NATIONAL CENTER For HEALTH STATISTICS | Series 2 Number 13

**VITAL and HEALTH STATISTICS**

**DATA EVALUATION AND METHODS RESEARCH**

# Computer Simulation of Hospital Discharges

Micro-simulation of measurement errors in hospital discharge data reported in the Health Interview Survey.

Washington, D.C. February 1966

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE
John W. Gardner
Secretary

Public Health Service
William H. Stewart
Surgeon General

Public Health Service Publication No. 1000-Series 2-No. 13

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C., 20402 - Price 35 cents

# NATIONAL CENTER FOR HEALTH STATISTICS

FORREST E. LINDER, PH. D., *Director*

THEODORE D. WOOLSEY, *Deputy Director*

OSWALD K. SAGEN, PH. D., *Assistant Director*

WALT R. SIMMONS, M.A., *Statistical Advisor*

ALICE M. WATERHOUSE, M.D., *Medical Advisor*

JAMES E. KELLY, D.D.S., *Dental Advisor*

LOUIS R. STOLCIS, M.A., *Executive Officer*

**OFFICE OF HEALTH STATISTICS ANALYSIS**

IWAO M. MORIYAMA, PH. D., *Chief*

**DIVISION OF VITAL STATISTICS**

ROBERT D. GROVE, PH. D., *Chief*

**DIVISION OF HEALTH INTERVIEW STATISTICS**

PHILIP S. LAWRENCE, SC. D., *Chief*

**DIVISION OF HEALTH RECORDS STATISTICS**

MONROE G. SIRKEN, PH. D., *Chief*

**DIVISION OF HEALTH EXAMINATION STATISTICS**

ARTHUR J. MCDOWELL, *Chief*

**DIVISION OF DATA PROCESSING**

SIDNEY BINDER, *Chief*

**Public Health Service Publication No. 1000-Series 2-No. 13**

*Library of Congress Catalog Card Number 65-62273*

# PREFACE

The purpose of the study described in this report was two-fold: (1) the underlying consideration was methodology, with emphasis on model building and on experience to be gained in the use of computer simulation techniques employed in analysis of health statistics; and (2) the immediate target was a better understanding of the impact of certain measurement deficiencies present in health interview surveys.

The specific problems studied are set forth in sections I and II of the report. The subject matter is hospital discharges, and more especially the discrepancies between the number of discharges as reported by household respondents to interview and those that actually occur. The Health Interview Survey of the National Center for Health Statistics in its household inquiry includes questions asking for the number and characteristics of hospital discharges experienced by household members in the year prior to interview. There are many reasons for discrepancy between the reported number of discharges and the true number. Two of these causes have been given particular attention. One is that hospital experience during the reference period for persons not living at the time of interview is not reported in a survey of living persons. This deficiency is relatively more important the longer the reference period. A second principal cause of discrepancy between reported and true data is the response error in the report for a living person. Empirical data and theory have indicated that this error, too, increases with length of reference period.

The interaction of these factors and their impact on reported data have been explored previously in a variety of ways, using record-check techniques, internal analysis of reported data, and hypothetical models. This research has contributed substantially to better knowledge of the subject but has left several questions unanswered. It seemed likely that understanding would be further promoted, and especially that better judgments could be made of the effect of changes in interview procedure, if the process were to be studied through a technique for simulating on a computer the hospital experience of a model population of individual persons, and subsequently simulating interviews of this population. Such an undertaking might have particular merit since the main threads of logic for the hospital problem might have considerably wider potential application—for example, a close analogy can be made between periods of unemployment and hospital episodes.

Accordingly, through a contractual arrangement the present study was carried out by Research Triangle Institute, Durham, N.C., in close cooperation with staff members of the National Center for Health Statistics. Dr. D. G. Horvitz of the Research Triangle Institute was the project director and principal author of this report. He was assisted by Dr. D. T. Searls, formerly on the Research Triangle Institute staff, and by Irving Drutman (deceased) of North Carolina State University. Mr. Drutman did most of the computer programming. Other contributors to the study were Mr. Joseph Snavely of the North Carolina State University Computing Center and Mr. Francis Giesbrecht of the Research Triangle Institute, who developed appropriate expected values and variances for the computer-generated discharge rates. Walt R. Simmons prepared an initial outline of the problem, proposed the simulation approach, and coordinated contributions of the staff of the Center to the project. Wilbur M. Sartwell of the Center staff supervised much of the computer calculation.

# CONTENTS

*IN THIS REPORT a study is presented on computer micro-simulation of discharges from short-stay hospitals, and on the associated measurement errors that occur in household interview surveys, as set forth in the preface. A synthetic universe of 10,000 persons was established with demographic characteristics similar to those of the U.S. civilian, noninstitutional population. On the basis of earlier theoretical work and empirical record-check studies, this universe was subjected to a series of stochastic operations to simulate hospital experience, and the reporting of that experience in household interviews.*

*Each individual person was moved from one state to another—e.g., from not-in-a-hospital to in-a-hospital, or from in-a-hospital to discharged-alive—by a random process with probabilities which varied by such factors as age, sex, distance from death, number of days already in the hospital, and a general health index. Thus it was possible to count the simulated hospital discharges over a 12-month period, and to tabulate them in a variety of ways.*

*At monthly intervals the living persons in the synthetic population then were "interviewed" by the computer and reported their hospital experience over the previous year. Two sets of simulated interview data were tabulated. In one, respondents reported without error. For this set, comparison with total experience reflected the impact on discharge statistics of the missing data for persons not living at the time of interview. In the other, response was conditioned by probabilities of reporting correctly, which varied by distance between interview and discharge, length of stay, reason for hospitalization, and other less significant factors. Comparison of this latter set of data with total experience gives a mechanism for studying a wide range of problems found in the interview data.*

*Throughout the study, emphasis was placed on the development and use of a flexible method of analysis. The report is not an evaluation of the reporting of hospital discharges in the Health Interview Survey.*

## SYMBOLS

| | |
|---|---|
| Data not available------------------------ | --- |
| Category not applicable------------------- | ... |
| Quantity zero----------------------------- | - |
| Quantity more than 0 but less than 0.05----- | 0.0 |
| Figure does not meet standards of reliability or precision------------------ | * |

# COMPUTER SIMULATION OF HOSPITAL DISCHARGES

## I. INTRODUCTION

The Health Interview Survey of the National Center for Health Statistics provides estimates of the number of discharges from hospitals on an annual basis for the living, civilian, noninstitutional population. The data are gathered in a household interview survey by means of personal interviews conducted each week, during a 52-week period, in area probability samples of households throughout the United States. The information on discharges (along with hospital utilization) is obtained for each resident in the sample households for a reference period of 12 months prior to the week of interview.

There are some readily recognized factors in the survey procedure which cause the number of discharges reported by the respondents to differ from the actual number which occurred in hospitals during the reference year. One important factor is the failure of the respondents to report correctly each hospital episode during the reference year. A second factor is that the survey covers only persons living on the date of interview. The hospital experience of persons who died in the year prior to interview is not included.

If the difference between reported discharges and all discharges taking place during the reference year is examined on a weekly or monthly basis, a definite decreasing trend or decay, moving backward in time from the date of interview, of the number of discharges reported by the respondents in the Health Interview Survey is observed. Explanations for this decay curve include the following factors.

1. *Response errors.*—Underreporting can be expected to increase with increasing length of the recall period. In other words, recent discharges are more likely to be recalled and reported accurately than discharges which occurred earlier in the reference year.

2. *Persons in their last year of life.*—A study of hospital utilization in the last year of life reports that the "daily discharge rate per 1,000 deaths increases gradually from less than 1 during the twelfth month before death to about 3 on the day before death."[1] The Health Interview Survey obtains information from persons who *will* die in the year following the date of interview. The discharges for these persons for the reference year are more frequent for the period immediately prior to the date of interview than for earlier periods in the reference year, thus contributing to the observed decay curve.

3. *Population growth.*—Only living persons residing in the sample households on the date of the interview are eligible for the survey. The size of this population is probably at least 1.5 percent smaller 12 months prior to the date of interview, since during this period there are births and other additions to the household population such as returnees from mental and penal institutions. During this same period, losses in the household population occur, but these are not recorded since they involve persons who died or were institutionalized.

4. *Hospital discharge trend.*—A portion of the observed trend may be a legitimate consequence of natural phenomena related to the hospitalization needs of the population. If there is an increasing trend in hospital admission rates, then the same trend will be present in the discharge rates. Such a trend is not expected to be very great during a period as short as 1 year.

Response errors in reported hospital discharges have been studied by the Survey Research

Center, University of Michigan, in cooperation with the Bureau of the Census and the National Center for Health Statistics. The first study employed a sample of individuals with known hospitalization records.[2] These persons were interviewed concerning their hospital experience, and the results were compared with the records obtained from hospitals. The comparisons confirmed that underreporting of hospitalization increases with length of recall period. For discharges occurring near the beginning of the 12-month period prior to interview such underreporting was particularly serious. The study estimated underreporting of hospital episodes for the reference year to be 10 percent.

A second study compared three survey procedures for obtaining hospital episode data, including the Health Interview Survey procedure which was used as the standard.[3] Reporting accuracy was found to be significantly improved by using a revised interview schedule with a mail followup to obtain information concerning hospital stays that had been overlooked in the interview.

With respect to decedents during the reference year, the Division of Vital Statistics of the Center conducted a study of hospitalizations during the last year of life from the records of a sample of deaths in the Middle Atlantic States, i.e., New York, New Jersey, and Pennsylvania.[1] The study estimated that the hospital discharges reported in the Health Interview Survey for the Middle Atlantic States needed to be adjusted upward by approximately 8 percent to include the experience of decedents. A similar study on a national scale is now nearing completion.

The Health Interview Survey collects data from a new sample of households each week.[4] It is therefore possible to compare the hospital discharges reported for a particular calendar period by two or more of these weekly samples. For example, consider the number of hospital discharges reported for the month prior to interview of each weekly sample and compare this with the number of hospital discharges reported for the *same* month by each sample interviewed 4 weeks later. The average discrepancy for the paired weekly samples represents an estimate of the combined effects of mortality and response errors for the second month prior to interview. Such factors as population growth or hospitalization trends are not included in the observed difference.

Analyses of this type have been carried out with Health Interview Survey data to estimate the relationship between underreporting (including mortality and response errors) and the time interval between discharge and date of interview. Simmons and Bryant derived adjustment factors based on these internal analyses by which hospital discharges reported in the Health Interview Survey need to be inflated according to the distance between discharge and interview to produce an estimate of total hospital discharges, including discharges for persons dying during the reference year.[5] Although so extensive an adjustment procedure has not been adopted, publication of hospital discharges reported in the Health Interview Survey is now based on data for the most recent 6 months of the reference year. The 12-month-reference period is retained in the interview.

While research has resulted in greater understanding and knowledge of the role played by various factors affecting observed discrepancies, this understanding and knowledge is still insufficient for specification of a completely satisfactory procedure of data collection and estimation. Part of this difficulty might be explained by the fact that the major studies of response error and mortality factors have been carried out independently. An ideal research design might conduct a prospective study on a large population sample for 1 year, observe (independently) the actual hospitalization experience of this sample, and interview those persons living at the end of the year. The required data for a fuller understanding would probably result from such a study. However, this is not considered a feasible research project; it might be impossible to carry it out satisfactorily.

An alternative research approach is to simulate this prospective study on a computer. This implies specifying a population to be followed over time, with the initial state of each individual known, such as age, whether or not in a hospital, and if so, the number of days the individual has already spent in a hospital. It also requires the specification of the transition probabilities for each pair of possible states for each time period (such as a week), including mortality. The division of the population into the various states for each time period is then generated successively by means of the transition probabilities. In this way the hospital discharges can be counted for each time period, including those of individuals discharged dead as well as those of individuals who die in subsequent time periods.

The household interview among living persons in the generated population at the end of 1 year can also be simulated. This simulation uses a probability function relating failure to report hospital episodes to the number of weeks between discharge and interview. The simulated interview data can then be compared with the generated hospital discharge data and the distribution of the discrepancy among the contributing factors determined for each time period.

The computer simulation approach was used in this project.

# II. PROJECT OBJECTIVES

The major purpose of this project was to develop a research tool for comparison of alternative hospital episode interview survey procedures. It was expected that the computer simulation approach could lead to relatively inexpensive evaluation of the effects of alternative procedures and eventually to more efficient and accurate procedures for the continuous collection and estimation of hospital discharge statistics.

Specific objectives of the project were:

1. To develop probability models for generating (a) hospital admissions and durations of stay for a given population, and (b) interview data on hospital episodes as collected in the Health Interview Survey.
2. To determine suitable parameter inputs for the models from existing data.
3. To program an IBM 1410 computer for experimental simulation under the models.
4. To estimate, through computer simulations, the specific effects of the various factors related to the discrepancy between hospital discharges reported in the interview survey and all discharges.
5. To suggest, on the basis of the research results, a method for continuous collection and adjustment of hospital discharge data.

# III. PROCEDURES

## SUMMARY

The initial phase of this project was concerned primarily with developing a probability model for generating hospital episodes for individuals on a computer. The model adopted assumes that each individual in the population of interest has a particular probability of being hospitalized each week. It further assumes that this weekly hospital admission probability remains constant for a given individual over the time period of interest (provided he is not in his last year of life), but varies from individual to individual. Based on empirical studies of data available from the Health Interview Survey and on theoretical considerations, it was determined that the generalized gamma distribution provides a suitable and consistent model for the distribution of the weekly admission probabilities over the population. Once an individual is hospitalized, the model provides for discharge from the hospital on a daily probability basis with the chance of discharge conditional on the number of days already hospitalized. The log-normal distribution was adopted as the duration-of-stay model, following empirical analysis of length-of-stay data available from the Health Interview Survey.

A computer program was developed in the second phase of this project to generate hospitalization histories for each individual in a model U.S. population. The weekly admission probabilities and daily discharge probabilities employed in the computer program were estimated for individuals in each of 12 age-sex groups consistent with the hospital episodes model developed in the first phase. In brief, the computer program generates uniform random numbers to compare with the appropriate weekly hospital admission probability for an individual during each week that the individual is not hospitalized. When an individual is hospitalized by the computer, it then generates uniform random numbers to compare with the appropriate daily discharge probabilities until the individual is discharged. The computer records the day of admission and day of discharge for each hospital episode generated.

This basic computer program, with some modifications, was carried out for an initial population of 10,000 individuals, distributed by age and sex to represent the U.S. civilian, noninstitutional population, for a period of 108 weeks or 756 days. The modifications included introducing births and deaths in order to give a dynamic dimension to the population and using a separate set of daily hospital admission probabilities for individuals in their last year of life. These latter probabilities increased gradually as the day of death approached. Except for deliveries, reasons for hospitalization were not assigned in the computer simulation program. The computer determined on a random basis those deliveries which were to occur in a hospital.

In the third phase of the project a relatively simple model was devised to simulate the responses obtained in household interviews for individuals experiencing one or more hospital episodes in the year prior to interview. For each hospital episode, the model simulates on a probability basis failure to report the episode, reported length of stay (if the episode is reported), and reported month of discharge. The model treats reporting of each hospital episode as a random event dependent on length of the recall period and length of hospital stay for the episode. The distribution of errors in reported length of stay is approximated in the model by a normal or Gaussian distribution. Response errors in the reported month of discharge are simulated in the model by first approximating errors in the reported date of admission by a normal distribution. The reported length of stay is then added to the reported date of admission to obtain the reported discharge date.

A computer program to generate interview results consistent with the interview simulation model was developed in the fourth phase of the project. The input data for this program consisted of the 108 weeks of hospital episode data generated by the first computer program together with parameter values for the interview simulation model. Estimates of the necessary parameters were based on evidence from exploratory work which had been done in the National Center for Health Statistics and especially on the results obtained in the previously mentioned response error study conducted by the Survey Research Center, University of Michigan. This interview simulation computer program was run for 13 separate interview dates 4 weeks apart beginning with week 60 of the 108-week period for which hospital episode data had been generated. The results were tabulated in three separate categories by the computer for each interview date. These results included number of discharges and number of hospital days, by sex, age, and each of 13 four-week periods prior to the interview date. The three tabulation categories were "interview reported" results for persons alive on the date of interview, which include simulated interview reporting errors; "perfect interview" results for persons alive on the date of interview, which simulate the results which would be obtained by the household interviews if there were no response errors of any kind; and "all discharges" which consist of the actual results for all hospital episodes generated by the first computer program for the year prior to the interview date for all persons, whether alive or dead on the interview date.

