SAFER • HEALTHIER • PEOPLE™

CDC™

Monitoring the
Nation's Health

Vital and Health Statistics

Series 2, Number 146

May 2008

Estimating Healthy Life Expectancies Using Longitudinal Survey Data: Methods and Techniques in Population Health Measures

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

Suggested citation

Molla MT, Madans JH. Estimating healthy life expectancies using longitudinal survey data: Methods and techniques in population health measures. National Center for Health Statistics. Vital Health Stat 2(146). 2008.

Library of Congress Cataloging-in-Publication Data

Estimating healthy life expectancies using longitudinal survey data : methods and techniques in population health measures.
p. ; cm. — (Vital and health statistics. Series 2, Data evaluation and methods research ; no. 146)
″May 2008″
Includes bibliographical references.
ISBN 0-8406-0620-6
1. Life expectancy—United States. 2. Demography—United States. 3. Longitudinal method—United States. I. National Center for Health Statistics (U.S.) II. Series. [DNLM: 1. Life Expectancy—United States. 2. Demography—United States. 3. Longitudinal Studies—United States. 4. Models, Statistical—United States. W2 A N148vb no.146 2008]
HB1335.E88 2008
304.6’245072—dc22
2008007102

For sale by the U.S. Government Printing Office
Superintendent of Documents
Mail Stop: SSOP
Washington, DC 20402-9328
Printed on acid-free paper.

Vital and Health Statistics

Series 2, Number 146

Estimating Healthy Life Expectancies Using Longitudinal Survey Data: Methods and Techniques in Population Health Measures

Data Evaluation and Methods Research

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

Hyattsville, Maryland
May 2008
DHHS Publication No. (PHS) 2008–1346

National Center for Health Statistics

Edward J. Sondik, Ph.D., *Director*

Jennifer H. Madans, Ph.D., *Acting Co-Deputy Director*

Michael H. Sadagursky, *Acting Co-Deputy Director*

Jennifer H. Madans, Ph.D., *Associate Director for Science*

Jennifer H. Madans, Ph.D., *Acting Associate Director for Planning, Budget, and Legislation*

Michael H. Sadagursky, *Associate Director for Management and Operations*

Lawrence H. Cox, Ph.D., *Associate Director for Research and Methodology*

Linda B. Torian, *Acting Director for Information Technology*

Linda B. Torian, *Acting Director for Information Services*

Linda T. Bilheimer, Ph.D., *Associate Director for Analysis and Epidemiology*

Charles J. Rothwell, M.S., *Director for Vital Statistics*

Jane E. Sisk, Ph.D., *Director for Health Care Statistics*

Jane F. Gentleman, Ph.D., *Director for Health Interview Statistics*

Clifford L. Johnson, M.S.P.H., *Director for Health and Nutrition Examination Surveys*

Office of Analysis and Epidemiology

Linda T. Bilheimer, Ph.D., *Director*

Diane Makuc, Dr.P.H., *Associate Director for Science*

Linda T. Bilheimer, Ph.D., *Acting Chief, Aging and Chronic Disease Statistics Branch*

Contents

Text Figures

Appendix Figures

Text Tables

Abstract

Objective

Summary measures of population health are statistics that combine mortality and morbidity to represent overall population health in a single index. Such measures include healthy life expectancy, also called disability-free life expectancy and active life expectancy. Healthy life expectancy can be calculated using cross-sectional or longitudinal survey data. This report presents a comprehensive discussion of a method for calculating healthy life expectancy using data from longitudinal surveys.

Methods

Healthy life expectancies are calculated using the multistate life table model. Expected life in various states of health is estimated using data from the Second Longitudinal Study of Aging and the Medicare Current Beneficiary Survey to illustrate the calculation of the statistics and the discussion of data and methodology related issues.

Results

The study shows that estimating summary measures of population health using longitudinal survey data provides the opportunity of using incidence rather than prevalence rates. Health measures estimated based on incidence reflect the most recent health status of the population. Models that use longitudinal survey data measure transitions from good to poor health as well as poor to good health. That is, the models account for recovery from morbidity or illness. Longitudinal survey data can also be used to calculate healthy or active life expectancies by initial health states.

