy. This classification is compared with their self-reported consumption of alcoholic beverages. The consumption of alcohol by beverage type is calculated to derive mean daily absolute ounces of ethanol consumed. This is accomplished by translating the number of drinks into ounces of absolute ethanol consumed per day, using 0.45 ounces of ethanol per 12 ounces of beer, 0.40 ounces per glass of wine, and 1.0 ounce per drink of liquor as mean ethanol equivalents.. This mean daily absolute ethanol (MDAE) intake is then compared with the self-described drinker classification.

Almost 64 percent of the current drinkers self-classified themselves as light drinkers, 31 percent as moderate, 3 percent as heavy, and almost 2 percent classified themselves as abstainers (Figure 14). When these same drinkers are classified according to

Figure 14



the objective MDAE standards (Abstainer, <0.01 ounces; Light, 0.01-0.21 ounces; Moderate, 0.22-0.99 ounces; and Heavier, >1.00 ounce), only 47 percent were actually light drinkers, 35 percent moderate, and 18 percent heavier drinkers. In essence, Mexican Americans tend to perceive their drinking as considerably less than objectively constructed standards.

Closer inspection of self-reported and objectively classified drinkers shows the extent of incongruity. Among those objectively classified as light drinkers, about 83 percent self-classified themselves similarly as measured by the objective criteria, but 14 percent considered themselves moderate, and 3 percent as abstainers. Among the moderate drinkers, 41 percent classified themselves as such, but the majority (55 percent) perceived themselves as light drinkers. In the case of heavier drinkers, only 11 percent self-classified themselves as such; 61 percent self-classified themselves as moderate, and surprisingly, almost 28 percent considered themselves light drinkers. Clearly, the data show that Mexican Americans are more likely to see themselves as lighter drinkers irrespective of the amount of alcohol consumed.

An assessment of the self-reported average daily alcohol consumption between male and female current drinkers shows that females, overall, drink considerably less than their male counterparts (Table 2). Among self-described abstainers, very light, light, and moderate drinkers, the MDAE intake among females is considerably less than for males. For self-described heavy drinkers, their MDAE consumption is only slightly less than that for males.

The MDAE consumption among males and females according to objectively constructed drinker categories differs only slightly

Table 2

MEAN DAILY CONSUMPTION OF ETHANOL OF SELF-REPORTED AND OBJECTIVELY CLASSIFIED DRINKER TYPES (by Sex)

SELF-REPORTED DRINKERS	Means Males	Means Females	OBJECTIVELY CONSTRUCTED	Means Males	Means Females
Abstainer	.48	.12	Abstainer	NA	NA
Very Light	.16	.05	Very Light	NA	NA
Light	.47	.17	Light	.10	.08
Moderate	1.11	.68	Moderate	.51	.43
Heavier	2.73	2.50	Heavier	2.36	2.96

Note: Weighted means.

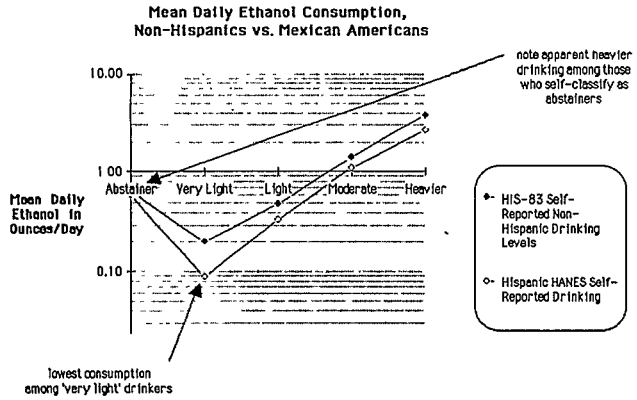
Note: Among Current Drinkers, there are no abstainers among objectively constructed Drinker Types.

Note: "Very Light" is not an objectively constructed Drinker Type category and thus no drinkers are listed in this category.

among light, moderate, and heavier drinkers. Perhaps the most revealing finding is among heavier drinkers where the MDAE intake of females is higher than that of males (2.96 versus 2.36 ounces, respectively). It is emphasized that the standard deviation of the mean for females is quite large; in fact, larger than the mean itself. For males, the standard deviation is also high, but smaller than the mean. Such variability around the mean denotes very high levels of alcohol consumption among some women. In fact, the MDAE range exceeds 18 ounces for females, but is less than 12 ounces for males. Certainly, the large standard deviations for both sexes suggest the influence of respondents with relatively high MDAE consumption.

A final analysis of MDAE intake entails a comparison of this index with that of non-Hispanics using the Health Interview Survey, 1983 (HIS 83) "self-reported non-Hispanic drinking levels" and Hispanic HANES "self-reported Mexican American drinking levels." Our analysis shows clearly that although some difference is found among the MDAE consumed by these two groups (collected via different samples), a similar overall pattern exists (Figure 15). Hispanic MDAE intake among all categories of

Figure 15



drinkers is lower than that of non-Hispanics, particularly among very light drinkers. Non-Hispanics show heavier drinking among some individuals who classify themselves as abstainers. In sum, it is noted that the patterns are quite similar, thus indicating a high degree of reliability associated with the data.

SUMMARY AND CONCLUSIONS

Although the findings of this research are preliminary, it is noted that about one-half of all Mexican Americans between the ages of 12 and 74 years are defined as abstainers--a finding reinforced by previous research. In general, abstainers do not drink because they dislike the taste of alcohol and/or have no need for it. Former drinkers, on the other hand, do not drink because of health problems and religious and moral reasons.

More than 80 percent of Mexican American respondents indicate that they started drinking between 14 and 21 years of age. The largest percentage started drinking alcohol at age 18. Beer is the most favored drink among Mexican American respondents, followed by liquor, and then wine, but as one becomes a heavier drinker, the probability of consuming all alcoholic beverage combinations becomes greater.

Several socio-demographic characteristics tend to influence drinking behavior among Mexican Americans. The proportion of current drinkers tends to be highest in age groups between 18 and 54 years and tapers off significantly after these years. Males are considerably more likely to be current drinkers than females, yet the more acculturated females tend to have a higher proportion of female current drinkers. Alcohol consumption tends to vary with education and income levels. As both education and income level increases, the greater the proportion of current drinkers.

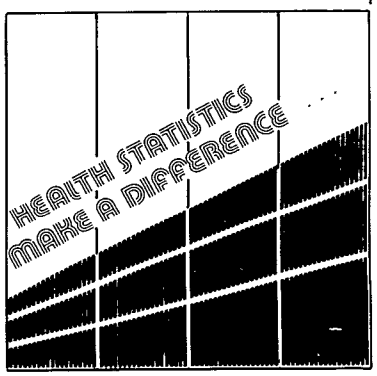
Self-reported drinker categories differ from those constructed objectively. Accordingly, Mexican American current drinkers tend to perceive their alcohol consumption as considerably less than the objectively constructed standards. Also, average daily ethanol consumed by Mexican American heavier drinkers varies quite considerably, and females have a slightly higher mean daily alcohol intake than males.

REFERENCES

1. U.S. Department of Health and Human Services, Fifth Special Report to the U.S. Congress on Alcohol and Health (Washington, D.C.: U.S. Department of Health and Human Services, December, 1983).
2. Schaefer, J.M. "Ethnic and Racial Variations in Alcohol Use and Abuse," in National Institute on Alcohol Abuse and Alcoholism, Alcohol and Health Monograph 4: Special Population Issues (Washington, D.C.: U.S. Government Printing Office, 1982), pgs. 293-311.
3. Institute of Medicine, National Academy of Sciences, Alcoholism, Alcohol Abuse, and Related Problems: Opportunities for Research (Washington, D.C.: National Academy of Sciences, 1980), p. 17.
4. Cahalan, D. and Room, R. Problem Drinking Among American Men: Monograph No. 7 (New Brunswick, N.J.: Rutgers Center of Alcohol Studies, 1974). and Cahalan, D. and Treiman, B. Drinking Behavior, Attitudes and Problems in San Francisco (Berkeley, Ca.: Social Research Group, School of Public Health, University of California, 1976).
5. Gilbert, M.J. Five Week Alcoholism Ethnography Conducted in Three Spanish Speaking Communities (Sacramento: California State Office of Alcoholism, 1978).
6. Williamson, G. "Out of the Garbage Rises the Truth," San Francisco Chronicle, February 12, 1976.
7. Alcocer, A.M. Final Report of Drinking Practices and Alcohol-Related Problems of Spanish-Speaking Persons in Three California Locales. Volume II (Alhambra, Ca.: Technical Systems Institute, 1977).
8. Cahalan, D., Roizen, R, and Room, R. "Alcohol Problems and Their Prevention: Public Attitudes in California," in Room, R. and Sheffield, S., eds. The Prevention of Alcohol Problems (Sacramento: California Office of Alcoholism, 1976), pp. 354-403.
9. Maril, R.E. and Zavaleta, A.N. "Drinking Patterns of Low-Income Mexican American Women," J Stud Alcohol 40(5):480-484, 1979.
10. de Rios, M.D. and Feldman, D.J. "Southern California Mexican American Drinking Patterns: Some Preliminary Observations," J Psychedelic Drugs 9(2):151-158, 1977.