The data generated by the computer for the 13 interview dates were averaged and estimates of annual hospital discharge rates and annual hospital days per 1,000 persons by age and sex were derived for each of the three tabulation categories. Using these results, both separate and combined estimates of the effects of interview response errors and of exclusion of persons who died during the reference year on hospital discharge data collected in the Health Interview Survey can be derived.

# A STOCHASTIC MODEL FOR HOSPITAL EPISODES

## Hospital Admissions Model

The model for hospital admissions was determined soon after the project was initiated. This was primarily due to a fortunate exposure to research on a mathematical model of an index of health by Dr. Chin Long Chiang, University of California at Berkeley.[6] The hospital admissions of an individual during a time interval of length $t$ can be treated as random events in time, that is, as a stochastic process. A simplified model assumes that the probability of the individual being hospitalized during a small time interval $dt$ is given by $\lambda dt$, where $\lambda$ is a positive constant.[a] If it is further assumed that this probability $\lambda dt$ is independent of the number of previous hospital admissions for the individual, then the process is a Poisson process. It follows that the probability of exactly $x$ admissions of the individual occurring during the time $t$ is given by

$$P_x(t) = \frac{e^{-\lambda t}(\lambda t)^x}{x!} \qquad x = 0, 1, 2, \ldots \qquad (1)$$

If the time interval $t$ is taken as 1 year (i.e., $t = 1$), then the probability density function for the number of hospitalizations annually for the individual is Poisson, where the parameter $\lambda$ is the expected number of hospital episodes during this period.

Suppose now that the probability of being hospitalized in a small time interval varies from individual to individual in a population so that $\lambda$ varies over the population. If the distributions of the $\lambda$'s is gamma, then the distribution of the population by number of hospital episodes yearly is negative binomial, derived as follows.

[a]More rigorously, the probability of one or more hospital admissions for an individual in the small interval $dt$ is given by $\lambda dt + o(dt)$ where the term $o(dt)$ denotes a quantity which is of smaller order of magnitude than $dt$ and is the probability that more than one admission occurs.

From equation (1) above, the distribution of admissions annually for an individual with parameter $\lambda$ is

$$f(x|\lambda) = \frac{e^{-\lambda}\,\lambda^x}{x!} \qquad x = 0, 1, 2, \ldots \qquad (2)$$

For all individuals in the population, the distribution of $\lambda$'s is assumed to be a gamma distribution, i.e.,

$$g(\lambda) = \frac{\beta}{\Gamma(\alpha)}(\beta\lambda)^{\alpha-1}e^{-\beta\lambda}, \qquad \alpha > o, \beta > o. \qquad (3)$$

Then the joint distribution of $x$ and $\lambda$ is

$$f(x|\lambda)g(\lambda) = \frac{\beta^\alpha}{x!\,\Gamma(\alpha)}\, e^{-\lambda(\beta+1)}\,\lambda^{\alpha+x-1}. \qquad (4)$$

The distribution of the population by number of hospital episodes annually, that is $f(x)$, is found by integrating equation (4) with respect to $\lambda$. Thus,

$$f(x) = \int_\lambda f(x|\lambda)g(\lambda)\,d\lambda$$

$$= \binom{\alpha+x-1}{x}\left(\frac{\beta}{1+\beta}\right)^\alpha\left(\frac{1}{1+\beta}\right)^x, \qquad x = 0, 1, 2, \ldots \qquad (5)$$

which is the negative binomial distribution.

Data available from the Health Interview Survey for the period July 1958-June 1960 were used to determine the goodness of fit of the negative binomial distribution to the observed frequencies of persons with 0, 1, 2, 3, and 4 or more hospital episodes in the average year. A separate fit was made for males and females in each of the following six age groups: under 15 years, 15-24, 25-34, 35-44, 45-64, and 65 years and older. Each fit was accomplished by estimating the parameters $\alpha$ and $\beta$ by the method of moments, that is, from the relations

$$\bar{x} = \alpha/\beta$$

$$s^2 = \alpha(1+\beta)/\beta^2$$

where $\bar{x}$ and $s^2$ are the observed mean and variance respectively. The comparisons of the observed and expected frequencies for the 12 age-sex groups were considered to be fairly good.

While a satisfactory fit of the negative binomial distribution is not sufficient evidence to claim the model to be valid, it does indicate that the model provides an excellent basis for generating hospital episodes reasonably consistent with observation.

## Duration-of-Stay Model

Once an individual is hospitalized, his length of stay depends largely on the reason for the hospitalization. Each diagnosis can be considered to generate its own length-of-stay distribution; for example, the length-of-stay distribution for tonsillectomies will be different from that for pneumonia cases. Since the overall length-of-stay distribution is a mixture of many different distributions, it is not expected that any one distribution will fit well. For purposes of computer simulation, the distribution of duration of stay observed in the Health Interview Survey could have been used, except that the data had been grouped into fairly large intervals, particularly for the upper tail of the distribution. A smoothed distribution was preferred.

In order to obtain some insight into an appropriate theoretical distribution for duration of stay, the conditional probabilities of discharge on a particular day, given that the individual has been hospitalized up to that day, were computed for the July 1958-June 1960 Health Interview Survey data for grouped periods on an average daily basis. The rise and fall of these conditional probabilities as duration of stay increased was characteristic of the log-normal distribution. Accordingly, this distribution was fitted to the available duration-of-stay data separately within age and sex groups. Since the agreement between these expected and observed proportions was considered satisfactory, the log-normal distribution was adopted as the duration-of-stay model.

## Computer Simulation of Hospital Episodes

The stochastic models for hospital admission and duration of stay developed above suggest that hospital episodes for the U.S. civilian, noninstitutional population can be readily simulated on a computer by means of a set of daily (or weekly) transition probabilities for each individual. These probabilities are assumed to remain constant over time for an individual, at least for periods up to 2 years, but to vary from individual to individual.

On a given day, say $i$, an individual can be in one of $S+1$ states. These states are:

$\bar{H}$ = not in hospital

$H_j$ = in hospital $j$ days for a particular episode, $j = 1, 2, \ldots, S$.

For each state on day $i$, transition probabilities are specified for the two eligible states for the individual on day $i+1$. Thus, for individual $k$ in state $\bar{H}$ on day $i$:

$P_k$ = the probability of being hospitalized on day $i+1$

$1-P_k$ = the probability of remaining out of the hospital on day $i+1$.

Similarly, for individual $k$ in state $H_j$ on day $i$:

$P_{jk}$ = the probability of being discharged on day $i+1$ (i.e., going to state $\bar{H}$)

$1-P_{jk}$ = the probability of remaining in the hospital on day $i+1$ (i.e., going to state $H_{j+1}$).

In brief, then, by specification of $S+1$ probabilities ($P_k$ and $P_{jk}$, $j = 1, 2, \ldots, S$) for individual $k$, a computer can be programmed to generate a hospitalization history for this individual during a designated time period. If the individual is not in the hospital initially, the computer generates a uniform random number $R_1$ between zero and one to compare with $P_k$. If $R_1 \le P_k$, individual $k$ is hospitalized on the first day (i.e., transferred from state $\bar{H}$ to state $H_1$). The computer then generates a second uniform random number $R_2$ to compare with $P_{1k}$. If $R_2 \le P_{1k}$, individual $k$ is discharged on the second day; otherwise individual $k$ remains in the hospital for a second day and a third uniform random number $R_3$ is generated for comparison with $P_{2k}$, etc., until discharge occurs. Following discharge, the next uniform random number is again compared with $P_k$. If the initial random number $R_1 > P_k$, individual $k$ remains in state $\bar{H}$ and $R_2$ is compared with $P_k$, etc., until hospitalization occurs or the designated time period is exhausted. The computer is programmed to record the day of admission and the day of discharge for each hospital episode generated.

The probability of hospital admission ($P_k$) was specified on a weekly basis rather than a daily basis, except for individuals in their last year of life. This change was necessary in order to reduce computer time. If an individual was admitted to the hospital in a given week, the computer assigned the specific day of the week, and hence the day of admission, by means of a random sequence.

The weekly admission probabilities were estimated by first fitting a negative binomial distribution to the distribution of the population by number of hospital episodes annually, as observed in the July 1958-June 1960 Health Interview Surveys, for each of 12 age-sex groups. Delivery episodes were excluded from the female age groups. The $\alpha$ and $\beta$ parameters estimated in the fitting process for a particular age-sex group (table A) are also, in accordance with the hospital admissions model, the parameters of the gamma distribution of $\lambda$ (equation 3), where $\lambda$ is the expected annual number of hospital episodes for a given individual in the group. While it would have been possible to determine a $\lambda$ for each individual in a group by sampling the appropriate gamma distribution at random, this was not considered necessary. Rather, each of the 12 age-sex groups was divided further into 10 equal subgroups. It was planned initially to assign the first subgroup in each age-sex group a value of $\lambda$ corresponding to the 5% point of the appropriate gamma distribution, the second a $\lambda$ corresponding to the 15% point, and so on to the $\lambda$ corresponding to the 95% point for the 10th subgroup. Since the gamma distributions of interest were highly skewed, the tables of the incomplete gamma-function used to determine these $\lambda$ values were lacking in some detail.[7] The tables are entered for arguments $u$ and $p$ where

$$u = \beta\lambda/\alpha^{1/2}$$

$$p = \alpha - 1 .$$

However, the tables did not give values of the argument $u$ below the 40th percentile in all cases of interest and below the 50th percentile in a few cases. Thus, the first four or five subgroups in each age-sex group were assigned $\lambda$'s corresponding to the interpolated 20th percentile (or 25th percentile) values of $u$. The average value of the assigned $\lambda$'s in each age-sex group was adjusted to the observed mean of the distribution of hospital episodes annually by adjusting the $\lambda$ corresponding to the 95% point.

The constant weekly admission probability $P_k$, which applied to all individuals in a subgroup, was obtained by dividing each assigned $\lambda$ by 52. These weekly admission probabilities for the 120 subgroups are given in table B. Each newborn individual was assigned to one of the 10 subgroups in the "under 15 years" age group of the same sex.

Table A. $\alpha$ and $\beta$ parameters of the negative binomial distributions fitted to the distribution of the population in 12 age-sex groups by number of annual hospital episodes

[See equation (5)]

| Age | Male | | Female | |
|---|---|---|---|---|
| | $\alpha$ | $\beta$ | $\alpha$ | $\beta$ |
| Under 15 years | 0.3097 | 4.9090 | 0.2432 | 4.7083 |
| 15-24 years | 0.2369 | 3.7665 | 0.1398 | 1.4480 |
| 25-34 years | 0.2824 | 4.2410 | 0.2290 | 1.7924 |
| 35-44 years | 0.2834 | 3.6292 | 0.3901 | 3.3889 |
| 45-64 years | 0.2833 | 2.6920 | 0.3622 | 3.2396 |
| 65+ years | 0.3906 | 2.6129 | 0.3569 | 2.8701 |

Table B. Estimated weekly hospital admission rates per 1,000 persons not in their last year of life, excluding deliveries, by age, sex, and 10 percent subgroups, and average weekly and annual hospital admission rates for all subgroups combined (computer input probabilities x $10^3$)

| Subgroup | Age groups: Under 15 years | 15-24 years | 25-34 years | 35-44 years | 45-64 years | 65+ years |
|---|---|---|---|---|---|---|
| | Male | | | | | |
| 1 | 0.146 | 0.233 | 0.273 | 0.318 | 0.430 | 0.548 |
| 2 | 0.146 | 0.233 | 0.273 | 0.318 | 0.430 | 0.548 |
| 3 | 0.146 | 0.233 | 0.273 | 0.318 | 0.430 | 0.548 |
| 4 | 0.146 | 0.233 | 0.273 | 0.318 | 0.430 | 0.548 |
| 5 | 0.219 | 0.233 | 0.273 | 0.318 | 0.430 | 0.822 |
| 6 | 0.439 | 0.350 | 0.409 | 0.477 | 0.644 | 1.370 |
| 7 | 0.768 | 0.700 | 0.793 | 0.927 | 1.250 | 2.284 |
| 8 | 1.382 | 1.325 | 1.202 | 1.404 | 1.894 | 3.655 |
| 9 | 2.522 | 2.574 | 2.788 | 3.257 | 4.395 | 6.076 |
| 10 | [1]5.549 | 6.074 | 6.144 | 7.177 | 9.685 | 12.336 |
| Average weekly rate for all subgroups combined | 1.1463 | 1.2188 | 1.2701 | 1.4832 | 2.0018 | 2.8735 |
| Average annual rate for all subgroups combined | 59.608 | 63.378 | 66.045 | 77.126 | 104.094 | 149.422 |
| | Female | | | | | |
| 1 | 0.187 | 0.327 | 0.344 | 0.422 | 0.332 | 0.375 |
| 2 | 0.187 | 0.327 | 0.344 | 0.422 | 0.332 | 0.375 |
| 3 | 0.187 | 0.327 | 0.344 | 0.422 | 0.332 | 0.375 |
| 4 | 0.187 | 0.327 | 0.344 | 0.422 | 0.332 | 0.375 |
| 5 | 0.187 | 0.327 | 0.344 | 0.633 | 0.499 | 0.563 |
| 6 | 0.280 | 0.327 | 0.516 | 1.055 | 0.890 | 1.005 |
| 7 | 0.560 | 0.491 | 1.238 | 1.759 | 1.531 | 1.729 |
| 8 | 1.060 | 1.325 | 2.475 | 2.814 | 2.564 | 2.895 |
| 9 | 2.061 | 3.435 | 5.054 | 4.678 | 4.452 | 5.025 |
| 10 | 4.904 | 11.354 | 13.709 | 9.497 | 10.102 | 11.404 |
| Average weekly rate for all subgroups combined | 0.9800 | 1.8567 | 2.4712 | 2.2124 | 2.1366 | 2.4121 |
| Average annual rate for all subgroups combined | 50.960 | 96.548 | 128.502 | 115.045 | 111.103 | 125.429 |

[1]This rate was incorrectly computed. The error was not discovered until after the computor runs. The correct value is 6.227. The expected annual rate for the computer generated episodes would have been raised from 59.6 to 63.1 per 1,000 persons by use of the correct value.

A slightly different model was used to generate the hospital histories of persons in their last year of life. Prior to generating a random number to determine if an individual would be hospitalized in the week of interest, the computer first checked whether or not the individual had entered his last year of life. If so, the computer changed to a set of daily probabilities of being hospitalized which increased gradually as the day of death approached. These probabilities were estimated from data on hospital utilization during selected time periods prior to death reported in the Middle Atlantic States study.[1] First, rough estimates of admission rates per 1,000 deaths and number of

Table C. Estimated daily hospital admission probabilities for persons in their last year of life as a function of time period to death

| Period prior to death | Daily admissions per 1,000 deaths | Persons not in hospital per 1,000 deaths | Daily admission probabilities[1] |
|---|---|---|---|
| 1 and 2 days | 41.8 | 674.9 | 0.061935 |
| 2 and 3 days | 30.4 | 702.3 | 0.043286 |
| 3 and 4 days | 31.8 | 731.1 | 0.043496 |
| 4 and 5 days | 18.8 | 746.9 | 0.025171 |
| 5 and 6 days | 23.1 | 767.0 | 0.030117 |
| 6 and 7 days | 27.5 | 791.5 | 0.034744 |
| 1 and 2 weeks | 7.3 | 818.5 | 0.008919 |
| 2 and 3 weeks | 6.2 | 845.9 | 0.007329 |
| 3 and 4 weeks | 7.0 | 880.2 | 0.007953 |
| 1 and 2 months | 3.1 | 915.2 | 0.003387 |
| 2 and 3 months | 2.6 | 949.3 | 0.002739 |
| 3 and 4 months | 1.8 | 963.9 | 0.001867 |
| 4 and 5 months | 1.1 | 967.1 | 0.001137 |
| 5 and 6 months | 1.3 | 977.4 | 0.001330 |
| 6-12 months | 0.65 | 985.1 | 0.000660 |

[1]Ratio of first to second column.