Keywords: summary measures of population health • health states • life expectancy • healthy life expectancy • active life expectancy • multistate life table • Markov chain • multinomial logistic regression

Estimating Healthy Life Expectancies Using Longitudinal Survey Data: Methods and Techniques in Population Health Measures

by Michael T. Molla, Ph.D., Office of Analysis and Epidemiology, and Jennifer H. Madans, Ph.D., Office of the Center Director

Introduction

For the populations of the industrialized nations of the world, the 20th century has been a period of both demographic and epidemiological transitions. In addition to the substantial fall in fertility during this period, all industrialized countries experienced a huge decline in crude death rates and infant mortality rates that led to an impressive rise in the average expectation of life. For example, during the first half of the 20th century, the average expectation of life at birth in the United States for the total population increased by nearly 45 percent; from 47.3 years (46.3 for males and 48.3 for females) in 1900 to 68.2 years (65.6 for males and 71.1 for females) by 1950, a gain of nearly 21 years. In the second half of the century, mortality in the United States continued to decline, though at a much slower rate compared with the first half of the century. In the period from 1950 to 2000, for example, the average expectation of life at birth for the total population increased from 68.2 years to 77.0 years (74.3 for males and 79.7 for females), a gain of about 8.8 years (1). In Japan, from 1950 to 2000, female life expectancy at birth increased by 21 years, whereas male life expectancy increased by 17.9 years (2). By the year 2000, the estimated Japanese average life expectancy was 77.5 years for males and 84.7 years for females (3).

The substantial gain in life expectancy both in the first as well as the second half of the century caused a dramatic change in the age structures of the populations of the industrialized countries. The longer average life span meant a sharp rise in the population aged 65 years and over (referred to as the ‘‘older population’’ in this text) as a percentage of the total population. There was also a sharp rise in noncommunicable degenerative diseases and chronic conditions largely affecting this older population. In the United States, the combined effect of these changes has encouraged a shift in focus from longevity-related issues to preventing disability, improving functioning, and relieving physical pain and emotional distress (4). This also means that population health measures need to account not only for longevity but also for morbidity (5).

Measures of Health that Combine Mortality and Morbidity

The development of health measures that take into account both mortality and morbidity of the population has been a focus of study since the 1960s. After Sanders published the results of his research on measuring community health levels (6), healthy life expectancy estimates were published by the U.S. Department of Health, Education, and Welfare for the first time in 1969 (7). In the late 1960s and early 1970s, Sullivan (8–10) published articles explaining a method of calculating healthy life expectancy using a standard single-decrement life table and data from cross-sectional health surveys. In 1974, the World Health Organization restated the definition of health as ‘‘a state of complete physical, mental and social well being’’ and recommended that person years of life in good health be calculated and be compared with the total person years of life (11). At about the same time, the Organisation for Economic Cooperation and Development began producing reports that included measures of healthy life (12).

In the decades that followed, concerns have been expressed about the conceptual, methodological, and ethical issues related to summary measures of health that combine mortality and morbidity. Although there is general agreement that mortality data alone are not sufficient to describe health, the conceptual and measurement issues associated with mortality are generally straightforward and well established. This is not the case for measures of morbidity or health status. Health status is a broad concept that encompasses many interrelated dimensions. Morbidity and disability are under the general umbrella of health status, but there is not universal agreement on how they are defined. There is even less agreement on how they should be measured. This has major implications for measures of health that combine aspects of health status, morbidity, or disability. The resulting measures are very dependent on how health status is operationalized, with different decisions leading to different conclusions about the health of a population and about changes over time and differentials across countries and subgroups within countries. The health status component of composite measures can include survey-based responses, single questions asking respondents to assess their health, or indicators that are constructed by combining a person’s characteristics in a few or many health domains. Health domains can be defined in different ways but often include aspects of physical functioning, sensory functioning, cognitive functioning, psychological functioning, pain, and fatigue. Great variation exists in the domains selected for use and in the manner in which information about domains is collected. In order to combine the individual descriptors into a single metric, they must be weighted in some manner. Weights that represent societal preferences for that health state are often used for this purpose. The manner in which the preferences are determined has raised ethical as well as methodologic questions that continue to be debated. Given these complexities, a user should have a clear understanding of how the health metric was developed. Measures cannot be interpreted without this knowledge.