Program Backup Speakers

Vital Statistics Methodology and Technology



PLACING PARENTAL OCCUPATION AND INDUSTRY ON THE BIRTH CERTIFICATE: THE NEW HAMPSHIRE EXPERIENCE

Betty L. DeAngelis, Vital Records and Health Statistics, New Hampshire

The demand for reliable maternal and paternal occupation-industry information increases as investigators continue to probe the linkage between parental exposure to work place hazards and such adverse pregnancy outcomes as fetal damage, low birth-weight and congenital malformations. Concern about the reproductive health of unprecedented numbers of women in the work force has heightened this demand and yet there is a dearth of parental occupation industry data that can be applied in analyses of pregnancy outcomes. One potential instrument for obtaining such data is the birth certificate which already contains information about pregnancy outcome, for example, birth-weight and the existence of congenital malformations. There is obvious utility in collecting parental occupation-industry in concert with outcome data items. One document can be used to capture key variables for specifying associations between work in a variety of occupations-industries and pregnancy outcomes, for detecting possible excess risk in connection with various occupations-industries and for monitoring trends in adverse outcomes as they relate to the workplace. With the addition of parental work items, the birth certificate parallels the death certificate in data function, providing occupation-industry information in conjunction with medical data.

Although the potential value of the data in question is widely acknowledged, a minority of states currently collects them through the medium of the birth certificate. At the national level, the 1980 National Natality Survey and National Fetal Mortality Survey contained parental occupation and type of industry items, but there is as yet, no mechanism for collection of these on a systematic, periodic basis nationwide.

In New Hampshire as elsewhere, we have been aware of the potential significance of the information and felt that eventually we would move to collect it. In January, 1985, with the support of the state's new epidemiologist, we made the decision to place occupation and industry items on the birth certificate. This decision was not taken lightly. While we felt that in the abstract few would dispute the need for the data, there were practical issues to consider in adding more detail to a certificate already crowded with items. First, we had to determine what specifically we would ask about occupation and industry. Ideally, complete parental occupation-type of industry histories would be most desirable from an analytical perspective but the taking of extended histories within the context of obtaining birth information is not practical. After assessing existing state certificate occupation-industry items and in consultation with our state epidemiologist, we decided upon the following:

- Mother and Father - Usual Occupation and kind of Business or Industry
- Mother and Father - Occupation and Kind of Business or Industry over the past 12 months

These items are intended to provide indicators for both extended occupational and industrial exposure and exposure during the critical period of pregnancy in the event of an occupation-industry shift at that time. Of course, they also reflect a compromise between the ideal of obtaining complete information and the practicalities of what is feasible when using the birth certificate as the collection instrument.

Having specified the data items for the certificate, we turned our attention to the key persons who would bear the heaviest burden in acquiring the new information, hospital personnel. In New Hampshire, the hospital staff member that has over-all responsibility for completion and submission of birth certificates is the hospital's director of medical records. (Medical record directors in this state also have major responsibilities in connection with the completion of death certificates and in providing hospital discharge data to the state).

It is trite but worth repeating that state success in collecting vital events data depends greatly on many individuals who are located far from the central state office and who are in contact with informants. If these individuals do not understand the uses and importance of information sought by the state, collected data are likely to be characterized by high rates of inaccuracy and incompleteness. In New Hampshire, we operate a continuous field program to advise and consult with those who are in effect, our "collection agents" and in this context, we held a workshop for medical record directors to discuss the new birth data items, prior to the introduction of a revised birth certificate.

Because of his high interest in the acquisition of the data we asked our state epidemiologist (an M.D.) to open the workshop with an address to the participants explaining the need for the data and indicating potential epidemiological usage. This discourse was well-received and was followed by detailed instructional information on completion of the new certificate items from our agency's supervisor of the vitals registration section which receives all vitals documents for manual processing. Oral presentations at the workshop were completed with an address by the state registrar.

We provided each workshop participant with a supply of the newly revised birth certificate and a manual "Guidelines for Reporting Occupation and Industry on the New Hampshire Birth Certificate", prepared by the agency as reference material for hospital personnel. (For the manual, we drew heavily from existing NCHS publications explaining the reporting of occupation and industry on death certificates). In addition, we indicated

that we were prepared to supply hospital personnel with copies of the U.S.Census Bureau Alphabetical Index of Industries and Occupations which can be used as a guide to acceptable, detailed occupation and industry titles. The medical record directors were receptive to this idea and we distributed this publication a few months later. Based on our experiences with funeral directors in processing death certificate occupation-industry, we knew that hospital personnel would profit from readily available reference information when reporting the data.

Approximately one month subsequent to the hospital medical record directors' workshop, we met with the state's town clerks who are the local registrars of vital events in New Hampshire. Informing them of the birth certificate change, we stressed their role as intermediaries who would be responsible for scanning certificate entries for completeness and where feasible, for clarity, upon receipt of documents from the hospital and prior to submission of copies for the state. The clerks were supplied with the guidelines manual and printed material explaining the significance of the new data items.

To allow for adaptation to the collection of additional birth certificate data, we did not institute mandatory reporting immediately. Instead, we scheduled a phase-in period of three months to give medical record directors time to instruct their personnel, to practice collection and to observe any particular problems which might arise in connection with obtaining the new data. During that three month period, we received no indication from hospital personnel that they encountered any major problems in asking parents to provide the information.

In our state offices we were able to initiate processing of the additional certificate data without delay even though we had only one part-time, untrained coder available for the job at the outset. We have two experienced death certificate occupation and industry coders who gave the new coder some basic in-house training and an assist in getting the coding started. Of course, they also provided guidance whenever coding questions arose. Our part-time birth coder subsequently received formal training at a Cincinnati workshop conducted by the NIOSH and the NCHS jointly.

At this juncture, the phase-in period is over, the reporting of parental occupation and industry is mandatory and we have a fair notion of some of the problems we face in obtaining accurate data.

Our coder reports that industry entries are more likely than occupation entries to be deficient in the detail required for proper classification. The term "textile mill" for example must be further qualified (cotton cloth, woolen cloth et.al.) for precise code assignment. We also find that the names of firms are sometimes furnished in lieu of type of industry. These problem entries indicate that hospital interviewers of parents need to probe a bit more to acquire detail and our personnel have provided informal feedback to hospitals on this issue. They are also comparable to inadequate entries we have experienced in death certificate coding. In the case of the death certificate, some inadequate entries must be tolerated. The funeral director working in a difficult situation may not be in a position to pursue details, nor is adequate information available in many instances where a decedent was elderly and/or without surviving relatives. Birth certificate information however, is obtained under favorable circumstances usually and we feel that hospital personnel can tactfully pursue detail without offending the sensibilities of informants. Indeed, hospital personnel report that parents are quite cooperative on the whole. Admittedly, we have also received reports of several instances where parents have shown great reluctance to provide information.

To date, we feel that implementation of the birth certificate collection of parental occupation-industry has gone well. As a matter of record, we have received no blanks for the new data items and "refused" has appeared just once in approximately 5,000 certificates. We have received no complaints or reports of major difficulties from medical record directors or other hospital personnel. Informal contact with town clerks appears to indicate that they are performing their roles as expected: in a few cases they have asked hospitals for data clarification before sending information on to the state. While some certificate entries for type of industry lack needed detail, our personnel provide feedback to hospitals to improve this situation. We believe that the lack of major problems thus far, is largely attributable to the use of our standard field program approach which involves:

- consultation with principals prior to changes in data requirements and registration procedures
- workshops for the exchange of information which are scheduled in advance of the formal implementation of changes
- the early provision of instructional and reference material to reporting agents
- the periodic distribution of newsletters and special "bulletins" which inform and advise system participants about registration issues and policies.