Table D. Probability of birth occurring in a hospital, by age of mother, 15-44 years

| Age of mother | Total annual births per 1,000 females, 1960 | Annual births in hospital per 1,000 females | Probability of delivery in hospital |
|---|---|---|---|
| 15-24 years | 166.32 | 135.86 | 0.816859 |
| 25-34 years | 152.86 | 145.79 | 0.953749 |
| 35-44 years | 36.60 | 31.40 | 0.857923 |

persons per 1,000 deaths not in the hospital as a function of the time period prior to death were derived from changes (first differences)in the nights of care rates and from the discharge rates. The ratio of these two quantities provided the required estimates of daily admission probabilities as a function of days to death. These estimates, shown in table C, were then plotted and the function smoothed graphically. The smoothed function provided 365 admission probabilities, one for each day in the last year of life.

Except for deliveries, reasons for hospitalization were not assigned in the simulation program. Females with delivery dates less than 31 days away from the day of interest were not admitted to hospital during this period. On the assigned delivery dates, the computer determined on a random basis which deliveries were to occur in hospitals. The probability of a delivery taking place in a hospital was estimated for three age groups of mothers by dividing the number of births in hospitals per 1,000 females[8] by the rate for all births. These probabilities are shown in table D.

The log-normal distribution

$$f(t) = \frac{1}{(2\pi)^{1/2}\, t\sigma}\, e^{-(\ln t - \mu)/2\sigma^2}, \quad t > 0 \qquad (6)$$

was fitted to the observed distribution of length of hospital stay (excluding deliveries) for each of the 12 age-sex groups using unpublished Health Interview Survey data for the period July 1958-June 1960. The parameters, $\mu$ and $\sigma$, in $f(t)$ were estimated from the equations

$$\bar{x} = e^{\mu + \sigma^2/2}$$

$$s^2/\bar{x}^2 = e^{\sigma^2} - 1,$$

where $\bar{x}$ and $s^2$ are the mean and variance of the observed duration-of-stay distribution. The conditional probabilities ($P_{jk}$) of discharge on day $t$, given that the individual had been hospitalized for the previous $t-1$ days, were then estimated from the fitted log-normal duration-of-stay distributions. The computer program limited length of stay to a maximum of 100 days so that $P_{100k}$ was set equal to .999999.

Separate sets of discharge probabilities were estimated for females 15-24, 25-34, and 35-44 years of age hospitalized for deliveries. The estimates were derived in the same manner as discussed above, using unpublished length-of-stay data for deliveries obtained from the Health Interview Survey, July 1958-June 1960. Length of stay was limited to a maximum of 21 days for females 15-24 years, 24 days for females 25-34 years, and 30 days for females 35-44 years.

Duration-of-stay distributions were not available for persons in their last year of life. However, average length-of-stay estimates by sex in age classes under 45, 45-64, and 65 years and over were obtained from the study of hospital utilization by decedents in the Middle Atlantic States.[1] The variances of the duration-of-stay distributions for these age-sex classes were imputed by using the relationship observed between $s^2$ and $\bar{x}$ for these distributions among persons not in their last year of life. Thus, estimates of the conditional discharge probabilities were derived as above with length of stay limited to a maximum of 100 days.

The estimates of the parameters $\mu$ and $\sigma$ for the log-normal fit of the duration-of-stay distributions in each of the above cases are given in table E.

The computer operations for generating hospitalization histories for persons *not* in their last year of life (Phase I) and for persons in their last year of life (Phase II) are given in detail in the Appendix.

The basic computer program, with modifications as discussed below, was carried out for an initial population of 10,000 individuals for 108 weeks or 756 days. This population was distributed by age and sex to represent the U.S. civilian, noninstitutional population.

The initial population was given a dynamic dimension by introducing births and deaths. The births were distributed over a 2-year period according to 1960 monthly birth rates and then assigned specific days within months at random. A total of 237 births (121 male and 116 female) were assigned the first year and 240 (123 male and 117 female) the second year. Coinciding with the birth dates, deliveries were assigned to females in the 15-24, 25-34, and 35-44 years of age groups.

A simple three-digit code was used to record dates on the computer, with the first day of the 108-week period coded 001. The first 26 days of the hospital episodes simulation program were utilized to establish the appropriate initial distribution of the population over the states $\bar{H}$ and $H_j$. This was necessary since all individuals were in state $\bar{H}$ (i.e., not in hospital) on day 001. An alternative procedure would have required assignment of about 22 individuals to the hospital states $H_j$ on day 001. Since the average length of stay in short-term hospitals is approximately 8 days and less than 10 percent of the episodes exceed 15 days, allowing the computer 26 days to establish an equilibrium distribution over the states $\bar{H}$ and $H_j$ is considered adequate. There were no additions to the population from births assigned prior to day 027. Hospitalization histories for newborn infants were generated by the computer only for the days following birth.

In order to introduce appropriate hospital admission rates for individuals entering their last year of life, death dates were assigned by age and sex covering a 3-year period. A total of 93 deaths were assigned in the first year, 94 in the second, and 89 in the third. As with the birth dates, these were distributed first according to 1960 monthly death rates and then were assigned specific days within months at random. The third year death dates were necessary since individuals scheduled to die in that year enter last year of life sometime during the second year.

Table E. Estimates of the parameters $\mu$ and $\sigma$ for log-normal distributions fitted to duration-of-stay distributions, by sex and age

| Age | Persons not in their last year of life | | | | | |
|---|---|---|---|---|---|---|
| | Male | | Female | | | |
| | | | Deliveries excluded | | Deliveries only | |
| | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ |
| Under 15 years | 1.22 | 1.12 | 1.16 | 1.15 | ... | ... |
| 15-24 years | 1.51 | 1.10 | 1.19 | 1.01 | 1.32 | 0.47 |
| 25-34 years | 1.63 | 1.02 | 1.46 | 0.94 | 1.33 | 0.53 |
| 35-44 years | 1.74 | 1.00 | 1.65 | 0.90 | 1.37 | 0.69 |
| 45-64 years | 2.08 | 0.93 | 1.94 | 0.91 | ... | ... |
| 65+ years | 2.30 | 0.89 | 2.33 | 0.85 | ... | ... |

| Age | Persons in their last year of life | | | |
|---|---|---|---|---|
| | Male | | Female | |
| | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ |
| Under 44 years | 2.21 | 0.90 | 1.96 | 0.93 |
| 45-64 years | 2.53 | 0.82 | 2.94 | 0.70 |
| 65+ years | 2.64 | 0.80 | 2.36 | 0.84 |

A four-digit number was used to code the day of death for computer purposes; all individuals not in their last year of life at the end of the second year were assigned 9999 as their day of death. No deaths were assigned prior to day 0027.

## INTERVIEW SIMULATION MODEL

A relatively simple model was devised for simulating the responses obtained in interviews with individuals experiencing one or more hospital episodes during the 12 months prior to the date of interview. For each hospital episode, the model simulates on a probability basis failure to report the episode, reported length of stay (if the episode is reported), and reported month of discharge.

### Underreporting of Hospital Episodes

The response error study by the Survey Research Center, University of Michigan, reported three major factors related to underreporting of hospital episodes.[2] It was found that underreporting increases with increasing time between discharge and interview, decreases with increasing length of stay, and increases for personally embarrassing or threatening types of illness. Only the first two factors are included in the interview simulation model. The Michigan study reported percent underreporting by number of weeks between hospital discharge and interview for three length-of-stay groups.[2] The Center also had produced, through internal analysis of reported data, rough distributions of underreporting by number

Table F. Probability of failure to report hospital episodes, by length of stay and number of weeks between discharge and interview, and average probability of failure

| Weeks between discharge and interview | Length of stay | | |
|---|---|---|---|
| | 1 day | 2-4 days | 5+ days |
| | Nondelivery episodes | | |
| 1-4 weeks---------- | 0.07 | 0.04 | 0.01 |
| 5-8 weeks---------- | 0.13 | 0.05 | 0.02 |
| 9-12 weeks--------- | 0.18 | 0.06 | 0.04 |
| 13-16 weeks-------- | 0.22 | 0.07 | 0.05 |
| 17-20 weeks-------- | 0.24 | 0.08 | 0.06 |
| 21-24 weeks-------- | 0.26 | 0.09 | 0.07 |
| 25-28 weeks-------- | 0.28 | 0.11 | 0.08 |
| 29-32 weeks-------- | 0.29 | 0.14 | 0.09 |
| 33-36 weeks-------- | 0.30 | 0.18 | 0.09 |
| 37-40 weeks-------- | 0.30 | 0.22 | 0.10 |
| 41-44 weeks-------- | 0.31 | 0.27 | 0.10 |
| 45-48 weeks-------- | 0.32 | 0.33 | 0.11 |
| 49-52 weeks-------- | 0.32 | 0.39 | 0.46 |
| 53-56 weeks-------- | 0.32 | 0.39 | 0.46 |
| 57-60 weeks-------- | 0.32 | 0.39 | 0.46 |
| Average probability of failure---- | 0.257 | 0.187 | 0.147 |
| | Delivery episodes | | |
| 1-4 weeks---------- | 0.00 | 0.00 | 0.00 |
| 5-8 weeks---------- | 0.00 | 0.00 | 0.00 |
| 9-12 weeks--------- | 0.01 | 0.01 | 0.00 |
| 13-16 weeks-------- | 0.01 | 0.01 | 0.00 |
| 17-20 weeks-------- | 0.02 | 0.02 | 0.01 |
| 21-24 weeks-------- | 0.02 | 0.02 | 0.01 |
| 25-28 weeks-------- | 0.03 | 0.03 | 0.03 |
| 29-32 weeks-------- | 0.03 | 0.03 | 0.03 |
| 33-36 weeks-------- | 0.03 | 0.03 | 0.03 |
| 37-40 weeks-------- | 0.04 | 0.04 | 0.04 |
| 41-44 weeks-------- | 0.05 | 0.04 | 0.04 |
| 45-48 weeks-------- | 0.05 | 0.05 | 0.05 |
| 49-52 weeks-------- | 0.06 | 0.05 | 0.05 |
| 53-56 weeks-------- | 0.07 | 0.06 | 0.06 |
| 57-60 weeks-------- | 0.07 | 0.06 | 0.06 |
| Average probability of failure---- | 0.033 | 0.030 | 0.027 |

of weeks between discharge and interview for four length-of-stay classes. After study of data from these sources, smooth curves were fitted for each of the length-of-stay groups, and estimates of underreporting rates for hospital episodes as a function of the time interval between discharge and interview (in 4-week periods) were obtained for the model. The model treats reporting of each hospital episode as a random event dependent on length of the recall period and length of the hospital stay for the episode.

These estimated underreporting rates were used for nondelivery episodes only. Since the data upon which they were based included all episodes, these estimates are slightly optimistic. The response error study mentioned above found only 3 percent underreporting of deliveries, whereas the average underreporting for all diagnoses was 10 percent. A separate set of underreporting rates, averaging 3 percent, was constructed for delivery episodes. These were also made dependent on length of recall period and length of hospital stay.

The estimated rates of underreporting of nondelivery and delivery episodes were treated as probabilities in the computer simulation. They are shown in table F for 15 four-week periods prior to interview. The last two intervals (53-56 weeks and 57-60 weeks) were included to allow for overreporting of episodes occurring more than 12 months prior to interview. These were included in the model by telescoping forward, again on a probability basis as discussed below, episodes reported by the respondent with actual discharge dates in the 14th or 15th 4-week periods prior to interview. The same underreporting rates were used for these latter two periods as were estimated for weeks 49-52 (the 13th 4-week period).

## Length-of-Stay Response Errors

The Michigan study found the average length of stay reported in household interviews to be slightly greater than the average calculated from hospital records.[2] One explanation given for this is that underreporting is more likely for short-stay episodes than for longer episodes, so that the average of reported episodes has an upward bias. Thus, it is quite possible that duration-of-stay response errors are symmetrically distributed about zero. The model for interview simulation in this study made use of this hypothesis, but also introduced a slight positive shift in the mean of the distribution of reporting errors in length of hospital stay.

The model approximates the distribution of length-of-stay response errors by a normal or

Gaussian distribution with a mean error of zero in an expected 95 percent of the responses and a mean error of 2 days in the remaining 5 percent. Thus, the overall distribution of errors is assumed normal with mean equal to 0.05 x 2.0 or 0.1 day. Unit variance was assigned these normal error distributions; this is considered a conservative value for this parameter.

A reported length of stay for a given episode is generated in two steps according to this model. First, a uniform random number between zero and one is compared with 0.05. If it is less than 0.05, 2 days are added to the actual length of stay; otherwise the actual length of stay is left unchanged. Second, a random normal deviate is generated and added to either the adjusted length of stay or the actual length of stay, depending on the previous comparison of the random number with 0.05. The resulting length of stay in days is accepted as the reported duration of stay.

### Month-of-Discharge Response Errors

The first Michigan study found that for 82 percent of the episodes, the respondent correctly reported the month of admission; about 11 percent were reported 1 or more months later than shown in the hospital records, and 7 percent were earlier by 1 or more months.[2] The later study, comparing three alternative hospitalization survey procedures, showed 14 percent reported the month of discharge later, 9 percent earlier, and 77 percent correctly, using the Health Interview Survey procedure.[3] The month of discharge is calculated by use of the reported admission date and the reported length of hospitalization. The evidence in these two studies indicates a greater tendency to telescope the hospital episode forward rather than backward in time, although the shift is a modest one. The bulk of the inaccurate reports were plus or minus 1 month of the correct month.

The model adopted for simulation of response errors leading to incorrect classification of the month of discharge also approximates errors in the date of admission by a normal distribution. As with the length-of-stay response errors, this distribution is a weighted combination of two normal distributions, the first with mean zero to apply in an expected 95 percent of the episodes and the second with a mean of 10 days applicable to the remaining 5 percent. The overall error distribution has mean equal to 0.05 x 10 or 0.5 days. The variance assigned these distributions depended on the number of weeks between date of interview and date of admission. This interval was divided into 4-week periods and the assigned standard deviation was set equal to 0.4 times the number of 4-week periods in the interval. Thus, the model permits larger errors in reported date of admission with increasing length of recall period. As with the length-of-stay model, these parameters are considered conservative.

A reported month of discharge for a given episode is generated in three steps. In the first step a uniform random number between zero and one is compared with 0.05. If it is less than 0.05, 10 days are added to the actual admission date; otherwise the actual admission date is left unchanged. In the second step, a random normal deviate is generated and multiplied by a standard deviation $\sigma$ depending on the number of weeks between the interview date and the date of admission. This product is added to either the adjusted admission date or the actual admission date, depending on the prior comparison of the random number with 0.05. In the third step, the reported length of stay is added to the adjusted admission date obtained in step two to yield the reported discharge date and hence the reported month of discharge.

### Computer Simulation of Interviews

The output of each computer-generated hospitalization includes the day admitted, whether the episode was for a delivery or not, and the day discharged. The output also includes the age, sex, and day of death for each individual experiencing one or more episodes during the 108 weeks of interest. These data make up the input for computer simulation of interviews on a specified interview date. The basic steps in the computer program for this simulation are outlined below.

1. The death date for each individual is compared with the interview date to determine if the individual is alive and hence eligible for interview. If the indi-

vidual has died the computer proceeds to the next individual.

2. If the individual is alive on the interview date, the computer determines whether the admission date for the first episode occurred prior to the interview date. If not, the next episode is examined.

3. If the admission date is earlier than the interview date, the discharge date for the episode is checked to determine if it is a completed episode. If not, the computer records an incomplete episode and proceeds to the next episode.

4. If the episode is completed prior to the interview date, the number of days between interview and discharge is computed to determine if discharge occurred more than 420 days prior. If so, the computer proceeds to the next episode.