Despite the difficulty of constructing measures that combine mortality and morbidity, such measures reinforce the importance of expanding the evaluation of health beyond mortality. In the United States, these measures have also been used to investigate possible relationships between changes in mortality and nonfatal health outcomes in general and morbidity in particular. One issue that has been debated in great detail is whether the factors responsible for the reductions in mortality had a similar effect on morbidity. Some argued that most of the years of life that the older population gained due to the decline of mortality were healthy years because morbidity was being pushed to the last few years of life—the compression of morbidity (13,14). Others argued that the mortality decline observed over the past several decades was more the result of improvements in health care that postponed death rather than disease prevention that increased healthy states, and therefore the net effect was a rise in the proportion of persons with serious disability and morbidity, especially at later ages (15).

At the same time, Manton (16) introduced the concept of ‘‘dynamic equilibrium,’’ in which he argued that chronic diseases had become less of a problem because deterioration of health due to disease had slowed down. He also suggested that although the decrease in mortality might have given rise to an increase in the prevalence of chronic diseases, the milder character of these diseases might have led to an overall better quality of life. The aging of the population and the issues related to the compression (expansion) of morbidity and the interaction of these demographic and epidemiological phenomena are indicative of the fact that the health of a population is multidimensional. Although summary measures that combine mortality and morbidity are useful for characterizing the joint trajectories, detailed information on mortality and morbidity are needed to understand the dynamics of population health.

As pointed out earlier, mortality has been declining continuously (though at a decreasing rate) over the last several decades not only in the economically advanced countries, but also in many of the developing countries. On the other hand, trends in morbidity have not been as clear as trends in mortality. In the developed countries, trends in morbidity have been influenced by the changes in the structure of the population due to aging as well as the transition from communicable diseases or acute conditions to degenerative diseases or chronic conditions.

The cumulative effects of these and other demographic and epidemiological phenomena have been mixed. In the United States, for example, expected years free of activity limitation among the older population varies widely by sex and race. At the age of 65 years, on average, a white female could expect to live longer than a white male of the same age. But the percentage of life that she expects to spend free of activity limitation will be smaller than the percentage of life free of activity limitation for a white male of the same age. On the other hand, a black adult male of the same age on average will expect to live not only a shorter life, but also spend a smaller proportion of his shorter life free of activity limitation (17).

According to some health experts, these already multifaceted health outcomes will become even more complex in the next few decades. With an aging population and rapid increase of the population aged 85 years and over, some health experts argue more health disparities by age are going to be expected. This is partly due to a cohort effect, because younger cohorts entering their senior years will be healthier than those who have done so before them. Also, advances in the pharmaceutical industries have helped reduce the effect of many chronic conditions among the older population.

Methods

Methods for Calculating Healthy Life Expectancies

A review of the literature on summary measures of population health indicates that health measures that have been used in almost all low-mortality countries and in most other regions of the world fall into the following two broad categories: measures of health gap and measures of health expectancies (18). The health gap measure is the difference between the actual health of a population and a predetermined goal for population health. Measures such as disability-adjusted life years (DALYs) and healthy life years (HeaLYs) fall in this category of population health measures (19). DALYs comprise years of life lost due to premature mortality and years lived with disability adjusted for the severity of disability (20). HeaLYs per 1,000 population per year is calculated on the basis of a set of components that take into account two value judgments about life lived at any given age compared with life lived at any other age, as well as healthy life lost because of premature death or disability. The method also discounts expectation of life lost in future years at a rate of 3 percent per annum. The components of the measure include the disease incidence rate per 1,000 population per year, average age at the onset of a disease, average age at death, expectation of life at age of onset, expectation of life at age of death, case fatality ratio, extent of disability, and average duration of disability for those disabled by the disease (21,22).