In-house, we were able to incorporate the additional data into our processing procedures without delay by relying on the existing resource represented by our experienced death certificate occupation-industry coders who trained our new birth coder in the basics. Ordinarily, we would have had to defer processing until personnel were formally trained at federally sponsored workshops or seminars.

We continue of course, to maintain contact with hospital personnel regarding certificate entry problems. For the long run, we will consider the possibility that those personnel might under-

take the coding of uncomplicated occupation and industry entries while reserving more difficult cases for experienced state coders.

The collection of parental occupation and industry data will make a difference in New Hampshire. Our state epidemiologist will apply the data in a statewide occupational-industrial hazard surveillance/monitoring program and the data will be added to the records maintained in our linked infant death-birth file, providing another important variable for analysis in relation to infant death.

HEALING STATISTICS

Warren Schonfeld,
New College of California

As back-up speaker I have assumed certain liberties and constraints that, although not officially mandated, have guided my presentation. Not knowing when I would speak, or to which session, I have worked to make my remarks relevant to most of us here, assuming that we are health statisticians or at least users of health data. This paper is particularly important for those who feel that the goal of their work is just to get the facts. Since I am speaking on the last day of the conference, I have had the opportunity to weave what others have said into the substance of my presentation. In short, I am attempting to connect what I am doing with what you are doing. With this intention I refer more often to other people's research than to my own.

There have been many good uses of statistics and research findings presented at this conference. I need not repeat what is being done. Dr. Manning Feinleib and other plenary session speakers, as well as individual presenters, have clearly illustrated the value of health statistics for giving us information on which to make policy and administrative decisions and for influencing the public as a group, at Federal, state, and local levels.

I want to talk about something different: how to influence directly the individual people from whom we obtain data at the very time of their participation in the process. After all, individual people are the smallest level of geopolitical organization. They are already affected by our interaction with them. How can we consciously increase the health-promoting impact upon them of this interaction? The answer lies in our intention.

Intention directs our thoughts, our behavior, and the outcomes we achieve. It is my intention to suggest a more direct role of health statistics and research to promote healing which brings me here today. It is your intention to hear what I say which has brought us together.

When we do research, our intention sets forth the design and defines the context within which we observe and interpret results. In this sense intention clearly affects what we find, although we sometimes stumble onto things not consciously intended. When we design statistical systems, it is our intention and intended use of data which we translate into the operating specifications.

It is a myth to think that we, as scientists, and statisticians, and policy makers, can keep ourselves separate from the data we collect, that our research and systems are objective. We cannot be objective unless we have no objective because our intention, and our assumptions, so much set our directions that what we find is almost always interpreted in that light. It's a little like the joke about the person who has lost his keys way over there but is looking under the lamp post where the light is better!

Yet this conclusion, or is it merely an assumption, does not detract from the potential of research or the use of statistics. It adds by opening up a whole new area that we may have avoided in the past under the assumption that it was taboo--the intentional and direct use of our statistical systems to accomplish our objectives. As Karl Yordy phrased it during the opening session Tuesday morning, the use of statistics is an important means to policy ends.

One of our assumptions as health statisticians is--and please recognize that what I really mean here is that one of my assumptions was--that our job is primarily to use the "objective" data provided by our research and statistical systems to develop information from which decisions can be made and actions taken to improve the health of people. There is nothing wrong with this. However, accumulating information itself is not particularly useful. To know the cause of a problem we want to change is not the most important thing; it is making the change which is important. Our intention can make that change. To develop statistical systems to find causes so we can make changes is less efficient than if the systems themselves can bring about the changes more directly.

We must recognize the power of intention. Here I will cite three different examples of that power and then proceed to be more specific about what this means in a practical way for us as health statisticians.

1) The placebo effect

The power of intention and belief has long been recognized in health under the label "placebo". The existence of the placebo effect prompted the development of a classical and well-accepted research design, the controlled clinical trial, in which treatments being tested are compared with a control group. Only recently, however, has it come to be appreciated that this effect, rather than being a methodological nuisance, is an effective healing force.[1]

We can utilize the power of belief to promote healing while doing research,

without jeopardizing the logic and control of the clinical trial, by adopting a slight variation in our experimental techniques appropriate when two or more approaches to healing are being tested. Rather than being assigned at the outset randomly to treatment groups, participants in the research can be informed about the different methods to be used and given a choice as to which group they will enter. For those who have no preference, randomization can be used; these participants will furnish the information required for classical clinical trial analysis. For those with a preference, the combined impact of choice and specific treatment can be assessed. This design creates a research model that reflects more realistically how people actually go about healing in life situations. It has already been suggested for use in studying what approaches can help people with epilepsy control seizures,[2] and it may be appropriate in providing people with AIDS actual care while simultaneously exploring the relative merits of medical and non-medical alternatives to healing AIDS.[3]

2) Healing effects of intention

A landmark research study, funded by the Division of Nursing of the U.S. Department of Health and Human Services, is documenting the power of intention to heal. Dr. Janet Quinn, head of research at the University of South Carolina School of Nursing, has been verifying the value of therapeutic touch using measurable effects in a controlled clinical trial setting.[4] The treatment group receives therapeutic touch, including the intention to heal, from trained nurses; the control group receives the same observable treatment from trained nurses, but without the intention to heal. The nurses' intention to heal results in reduction in stress among people receiving therapeutic touch; the control group does not experience the same effect.

3) Effects of intention on physical systems

There is also documented evidence that intention can affect physical systems as well as biological systems. Furthermore, those who are producing effects through their intention do not have to be trained and can use different personal approaches to achieve results. These conclusions are based on evidence accumulated at Princeton University, Department of Engineering, where experimental subjects are seemingly affecting the outcome of a random event generated by both electronic and physical apparatus.[5] According to Dr. Roger Nelson,[6] in testing the null hypothesis that the outcomes of the experiment are the results of a random process, the experimenters have observed results with a P-value less than .0001. Since this finding is derived from a large data base accumulated over six years, one which includes all subjects tested and which combines subjects whose performance could reasonably be expected by chance with those whose performance cannot reasonably be explained by chance alone, the P-value is even more striking. And the results in this carefully controlled experiment are correlated specifically with the intention of the subjects to affect the process!

Now, what does this mean for us?

We have worked hard to make our profession a science, using the power of logic in the development of our methodology. Ours is a useful technology, of which we can be proud, and yet we have a professional blind spot. In my opinion we often operate in practice with a limiting assumption, even if this is never made explicit. And this assumption is that through our logical analysis we can take into account all that is relevant to health and appropriate for health statistics. We assume we have no blind spot, that our findings are not affected by our intentions and assumptions, even those assumptions of which we are unaware.

I have some experimental survey results which may make this point clearer. These results are part of a larger investigation of the role of belief systems on healing which we are exploring at New College of California. Since I also teach within the Health Education Department at San Francisco State University (SFSU) and at the McLaren College of Business at the University of San Francisco (USF), parts of the study have involved students there as well.

The most relevant information comes from an economics class of 35 undergraduate business students at USF. At the beginning of class I distributed a ten-question questionnaire about health beliefs to the class. There were two versions of the questionnaire, which differed only in the response categories allowed in Questions 9 and 10, shown below. The questionnaires were physically mixed and distributed so that, in essence, the class was randomly divided into two groups, with 18 students in Group One receiving version one of the questionnaire and 17 students in Group Two receiving version two.

Here is Question 9 along with the two versions of the response categories:

Question 9

If someone has a sore throat and gargles with warm salt water, vinegar, and honey each morning for a week, and the sore throat goes away, what is

responsible for this improvement? (Check one answer in the space provided.)

Response Categories in Version One
___ the salt
___ the vinegar
___ the honey
___ all of the above
___ the warmth
___ the act of gargling
___ all of the above
___ I don't know
___ I don't know, but probably the person was taking some medicine
___ I don't know, but probably the person would have gotten better anyway
___ there is no way to tell what helped

Response Categories in Version Two
___ the salt
___ the vinegar
___ the honey
___ all of the above
___ the warmth
___ the act of gargling
___ all of the above
___ Other (please indicate: _____)

Group One, explicitly presented with several variations of the "I don't know" response categories as possible answers to Question 9, had 6 students list a specific cause of the improvement and 12 students select "I don't know" responses. Group Two, presented only with specific causes and an "Other" category as possible responses, had 14 students list specific causes and only 3 select the "Other" category, with one answer of "all", one "don't know", and one "time and the immune system".