5. If the episode is completed less than 420 days prior to the interview date, a uniform random number is generated and compared with the appropriate probability of failure to report the episode (based on the number of weeks between interview and discharge dates, length of stay, and reason for hospitalization as shown in table F). If the generated random number is less than this probability, the episode is recorded as nonrecalled and the computer proceeds to the next episode.

6. If the episode is recalled, a second uniform random number is generated and compared with 0.05. If it is less than 0.05, the computer adds 10 days to the actual admission date and continues. If not, the computer continues.

7. A random normal deviate is generated and multiplied by the appropriate standard deviation $\sigma$ (based on number of weeks between interview and admission dates). The resulting product is added to the adjusted or actual admission date, whichever is appropriate as per step (6), to obtain the reported admission date of the episode.

8. A third uniform random number is generated and compared with 0.05. If it is less than 0.05, the computer adds 2 days to the actual length of stay for the episode and continues. If not, the computer continues.

9. A second random normal deviate is generated and added to the adjusted or actual length of stay, whichever is appropriate as per step (8), to obtain the reported length of stay.

10. The reported length of stay is added to the reported admission date to determine the reported discharge date.

11. The interval between the interview date and reported discharge date is compared with 364 to determine if the episode is reported with discharge date in the year prior to interview. If so, the computer records the appropriate output data for the reported episode and proceeds to obtain "interview data" for the next episode. If the reported discharge date is more than 364 days prior to the interview date, the computer proceeds to the next episode.

This interview simulation program (Phase III) was carried out for 13 interview dates 28 days apart beginning with day 418. The hospitalization histories for the 1,870 individuals with one or more episodes generated by the hospital simulation program (Phases I and II) over the 108-week period provided the interview simulation input data. The results of the simulation for each interview date were tabulated by the computer and the following tables printed out.

1. Number of nonrecalled discharges by sex and age in each of 13 four-week periods prior to the interview date.

2. Number of nonrecalled delivery discharges for females by age in each of the 13 four-week periods.

3. Number of incomplete episodes by sex, age, and type of episode (i.e., nondelivery and delivery).
4. Number of reported discharges of 1-day stays by sex and age for the 13 four-week periods.
5. Number of reported discharges of 2-4-day stays by sex and age for the 13 four-week periods.
6. Number of reported discharges of 5-or-more-day stays by sex and age for the 13 four-week periods.
7. Number of reported discharges by sex and age for the 13 four-week periods.
8. Number of reported delivery discharges for females by age for the 13 four-week periods.
9. Number of reported hospital days associated with reported discharges in the 13 four-week periods by sex and age.
10. Number of persons by sex and age and reported number of completed episodes in the year prior to interview.
11. Number of persons by sex and age and reported number of completed nondelivery episodes in the year prior to interview.
12. Number of reported days in hospital in each of 17 four-week periods prior to interview for reported discharges by sex and age.
13. Number of reported days in hospital in each of 17 four-week periods prior to interview for reported delivery discharges for females by age.

The computer print-out of these tables is designated by the heading "interview reported." The computer program also tabulated this same set of tables using actual results for all episodes with discharge in the year prior to interview experienced by the persons alive on the date of interview, that is, with no response errors of any kind. These tables are designated in the computer print-out by the heading "perfect interview." Finally, the results for persons who died in the year prior to the interview date were tabulated by the computer and added to the "perfect interview" tables. The computer print-out of these tables is designated by the heading "all discharges."

## SIMULATION ESTIMATES OF ERRORS IN HOSPITAL DISCHARGE DATA

The computer-generated data for the 13 interview dates were averaged and estimates of annual hospital discharge rates by age and sex derived for the "interview reported," "perfect interview," and "all discharges" data tabulation categories. Similar sets of estimates were also derived for discharge rates excluding deliveries, annual hospital days per 1,000 persons with and without deliveries included, and average length of stay. These estimates are given in tables 1-5. The population bases for these rate estimates are given in table 6.

Estimates of the effects of interview response errors (using data for the full 12 months prior to interview) and of exclusion of persons who died during the reference year on hospital discharge data can be derived from tables 1-5. For example, interview response errors are estimated to reduce the annual discharge rate per 1,000 living persons by 106.0 - 94.0 = 12.0 or 11.3 percent (table 1). In addition, exclusion of persons who died during the reference year reduces the annual discharge rate by an estimated additional 6.6 discharges per 1,000 persons (112.6 - 106.0) or 5.9 percent. The overall annual rate based on the interview procedure is estimated to be less than the actual annual discharge rate by 112.6 - 94.0 = 18.6 per 1,000 persons or 16.5 percent. Similar estimates of effects of procedural errors on hospital discharge data can be determined from the tables for specific age-sex groups. Although input parameters for this study were based in part on empirical data, the specific output estimates of underreporting should be considered illustrative rather than necessarily a reflection of the situation which prevails in the Health Interview Survey.

Estimates of the percent underreporting of hospital discharges by number of weeks between discharge and interview for all discharges, deliveries only, and discharges excluding deliveries were computed for "interview reported" versus "perfect interview," "perfect interview" versus "all discharges," and "interview reported" versus "all discharges." These estimates are given in tables 7-9. A similar set of percent underreporting estimates was computed for hospital discharges by recall period and actual length of stay and are shown in tables 10-12.

# IV. RESULTS

## EVALUATION OF HOSPITAL EPISODES SIMULATION

Several aspects of the computer-generated hospital episode data were examined in order to evaluate the accuracy of the simulation. First, the generated distributions of the persons in each of the 12 age-sex groups by number of annual nondelivery episodes (perfect interview data) were compared with the expected distributions. With but minor exceptions, the computer simulation program generated distributions of the number of nondelivery episodes equivalent to the expected negative binomial distributions.

It is noted that, except for females 35-44 years of age, the expected frequencies of two or more episodes were higher than generated. This tendency on the low side could be due to inadequate representation of the upper tail of the gamma distribution of the weekly admission probabilities (i.e., the $\lambda$ values). It is possible that this aspect could be improved by subdividing the 10th subgroup in order to include $\lambda$ values corresponding, for example, to the 99th percentile. An alternative explanation of the observed deficiency of persons with two or more episodes is that the uniform random number subroutine, used in the computer program, failed to generate small random numbers in close order proximity as frequently as expected statistically.

The second aspect examined was a comparison of the generated annual discharge rates by age and sex, excluding deliveries, with the expected rates (table G). The sampling errors indicate that the differences in these rates are not statistically significant. The annual discharge rates generated by the computer for males and females 65 years and older are greater than the expected rates shown in table G since they include persons in their last year of life who were alive on the interview date (and hence subject to higher admission rates). The expected rates were not adjusted for the higher admission probabilities assigned to persons in their last year of life.

The Health Interview Survey annual discharge rates, excluding deliveries, reported for the period July 1958-June 1960 are higher than the expected rates for the computer simulation since the published rates are based on data reported for the most recent 6 months of the year prior to interview. On the other hand, the weekly admission probabilities were derived from unpublished Health Interview Survey data on the distribution of the population by number of annual nondelivery episodes based on reported experiences for the 12 months prior to interview.

The third aspect examined in evaluating the computer simulation of hospitalization histories was the distribution of persons in the hospital on the interview date by age in comparison with the unpublished Health Interview Survey distribution for the Sunday prior to interview. The data, given in Table H, show the two distributions to be in close agreement.

Fourth, the average length of stay in days by sex and age for the computer episodes (perfect interview data) are compared with the July 1958-June 1960 Health Interview Survey results in table J. Agreement, slightly better for females than males, is fairly good. The sample size (episodes) for males 15-24, 25-34, and 35-44 years of age, is only about 30 for each of these age classes, accounting in part for the variability observed in their length-of-stay averages.

The distribution of the generated lengths of stay has not been tabulated in detail. However, the distribution for 1-day, 2-4-day, and 5-or-more-day stays is available from table 10. This distribution is compared with the distribution

Table G. Comparison of computer generated and expected number of nondelivery episodes per 1,000 persons per year, and simulated population base and standard deviation of observed rate, by sex and age

| Sex and age | Observed number[1] | Expected number | Simulated population base | Standard deviation of observed rate |
|---|---|---|---|---|
| Male | | | | |
| Under 15 years | 64.9 | 62.1 | 1,740 | 5.20 |
| 15-24 years | 58.7 | 62.2 | 608 | 8.77 |
| 25-34 years | 54.0 | 64.5 | 613 | 8.87 |
| 35-44 years | 68.5 | 75.1 | 641 | 9.44 |
| 45-64 years | 102.0 | 100.1 | 965 | 9.04 |
| 65+ years | 168.1 | 142.0 | 345 | 18.11 |
| Female | | | | |
| Under 15 years | 51.5 | 50.2 | 1,675 | 4.70 |
| 15-24 years | 97.8 | 95.0 | 683 | 10.95 |
| 25-34 years | 105.8 | 125.8 | 669 | 12.85 |
| 35-44 years | 122.3 | 112.4 | 695 | 11.22 |
| 45-64 years | 105.0 | 97.4 | 1,043 | 8.63 |
| 65+ years | 135.2 | 119.3 | 429 | 14.92 |

[1]The observed rates are inflated slightly by the experience of persons in their last year of life. These persons are not included in the expected number.

Table H. Number and percent distribution of persons in hospital on day of interview, by age: computer simulation[1] versus Health Interview Survey[2]

| Age | Computer simulation | | Health Interview Survey | |
|---|---|---|---|---|
| | Number | Percent distribution | Number in thousands | Percent distribution |
| All ages | 344 | 100.0 | 367 | 100.0 |
| Under 15 years | 43 | 12.5 | 48 | 13.1 |
| 15-24 years | 37 | 10.8 | 42 | 11.4 |
| 25-34 years | 40 | 11.6 | 43 | 11.7 |
| 35-44 years | 42 | 12.2 | 54 | 14.7 |
| 45-64 years | 110 | 32.0 | 106 | 28.9 |
| 65+ years | 72 | 20.9 | 74 | 20.2 |

[1]Total of incomplete episodes for 13 interview dates.
[2]Average number of persons in short-stay hospitals last Sunday night, United States, July 1959-June 1960.

Table J. Comparison of average length of stay in days, by sex and age: computer generated[1] versus Health Interview Survey[2].

| Sex and age | Computer generated | Health Interview Survey |
|---|---|---|
| | Length of stay in days | |
| **Male** | | |
| All ages----- | 10.1 | 10.5 |
| Under 15 years----- | 6.0 | 6.1 |
| 15-24 years-------- | 9.6 | 8.2 |
| 25-34 years-------- | 10.7 | 9.3 |
| 35-44 years-------- | 8.4 | 11.8 |
| 45-64 years-------- | 13.3 | 12.2 |
| 65+ years---------- | 13.7 | 15.9 |
| **Female** | | |
| All ages----- | 6.9 | 7.2 |
| Under 15 years----- | 5.6 | 5.8 |
| 15-24 years-------- | 4.4 | 4.5 |
| 25-34 years-------- | 4.6 | 5.2 |
| 35-44 years-------- | 6.6 | 6.7 |
| 45-64 years-------- | 10.9 | 11.4 |
| 65+ years---------- | 15.4 | 14.0 |

[1]Perfect interview data; average of 13 interview dates.
[2]See table 1, p. 14, in reference 8.

of discharge rates for these same length-of-stay groups as derived from unpublished July 1958-June 1960 Health Interview Survey data in table K. Agreement is quite good.

It seems clear from the above analysis that the hospital episodes simulation model and computer program are quite satisfactory. Further improvements, one of which has already been mentioned, are possible. It would be desirable that the various hospitalization statistics within age-sex groups generated by the computer have greater reliability than can be obtained with a population run of 10,000. The computer program should also be revised to permit individuals to shift over time from their initial age group to the next higher age group. This is particularly important for the two older age groups, as will be made clear from results discussed in later sections. For example, under the present program when 2-year histories are generated, the number of persons 65 years and older for the second year is reduced significantly due to deaths during the first year. The assignment of reasons for hospitalization within age-sex groups can be added to the computer program with relatively little difficulty. Length-of-stay distributions for each reason or condition would be more realistic if this change were made in the program.

Table K. Comparison of length-of-stay distributions: computer generated discharges[1] versus Health Interview Survey discharges[2]

| Length of stay | Computer generated discharges | | Health Interview Survey discharges | |
|---|---|---|---|---|
| | Number | Percent distribution | Rate per 1,000 persons | Percent distribution |
| Total---------- | 1,071.1 | 100.0 | 114.5 | 100.0 |
| 1 day---------- | 131.8 | 12.3 | 12.6 | 11.0 |
| 2-4 days---------- | 383.5 | 35.8 | 41.0 | 35.8 |
| 5+ days---------- | 555.8 | 51.9 | 60.9 | 53.2 |

[1]Perfect interview data; average of 13 interview dates.
[2]Unpublished data, July 1958-June 1960.

## EVALUATION OF INTERVIEW SIMULATION

The interview simulation model introduced errors due to failure to report hospital discharges which occurred in the year prior to interview, failure to report discharge dates accurately, and failure to report length of stay accurately. As discussed previously, the parameters for generating these errors were based largely on results obtained in the Michigan study. Percent underreporting of hospital discharges as generated by the computer is compared with the Michigan study data in table L separately by length of stay and by weeks between discharge and interview. As expected, since the assigned probabilities were based on these two factors, the generated results essentially reproduced the Michigan study data. A more detailed comparison of the computer-generated underreporting rates with the assigned rates jointly by length of stay and interval between discharge and interview is given in table M. As in table L, the generated underreporting rates include the effect of reporting the discharge date inaccurately. Thus, the computer overreported 2-4-day stays and 5-or-more-day stays for the 4-week period immediately prior to interview. The agreement between the observed and expected results in table M is fairly good, but not outstanding. The total number of episodes for each cell was not large for any one interviewing date, ranging from 10 for the 1-day stays to 30 for the 2-4-day stays and 40 for the 5-or-more-day stays. However, the generated results shown are averages for 13 interviewing dates, and hence are based on fairly substantial numbers of cases. The effect of inaccurately reported discharge dates may be responsible for the several instances of somewhat larger differences than expected.

The computer simulations of failure to report the discharge date and/or the length of stay accurately have not been evaluated in detail. As discussed in the next section, the net shifting of discharge dates by the computer was essentially negligible. The proportion of discharge dates reported accurately (i.e., within the same 4-week period as the actual discharge date) has not been determined. The average length of stay for the interview reported discharges was 0.3 of a day greater than for the perfect interview discharges, which agrees with the Michigan study.[2] The distributions of reported length of stay by actual length of stay have not been tabulated, however.

Table L. Percent underreporting of hospital discharges, by length of stay and number of weeks between discharge and interview: computer generated[1] versus Michigan study[2]

| Length of stay and weeks between discharge and interview | Computer generated | Michigan study |
|---|---|---|
| Length of stay | | |
| Total | 11.3 | 12.0 |
| 1 day | 23.2 | 26.0 |
| 2-4 days | 11.3 | 14.0 |
| 5+ days | 8.5 | 9.0 |
| Weeks between discharge and interview | | |
| Total | 11.3 | 12.0 |
| 1-20 weeks | 4.9 | 5.0 |
| 21-40 weeks | 10.7 | 9.0 |
| 41-52 weeks | 23.0 | 24.0 |

[1]Interview reported versus perfect interview; average of 13 interview dates. Includes errors in reported discharge dates.

[2]See table 15, p. 21, and table 40, p. 36, in reference 2.

Based on this limited evaluation, the interview simulation program appears to have been fairly successful. Further analysis is necessary before any suggestions regarding revisions in the model and computer program can be made.