Health expectancy is a generic term that refers to a collection of measures defined in terms of life expectancy in a predetermined state of health. This category of measures includes healthy life expectancy, active life expectancy, disability-free life expectancy, disease-free life expectancy, and impairment-free life expectancy. A classification system of health expectancies developed by the Network on Health Expectancy and the Disability Process, also known as Réseau Espérance de Vie en Santé (REVES), in 1994 laid down the concepts and health expectancy measures that need to be included in such a system. The major concepts included in this particular system were disease, impairment, disability, handicap, and self-perceived health (or self-assessed general health). The health expectancy measures based on these measures were health expectancies with or without disease, with or without impairment, with or without functional limitation, with or without activity restriction, with or without handicap, and in good or better health (23). Each of these measures includes both the positive and the negative health states, and the sum of the complete complementary health states is expected to add up to total life expectancy.

Healthy life expectancy, also called life expectancy in good health, is the average number of years a person is expected to live in a health state defined as the ‘‘favorable part of the distribution of perceived health status’’ (24). Active life expectancy measures the number of years a person is expected to live without restrictions in activities of daily living (ADLs) or instrumental activities of daily living (IADLs) if current patterns of mortality, ADLs, and IADLs continue into the future unchanged. Disability-free life expectancy (DFLE) is the average number of years a person is expected to live free of disability if current patterns of mortality and disability continue to extend into the future unchanged. DFLE is differentiated into functional limitation-free life expectancy and activity restriction-free life expectancy. Although functional limitation is a function of the Nagi physical activity status, activity restriction is a function of ADLs and IADLs. Healthy life expectancy and active life expectancy estimates will be presented and discussed in more detail in this report later. Healthy life expectancy is estimated using four different methods that make use of models that are based on demographic and epidemiological concepts and assumptions. All four methods are either totally or partially dependent on the life table technique, and the frequency of their use is determined mainly on the basis of the kind of data that are available. The first method is known as the ‘‘prevalence-rate’’ method or the ‘‘Sullivan method’’ (9). The second method, first used by Katz (25), is based on the ‘‘multiple decrement’’ life table technique. The third is a method that uses microsimulation to estimate individual healthy life cycle trajectories (26), and the fourth is a method known as the ‘‘increment-decrement’’ or the ‘‘multi-state’’ life table technique (27).

Of the four methods that are used to estimate healthy life expectancy or expected years of active life, the prevalence-rate, or Sullivan method, is used most frequently. This is partly because the data that are required to use the method are abundantly available and also because the demographic model used in this method is simpler than the ones used in the other three methods. The prevalence-rate method, originally used to calculate expected years of working life by labor economists (28), was first adapted as a health measure by Sullivan (9). The method uses the standard single-decrement life table technique to estimate life expectancy and combines life expectancy values with health-related prevalence rates from cross-sectional survey data to estimate health expectancies, including expected years of healthy or active life.

The details of the Sullivan method as a health measure, including the underlying assumptions and the input data needed to estimate healthy life expectancy, are discussed in a previous publication (29). The method first uses vital statistics data to estimate single-decrement life table values, assuming that current mortality conditions will continue into the future unchanged. This is followed by the calculation of prevalence rates of impairment, disease, chronic conditions, or limitations due to health problems by age from cross-sectional survey data. The age-specific prevalence rates are then used to divide the ''total life years'' to be lived by the life table population of a given age into a state defined as ''healthy'' and ''not healthy'' or ''active'' and ''inactive.'' Healthy life expectancy is calculated on the basis of the total number of years lived in a healthy state. The method also assumes that current morbidity conditions will continue unchanged into the future.

The second method uses the ''multiple-decrement'' life table model. This method was first used by Katz and his colleagues to estimate active life expectancy using data from the first and second waves of the Massachusetts Health Care Panel Study (25). They used their model to measure changes from living in the community and being independent in ADLs to being dependent in at least one ADL or being institutionalized or deceased. They proceeded further to estimate what they called active life expectancy, or ''remaining years of independent ADL.''