Allowing for your own interpretation of these results, I will only suggest that much of our research which assumes that there are specific causes responsible for improvement in health may show such results only because we have forced ourselves to eliminate the "I don't know" categories. At least these 35 students were nowhere near as certain that a specific agent was responsible for health improvement as research using version two of the questionnaire might have shown.

Does this mean that in the minds, and maybe in the reality, of lay people the specific agent of cure is less important and less definite than it is to us health statisticians, or that there is less belief that there is a specific agent of cure, or that they are just not as knowledgeable? Please indulge my speculation. I realize that I have only presented results based on one sample of size 35 divided into two groups.

But the point I am making is that there may be a danger in forcing people in the direction of so-called factual health knowledge when they may not really believe it at all, or when it may not be real for them. Even more dangerous is the belief developed by professionals, after much experience and conditioning, that we can have answers that are objective, that are true, and which by implication fault other people for beliefs, choices, and behavior which seem irrational or contrary to those which the data suggest. In truth, facts show variability; and individual differences in belief, as well as freedom in choice of actions based upon those beliefs, may be as important in healing as conformance to what the "facts" show to be true in general or on the average.

Any applied statistician knows that the way questions are asked and the way respondents are prompted, requested, or allowed to answer questions affects the specific results and interpretation of almost any data collection activity. What are the "facts"?

Question 10 asked people to indicate "What is our greatest source of health?"; several possible answers were provided with instructions for one answer to be checked. The two versions of the questionnaire differed in that version two omitted the last response category which was listed in version one and replaced it with the category "Other", as shown below:

Question 10
What is our greatest source of health? (Check one answer.)

Response Categories in Version One
___ moderation and balance
___ nutrition and diet
___ getting exercise and fresh air
___ a clean, safe, and nourishing environment
___ public health and sanitation
___ medicine, health care, and alternative health care practices
___ modern science and the knowledge we have gained from science
___ the truth and splendor of our being

Response Categories in Version Two
___ moderation and balance
___ nutrition and diet
___ getting exercise and fresh air
___ a clean, safe, and nourishing environment
___ public health and sanitation
___ medicine, health care, and alternative health care practices
___ modern science and the knowledge we have gained from science
___ Other (please indicate: _____)

Out of 18 business students responding to version one of the questionnaire, 2 selected "the truth and splendor of our being" as our greatest source of health. None of the 17 students responding to

version two, in which this response category was eliminated, specified the answer "Other". I recognize that the difference in the proportions 2 out of 18 and 0 out of 17 is not statistically significant,[7] nor are the differences of any real practical significance, except in one aspect: although it is not my point to elaborate on my interpretation of the meaning of "the truth and splendor of our being", I will suggest that it represents a dimension, a source of health, different from the other response categories--one which would have seemed to be nonexistent if version two of the questionnaire were the only form used. Could lay people have different views of health than we currently emphasize in our officially sanctioned health statistics activities?[8]

And here's my point: if health is affected by some factors beyond those that we understand scientifically, rationally, or even if many people believe this to be the case, are we not missing the whole picture when we ignore these factors and their impact in doing health research and operating statistical systems? I suggest that the power of intention is one of those factors which affects outcomes and yet is often overlooked, disregarded, discredited, or defined as not being within the legitimate framework of the scientific or rational, even though in actuality much of what we do stems initially from our intention.

Has our intention to understand things rationally, to answer questions objectively, insidiously taken priority over our intention to heal? Have we forgotten that as our knowledge grows the things that we do not now understand or that currently seem irrational may soon be understood and considered rational? Wisdom suggests that we acknowledge such possibilities now rather than waiting until later to use the healing power of intention. Have we become so enamored with our statistics, our ideology, our technology, and so focused on a cautious, conservative, mainstream accepted viewpoint, that we have lost sight of, and hence undermined, the broader intention to heal by using our human energies in as effective ways as possible?

This issue touches us as health statisticians directly in at least two ways. First, we must recognize the power of intention as a healing force so that we can look at it in our research and statistical systems. There is energy in looking at something, and the process of observing imparts some interactive effect on what is observed.[9] The very act of studying intention in healing lends the kind of attention, credibility, and resources which can amplify its effect so we can utilize it more effectively as a source of healing. Second, we must appreciate that the intention of our activities and statistical investigations affects what we find and accomplish.

More specifically, how can we use intention to tap into the greater potential of health research and statistical systems to promote healing? Here are seven ideas.

1) INTEND:
Consciously and explicitly acknowledge the intention to promote healing. This is most important.

2) RESPECT:
Respect the beliefs and realities of individuals participating in health research or providing the data collected by statistical systems more than the assumption that we as professional health statisticians know better.

3) EMPOWER:
Incorporate into health research and statistical systems validation, even encouragement, of the individual's own ability to be healthy and promote healing rather than suggesting that the individual is less able to be healthy without some form of external help. For example, where possible in clinical trial research, adopt a design that gives the individual who has a preference some choice in the approach to healing, as described previously in the discussion of the placebo effect. In survey research give attention to people's feelings about themselves and sense of their own health. The work by John Ware using questions about self-perceived health and showing a strong association between what people report themselves and other more "objective" measures of health is a good example of this.[10] When reporting findings, use statistics describing variability as well as statistics describing averages. Just citing averages often suggests a norm which is desirable rather than honoring and validating that personal variation can be as healthy.

4) ILLUMINATE:
Include the full range of our wisdom in our work. Include aspects of health in addition to those easily measured. If we only measure, or use in our judgment, those characteristics that are easily measured, we are ignoring aspects of life we know to exist. There are keys to be found in many places, not just under the lamp post of incandescent light.

Take a multi-dimensional view. For example, length of life is not necessarily the most important characteristic of life just because it is more easily measurable than depth of life, or breadth, or volume, or substance, or meaning, or value, or joy, or quality of

life. Quality of life is a necessary consideration, unless we choose to make it irrelevant by our continued emphasis on quantity to the exclusion of quality.

Expose people to new ideas and opportunities for promoting health. The new "Health Promotion and Disease Prevention Supplement Booklet" of the National Health Interview Survey[11] is a step in this direction; it presents options for health that many respondents may not previously have considered. This educational function of our work is important and can be developed. Recent data from the National Center for Health Statistics confirms that education is a key correlate of health.[12]

5) INVOLVE:
When possible, provide feedback of research results or data collected to the participants in understandable summary form. This facilitates communication and opens up the possibility of more direct use of our statistical systems to promote healing. Just as biofeedback has been shown to be a useful technique, our systems of biostatistical feedback can be used to promote health education, healthy behavior, and health. All people returning questionnaires in the New College, SFSU, and USF studies have been asked if they want to receive copies of the findings. Involve participants, at least in the pilot study or developmental stage of statistical systems, by asking for their comments and suggestions.

6) RECONSIDER:
What appear from one viewpoint to be problems may in fact be opportunities when approached from a broader or different perspective. This suggests that we reconsider our objectives.

An example will illustrate this idea. In an article about data requirements for measuring our progress towards achieving health promotion and disease prevention goals,[13] the authors mention an inherent methodological problem of follow-up surveys: that the answers which people give to questions on the follow-up survey may not be representative of the general public since respondents may be influenced by their participation in the initial round of data collection. In other words, the exposure to the first questionnaire may make them different from those not so exposed. If our objective is to make unbiased estimates for the population, this is a problem. If our objective is to use the statistical system directly to promote health, this phenomenon may be wonderful; the methodological problem becomes a useful effect which might be enhanced in a similar way that we might capitalize on the placebo effect for healing in clinical trial research. It is reasonable that we may want to strive for both objectives.

7) EXPERIMENT:
Be willing to explore new areas, use new ideas, and experiment with new approaches. Use multiple measures of the variables and characteristics deemed important. By approaching questions from different viewpoints, our fields of vision may overlap sufficiently to see past any blind spots inherent in a single perspective. For example, it may be more productive to think about AIDS as some combination of chronic, acute, and self-selected conditions rather than as merely an infectious disease; this suggestion was made during the discussion following Dr. Alan Kristal's presentation Wednesday afternoon.

At the same time as we invest effort in developing more technically correct systems, we must create new protocols. So as not to lock ourselves into the self-limiting aspects of our current systems and assumptions, we must dare to risk, to do some things beyond or outside what we all agree at present is technically correct. On an experimental and carefully observed conscious basis at least, we need to violate the principles and assumptions of existing systems; otherwise, all our findings will be dependent upon those principles and assumptions. It was Karl Yordy again who suggested we need to maintain a balance between existing systems and new approaches appropriate as our world changes. In that same Tuesday morning session, Dr. Manning Feinleib proudly described the addition of new areas and methods of investigation to the operating systems of the National Center for Health Statistics.