## ESTIMATES OF SPECIFIC ERROR COMPONENTS

As mentioned in the introduction, a definite decreasing trend can be observed in the number of discharges reported in the Health Interview Survey when tabulated by month prior to interview. It is of considerable interest to determine the fac-

Table M. Percent underreporting of hospital discharges by actual length of stay and number of weeks between hospital discharge and interview: computer generated[1] versus assigned rates[2]

| Weeks between discharge and interview | 1-day stay | | 2-4-days stay | | 5+-days stay | |
|---|---|---|---|---|---|---|
| | Computer generated | Assigned rate | Computer generated | Assigned rate | Computer generated | Assigned rate |
| Total | 23.2 | 24.8 | 11.3 | 15.6 | 8.5 | 9.8 |
| 1-4 weeks | 3.2 | 7.0 | [3]0.3 | 4.0 | [3]0.5 | 1.0 |
| 5-8 weeks | 16.3 | 13.0 | 4.9 | 5.0 | 2.3 | 2.0 |
| 9-12 weeks | 20.0 | 18.0 | 2.8 | 6.0 | 1.2 | 4.0 |
| 13-16 weeks | 21.6 | 22.0 | 3.1 | 7.0 | 5.2 | 5.0 |
| 17-20 weeks | 18.8 | 24.0 | 11.4 | 8.0 | 5.1 | 6.0 |
| 21-24 weeks | 23.5 | 26.0 | 4.4 | 9.0 | 4.9 | 7.0 |
| 25-28 weeks | 22.8 | 28.0 | 10.3 | 11.0 | 9.1 | 8.0 |
| 29-32 weeks | 31.0 | 29.0 | 8.1 | 14.0 | 9.6 | 9.0 |
| 33-36 weeks | 25.5 | 30.0 | 13.3 | 18.0 | 10.1 | 9.0 |
| 37-40 weeks | 32.1 | 30.0 | 11.1 | 22.0 | 5.0 | 10.0 |
| 41-44 weeks | 20.4 | 31.0 | 21.7 | 27.0 | 6.7 | 10.0 |
| 45-48 weeks | 31.5 | 32.0 | 26.4 | 33.0 | 15.0 | 11.0 |
| 49-52 weeks | 32.4 | 32.0 | 29.0 | 39.0 | 37.2 | 46.0 |

[1]Interview reported versus perfect interview; average of 13 interview dates. Includes errors in reported discharge dates (see table 10).
[2]Nondelivery episodes only.
[3]Percent overreported.

tors contributing to this decay curve and the magnitude of their respective effects. Accordingly, estimates have been derived of the component parts of the discrepancy between the interview reported discharges and all discharges in 4-week intervals prior to interview, using the computer generated hospital episode and interview simulation data. These estimates are given in absolute numbers of discharges (average of 13 interview dates) and also as a percent of all discharges in each of the 13 four-week periods in the year prior to interview in table N. The average estimates for 12, 24, 36, and 52 weeks prior to interview are also shown in this table.

The observed decay curve is shown in the column headed "interview reported." The discrepancy (i.e., all discharges less interview reported discharges) increases as the interval between discharge and interview increases, as does the number of not reported discharges and also the number of discharges of persons who died in the year prior to interview (all discharges less perfect interview discharges). The error component due to shifting of discharge dates fluctuates from positive (back in time) to negative (forward in time), but remains at a fairly low level; the average of this component is essentially zero for the year prior to interview.

It is clear that the number of discharges of persons who died in the year prior to interview should increase as the interval between discharge and interview increases, since this group is somewhat larger numerically at the beginning of the year of interest and decreases in size as the interview date is approached. This might suggest that the total number of discharges should also increase as the interval between discharge and interview increases. This is incorrect, although the average of the generated "all discharges" over the 13 interview dates does exhibit this incorrect relationship in table N and also in table 8. This error is due to the unfortunate oversight of failing to age the population in the computer simulation program. Since the living population is aging and also increasing in size during the year and since the number of persons living on the date

Table N. Estimated contribution of error components to discrepancy between interview reported and all discharges, by number of weeks between discharge and interview

[Average of 13 interview dates]

| Weeks between discharge and interview | All discharges | Perfect interview | Interview reported | Discrepancy: all less interview reported | Not reported | All discharges less perfect interview[1] | Net shifting of discharge date[2] |
|---|---|---|---|---|---|---|---|
| | Number of discharges | | | | | | |
| 1-4 weeks------ | 85.5 | 82.2 | 82.2 | 3.3 | 1.5 | 3.3 | -1.5 |
| 5-8 weeks------ | 86.2 | 81.8 | 77.8 | 8.4 | 3.8 | 4.4 | 0.2 |
| 9-12 weeks----- | 86.5 | 81.7 | 78.4 | 8.1 | 4.2 | 4.8 | -0.9 |
| 13-16 weeks---- | 88.6 | 83.5 | 78.1 | 10.5 | 5.3 | 5.1 | 0.1 |
| 17-20 weeks---- | 88.1 | 82.3 | 74.9 | 13.2 | 5.8 | 5.8 | 1.6 |
| 21-24 weeks---- | 87.6 | 82.1 | 76.3 | 11.3 | 7.2 | 5.5 | -1.4 |
| 25-28 weeks---- | 87.8 | 82.2 | 73.0 | 14.8 | 8.2 | 5.6 | 1.0 |
| 29-32 weeks---- | 88.3 | 82.4 | 72.8 | 15.5 | 9.5 | 5.9 | 0.1 |
| 33-36 weeks---- | 88.1 | 82.1 | 71.2 | 16.9 | 9.9 | 6.0 | 1.0 |
| 37-40 weeks---- | 89.1 | 82.7 | 73.8 | 15.3 | 10.4 | 6.4 | -1.5 |
| 41-44 weeks---- | 89.3 | 82.9 | 71.4 | 17.9 | 11.3 | 6.4 | 0.2 |
| 45-48 weeks---- | 89.0 | 82.5 | 64.8 | 24.2 | 16.2 | 6.5 | 1.5 |
| 49-52 weeks---- | 89.4 | 82.7 | 54.8 | 34.6 | 28.5 | 6.7 | -0.6 |
| Average estimate for: | | | | | | | |
| 1-12 weeks----- | 86.1 | 81.9 | 79.5 | 6.6 | 3.2 | 4.1 | -0.7 |
| 1-24 weeks----- | 87.1 | 82.3 | 78.0 | 9.1 | 4.6 | 4.8 | -0.3 |
| 1-36 weeks----- | 87.4 | 82.3 | 76.1 | 11.3 | 6.1 | 5.2 | 0.02 |
| 1-52 weeks----- | 88.0 | 82.4 | 73.0 | 15.0 | 9.4 | 5.6 | -0.02 |
| | Percent distribution of all discharges | | | | | | |
| 1-4 weeks------ | 100.0 | 96.1 | 96.1 | 3.9 | 1.8 | 3.9 | -1.8 |
| 5-8 weeks------ | 100.0 | 94.9 | 90.3 | 9.7 | 4.4 | 5.1 | 0.2 |
| 9-12 weeks----- | 100.0 | 94.5 | 90.6 | 9.4 | 4.9 | 5.5 | -1.0 |
| 13-16 weeks---- | 100.0 | 94.2 | 88.1 | 11.9 | 6.0 | 5.8 | 0.1 |
| 17-20 weeks---- | 100.0 | 93.4 | 85.0 | 15.0 | 6.6 | 6.6 | 1.8 |
| 21-24 weeks---- | 100.0 | 93.7 | 87.1 | 12.9 | 8.2 | 6.3 | -1.6 |
| 25-28 weeks---- | 100.0 | 93.6 | 83.1 | 16.9 | 9.3 | 6.4 | 1.2 |
| 29-32 weeks---- | 100.0 | 93.3 | 82.4 | 17.6 | 10.8 | 6.7 | 0.1 |
| 33-36 weeks---- | 100.0 | 93.2 | 80.8 | 19.2 | 11.2 | 6.8 | 1.2 |
| 37-40 weeks---- | 100.0 | 92.8 | 82.8 | 17.2 | 11.7 | 7.2 | -1.7 |
| 41-44 weeks---- | 100.0 | 92.8 | 80.0 | 20.0 | 12.6 | 7.2 | 0.2 |
| 45-48 weeks---- | 100.0 | 92.7 | 72.8 | 27.2 | 18.2 | 7.3 | 1.7 |
| 49-52 weeks---- | 100.0 | 92.5 | 61.3 | 38.7 | 31.9 | 7.5 | -0.7 |
| Average estimate for: | | | | | | | |
| 1-12 weeks----- | 100.0 | 95.2 | 92.3 | 7.7 | 3.8 | 4.8 | -0.9 |
| 1-24 weeks----- | 100.0 | 94.5 | 89.5 | 10.5 | 5.4 | 5.5 | -0.4 |
| 1-36 weeks----- | 100.0 | 94.1 | 87.1 | 12.9 | 7.0 | 5.9 | 0.02 |
| 1-52 weeks----- | 100.0 | 93.7 | 82.9 | 17.1 | 10.7 | 6.4 | -0.02 |

[1]Discharges of persons who died during the year prior to interview.
[2]A negative value means discharge date shifted forward in time.

of interview, but already in their last year of life, is somewhat larger on the date of interview than at the beginning of the reference year, the number of discharges of persons alive on the interview date (perfect interview discharges) should decrease as the time interval between discharge and interview increases. This is the key phenomenon previously stated in the introduction. Hence "all discharges" should either decrease or remain constant as the interval between discharge and interview increases.

The computer incorrectly generated a relatively constant monthly number of discharges during the reference year for persons alive on the interview date (perfect interview discharges), at least on the average for the 13 interview dates (see table N), because persons 65 years and older who died were not replaced by new persons from the 45-64 year age group. This reduced the 65 years and over age group over time. The number of discharges of living persons was reduced from 1,088 in the year prior to the first interview date to 1,049 in the year prior to the last interview date. Similarly, the number of all discharges was reduced from 1,162 in the year prior to the first interview date to 1,111 in the year prior to the last interview date. Without these decreases (which should not have occurred) the total number of discharges by weeks between discharge and interview would have remained approximately constant and the number of discharges among persons living on the date of interview would have decreased with increasing time interval between discharge and interview.

While the average levels shown in table N (and in table 8) for all discharges, perfect interview discharges, and interview reported discharges are not correct as to level, the estimates of the error components and of the discrepancy itself are considered satisfactory. This should be clear, since the weaknesses in the generation model tend to be compensating when the discrepancy and its components are computed.

Table N shows the underestimate of all discharges from an interview procedure using data reported for the entire reference year to be 17.1 percent. If only the data reported for the 24 weeks (approximately 6 months) immediately prior to interview are used, the underestimate of all discharges is reduced to 10.5 percent. The major source of this reduction is the not reported error component which is cut in half (5.4 versus 10.7 percent). It is of interest to note that, even if no response errors were made, the number of reported discharges in the interview is estimated to be lower than all discharges by approximately 4 percent if reporting is confined to the 4 weeks immediately prior to interview and 6.4 percent when reporting for the year prior to interview.

## METHODS FOR INCREASING ACCURACY

Inspection of tables 1-4 shows that the average annual hospital discharges and hospital days for persons alive on the interview date within each age-sex group are underestimated by approximately 11 percent when a procedure using all data reported for the 12 months prior to interview is employed. The estimates are improved when they are based only on the episodes with reported discharge dates occurring in the most recent 6 months prior to interview. The generated data have not been tabulated on this basis so that the improvement for each of the age-sex groups has not been ascertained. However, the average underestimate is reduced by a factor of two, approximately, with this procedure. It is doubtful that basing the estimates of interest only on hospitalizations reported within a shorter time interval than 6 months between interview and discharge would be economically efficient. Apparently it is possible to further increase accuracy by use of Procedure B as reported in the study by the University of Michigan in which three alternative survey procedures were compared.[3] The relative biases in the average annual number of discharges and hospital days by age and sex with this procedure can be estimated by means of the interview simulation program on the computer. The program would require a set of parameters (i.e., probabilities of failure to report the episode, etc.) appropriate to Procedure B. Apparently, the data for estimating these parameters are available from the study which compared Procedure B with the standard procedure used in this project.

Further improvement in the accuracy of the hospital statistics based on the Health Interview Survey through changes in the interview procedure is doubtful. A method of adjusting the survey statistics is necessary. One such method, discussed briefly in the introductory section, uses the J-analysis technique of Simmons and Bryant to derive inflation factors by which reported hospital discharges are weighted to estimate total actual discharges, including those of persons not alive on the interview date. Because of limited time, evaluation of the Simmons and Bryant approach by means of the generated data was not carried out.

Estimation of inflation factors to improve the accuracy of published hospital statistics based on the Health Interview Survey appears both feasible and desirable. Using the observed data to derive the adjustment factors has considerable appeal. It seems advisable to explore alternative methods of estimating adjustment factors using simulation models.

# V. CONCLUSIONS

A probability model for generating hospital admissions and duration of stay for the U.S. population together with an IBM 1410 computer program for simulation of hospitalization histories under the model were developed in this project. The simulation program was carried out for an initial population of 10,000 individuals for a period of 108 weeks; while the results were judged very satisfactory, there is room for improvement in several aspects. These are:

> Estimation of weekly admission probabilities should, at the very minimum, be based on data obtained in the Health Interview Survey for the most recent 6 months prior to interview. These probabilities should be improved further by appropriate adjustment of the observed episodes distributions to reflect all hospitalizations rather than reported hospitalizations.
>
> The estimated daily admission probabilities for persons in their last year of life were based on sketchy data and should be improved, using data obtained from a national study.
>
> The simulation program should permit individuals in specific age-sex groups to shift to the next older group over time. This is particularly essential for the 45-64 and 65 years and over age groups, since deaths reduce these groups significantly over time if the population is age-static. This could be accomplished, with relatively little change in the existing program, by adding an age-shifting date to be treated in a manner similar to the birth and death dates already in the program.
>
> Reasons for hospitalization should be included in the program, to be assigned on a probability basis, provided sufficient data are available for developing length-of-stay distributions by reason.

A probability model and computer program for simulating interview data on hospital episodes as collected in the Health Interview Survey were also developed in this project. The computer program was carried out for 13 interview dates 28 days apart using the data generated by the hospital episodes simulation program as input. The generated interview data were also judged satisfactory, providing estimates of the relative biases due to measurement errors for each of the principal hospitalization statistics obtained in the Health Interview Survey. It is noted that the estimated relative biases are fairly substantial.

The interview simulation model was not analyzed intensively, due to limited time available to complete this project. The parameters associated with errors in reporting length of stay and discharge date are considered conservative. Further study and analysis is necessary before any suggestions on revisions in the model and computer program can be made.

It is doubtful that further significant reductions in the measurement errors of hospitalization data collected in the Health Interview Survey are possible without adding unduly to the cost. The survey design suggests that satisfactory adjustment factors can be estimated from the collected data. The simulation models and computer programs developed in this project provide a useful research tool for studying alternative methods of adjustment.

The computer program for generating hospitalization histories is essentially a program for distributing episodes in the population consistent with the negative binomial distribution. Hence, it should be useful, with but minor revisions, for simulating the distributions of other events which have been observed to be negative binomial. These include, for example, the distribution of the population by number of colds annually and by number of doctor visits annually. Undoubtedly there are other health variables in this class.

The hospital episodes computer program, revised as suggested, should also be useful for studies of the effects on the demand for hospital beds of trends in such variables as age, sex, reasons for hospitalization, and duration of stay.

## REFERENCES

[1]National Center for Health Statistics: Hospital utilization in the last year of life. *Vital and Health Statistics*. PHS Pub. No. 1000-Series 2-No. 10. Public Health Service. Washington. U.S. Government Printing Office, July 1965.

[2]National Center for Health Statistics: Reporting of hospitalization in the Health Interview Survey. *Vital and Health Statistics*. PHS Pub. No. 1000-Series 2-No. 6. Public Health Service. Washington. U.S. Government Printing Office, July 1965.

[3]National Center for Health Statistics: Comparison of hospitalization reporting in three survey procedures. *Vital and Health Statistics*. PHS Pub. No. 1000-Series 2-No. 8. Public Health Service. Washington. U.S. Government Printing Office, July 1965.

[4]U.S. National Health Survey: The statistical design of the Health Household-Interview Survey. *Health Statistics*. PHS Pub. No. 584-A2. Public Health Service. Washington. U.S. Government Printing Office, July 1958.

[5]Simmons, Walt R., and Bryant, E. E.: An evaluation of hospitalization data from the Health Interview Survey. *Am.J. Pub.Health* 52(10):1638-1647, Oct. 1962.

[6]National Center for Health Statistics: An index of health, mathematical models. *Vital and Health Statistics*. PHS Pub. No. 1000-Series 2-No. 5. Public Health Service. Washington. U.S. Government Printing Office, May 1965.