This method differs from the prevalence-rate or Sullivan's method in two ways. First, whereas the prevalence-rate method calculates active life expectancy using a single-decrement life table technique, this model calculates expected years of active life using the multiple-decrement life table technique. Second, the prevalence-rate method uses health data collected through cross-sectional surveys or censuses, whereas the model developed by Katz and his colleagues estimates active life expectancy using longitudinal panel data. However, in spite of its use of longitudinal panel data, both this method and the prevalence-rate method assume illness or health status affected by morbid conditions as irreversible. That is, the possibility of recovery from health states affected by morbidity or disability does not exist in either of the two models.

The third method was first introduced by Laditka and Wolf (26). Unlike the other three models, this method calculates life expectancies in different health states using a model of health status transitions and microsimulation. The microsimulation part of the analytical process is used to identify the underlying Markov chain that closely replicates the one in the actual data.

Laditka and Wolf summarized the method using a model of transition processes that occur between discrete states in discrete time. The time unit used in their model was a month, and the actual panel of functional status information used was from the Longitudinal Study of Aging (LSOA). Transition probabilities were calculated using a multinomial logistic regression with age, race, and education as covariates. The unknown parameters of the model were estimated through an iterative maximization of the log-likelihood function.

Matrices of probabilities were then used to simulate monthly functional status histories. For each person, the simulation of functional status history ends at the time of his or her death. This process was repeated until a large sample of monthly functional status histories was estimated. The sex and functional status composition of the initial cohorts are determined on the basis of observed average characteristics of persons from actual survey data.

The key steps of calculating life expectancies in the different health states include the calculation of health state transition probabilities using a multinomial logistic regression, the estimation of a large sample of health status histories using simulation that is based on the transition probabilities estimated in the first step, the selection of a starting age for age of analysis, and the selection of actual survey data that can serve as the base for the sex and health status composition of each synthetic cohort at the initial age of the analysis. The approach provides an approximate distribution of the population across health states, the length of time spent in each health state, the number of health transitions, and the distribution of the episodes by the duration (length) of the episodes (26,30).

The fourth method used to estimate healthy life expectancy uses the increment-decrement life table or multistate life table technique. This approach is different from the first two in its underlying assumptions about the dynamics of health states. Unlike the other two models, which are based on the life table technique, this method's underlying assumptions and its use of data on incidence of health states makes it flexible enough to account for recovery from illnesses and other health states affected by morbid conditions. The model was originally developed by demographers to analyze transitions between states of life over a person's life course and has been used in various areas of social science research. The model was used, for example, to analyze intergenerational social class transitions (31), transitions between different areas of residence (26), marital status (32), the labor force (33,34), working life (35), active life expectancy (36–38), life in different health states (39–42), school life (43), career change over the life course of a cohort (44), obesity, and active life expectancy (45).

The multistate life table model is designed to accommodate multiple changes of health states. Within each age interval, the model also allows multiple exits from and reentries into any given health state (38). The model uses data from two or more waves of a longitudinal survey. The data analysis begins with the calculation of age-specific transition probabilities that describe the movements between the various transient (nonabsorbing) health states and the probabilities of dying (an absorbing state) of persons in different health states within a given age interval. Like the microsimulation approach, the model also has the advantage of generating life expectancies in different health states for the total population or population subgroups by their initial health states (35).

Other key characteristics that distinguish the multistate life table model from the first two include the assumptions made about the underlying stochastic processes that determine transitions between the various health states. These transitions and the underlying stochastic processes can be analyzed by constructing a Markov model that specifies health states and state transitions with an underlying Markov assumption. In such a model, health states are assumed to be mutually exclusive and collectively exhaustive, and the transition probabilities between current and future health states is assumed to be independent of previous health states. These two sets of characteristics jointly define a Markov chain. If time intervals between two consecutive health states are uniform and the transition probabilities are time invariant, then the Markov chain is characterized as homogenous. Not all Markov chains are homogenous. If a Markov chain is homogenous, then ‘‘piecewise-constant transition rates’’ can be specified, and duration in each health state and life expectancy can be calculated using matrix algebra (38,42).