We must open our minds not to insist that health, or the benefits of health statistics systems, can only come through what we now consider to be rational means. Expand our minds, redefining the logic of health statistics to include the very rational idea that, if the ultimate intention of our systems is healing, then a more explicit intention and design of our statistical systems to bring about healing is certainly in our greater interests and will be forthcoming. If we cannot embrace all those things that we do not understand rationally or that we have not personally experienced as real, at least we can hold open the possibility of their existence rather than systematically excluding them from our science. Begin as scientists by entertaining the hypothesis that exercising our informed choices and developing a nurturing belief system can be as powerful forces as physical exercise in promoting health, however defined.

Where we put our intention and the

questions we ask are as important, if not more so, than the answers we get because they provide the initial direction for our energy. We might even dare to stop looking for answers as if they are so important. It might be more interesting to look at the questions we are asking, the ones we seem to need to ask. They can tell a lot about our world view, those assumptions so subtle they are often hidden from us.

As an example, consider health research, whether experimental or survey, looking at the question: How can we make something, like health, better? We are constantly making comparisons involving the concepts of good, better, and best; that is one reason we need measurement, or at least one way we use it. We are searching and researching for remedies, treatments, medicines, cures, techniques that work better. There is a certain kind of familiar and comfortable logic in this process, but is it really progress? If we take a larger perspective, being caught in the rational mentality of looking for better may not be better, may not make us better, may in fact be creating new disease, at least dissatisfaction, with what is. If you felt there was something better to be doing with the last fifteen minutes of your time rather than hearing what I have said, you might feel worse than if I didn't even suggest the comparison.

In our search, or research, for better we may lose sight of what is really powerful: that the way we look at things and our intentions determine what we find as much as what is actually there externally. This insight must balance our interpretation of what we see by looking out.

Many of the seven suggested Ideas for Renewing our Expectations (or I.R.E.'s) about the healing potential of statistics are equally as appropriate, in their general intent and content, for providers of health care. They also express in a positive way, one which capitalizes upon the ire of people who may consider themselves to be nothing more than a vital statistic in today's world, some of the negative aspects of our profession which often leave others, and ourselves, frustrated, angry, or disillusioned about the human limitations of our rational mind and technological culture.

I offer these as thoughts for healing statistics. Those of you with comments, questions, or suggestions please write to me at New College of California, 777 Valencia Street, San Francisco, CA 94110.

---

[1] Brendan O'Regan, "Placebo: The Hidden Asset in Healing", Investigations, Volume 2, Number 1, Research Bulletin of the Institute of Noetic Sciences, 475 Gate Five Road, Suite 300, Sausalito, CA 94965.

[2] Warren H. Schonfeld, "A New Design for Healing Research in Epilepsy", Paper Prepared for the International Epilepsy Congress, September 7, 1985, Hamburg, Germany.

[3] Warren H. Schonfeld, Letter and Draft Proposal, "Research on Methods of Healing AIDS", Addressed to the San Francisco AIDS Alternative Healing Project, August 28, 1985.

[4] Janet F. Quinn, An Investigation of the Effects of Therapeutic Touch, Done without Physical Contact, on State Anxiety of Hospitalized Cardiovascular Patients, Unpublished Doctoral Dissertation, 1982.

[5] R.G. Jahn, B.J. Dunne, and R.D. Nelson, "Princeton Engineering Anomalies Research", Technical Note PEAR 84002, Princeton University, School of Engineering/Applied Science, November, 1983.

[6] Roger D. Nelson, Presentation to the Third Annual Retreat of the Raphaelite Institute, Marriottsville, MD, May, 1985.

[7] There are additional data available from another survey in which the proportional differences are much greater: 4 out of 13 selected "the truth and splendor of our being" when responding to version one of the questionnaire and 0 out of 7 suggesting this answer in their responses to version two. This was a survey mailed to people who had expressed interest in becoming a student in the Integrated Health Studies program at New College. A total of 75 prospective students were randomly divided into two groups, with 38 receiving version one and 37 receiving version two of the questionnaire. Here the response rates were much too low to reach any valid conclusions.

[8] Shirley C. Laffrey, "Development of a Health Conception Scale", Research in Nursing and Health, July, 1984.

[9] Fritjof Capra, The Tao of Physics, Shambhala Publications, Boulder, CO, 1975.

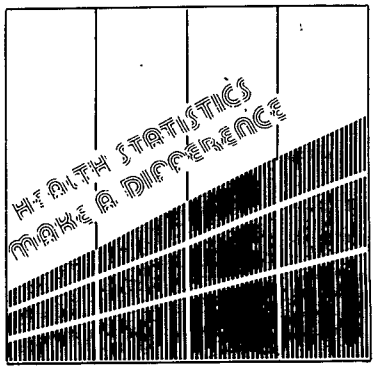
[10] J.E. Ware, Jr., R.H. Brook, A.R. Davies, et al., "Choosing Measures of Health Status for Individuals in General Populations", AJPH 71, 620-625, 1981.

[11] "Health Promotion and Disease Prevention Supplement Booklet", National Health Interview Survey, Form HIS-1(SB)(1985), U.S. Public Health Service, Washington, D.C.

[12] "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, DHHS Pub. No. (PHS) 81-1162, National Center for Health Statistics, Hyattsville, MD, June, 1981.

[13] Lawrence W. Green, Ronald W. Wilson, and Katherine G. Bauer, "Data Requirements to Measure Progress on the Objectives for the Nation in Health Promotion and Disease Prevention", AJPH 73: 1, 18-24, January, 1983.

Appendix



PROGRAM SESSION ORGANIZERS

Martin A. Baum, Ph.D. (H)
Deputy Associate Director
Office of Data Processing and Services*

Robert J. Casady, Ph.D. (M)
Chief, Statistical Methods Staff
Office of Research and Methodology*

Bruce B. Cohen, Ph.D. (G)
Health Status Measurement Branch
Division of Epidemiology & Health Promotion
Office of Analysis and Epidemiology*

Lester R. Curtin, Ph.D. (T)
Statistical Methods Staff
Office of Research and Methodology*

Robert Fuchsberg (V)
Director, Division of Health Interview Survey
Office of Interview and Examination
Statistics Program*

George A. Gay (F)
Chief, Registration Methods Branch
Division of Vital Statistics
Office of Vital and Health Care Statistics*

Kenneth W. Harris (L)
Special Assistant for Program
Coordination and Statistical Standards
Office of Research and Methodology*

Robert L. Heuser (K)
Chief, Natality Statistics Branch
Division of Vital Statistics
Office of Vital and Health Care Statistics*

Peter Hurley (1st, 2nd, 3rd Plenary)
Associate Director
Office of Vital and Health Care Statistics*

Deborah Ingram, Ph.D. (U)
Analytical Coordination Branch
Division of Analysis
Office of Analysis and Epidemiology*

James C. Jacks, Ph.D. (I)
Director, Division of Data Services
Office of Data Processing and Services*

Joel C. Kleinman, PH.D. (X)
Director, Office of Analysis
Office of Analysis and Epidemiology*

Mary Grace Kovar, Dr.P.H. (Q)
Special Assistant for Data Policy
Office of Interview and Examination Statistics*

Alan K. Kreger (P)
Chief, Computer Users Staff
Office of Data Processing and Services*

Garrie J. Losee (S)
Associate Director
Office of Data Processing and Services*

Jennifer Madans, Ph.D. (N)
Deputy Director, Division of Analysis
Office of Analysis and Epidemiology*

Diane Makuc, Dr.P.H. (C)
Chief, Analytical Coordination Branch
Division of Analysis
Office of Analysis and Epidemiology*

Harry M. Rosenberg, Ph.D. (O)
Chief, Mortality Statistics Branch
Division of Vital Statistics
Office of Vital and Health Care Statistics*

George A. Schnack (E)
Director, Division of Data Processing
Office of Data Processing and Services*

James A. Smith (Special Sessions 3, 5)
Program Development Staff
Office of Program Planning, Evaluation
and Coordination*

Sandra Surber Smith (A)
Chief, Scientific and Technical
Information Branch
Division of Data Services
Office of Data Processing and Services*

William F. Stewart (Special Sessions 1, 2, 4)
Chief, Program Development Staff
Office of Program Planning, Evaluation
and Coordination*

Owen Thornberry, Ph.D. (W)
Deputy Director
Division of Health Interview Survey
Office of Interview and Examination Statistics*