[7]Pearson, K., ed: *Tables of the incomplete* $\Gamma$*-function*. Cambridge, England. Cambridge University Press, 1957 printing of original 1922 edition.

[8]U.S. National Health Survey: Hospital discharges, United States, 1958-1960. *Health Statistics*. PHS Pub. No. 584-B32. Public Health Service. Washington. U.S. Government Printing Office, Apr. 1962.

# DETAILED TABLES

Table 1. Average annual number, number per 1,000 persons, and percent distribution of patients discharged in year prior to interview for each of three types of simulation, by sex and age

[Average of 13 interview dates]

| Sex and age | For living persons | | | | | | All discharges | | |
|---|---|---|---|---|---|---|---|---|---|
| | Interview reported discharges | | | Perfect interview discharges | | | | | |
| | Number | Number per 1,000 persons | Percent distribution | Number | Number per 1,000 persons | Percent distribution | Number | Number per 1,000 persons | Percent distribution |
| Both sexes | | | | | | | | | |
| All ages-- | 949.6 | 94.0 | 100.0 | 1,071.0 | 106.0 | 100.0 | 1,143.4 | 112.6 | 100.0 |
| Under 15 years-- | 167.0 | 48.9 | 17.6 | 199.2 | 58.3 | 18.6 | 202.8 | 59.3 | 17.7 |
| 15-24 years----- | 176.8 | 136.9 | 18.6 | 196.6 | 152.3 | 18.4 | 196.8 | 152.3 | 17.2 |
| 25-34 years----- | 186.2 | 145.2 | 19.6 | 201.4 | 157.1 | 18.8 | 201.9 | 157.4 | 17.7 |
| 35-44 years----- | 133.0 | 99.6 | 14.0 | 149.9 | 112.2 | 14.0 | 157.6 | 117.7 | 13.8 |
| 45-64 years----- | 183.4 | 91.3 | 19.3 | 207.9 | 103.5 | 19.4 | 221.1 | 109.5 | 19.3 |
| 65+ years------- | 103.2 | 133.3 | 10.9 | 116.0 | 149.9 | 10.8 | 163.2 | 203.5 | 14.3 |
| Male | | | | | | | | | |
| All ages-- | 332.9 | 67.8 | 100.0 | 382.1 | 77.8 | 100.0 | 421.7 | 85.4 | 100.0 |
| Under 15 years-- | 95.5 | 54.9 | 28.7 | 113.0 | 64.9 | 29.6 | 116.3 | 66.8 | 27.6 |
| 15-24 years----- | 29.5 | 48.5 | 8.9 | 35.7 | 58.7 | 9.3 | 35.7 | 58.6 | 8.5 |
| 25-34 years----- | 30.2 | 49.3 | 9.1 | 33.1 | 54.0 | 8.7 | 33.1 | 53.9 | 7.8 |
| 35-44 years----- | 39.2 | 61.2 | 11.8 | 43.9 | 68.5 | 11.5 | 47.3 | 73.7 | 11.2 |
| 45-64 years----- | 87.3 | 90.5 | 26.2 | 98.4 | 102.0 | 25.8 | 104.9 | 107.9 | 24.9 |
| 65+ years------- | 51.2 | 148.4 | 15.3 | 58.0 | 168.1 | 15.1 | 84.4 | 235.1 | 20.0 |
| Female | | | | | | | | | |
| All ages-- | 616.7 | 118.7 | 100.0 | 688.9 | 132.6 | 100.0 | 721.7 | 138.4 | 100.0 |
| Under 15 years-- | 71.5 | 42.7 | 11.6 | 86.2 | 51.5 | 12.5 | 86.5 | 51.6 | 12.0 |
| 15-24 years----- | 147.3 | 215.7 | 23.9 | 160.9 | 235.6 | 23.4 | 161.1 | 235.9 | 22.3 |
| 25-34 years----- | 156.0 | 233.2 | 25.3 | 168.3 | 251.6 | 24.4 | 168.8 | 251.9 | 23.4 |
| 35-44 years----- | 93.8 | 135.0 | 15.2 | 106.0 | 152.5 | 15.4 | 110.3 | 158.5 | 15.3 |
| 45-64 years----- | 96.1 | 92.1 | 15.6 | 109.5 | 105.0 | 15.9 | 116.2 | 111.0 | 16.1 |
| 65+ years------- | 52.0 | 121.2 | 8.4 | 58.0 | 135.2 | 8.4 | 78.8 | 177.9 | 10.9 |

Table 2. Average annual number, number per 1,000 persons, and percent distribution of patients discharged in year prior to interview, excluding deliveries, for each of three types of simulation, by sex and age

[Average of 13 interview dates]

| Sex and age | For living persons | | | | | | All discharges | | |
|---|---|---|---|---|---|---|---|---|---|
| | Interview reported discharges | | | Perfect interview discharges | | | | | |
| | Number | Number per 1,000 persons | Percent distribution | Number | Number per 1,000 persons | Percent distribution | Number | Number per 1,000 persons | Percent distribution |
| | Excluding deliveries | | | | | | | | |
| Both sexes | | | | | | | | | |
| All ages-- | 741.4 | 73.4 | 100.0 | 858.4 | 84.9 | 100.0 | 930.0 | 91.6 | 100.0 |
| Under 15 years-- | 167.0 | 48.9 | 22.5 | 199.2 | 58.3 | 23.2 | 202.8 | 59.3 | 21.8 |
| 15-24 years----- | 84.9 | 65.8 | 11.5 | 102.5 | 79.4 | 11.9 | 102.6 | 79.4 | 11.0 |
| 25-34 years----- | 90.1 | 70.3 | 12.2 | 103.9 | 81.0 | 12.1 | 104.5 | 81.4 | 11.2 |
| 35-44 years----- | 112.8 | 84.4 | 15.2 | 128.9 | 96.5 | 15.0 | 135.8 | 101.4 | 14.6 |
| 45-64 years----- | 183.4 | 91.3 | 24.7 | 207.9 | 103.5 | 24.2 | 221.1 | 109.5 | 23.8 |
| 65+ years------- | 103.2 | 133.3 | 13.9 | 116.0 | 149.9 | 13.6 | 163.2 | 203.5 | 17.6 |
| Male | | | | | | | | | |
| All ages-- | 332.9 | 67.8 | 100.0 | 382.1 | 77.8 | 100.0 | 421.7 | 85.4 | 100.0 |
| Under 15 years-- | 95.5 | 54.9 | 28.7 | 113.0 | 64.9 | 29.6 | 116.3 | 66.8 | 27.6 |
| 15-24 years----- | 29.5 | 48.5 | 8.9 | 35.7 | 58.7 | 9.3 | 35.7 | 58.6 | 8.5 |
| 25-34 years----- | 30.2 | 49.3 | 9.1 | 33.1 | 54.0 | 8.7 | 33.1 | 53.9 | 7.8 |
| 35-44 years----- | 39.2 | 61.2 | 11.8 | 43.9 | 68.5 | 11.5 | 47.3 | 73.7 | 11.2 |
| 45-64 years----- | 87.3 | 90.5 | 26.2 | 98.4 | 102.0 | 25.8 | 104.9 | 107.9 | 24.9 |
| 65+ years------- | 51.2 | 148.4 | 15.3 | 58.0 | 168.1 | 15.1 | 84.4 | 235.1 | 20.0 |
| Female | | | | | | | | | |
| All ages-- | 408.5 | 78.6 | 100.0 | 476.3 | 91.7 | 100.0 | 508.3 | 97.5 | 100.0 |
| Under 15 years-- | 71.5 | 42.7 | 17.5 | 86.2 | 51.5 | 18.1 | 86.5 | 51.6 | 17.0 |
| 15-24 years----- | 55.4 | 81.1 | 13.6 | 66.8 | 97.8 | 14.0 | 66.9 | 98.0 | 13.2 |
| 25-34 years----- | 59.9 | 89.5 | 14.7 | 70.8 | 105.8 | 14.9 | 71.4 | 106.6 | 14.0 |
| 35-44 years----- | 73.6 | 105.9 | 18.0 | 85.0 | 122.3 | 17.8 | 88.5 | 127.2 | 17.4 |
| 45-64 years----- | 96.1 | 92.1 | 23.5 | 109.5 | 105.0 | 23.0 | 116.2 | 111.0 | 22.9 |
| 65+ years------- | 52.0 | 121.2 | 12.7 | 58.0 | 135.2 | 12.2 | 78.8 | 177.9 | 15.5 |

Table 3. Average annual number, days per 1,000 persons, and percent distribution of hospital days in year prior to interview for each of three types of simulation, by sex and age

[Average of 13 interview dates]

| Sex and age | For living persons | | | | | | All discharges | | |
|---|---|---|---|---|---|---|---|---|---|
| | Interview reported | | | Perfect interview | | | | | |
| | Number of days | Days per 1,000 persons | Percent distribution | Number of days | Days per 1,000 persons | Percent distribution | Number of days | Days per 1,000 persons | Percent distribution |
| | Hospital days | | | | | | | | |
| Both sexes | | | | | | | | | |
| All ages-- | 7,917.1 | 783.3 | 100.0 | 8,604.6 | 851.4 | 100.0 | 9,303.4 | 916.1 | 100.0 |
| Under 15 years-- | 1,066.3 | 312.1 | 13.5 | 1,164.4 | 340.9 | 13.5 | 1,186.2 | 346.8 | 12.8 |
| 15-24 years----- | 992.6 | 768.9 | 12.5 | 1,057.1 | 818.8 | 12.4 | 1,057.6 | 818.6 | 11.3 |
| 25-34 years----- | 1,082.0 | 844.0 | 13.7 | 1,133.9 | 884.5 | 13.2 | 1,135.2 | 884.8 | 12.2 |
| 35-44 years----- | 988.9 | 740.2 | 12.5 | 1,068.1 | 799.5 | 12.4 | 1,105.6 | 825.6 | 11.9 |
| 45-64 years----- | 2,271.2 | 1,131.1 | 28.7 | 2,496.1 | 1,243.1 | 29.0 | 2,678.3 | 1,326.5 | 28.8 |
| 65+ years------- | 1,516.1 | 1,958.8 | 19.1 | 1,685.0 | 2,177.0 | 19.5 | 2,140.5 | 2,669.0 | 23.0 |
| Male | | | | | | | | | |
| All ages-- | 3,497.4 | 711.9 | 100.0 | 3,844.8 | 782.6 | 100.0 | 4,238.2 | 857.9 | 100.0 |
| Under 15 years-- | 622.5 | 357.8 | 17.8 | 681.4 | 391.6 | 17.7 | 702.2 | 403.1 | 16.5 |
| 15-24 years----- | 303.7 | 499.5 | 8.7 | 342.9 | 564.0 | 8.9 | 342.9 | 563.1 | 8.1 |
| 25-34 years----- | 335.2 | 546.8 | 9.6 | 352.7 | 575.4 | 9.2 | 352.7 | 574.4 | 8.3 |
| 35-44 years----- | 337.0 | 525.7 | 9.6 | 368.1 | 574.3 | 9.6 | 392.7 | 611.7 | 9.3 |
| 45-64 years----- | 1,186.5 | 1,229.5 | 33.9 | 1,305.2 | 1,352.5 | 34.0 | 1,360.2 | 1,399.4 | 32.1 |
| 65+ years------- | 712.5 | 2,065.2 | 20.4 | 794.5 | 2,302.9 | 20.6 | 1,087.5 | 3,029.2 | 25.5 |
| Female | | | | | | | | | |
| All ages-- | 4,419.7 | 850.9 | 100.0 | 4,759.8 | 916.4 | 100.0 | 5,065.2 | 971.1 | 100.0 |
| Under 15 years-- | 443.8 | 265.0 | 10.1 | 483.0 | 288.4 | 10.2 | 484.0 | 288.6 | 9.6 |
| 15-24 years----- | 688.9 | 1,008.6 | 15.6 | 714.2 | 1,045.7 | 15.0 | 714.7 | 1,046.4 | 14.1 |
| 25-34 years----- | 746.8 | 1,116.3 | 16.9 | 781.2 | 1,167.7 | 16.4 | 782.5 | 1,167.9 | 15.4 |
| 35-44 years----- | 651.9 | 938.0 | 14.7 | 700.0 | 1,007.2 | 14.7 | 712.9 | 1,024.3 | 14.1 |
| 45-64 years----- | 1,084.7 | 1,040.0 | 24.5 | 1,190.9 | 1,141.8 | 25.0 | 1,318.1 | 1,259.0 | 26.0 |
| 65+ years------- | 803.6 | 1,873.2 | 18.2 | 890.5 | 2,075.8 | 18.7 | 1,053.0 | 2,377.0 | 20.8 |

Table 4. Average annual number, days per 1,000 persons, and percent distribution of hospital days in year prior to interview, excluding deliveries, for three types of simulation, by sex and age

[Average of 13 interview dates]

| Sex and age | For living persons | | | | | | All discharges | | |
|---|---|---|---|---|---|---|---|---|---|
| | Interview reported | | | Perfect interview | | | | | |
| | Number of days | Days per 1,000 persons | Percent distribution | Number of days | Days per 1,000 persons | Percent distribution | Number of days | Days per 1,000 persons | Percent distribution |
| | Hospital days excluding deliveries | | | | | | | | |
| **Both sexes** | | | | | | | | | |
| All ages-- | 7,042.0 | 696.7 | 100.0 | 7,740.9 | 765.9 | 100.0 | 8,439.7 | 831.1 | 100.0 |
| Under 15 years-- | 1,066.3 | 312.1 | 15.1 | 1,164.4 | 340.9 | 15.0 | 1,186.2 | 346.8 | 14.1 |
| 15-24 years----- | 618.5 | 479.1 | 8.8 | 688.9 | 533.6 | 8.9 | 689.4 | 533.6 | 8.2 |
| 25-34 years----- | 698.7 | 545.0 | 9.9 | 756.4 | 590.0 | 9.8 | 757.7 | 590.6 | 9.0 |
| 35-44 years----- | 871.2 | 652.1 | 12.4 | 950.1 | 711.2 | 12.3 | 987.6 | 737.6 | 11.7 |
| 45-64 years----- | 2,271.2 | 1,131.1 | 32.3 | 2,496.1 | 1,243.1 | 32.2 | 2,678.3 | 1,326.5 | 31.7 |
| 65+ years------- | 1,516.1 | 1,958.8 | 21.5 | 1,685.0 | 2,177.0 | 21.8 | 2,140.5 | 2,669.0 | 25.3 |
| **Male** | | | | | | | | | |
| All ages-- | 3,497.4 | 711.9 | 100.0 | 3,844.8 | 782.6 | 100.0 | 4,238.2 | 857.9 | 100.0 |
| Under 15 years-- | 622.5 | 357.8 | 17.8 | 681.4 | 391.6 | 17.7 | 702.2 | 403.1 | 16.6 |
| 15-24 years----- | 303.7 | 499.5 | 8.7 | 342.9 | 564.0 | 8.9 | 342.9 | 563.1 | 8.1 |
| 25-34 years----- | 335.2 | 546.8 | 9.6 | 352.7 | 575.4 | 9.2 | 352.7 | 574.4 | 8.3 |
| 35-44 years----- | 337.0 | 525.7 | 9.6 | 368.1 | 574.3 | 9.6 | 392.7 | 611.7 | 9.3 |
| 45-64 years----- | 1,186.5 | 1,229.5 | 33.9 | 1,305.2 | 1,352.5 | 34.0 | 1,360.2 | 1,399.4 | 32.1 |
| 65+ years------- | 712.5 | 2,065.2 | 20.4 | 794.5 | 2,302.9 | 20.6 | 1,087.5 | 3,029.2 | 25.6 |
| **Female** | | | | | | | | | |
| All ages-- | 3,544.6 | 682.4 | 100.0 | 3,896.1 | 750.1 | 100.0 | 4,201.5 | 805.5 | 100.0 |
| Under 15 years-- | 443.8 | 265.0 | 12.5 | 483.0 | 288.4 | 12.4 | 484.0 | 288.6 | 11.5 |
| 15-24 years----- | 314.8 | 460.9 | 8.9 | 346.0 | 506.6 | 8.9 | 346.5 | 507.3 | 8.2 |
| 25-34 years----- | 363.5 | 543.3 | 10.3 | 403.7 | 603.4 | 10.4 | 405.0 | 604.5 | 9.6 |
| 35-44 years----- | 534.2 | 768.6 | 15.1 | 582.0 | 837.4 | 14.9 | 594.9 | 854.7 | 14.2 |
| 45-64 years----- | 1,084.7 | 1,040.0 | 30.6 | 1,190.9 | 1,141.8 | 30.6 | 1,318.1 | 1,258.9 | 31.4 |
| 65+ years------- | 803.6 | 1,873.2 | 22.6 | 890.5 | 2,075.8 | 22.8 | 1,053.0 | 2,377.0 | 25.1 |