Although the increment-decrement life table model generates life expectancies in different health states (while at the same time accounting for the probability of recovery) and is the model of choice for researchers, it has not been used as widely as the model based on the single-decrement life table. This is because the multistate life table model is technically more complex, uses more stringent underlying assumptions, and needs longitudinal survey data that, compared with data from cross-sectional surveys, are assumed to be more expensive and time consuming to collect (46).

However, though the method could be used with data from a longitudinal survey with several waves, one should note that the method could also be used with data from a baseline and only one follow-up survey. Even more important, the multistate method has the advantage over the other two life table-based methods in that estimates from this model are based on ‘‘flows’’ rather than ‘‘stocks’’ of health states, and healthy life expectancies are estimated on the basis of incidence rather than health prevalence rates. Because incidence-based estimates reflect more recent pictures of health status, such estimates (unlike estimates based on prevalence rates) can be used to make projections about the future health status of the population. Incidence and prevalence as measures of morbidity are discussed in most of the standard epidemiology text books including one by Gordis (47) and another by Selvin (48).

Objectives and Contents of the Report

In this study, the method for estimating healthy and active life expectancy using the multistate life table model and data from two longitudinal surveys is described. First, data were used from LSOA II to estimate life expectancy with and without activity limitation (active life expectancy), and then data were used from the Centers for Medicare and Medicaid Services’ (CMS) Medicare Current Beneficiary Survey (MCBS) to estimate expected years in good (healthy life expectancy) and poor health. Healthy life expectancies, as well as active life expectancies, are calculated using a personal computer (PC) program called IMaCh. The data used for illustrating the method and its sources are discussed in the following section. The model specification is explained under ‘‘Model Specification,’’ and the section that follows introduces the IMaCh program and briefly explains the procedure for running it. Illustrative examples of active life expectancy and expected years of life in good and poor health are provided in subsequent sections. Finally, the ‘‘Discussion’’ presents a brief summary and conclusion of the report.

Data Sources for Illustrating the Method

The method is illustrated using data from two longitudinal surveys. Active life expectancy is calculated using data from LSOA II, and expected years in good and poor health is calculated using data from MCBS. LSOA II was a collaborative project of the National Center for Health Statistics and the National Institute on Aging and also was financially supported by the U.S. Department of Health and Human Services’ Office of the Assistant Secretary for Planning and Evaluation and the Centers for Disease Control and Prevention’s National Center for Chronic Disease Prevention and Health Promotion. The project was a multicohort study designed to examine changes in the health, functional status, living arrangements, and health care utilization of Americans aged 70 years and over.

LSOA II was conducted 10 years after the first LSOA, and LSOA II closely replicates the first LSOA in content as well as methodology. LSOA II included 9,447 sample persons who were 70 years of age and over in 1995. The baseline data were collected through personal interviews by the staff of the U.S. Census Bureau, and the follow-up surveys were conducted by the National Opinion Research Center at the University of Chicago. The baseline data source for each sample person for LSOA II came from the 1994 National Health Interview Survey (NHIS) core questionnaire, the Family Resources Supplement to the 1994 NHIS, Phase I of NHIS on Disability (NHIS-D), and Phase II of NHIS-D. The 1994 NHIS was a sample survey with complex sample design and oversampled persons of African American heritage. Finally, the baseline LSOA II data included only community dwellers and not the institutionalized (such as those who resided in nursing homes at the time of the baseline survey). However, the institutionalized persons who resided in a nursing, convalescent, or rest home were included in the follow-up surveys. Data were collected in three waves that were about 2 years apart (49).

The second data used for illustrating the method is provided by MCBS, which is a longitudinal panel survey that is sponsored by CMS. The annual sample size is about 12,000 completed interviews and includes both community dwellers as well as the institutionalized population. It is a continuous, multipurpose survey of a representative sample of the Medicare population. Each cohort is interviewed three times a year for a period of up to 4 years.

MCBS follows a rotating panel design with one-third of the sample replaced each year. The survey includes both the aged and disabled Medicare enrollees; the oldest old (aged 85 years and over) and the disabled (aged 64 years and under) are oversampled. The survey focuses on issues such as health care use and expenditures and ability to pay. The survey also includes data on demographic characteristics, health and functioning status, access to care, and insurance coverage. Data on the health status of respondents are collected every year in the September–December round (50,51).