James A. Weed, Ph.D. (R)
Deputy Director
Division of Vital Statistics
Office of Vital and Health Care Statistics*

Gerald Wheeler (B)
Division of Health Examination Statistics
Office of Interview and Examination Statistics*

P. Douglas Williams (J)
Chief, Data Applications and Research Staff
Office of Interview and Examination Statistics*

Ronald Wilson (Special Session 2)
Director, Division of Epidemiology
and Health Promotion
Acting Chief, Environmental Studies Branch
Office of Analysis and Epidemiology*

Andrew White, Ph.D. (D)
Survey Design Staff
Office of Research and Methodology*

* National Center for Health Statistics
3700 East-West Highway
Hyattsville, MD 20782

( ) Indicates Session

PROGRAM SESSION CHAIRPERSONS

Richard M. Biery, M.D., M.S.P.H.
(Third Plenary)
Director
Kansas City Department of Health
10th Floor City Hall
414 East 12th Street
Kansas City, Missouri 64106

Linda T. Bilheimer, Ph.D. (O)
Director
Division of Health Statistics & Epidemiology
Arkansas Department of Public Health
4815 West Markham Street
Little Rock, Arkansas 72205-3867

Jacob A. Brody, M.D. (Q)
Dean
University of Illinois at Chicago
School of Public Health
Box 6998
Chicago, Illinois 60680

Norbert W. Budde, Ph.D.
(Special Session 2)
Director
Division of Survey and Data Resources
American Medical Association
535 North Dearborn Street
Chicago, Illinois 60610

Ann Coulson, F.A.C.E. (A)
Lecturer and Research Epidemiologist
Division of Epidemiology, CHS 71-236
UCLA, School of Public Health
Los Angeles, California 90024

Richard J. David, M.D. (K)
Assistant Professor of Pediatrics
Division of Neonatology
Northwestern University Medical School
2300 Children's Plaza
Chicago, Illinois 60614

Manning Feinleib, M.D., Dr.P.H.
(First Plenary)
Director
National Center for Health Statistics
3700 East West Highway
Hyattsville, Maryland 20782

Gail F. Fisher, Ph.D.
(Special Session 4)
Executive Secretary
National Committee on Vital
and Health Statistics
3700 East-West Highway
Hyattsville, Maryland 20782

( ) Indicates Session

Floyd J. Fowler, Jr., Ph.D. (W)
Senior Research Fellow
Center for Survey Research
University of Massachusetts
100 Arlington Street
Boston, Massachusetts 02116

Claude Earl Fox, M.D., M.P.H. (F)
Deputy State Health Officer
Mississippi State Department of Health
P. O. Box 1700
Jackson, Mississippi 39215-1700

Marsha Gold, Sc.D. (N)
Director
Policy Analysis and Program Evaluation
Maryland Dept. of Health & Mental Hygiene
201 W. Preston Street
Baltimore, Maryland 21201

Robert S. Gold, Ph.D., Dr.P.H. (U)
Project Director
School Health Initiative
Office of Disease Prevention & Health
Promotion
330 C. Street, S.W., Rm. 2132
Washington, D.C. 20201

Merwyn R. Greenlick, Ph.D. (J)
Vice President (Research)
Kaiser Foundation Hospitals and
Director, Center for Research
4610 S. E. Belmont
Portland, Oregon 97215

Carol J. Rowland Hogue, Ph.D. (T)
Chief, Pregnancy Epidemiology Branch
Division of Reproductive Health
Centers for Disease Control
1600 Clifton Rd., Bldg. 1, RM. 4419
Atlanta, GA 30333

Woodie Kessel, M.D., M.P.H. (X)
Chief, Research & Training Branch
Division of Maternal & Child Health
Health Resources & Services Administration
5600 Fishers Lane
Rockville, Maryland 20857

Frederick L. King, B.S. (R)
State Registrar
Minnesota Department of Health
717 Delaware Street, S. E.
P.O. Box 9441
Minneapolis, Minnesota 55440

Monroe Lerner, Ph.D. (G)
Professor
School of Hygiene
The Johns Hopkins University
615 N. Wolfe Street
Baltimore, Maryland 21205

Robert Lewis, Ph.D. (S)
Professor
School of Public Health
University of South Carolina
Columbia, South Carolina 29208

Henry Malin (Special Session 5)
Acting Chief, Surveillance Branch
Division of Biometry and Epidemiology
National Institute on Alcohol Abuse
and Alcoholism
5600 Fishers Lane, Rm. 14C26
Rockville, Maryland 20857

Ronald W. Manderscheid, Ph.D.
(Special Session 1)
Chief, Survey and Reports Branch
National Institute of Mental Health
Room 18C-07 Parklawn Building
5600 Fishers Lane
Rockville, Maryland 20857

John Morgan, Ph.D. (I)
Consultant
Code Three, Health Information Systems 3-M
1141 East 3900 South, Suite A160
Salt Lake City, Utah 84124

David Morganstein, M.A. (D)
Vice President
WESTAT
1650 Research Blvd.
Rockville, Maryland 20850

Charles D. Palit, Ph.D. (H)
Professor
University of Wisconsin
110 Lowell Hall
610 Langdon Street
Madison, Wisconsin 53703

John L. Pendleton, M.A. (B)
Health Programs Consultant
4106 Parkedge Lane
Annandale, Virginia 22003

Earl Pollack, Sc.D. (M)
Pollack Associates, Ltd.
7504 Elmore Lane
Bethesda, Maryland 20817

Patricia W. Potrzebowski, Ph.D.
(Second Plenary)
Director
Division of Health Statistics & Research
Pennsylvania Department of Health
P.O. Box 90
Harrisburg, Pennsylvania 17198

H. Denman Scott, M.D., M.P.H. (V)
Director of Health
Rhode Island Dept. of Health
401 Cannon Building
75 Davis Street
Providence, Rhode Island 02908

William D. Spears, M.A. (P)
Faculty Associate
School of Public Health
University of Texas
P. O. Box 20186
Houston, Texas 77225

Howard V. Stambler, B.A.
(Special Session 3)
Director
Office of Data Analysis and Management
Bureau of Health Professions, HRSA
Parklawn Bldg., Rm. 8-41
5600 Fishers Lane
Rockville, Maryland 20857

James C. Story, Ph.D. (L)
Chairman
Dept. of Anatomical Sciences
Meharry Medical College
1005 D. B. Todd Boulevard
Nashville, Tennessee 37208

P. Douglas Williams (E, S)
Chief
Data Applications and Research Staff
Interview and Examination Statistics
Program
National Center for Health Statistics
3700 East West Highway
Hyattsville, Maryland 20782

Pamela Wolf, Dr.P.H. (C)
Biostatistician
National Institute for Child Health
and Human Development
4858 MacArthur Blvd.
Washington, D.C. 20007

Tim E. Aldrich, M.P.H. (L)
Oak Ridge National Laboratory
P.O. Box X H.A.S.R.D.
Oak Ridge, Tennessee 37831

Charles D. Baker (First Plenary)
Under Secretary of Health
and Human Services
Hubert H. Humphrey Building, 614G
200 Independence Avenue SW
Washington, DC 20201

Reginald P. Baker, Ph.D. (H)
Associate Director for Applications
and Systems Development
Information Resources Center
National Opinion Research Center
6030 South Ellis
Chicago, Illinois 60637

Robert H. Barnes, M.D. (Special 4)
Chairman
National Committee on Vital
and Health Statistics
1014 East Blaine Street
Seattle, Washington 98102

Winslow J. Bashe, Jr., M.D., M.P.H. (K)
Professor, Community Medicine
Wright State University School of Medicine
P.O. Box 927
Dayton, Ohio 45401

Rosalyn D. Bass (Special 1)
Chief
Developmental & Special Population
Research Branch
National Institute of Mental Health
Parklawn Building, Room 18C-07
5600 Fishers Lane
Rockville, Maryland 20857

R. Barker Bausell, Ph.D. (V)
Associate Professor
Evaluation and the Health Professions
University of Maryland
655 W. Lombard Street
Baltimore, Maryland 21201

Judith Baxter, Ph.D. (E)
Mortality Programming Director
Division of Epidemiology
University of Minnesota
Stadium Gate 27
Minneapolis, Minnesota 55455

Ronald G. Blankenbaker, M.D. (Special 4)
Vice President, Medical Affairs
St. Vincent Hospital
and Health Care Center
2001 West 86th Street
Indianapolis, Indiana 46260