Table 5. Average length of stay in days for each of three types of simulation, by sex and age

[Average of 13 interview dates]

| Sex and age | For living persons | | | | | | All discharges | | |
|---|---|---|---|---|---|---|---|---|---|
| | Interview reported | | | Perfect interview | | | | | |
| | Number of hospital days | Number of discharges | Average length of stay | Number of hospital days | Number of discharges | Average length of stay | Number of hospital days | Number of discharges | Average length of stay |
| **Both sexes** | | | | | | | | | |
| All ages-- | 7,917.1 | 949.6 | 8.3 | 8,604.6 | 1,071.0 | 8.0 | 9,303.4 | 1,143.4 | 8.1 |
| Under 15 years-- | 1,066.3 | 167.0 | 6.4 | 1,164.4 | 199.2 | 5.8 | 1,186.2 | 202.8 | 5.8 |
| 15-24 years----- | 992.6 | 176.8 | 5.6 | 1,066.9 | 196.6 | 5.4 | 1,057.6 | 196.8 | 5.4 |
| 25-34 years----- | 1,082.0 | 186.2 | 5.8 | 1,133.9 | 201.4 | 5.6 | 1,135.2 | 201.9 | 5.6 |
| 35-44 years----- | 988.9 | 133.0 | 7.4 | 1,068.1 | 149.9 | 7.1 | 1,105.6 | 157.6 | 7.0 |
| 45-64 years----- | 2,271.2 | 183.4 | 12.4 | 2,496.1 | 207.9 | 12.0 | 2,678.3 | 221.1 | 12.1 |
| 65+ years------- | 1,516.1 | 103.2 | 14.7 | 1,685.0 | 116.0 | 14.5 | 2,140.5 | 163.2 | 13.1 |
| **Male** | | | | | | | | | |
| All ages-- | 3,497.4 | 332.9 | 10.5 | 3,844.8 | 382.1 | 10.1 | 4,238.2 | 421.7 | 10.1 |
| Under 15 years-- | 622.5 | 95.5 | 6.5 | 681.4 | 113.0 | 6.0 | 702.2 | 116.3 | 6.0 |
| 15-24 years----- | 303.7 | 29.5 | 10.3 | 342.9 | 35.7 | 9.6 | 342.9 | 35.7 | 9.6 |
| 25-34 years----- | 335.2 | 30.2 | 11.1 | 352.7 | 33.1 | 10.7 | 352.7 | 33.1 | 10.7 |
| 35-44 years----- | 337.0 | 39.2 | 8.6 | 368.1 | 43.9 | 8.4 | 392.7 | 47.3 | 8.3 |
| 45-64 years----- | 1,186.5 | 87.3 | 13.6 | 1,305.2 | 98.4 | 13.3 | 1,360.2 | 104.9 | 13.0 |
| 65+ years------- | 712.5 | 51.2 | 13.9 | 794.5 | 58.0 | 13.7 | 1,087.5 | 84.4 | 12.9 |
| **Female** | | | | | | | | | |
| All ages-- | 4,419.7 | 616.7 | 7.2 | 4,759.8 | 688.9 | 6.9 | 5,065.2 | 721.7 | 7.0 |
| Under 15 years-- | 443.8 | 71.5 | 6.2 | 483.0 | 86.2 | 5.6 | 484.0 | 86.5 | 5.6 |
| 15-24 years----- | 688.9 | 147.3 | 4.7 | 714.2 | 160.9 | 4.4 | 714.7 | 161.1 | 4.4 |
| 25-34 years----- | 746.8 | 156.0 | 4.8 | 781.2 | 168.3 | 4.6 | 782.5 | 168.8 | 4.6 |
| 35-44 years----- | 651.9 | 93.8 | 6.9 | 700.0 | 106.0 | 6.6 | 712.9 | 110.3 | 6.5 |
| 45-64 years----- | 1,084.7 | 96.1 | 11.3 | 1,190.9 | 109.5 | 10.9 | 1,318.1 | 116.2 | 11.3 |
| 65+ years------- | 803.6 | 52.0 | 15.5 | 890.5 | 58.0 | 15.4 | 1,053.0 | 78.8 | 13.4 |

Table 6. Population changes during year prior to interview and population bases used in obtaining rates

[Average of 13 interview dates]

| Sex and age | Initial number of persons[1] | Births prior to first day of year | Deaths prior to first day of year | Births during year | Deaths during year | Final number of persons | Rate bases | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Interview reported | Perfect interview | All discharges |
| **Both sexes** | | | | | | | | | |
| All ages-- | 10,000 | 144.5 | 58.6 | 235.6 | 96.4 | 10,225 | 10,107 | 10,107 | 10,155 |
| Under 15 years-- | 3,167 | 144.5 | 5.4 | 235.6 | 7.9 | 3,534 | 3,416 | 3,416 | 3,420 |
| 15-24 years----- | 1,293 | ... | 0.3 | ... | 2.0 | 1,291 | 1,291 | 1,291 | 1,292 |
| 25-34 years----- | 1,286 | ... | 1.4 | ... | 2.2 | 1,282 | 1,282 | 1,282 | 1,283 |
| 35-44 years----- | 1,343 | ... | 1.5 | ... | 5.3 | 1,336 | 1,336 | 1,336 | 1,339 |
| 45-64 years----- | 2,045 | ... | 14.9 | ... | 22.6 | 2,008 | 2,008 | 2,008 | 2,019 |
| 65+ years------- | 866 | ... | 35.1 | ... | 56.4 | 774 | 774 | 774 | 802 |
| **Male** | | | | | | | | | |
| All ages-- | 4,866 | 74.0 | 32.8 | 119.5 | 53.3 | 4,973 | 4,913 | 4,913 | 4,940 |
| Under 15 years-- | 1,615 | 74.0 | 3.6 | 119.5 | 4.5 | 1,800 | 1,740 | 1,740 | 1,742 |
| 15-24 years----- | 610 | ... | 0.2 | ... | 1.8 | 608 | 608 | 608 | 609 |
| 25-34 years----- | 615 | ... | 0.6 | ... | 1.0 | 613 | 613 | 613 | 614 |
| 35-44 years----- | 645 | ... | 1.0 | ... | 2.6 | 641 | 641 | 641 | 642 |
| 45-64 years----- | 989 | ... | 9.4 | ... | 14.7 | 965 | 965 | 965 | 972 |
| 65+ years------- | 392 | ... | 18.0 | ... | 28.7 | 345 | 345 | 345 | 359 |
| **Female** | | | | | | | | | |
| All ages-- | 5,134 | 70.5 | 25.8 | 116.1 | 43.1 | 5,252 | 5,194 | 5,194 | 5,216 |
| Under 15 years-- | 1,552 | 70.5 | 1.8 | 116.1 | 3.4 | 1,733 | 1,675 | 1,675 | 1,677 |
| 15-24 years----- | 683 | ... | 0.1 | ... | 0.2 | 683 | 683 | 683 | 683 |
| 25-34 years----- | 671 | ... | 0.8 | ... | 1.2 | 669 | 669 | 669 | 670 |
| 35-44 years----- | 698 | ... | 0.5 | ... | 2.7 | 695 | 695 | 695 | 696 |
| 45-64 years----- | 1,056 | ... | 5.5 | ... | 7.9 | 1,043 | 1,043 | 1,043 | 1,047 |
| 65+ years------- | 474 | ... | 17.1 | ... | 27.7 | 429 | 429 | 429 | 443 |

[1]Distribution based on table 29, p. 42, of reference 8.

Table 7. Percent underreporting of hospital discharges, by type of discharge and number of weeks between discharge and interview: interview reported versus perfect interview

[Average of 13 interview dates]

| Weeks between discharge and interview | Delivery and nondelivery discharges | | | Delivery discharges | | | Discharges excluding deliveries | | |
|---|---|---|---|---|---|---|---|---|---|
| | Interview reported | Perfect interview | Percent underreported | Interview reported | Perfect interview | Percent underreported | Interview reported | Perfect interview | Percent underreported |
| Total----- | 949.5 | 1,071.1 | 11.4 | 208.3 | 212.7 | 2.1 | 741.2 | 858.4 | 13.7 |
| 1-4------------- | 82.2 | 82.2 | 0.0 | 16.8 | 16.2 | [1]3.7 | 65.4 | 66.0 | 0.9 |
| 5-8------------- | 77.8 | 81.8 | 4.9 | 16.4 | 16.5 | 0.6 | 61.4 | 65.3 | 6.0 |
| 9-12------------ | 78.4 | 81.7 | 4.0 | 16.2 | 16.5 | 1.8 | 62.2 | 65.2 | 4.6 |
| 13-16----------- | 78.1 | 83.5 | 6.5 | 16.3 | 16.5 | 1.2 | 61.8 | 67.0 | 8.8 |
| 17-20----------- | 74.9 | 82.3 | 9.0 | 15.7 | 16.5 | 4.8 | 59.2 | 65.8 | 10.0 |
| 21-24----------- | 76.3 | 82.1 | 7.1 | 16.4 | 16.7 | 1.8 | 59.9 | 65.4 | 8.4 |
| 25-28----------- | 73.0 | 82.2 | 11.2 | 15.8 | 16.5 | 4.2 | 57.2 | 65.7 | 12.9 |
| 29-32----------- | 72.8 | 82.4 | 11.7 | 16.0 | 16.4 | 2.4 | 56.8 | 66.0 | 13.9 |
| 33-36----------- | 71.2 | 82.1 | 13.3 | 15.8 | 16.2 | 2.5 | 55.4 | 65.9 | 15.9 |
| 37-40----------- | 73.8 | 82.7 | 10.8 | 15.6 | 16.0 | 2.5 | 58.2 | 66.7 | 12.7 |
| 41-44----------- | 71.4 | 82.9 | 13.9 | 16.1 | 16.2 | 0.6 | 55.3 | 66.7 | 17.1 |
| 45-48----------- | 64.8 | 82.5 | 21.5 | 15.0 | 16.2 | 7.4 | 49.8 | 66.3 | 24.9 |
| 49-52----------- | 54.8 | 82.7 | 33.7 | 16.2 | 16.3 | 0.6 | 38.6 | 66.4 | 41.9 |

[1]Percent overreported.

Table 8. Percent underreporting of hospital discharges, by type of discharge and number of weeks between discharge and interview: perfect interview versus all discharges

[Average of 13 interview dates]

| Weeks between discharge and interview | Delivery and nondelivery discharges | | | Delivery discharges | | | Discharges excluding deliveries | | |
|---|---|---|---|---|---|---|---|---|---|
| | Perfect interview | All | Percent underreported | Perfect interview | All | Percent underreported | Perfect interview | All | Percent underreported |
| Total----- | 1,071.1 | 1,143.5 | 6.3 | 212.7 | 212.7 | 0.0 | 858.4 | 930.8 | 7.8 |
| 1-4------------- | 82.2 | 85.5 | 3.9 | 16.2 | 16.2 | 0.0 | 66.0 | 69.3 | 4.8 |
| 5-8------------- | 81.8 | 86.2 | 5.1 | 16.5 | 16.5 | 0.0 | 65.3 | 69.7 | 6.3 |
| 9-12------------ | 81.7 | 86.5 | 5.5 | 16.5 | 16.5 | 0.0 | 65.2 | 70.0 | 6.9 |
| 13-16----------- | 83.5 | 88.6 | 5.8 | 16.5 | 16.5 | 0.0 | 67.0 | 72.1 | 7.1 |
| 17-20----------- | 82.3 | 88.1 | 6.6 | 16.5 | 16.5 | 0.0 | 65.8 | 71.6 | 8.1 |
| 21-24----------- | 82.1 | 87.6 | 6.8 | 16.7 | 16.7 | 0.0 | 65.4 | 70.9 | 5.5 |
| 25-28----------- | 82.2 | 87.8 | 6.4 | 16.5 | 16.5 | 0.0 | 65.7 | 71.3 | 7.9 |
| 29-32----------- | 82.4 | 88.3 | 6.7 | 16.4 | 16.4 | 0.0 | 66.0 | 71.9 | 8.2 |
| 33-36----------- | 82.1 | 88.1 | 6.8 | 16.2 | 16.2 | 0.0 | 65.9 | 71.9 | 8.3 |
| 37-40----------- | 82.7 | 89.1 | 7.2 | 16.0 | 16.0 | 0.0 | 66.7 | 73.1 | 8.8 |
| 41-44----------- | 82.9 | 89.3 | 7.2 | 16.2 | 16.2 | 0.0 | 66.7 | 73.1 | 8.8 |
| 45-48----------- | 82.5 | 89.0 | 7.3 | 16.2 | 16.2 | 0.0 | 66.3 | 72.8 | 8.9 |
| 49-52----------- | 82.7 | 89.4 | 7.5 | 16.3 | 16.3 | 0.0 | 66.4 | 73.1 | 9.2 |

Table 9. Percent underreporting of hospital discharges, by type of discharge and number of weeks between discharge and interview: interview reported versus all discharges

[Average of 13 interview dates]

| Weeks between discharge and interview | Delivery and nondelivery discharges | | | Delivery discharges | | | Discharges excluding deliveries | | |
|---|---|---|---|---|---|---|---|---|---|
| | Inter-view re-ported | All | Percent under-re-ported | Inter-view re-ported | All | Percent under re-ported | Inter-view re-ported | All | Percent under-re-ported |
| Total----- | 949.5 | 1,143.5 | 17.0 | 208.3 | 212.7 | 2.1 | 741.2 | 930.8 | 20.4 |
| 1-4------------- | 82.2 | 85.5 | 3.9 | 16.8 | 16.2 | [1]3.7 | 65.4 | 69.3 | 5.6 |
| 5-8------------- | 77.8 | 86.2 | 9.7 | 16.4 | 16.5 | 0.6 | 61.4 | 69.7 | 11.9 |
| 9-12------------ | 78.4 | 86.5 | 9.4 | 16.2 | 16.5 | 1.8 | 62.2 | 70.0 | 11.1 |
| 13-16----------- | 78.1 | 88.6 | 11.9 | 16.3 | 16.5 | 1.2 | 61.8 | 72.1 | 14.9 |
| 17-20----------- | 74.9 | 88.1 | 15.0 | 15.7 | 16.5 | 4.8 | 59.2 | 71.6 | 17.3 |
| 21-24----------- | 76.3 | 87.6 | 12.9 | 16.4 | 16.7 | 1.8 | 59.9 | 70.9 | 15.5 |
| 25-28----------- | 73.0 | 87.8 | 16.9 | 15.8 | 16.5 | 4.2 | 57.2 | 71.3 | 19.7 |
| 29-32----------- | 72.8 | 88.3 | 17.6 | 16.0 | 16.4 | 2.4 | 56.8 | 71.9 | 21.0 |
| 33-36----------- | 71.2 | 88.1 | 19.2 | 15.8 | 16.2 | 2.5 | 55.4 | 71.9 | 22.9 |
| 37-40----------- | 73.8 | 89.1 | 17.2 | 15.6 | 16.0 | 2.5 | 58.2 | 73.1 | 20.4 |
| 41-44----------- | 71.4 | 89.3 | 20.0 | 16.1 | 16.2 | 0.6 | 55.3 | 73.1 | 24.4 |
| 45-48----------- | 64.8 | 89.0 | 27.2 | 15.0 | 16.2 | 7.4 | 49.8 | 72.8 | 31.6 |
| 49-52----------- | 54.8 | 89.4 | 38.7 | 16.2 | 16.3 | 0.6 | 38.6 | 73.1 | 47.2 |

[1]Percent overreported.

Table 10. Percent underreporting of hospital discharges, by actual length of stay and number of weeks between discharge and interview: interview reported versus perfect interview

[Average of 13 interview dates]

| Weeks between discharge and interview | 1-day stay | | | 2-4-day stay | | | 5+-day stay | | |
|---|---|---|---|---|---|---|---|---|---|
| | Inter-view re-ported dis-charges | Perfect inter-view dis-charges | Percent under-re-ported | Inter-view re-ported dis-charges | Perfect inter-view dis-charges | Percent under-re-ported | Inter-view re-ported dis-charges | Perfect inter-view dis-charges | Percent under-re-ported |
| Total----- | 101.2 | 131.8 | 23.2 | 339.9 | 383.5 | 11.4 | 508.8 | 555.8 | 8.5 |
| 1-4------------- | 9.1 | 9.4 | 3.2 | 29.5 | 29.4 | [1]0.3 | 43.7 | 43.5 | [1]0.5 |
| 5-8------------- | 8.2 | 9.8 | 16.3 | 27.1 | 28.5 | 4.9 | 42.5 | 43.5 | 2.3 |
| 9-12------------ | 8.0 | 10.0 | 20.0 | 27.5 | 28.3 | 2.8 | 42.9 | 43.4 | 1.2 |
| 13-16----------- | 8.0 | 10.2 | 21.6 | 28.4 | 29.3 | 3.1 | 41.7 | 44.0 | 5.2 |
| 17-20----------- | 7.8 | 9.6 | 18.8 | 26.4 | 29.8 | 11.4 | 40.7 | 42.9 | 5.1 |
| 21-24----------- | 7.5 | 9.8 | 23.5 | 28.2 | 29.5 | 4.4 | 40.7 | 42.8 | 4.9 |
| 25-28----------- | 7.8 | 10.1 | 22.8 | 26.2 | 29.2 | 10.3 | 39.0 | 42.9 | 9.1 |
| 29-32----------- | 6.9 | 10.0 | 31.0 | 27.2 | 29.6 | 8.1 | 38.7 | 42.8 | 9.6 |
| 33-36----------- | 7.6 | 10.2 | 25.5 | 25.5 | 29.4 | 13.3 | 38.2 | 42.5 | 10.1 |
| 37-40----------- | 7.2 | 10.6 | 32.1 | 26.5 | 29.8 | 11.1 | 40.1 | 42.2 | 5.0 |
| 41-44----------- | 8.6 | 10.8 | 20.4 | 23.8 | 30.4 | 21.7 | 39.0 | 41.8 | 6.7 |
| 45-48----------- | 7.4 | 10.8 | 31.5 | 22.3 | 30.3 | 26.4 | 35.1 | 41.3 | 15.0 |
| 49-52----------- | 7.1 | 10.5 | 32.4 | 21.3 | 30.0 | 29.0 | 26.5 | 42.2 | 37.2 |

[1]Percent overreported.

Table 11. Percent underreporting of hospital discharges, by actual length of stay and number of weeks between discharge and interview: perfect interview versus all discharges

[Average of 13 interview dates]

| Weeks between discharge and interview | 1-day stay | | | 2-4-day stay | | | 5+-day stay | | |
|---|---|---|---|---|---|---|---|---|---|
| | Perfect interview discharges | All discharges | Percent underreported | Perfect interview discharges | All discharges | Percent underreported | Perfect interview discharges | All discharges | Percent underreported |
| Total----- | 131.8 | 136.1 | 3.2 | 383.5 | 405.4 | 5.4 | 555.8 | 602.2 | 7.7 |
| 1-4------------- | 9.4 | 9.6 | 2.1 | 29.4 | 30.4 | 3.3 | 43.5 | 45.5 | 4.4 |
| 5-8------------- | 9.8 | 10.2 | 3.9 | 28.5 | 30.0 | 5.0 | 43.5 | 46.1 | 5.6 |
| 9-12------------ | 10.0 | 10.3 | 2.9 | 28.3 | 30.0 | 5.7 | 43.4 | 46.2 | 6.1 |
| 13-16----------- | 10.2 | 10.5 | 2.9 | 29.3 | 31.1 | 5.8 | 44.0 | 47.1 | 6.6 |
| 17-20----------- | 9.6 | 9.8 | 2.0 | 29.8 | 31.6 | 5.7 | 42.9 | 46.6 | 7.9 |
| 21-24----------- | 9.8 | 10.0 | 2.0 | 29.5 | 31.2 | 5.4 | 42.8 | 46.5 | 8.0 |
| 25-28----------- | 10.1 | 10.4 | 2.9 | 29.2 | 30.8 | 5.2 | 42.9 | 46.6 | 7.9 |
| 29-32----------- | 10.0 | 10.3 | 2.9 | 29.6 | 31.2 | 5.1 | 42.8 | 46.8 | 8.5 |
| 33-36----------- | 10.2 | 10.5 | 2.9 | 29.4 | 31.0 | 5.2 | 42.5 | 46.6 | 8.8 |
| 37-40----------- | 10.6 | 11.0 | 3.6 | 29.8 | 31.7 | 6.0 | 42.2 | 46.4 | 9.1 |
| 41-44----------- | 10.8 | 11.2 | 3.6 | 30.4 | 32.3 | 5.9 | 41.8 | 45.8 | 8.7 |
| 45-48----------- | 10.8 | 11.3 | 4.4 | 30.4 | 32.2 | 5.6 | 41.3 | 45.5 | 9.2 |
| 49-52----------- | 10.5 | 11.0 | 4.5 | 30.0 | 31.9 | 6.0 | 42.2 | 46.5 | 9.2 |

Table 12. Percent underreporting of hospital discharges, by actual length of stay and number of weeks between discharge and interview: interview reported versus all discharges

[Average of 13 interview dates]

| Weeks between discharge and interview | 1-day stay | | | 2-4-day stay | | | 5+-day stay | | |
|---|---|---|---|---|---|---|---|---|---|
| | Interview reported discharges | All discharges | Percent underreported | Interview reported discharges | All discharges | Percent underreported | Interview reported discharges | All discharges | Percent underreported |
| Total----- | 101.2 | 136.1 | 25.6 | 339.9 | 405.4 | 16.2 | 508.8 | 602.2 | 15.5 |
| 1-4------------- | 9.1 | 9.6 | 5.2 | 29.5 | 30.4 | 3.0 | 43.7 | 45.5 | 4.0 |
| 5-8------------- | 8.2 | 10.2 | 19.6 | 27.1 | 30.0 | 9.7 | 42.5 | 46.1 | 7.8 |
| 9-12------------ | 8.0 | 10.3 | 22.3 | 27.5 | 30.0 | 8.3 | 42.9 | 46.2 | 7.1 |
| 13-16----------- | 8.0 | 10.5 | 23.8 | 28.4 | 31.1 | 8.7 | 41.7 | 47.1 | 11.5 |
| 17-20----------- | 7.8 | 9.8 | 20.4 | 26.4 | 31.6 | 16.5 | 40.7 | 46.6 | 12.7 |
| 21-24----------- | 7.5 | 10.0 | 25.0 | 28.2 | 31.2 | 9.6 | 40.7 | 46.5 | 12.5 |
| 25-28----------- | 7.8 | 10.4 | 25.0 | 26.2 | 30.8 | 14.9 | 39.0 | 46.6 | 16.3 |
| 29-32----------- | 6.9 | 10.3 | 33.0 | 27.2 | 31.2 | 12.8 | 38.7 | 46.8 | 17.3 |
| 33-36----------- | 7.6 | 10.5 | 27.6 | 25.5 | 31.0 | 17.7 | 38.2 | 46.6 | 18.0 |
| 37-40----------- | 7.2 | 11.0 | 34.5 | 26.5 | 31.7 | 16.4 | 40.1 | 46.4 | 13.6 |
| 41-44----------- | 8.6 | 11.2 | 23.2 | 23.8 | 32.3 | 26.3 | 39.0 | 45.8 | 14.8 |
| 45-48----------- | 7.4 | 11.3 | 34.5 | 22.3 | 32.2 | 30.7 | 35.1 | 45.5 | 22.9 |
| 49-52----------- | 7.1 | 11.0 | 35.5 | 21.3 | 31.9 | 33.2 | 26.5 | 46.5 | 43.0 |

# APPENDIX

## OUTLINE FOR COMPUTER SIMULATION OF HOSPITAL DISCHARGES

[Input data are found in table B for the MP1 matrix, in table C for the MP2 matrix, and in table D for the MP3 matrix. For other matrices in the computer program, data are not reproduced in this report because of their bulk]

Each age-sex group of $n$ individuals is assigned birth dates $b_k$, delivery dates $c_k$, and death dates $d_k$, where $k=1,2,\ldots,n$. The input data also includes:

1. Weekly admission probabilities $P_k$ appropriate to the $k$th individual according to his age, sex, and subgroup as per the $MP1$ matrix;
2. Daily discharge probabilities $P_{jk}$ appropriate to the $k$th individual according to his age, sex, and number of days already hospitalized as per the $MP2$ matrix;
3. Probabilities $P_o$ of being hospitalized for a delivery according to age as per the $MP3$ matrix;
4. Daily discharge probabilities $P_s$ for delivery hospitalizations according to age and number of days already hospitalized as per the $MP4$ matrix.

These probability matrices are all used in Phase I. In Phase II, the input data consists of birth dates, delivery dates, death dates, and the number of days to death $m$ for individuals determined in Phase I to be in their last year of life. The input data for Phase II also includes:

1. Daily admission probabilities $P_m$ according to the number of days of life remaining to the individual as per the $MP7$ matrix;
2. Daily discharge probabilities $P_g$ according to age, sex, and number of days already hospitalized as per the $MP8$ matrix.

Histories are generated separately for each of the $n$ individuals in an age-sex group. Starting with the first individual the basic steps in the computer program are as follows:

I. Determine whether $d_k - b_k > 364$. If no, set $m = 365 - (d_k - b_k)$ and day $i = 1$ and proceed to III (Phase II). If yes, set $i = b_k$ and:
  a. Generate uniform random numbers $R_i$ for each day from $b_k$ to 756 as outlined below. First, however, check, is $d_k - i < 365$? If yes, set $m = 365 - (d_k - i)$ and proceed to III. If no, is $c_k - 30 < i < c_k$? If yes, proceed to II. If no, proceed to I-b.
  b. Generate $R_i$. Is $R_i < P_k$? If no, loop back to I-a. taking $i = i + 7$. If yes, proceed to I-c.
  c. Select a $y$ value in order from the sequence 4, 0, 1, 3, 6, 5, 2. When the sequence is exhausted start over; do not start over for a new individual. Take $i + y$ equal to $i$ and proceed to I-d.
  d. Record $i$, the admit date, for the $k$th individual. Is $i \geq 756$? If yes, record $i + 1$ as the discharge date for this admission and loop back to I for the next ($k + 1^{st}$) individual. If no, proceed to I-e.
  e. Generate $R_{i+j}$ ($j = 1$ to 100) and proceed to I-f.
  f. Is $R_{i+j} < P_{jk}$? If no, proceed to I-g. If yes, proceed to I-h.
  g. Is $i + j = 756$? If yes, record 757 as the discharge date and loop to I for the next individual. If no, loop back to I-e. taking $j = j + 1$.
  h. Record $i + j$ as the discharge date for this admission. Is $i + j = 756$? If yes, loop back to I for the next individual. If no, loop back to I-a. taking $i = i + j + 1$.

II. Is $c_k = i$? If no, loop back to I-a. taking $i = i + 1$. If yes, proceed to II-a.
  a. Generate random number $R_o$. Is $R_o < P_o$? If no, loop back to I-a. taking $i + i + 1$. If yes, proceed to II-b.
  b. Record $i$ as a delivery admission date; then proceed to II-c.
  c. Generate $R_{i+s}$ ($s = 1$ to 30). Is $R_{i+s} < P_s$? If no, proceed to II-d. If yes, proceed to II-e.
  d. Is $i + s = 756$? If yes, record 757 as the discharge date and loop back to I for the next individual. If no, loop back to II-c. taking $s = s + 1$.
  e. Record $i + s$ as the discharge date. Is $i + s = 756$? If yes, loop back to I for the next individual. If no, loop back to I-a. taking $i = i + s + 1$.

III. Generate $R_{i+f-1}$ [$f = 1$ to $(366 - m)$].
  a. Is $R_{i+f-1} \leq P_m$? If no, loop back to III taking $f = f + 7$. If yes, proceed to III-b.
  b. Select an $x$ value in order from sequence 1, 6, 0, 3, 4, 2, 5. When the sequence is exhausted start over; do not start over for a new individual. Take $f + x$ equal to $f$ and record $i + f - 1$ as the admit date. Proceed to III-c.
  c. Is $i + f - 1 \geq d_k$? If yes, record $i + f$ as the discharge date and loop back to I for the next individual. If no, proceed to III-d.
  d. Generate $R_{i+f+g-1}$ ($g + 1$ to 100). Is $R_{i+f+g-1} < P_g$? If no, proceed to III-e. If yes, proceed to III-f.
  e. Is $i + f + g - 1 = d_k$? If yes, record $d_k + 1$ as the discharge date and loop back to I for the next individual. If no, loop back to III-d. taking $g = g + 1$.
  f. Record $i + f + g - 1$ as the discharge date for this admission. Is $i + f + g - 1 = d_o$? If yes, loop back to I for the next individual. If no, loop back to III taking $f = f + g + 1$.

———ooo———



☆ U.S. GOVERNMENT PRINTING OFFICE : 1966 O - 797-592

## OUTLINE OF REPORT SERIES FOR VITAL AND HEALTH STATISTICS

### Public Health Service Publication No. 1000

*Series 1.* *Programs and collection procedures.*—Reports which describe the general programs of the National Center for Health Statistics and its offices and divisions, data collection methods used, definitions, and other material necessary for understanding the data.

Reports number 1-4

*Series 2.* *Data evaluation and methods research.*—Studies of new statistical methodology including: experimental tests of new survey methods, studies of vital statistics collection methods, new analytical techniques, objective evaluations of reliability of collected data, contributions to statistical theory.

Reports number 1-13

*Series 3.* *Analytical studies.*—Reports presenting analytical or interpretive studies based on vital and health statistics, carrying the analysis further than the expository types of reports in the other series.

Reports number 1-4

*Series 4.* *Documents and committee reports.*—Final reports of major committees concerned with vital and health statistics, and documents such as recommended model vital registration laws and revised birth and death certificates.

Reports number 1 and 2

*Series 10.* *Data From the Health Interview Survey.*—Statistics on illness, accidental injuries, disability, use of hospital, medical, dental, and other services, and other health-related topics, based on data collected in a continuing national household interview survey.

Reports number 1-26

*Series 11.* *Data From the Health Examination Survey.*—Statistics based on the direct examination, testing, and measurement of national samples of the population, including the medically defined prevalence of specific diseases, and distributions of the population with respect to various physical and physiological measurements.

Reports number 1-12

*Series 12.* *Data From the Health Records Survey.*—Statistics from records of hospital discharges and statistics relating to the health characteristics of persons in institutions, and on hospital, medical, nursing, and personal care received, based on national samples of establishments providing these services and samples of the residents or patients.

Reports number 1-3

*Series 20.* *Data on mortality.*—Various statistics on mortality other than as included in annual or monthly reports—special analyses by cause of death, age, and other demographic variables, also geographic and time series analyses.

Reports number 1

*Series 21.* *Data on natality, marriage, and divorce.*—Various statistics on natality, marriage, and divorce other than as included in annual or monthly reports—special analyses by demographic variables, also geographic and time series analyses, studies of fertility.

Reports number 1-7

*Series 22.* *Data From the National Natality and Mortality Surveys.*—Statistics on characteristics of births and deaths not available from the vital records, based on sample surveys stemming from these records, including such topics as mortality by socioeconomic class, medical experience in the last year of life, characteristics of pregnancy, etc.

Reports number 1

For a list of titles of reports published in these series, write to: National Center for Health Statistics
U.S. Public Health Service
Washington, D.C. 20201