J. Michael Bowling, M.A. (W)
Statistician
North Carolina State Center
for Health Statistics
Division of Health Services
P.O. Box 2091
Raleigh, North Carolina 27602-2091

Lisa E. Brown, B.A. (X)
Research Analyst
Department of Public Health
Division of Family Health Services
150 Tremont Street, 4th Floor
Boston, Massachusetts 02111

Cynthia E. Burghard, B.A. (P)
Assistant Director
Massachusetts Health Data Consortium, Inc.
400-1 Totten Pond Road
Waltham, Massachusetts 02154

Joseph D. Carney, M.A. (B)
Manager
Oregon Center for Health Statistics
1400 S. W. 5th Avenue
Portland, Oregon 97201

Charles Christian, Ph.D. (Special 5)
113 North Rock Glen Rd. Suite A
Baltimore, Maryland 21229

Jonathan E. Conklin, M.A. (N)
Senior Research Manager
SysteMetrics/McGraw-Hill
104 West Anapamu Street
Santa Barbara, California 93101

Margaret Cooke, S.M. (U)
Statistician
National Center for Health Statistics
3700 East West Highway, Room 2-27
Hyattsville, Maryland 20782

Yahya A. H. Daoud, M.A. (Y)
Research Scientist
Detroit Department of Health
1151 Taylor, Room 210B
Detroit, Michigan 48202

Barry R. Davis, M.D., Ph.D. (H)
Assistant Professor of Biometry
School of Public Health
University of Texas
1200 Herman Pressler Street
Houston, Texas 77030

C. Howard Davis, Ph.D. (Special 3)
Economist
Division of Medicine
Bureau of Health Professions
Room 4C21, Parklawn Building
5600 Fishers Lane
Rockville, Maryland 20857

( ) Indicates Session

Betty DeAngelis (Backup Speaker)
Analyst-Data Systems
Division of Public Health
Health & Welfare Building
Hazen Drive
Concord, New Hampshire 03301

John Drabek, Ph.D. (Special 3)
Chief, Technical Analysis &
Coordination Branch
Bureau of Health Professions
Office of Data Analysis and Management
Parklawn Building, Room 8-41
5600 Fishers Lane
Rockville, Maryland 20857

Mary Dufour, M.D. (Special 5)
Alcohol, Drug Abuse and
Mental Health Administration
Parklawn Building, Rm. 14C26
5600 Fishers Lane
Rockville, Maryland 20857

Mary Ann Eiler, Ph.D. (Special 2)
Assistant Director
Department of Data Release Services
American Medical Association
535 North Dearborn Street
Chicago, Illinois 60610

Manning Feinleib, M.D., Dr.P.H.
(First Plenary)
Director
National Center for Health Statistics
3700 East-West Highway
Hyattsville, Maryland 20782

William R. Felts, Jr., M.D. (Special 4)
Professor of Medicine
George Washington University
Medical Center
2150 Pennsylvania Avenue NW
Washington, DC 20037

Richard G. Frank, Ph.D. (C, G)
Assistant Professor
Health Services Research &
Development Center
The Johns Hopkins University
624 North Broadway
Baltimore, Maryland 21205

Irvin G. Franzen, B.A., M.P.H. (F)
State Registrar
Kansas Department of Health & Environment
Forbes Field, Building 740
Topeka, Kansas 66620

Mary Anne Freedman, M.A. (F)
Director, Public Health Statistics
Vermont Department of Health
P.O. Box 70
Burlington, Vermont 05402

Andrew M. Friede, M.D. (Y)
Epidemic Intelligence Service Officer
Division of Reproductive Health
Centers for Disease Control
1600 Clifton Road NE, Room 4419
Atlanta, Georgia 30333

Alan S. Friedlob, M.S. (I)
Health Services Research Analyst
Health Care Financing Administration
Oak Meadows Building, 2-F-3
6325 Security Boulevard
Baltimore, Maryland 21228

Jay H. Glasser, Ph.D. (L)
Associate Professor
School of Public Health
University of Texas
P.O. Box 20186
Houston, Texas 77025

Marilyn K. Goldhaber, M.P.H. (T)
Environmental Epidemiologist
Medical Methods Research
The Permanente Medical Group, Inc.
3451 Piedmont Avenue
Oakland, California 94611

Harold F. Goldsmith, Ph.D. (J)
Chief, Demographic Research Section
National Institute of Mental Health
5600 Fishers Lane, Room 18-105
Rockville, Maryland 20857

Jeffrey B. Gould, M.D., M.P.H. (A)
Associate Professor and
IPODM Project Director
School of Public Health
University of California
309 Earl Warren Hall
Berkeley, California 94720

Philip L. Graitcer, D.M.D., M.P.H. (A)
Medical Epidemiologist
Centers for Disease Control
1600 Clifton Road NE
Atlanta, Georgia 30333

Paul D. Gunderson, Ph.D. (B)
Director
Minnesota Center for Health Statistics
717 Delaware Street SE
Minneapolis, Minnesota 55440

Shelia K. Hoar, Sc.D. (C)
Senior Staff Fellow
Occupational Studies Section
National Cancer Institute
Landow Building, Room 4C16
Bethesda, Maryland 20205

Gary C. Hogelin, M.P.A. (W)
Chief, Field Services Branch
Centers for Disease Control
Bldg. 3, SB12
1600 Clifton Road NE
Atlanta, Georgia 30333

Embry Martin Howell, M.S.P.H. (P)
Government Program Manager
SysteMetrics/McGraw-Hill
4733 Bethesda Avenue, Suite 600
Bethesda, Maryland 20814

Alan B. Humphrey, Ph.D. (S)
Associate Professor
College of Business Administration
University of Rhode Island
Kingston, Rhode Island 02881

Paul W.C. Johnson, M.S.P.H. (K)
Chief, Statistical Analysis Unit
State Center for Health Statistics
North Carolina Division of
Health Services
P.O. Box 2091
Raleigh, North Carolina 27602

Richard A. Johnson, M.S. (D)
Research Assistant
School of Public Health
University of Texas
1200 Hermann Pressler
Houston, Texas 77025

Rumaldo Z. Juarez, Ph.D. (R)
Associate Professor
Department of Sociology & Social Work
Pan American University
Edinburgh, Texas 78539

Mark S. Kamlet, Ph.D. (M)
Associate Profesor of Economics
and Public Policy
Department of Social Sciences
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

John Kiely, Ph.D. (K)
Assistant Professor
of Clinical Public Health
Columbia University
Sergievsky Center
630 West 168th Street
New York, New York 10032

Harold W. Kohl. M.S.P.H. (H)
Statistician, Epidemiology
Institute for Aerobics Research
12200 Preston Road
Dallas, Texas 75230

Alan R. Kristal, Dr.P.H. (O)
Director
Epidemiologic Surveillance & Statistics
New York City Department of Health
125 Worth Street Room 326
New York, New York 10013

James M. Lepkowski, Ph.D. (M)
Study Director
Institute for Social Research
University of Michigan
P.O. Box 1248
Ann Arbor, Michigan 48105

Robert Lewis, Ph.D. (S)
Professor
School of Public Health
University of South Carolina
Columbia, South Carolina 29208

Paul Lichtenstein, M.H.S. (I)
Health Services Research Analyst
Health Care Financing Administration
6325 Security Boulevard
Baltimore, Maryland 21207

Anthony J. Likovich, Jr., B.A. (N)
Senior Health Care Analyst
SysteMetrics/McGraw Hill
104 West Anapamu Street
Santa Barbara, California 93101

John D. Loft, Ph.D. (Special 2)
Assistant to the Division Director
Division of Survey and Data Resources
American Medical Association
535 North Dearborn Street
Chicago, Illinois 60610

N. Martin MacDowell, Dr.P.H. (V)
Visiting Assistant Professor
Graduate Program in Hospital and
Health Administration
2700 Steindler Building
University of Iowa
Iowa City, Iowa 52242

Sarah S. Marter, M.P.H. (D)
Lead Analyst
The Health Data Institute, Inc.
7 Wells Avenue
Newton, Massachusetts 02159

James W. Masterson, M.P.H. (I)
Director of Operations Research
Chicago Department of Health
55 West Washington, Room CL 99
Chicago, Illinois 60643

Andrew D. McBride, M.D., M.P.H.
(First Plenary)
Commissioner of Public Health
of the District of Columbia
1875 Connecticut Avenue NW, Suite 825
Washington, DC 20009

John M. McQuiston, Ph.D. (M)
Professor and Director
Center for Population Research
Department of Sociology
University of Montana
Missoula, Montana 59812

Allan M. Miller, M.A. (P)
System Administrator/Research Analyst
Inland Counties Health Systems Agency
1600 Chicago Avenue, Suite R-1
Riverside, California 92507

Stephen D. Minton, M.D. (R)
Department of Neonatology
Utah Valley Regional Medical Center
1034 North 5th West, P.O. Box 390
Provo, Utah 84603

Ross M. Mullner, Ph.D. (J)
Director, Dept. of Special Studies
American Hospital Association
840 North Lake Shore Drive
Chicago, Illinois 60611

Donald N. Muse, Ph.D. (Q)
9733 Lake Shore Drive
Gaithersburg, Maryland 20879

Barry E. Nangle, Ph.D. (O)
Information Analyst
Division of Family Health Services
Utah Department of Health
44 Medical Drive
Salt Lake City, Utah 84113

Kristine Olsen Powell, B.A. (A)
Survey Statistician
International Statistical Programs Center
Bureau of the Census
Scuderi 304, ISPC/ESB
Washington, D.C. 20233

Mary K. O'Mara, M.P.H. (C)
Administrator
Birth Defects Monitoring Program
Department of Social and Health Services
Mail Stop ET-14
Olympia, Washington 98504

Jan Ouren, Ph.D. (G)
Research Manager
SysteMetrics/McGraw-Hill
272 Santa Monica Way
Santa Barbara, California 93101

Richard Papel, M.A.(E)
Research Analyst
SysteMetrics/McGraw-Hill
4733 Bethesda Avenue, #600
Bethesda, Maryland 20814 (E)

Christine Pattee, Dr.P.H. (Q)
Acting Chief, Statistical Analysis Unit
Connecticut Department of Health Services
150 Washington Street
Hartford, Connecticut 06106

Edward B. Perrin, Ph.D. (Special NCHS)
Professor/Chair, Health Services, SC-37
School of Public Health
and Community Medicine
University of Washington
Seattle, Washington 98195

Denton R. Peterson, M.S. (V)
Information Systems Coordinator
Division of Disease Prevention
and Health Promotion
Minnesota Department of Health
717 Delaware Street SE
Minneapolis, Minnesota 55454

Frederick Pratter, Ph.D. (E)
Senior Analyst
Abt Associates, Inc.
55 Wheeler Street
Cambridge, Massachusetts 02138

Dorothy P. Rice (Special NCHS)
Professor
Aging Health Policy Center
University of California, N 631 Q
San Francisco, California 94143

Anthony M. Roman, M.A. (W)
Special Assistant
Demographic Surveys Division
Bureau of the Census Rm. 3335
Washington, DC 20233

Marilyn J. Rosenstein (Special 1)
Chief, Sample Survey Research Program
National Institute of Mental Health
Parklawn Building, Room 18C-07
5600 Fishers Lane
Rockville, Maryland 20857

Dennis Ross-Degnan, M.S.P.H. (D)
Director of Research
International Eye Foundation
124 Bourne Street
Roslindale, Massachusetts 02131

Karen Rudzinski, M.A. (Special 3)
Program Analyst
Office of Data Analysis and
Special Projects
Bureau of Health Professions
Parklawn Building, Rm. 4C-21
5600 Fishers Lane
Rockville, Maryland 20857

Gail Rushford (R)
Field Representative
Division of Public Health Statistics
Vermont Department of Health
60 Main Street, P.O. Box 70
Burlington, Vermont 05402

Warren Schonfeld, Ph.D. (Backup)
Coordinator
Integrated Health Studies
New College of California
777 Valencia Street
San Francisco, California 94110

Wayne F. Schramm, M.A. (X)
Director
Bureau of Health Data Analysis
Missouri Division of Health
1730 E. Elm, P.O. Box 570
Jefferson City, Missouri 65102

Lance Segars, M.S. (U)
Research Analyst
San Diego County Department of
Health Services - Alcohol Program
3851 Rosecrans Street
San Diego, California 92110

David J. Sencer, M.D., M.P.H.
(Third Plenary)
Commissioner of Health
City of New York
125 Worth Street
New York, New York 10013

Jack C. Smith, M.S. (X)
Chief, Research & Statistics Branch
Division of Reproductive Health
Centers for Disease Control
1600 Clifton Road NE
Bldg. 1, Rm. 4409
Atlanta, Georgia 30333

David A. Spivack, M.R.P. (Q)
Director of Planning
Mt. Sinai Medical Center
4300 Alton Road
Miami Beach, Florida 33140

Debora L. Stewart, B.G.S. (S)
Kansas Applied Remote
Sensing Program
University of Kansas
Space Technology Center
2291 Irving Hill Drive--Campus West
Lawrence, Kansas 66045-2969

Elliot M. Stone, M.U.A. (B)
Executive Director
Massachusetts Health Data Consortium, Inc.
400-1 Totten Pond Road
Waltham, Massachusetts 02154

Edward J. Tu, Ph.D. (O)
Research Scientist
Bureau of Biostatistics
New York State Health Department
Room 308, Tower Bldg.
Empire State Plaza
Albany, New York 12237

Timothy Tyson, M.S. (G)
Evaluation Specialist
Department of Urban and Regional Planning
University of Wisconsin
925 Lathrop Drive
Madison, Wisconsin 53706

Nicole Urban, Sc.D. (U)
Staff Scientist/Scientific Administrator
Cancer Prevention Research Unit
Fred Hutchinson Cancer Research Center
1124 Columbia Street, Mail Stop W-202
Seattle, Washington 98104

George Van Amburg, M.P.H. (F)
State Registrar and Chief
Office of the State Registrar and
Center for Health Statistics
Michigan Department of Public Health
P.O. Box 30035
Lansing, Michigan 48909

Bailus Walker, Jr., Ph.D., M.P.H.
(Second Plenary)
Commissioner of Public Health
Massachusetts Department of Public Health
150 Tremont Street
Boston, Massachusetts 02111

Mary Lou S. White, B.A. (Special 2)
Director
Survey Design and Analysis
American Medical Association
535 North Dearborn Street
Chicago, Illinois 60610

Jerome Wilson, Ph.D. (L)
Epidemiologist
National Cancer Institute
Landow Building, Room 3A-22
Bethesda, Maryland 20205

Ronald W. Wilson, M.A. (Special 5)
Director
Division of Epidemiology
and Health Promotion
National Center for Health Statistics
3700 East West Highway
Hyattsville, Maryland 20782

Michael Witkin, M.A. (Special 1)
Supervisory Statistician
Survey and Reports Branch
National Institute of Mental Health
Parklawn Building, Room 18C-07
5600 Fishers Lane
Rockville, Maryland 20857

William E. Wright, Ph.D. (J)
Chief, Facility Care Research Unit
Center for Health Statistics
California Department of Health Services
714 P Street, Room 1476
Sacramento, California 95814

Karl D. Yordy (First Plenary)
Senior Program Officer and Director
Division of Health Care Statistics
Institute of Medicine
National Academy of Sciences
2101 Constitution Avenue NW, Rm.317
Washington, DC 20418

1985 PHCRS PLANNING COMMITTEE

National Center for Health Statistics:

Mr. Robert Israel, Chairperson
Mr. Jack Anderson
Dr. Edward Bacon
Mr. E. Earl Bryant
Mr. Kenneth Harris
Dr. Joel Kleinman
Mr. Peter Hurley
Dr. Mary Grace Kovar
Mr. Garrie Losee
Mr. James Smith
Mr. P. Douglas Williams
Mr. Ronald Wilson

EXHIBITS AND PUBLICATIONS SUPPORT STAFF

Linda L. Bean
Geraldine K. Cherry
Jarmila G. Frazier
Annette F. Gaidurgis
Sarah M. Hinkle
John Jetter
Nancy L. Planet
Patricia L. Wilson

STAFF OF THE PHCRS SECRETARIAT

Program Development Staff
Office of Program Planning, Evaluation
and Coordination
National Center for Health Statistics

Mr. James A. Smith, Executive Secretary

Ms. Jacquelyn L. Adler
Ms. Kay Barrett
Ms. Lorraine Beatty
Ms. Doris Scherbak
Mr. William F. Stewart

U. S. GOVERNMENT PRINTING OFFICE: 1986—499-122 REGION NO 8

U.S. DEPARTMENT OF HEALTH AND
HUMAN SERVICES
Public Health Service
National Center for Health Statistics
3700 East-West Highway
Hyattsville, Maryland 20782

DHHS Publication No. (PHS) 86-1214