Data on age, sex, date of birth, date of death (for decedents), date and health status at first interview, and date and health status at the second interview were selected from each of the data sets. However, the health component measure chosen from the first data source is different from the health component measure selected from the second one. A measure of nonfatal health outcome based on ADLs was chosen for the analysis based on the LSOA II data, and self-assessed (or self-reported) health was chosen for the analysis based on the MCBS data. The analysis based on the MCBS data also included additional information on level of education. The questions on ADLs from LSOA II data and on self-assessed health from the MCBS data are presented in Appendix I. There was no designated reason for choosing the type of nonfatal health outcome from each of the two data sets. Because these measures are being used only for illustrative purposes, the measure based on ADLs could have been chosen from the MCBS data and the measure based on self-assessed health from the LSOA II data.

The responses to the set of questions on ADLs and self-assessed health were recoded for the analysis. Because of the recode, ‘‘activity limitation’’ is defined as having difficulty in at least one of the following six ADLs: bathing, dressing, eating, moving in and out of a bed or a chair, walking, or using the toilet. The recoded nonfatal health outcome measure of being in ‘‘good or better health’’ stands for being in good, very good, or excellent health, and the recoded ‘‘poor health’’ includes survey respondents who were in poor or fair health.

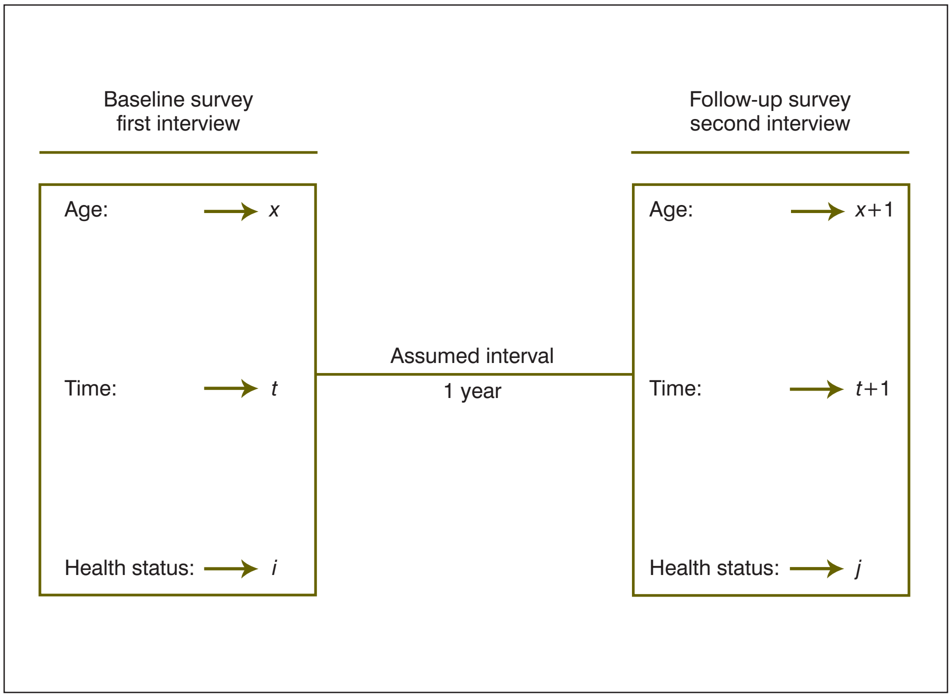


Figure 1. A schematic presentation of a survey respondent interviewed at baseline and at the first follow-up

Model Specification

The application of the multistate life table model in the estimation of healthy and active life expectancies is illustrated using a three-state model specification. The model assumes that respondents were asked the same question about their health status twice, once at the time of the first interview (such as a baseline survey) and again at the time of the second interview (such as a follow-up survey) about a year later. For a respondent of age x at the first interview, the age, time of interview, and health status at each of the two interviews can be presented schematically, as shown in Figure 1.

The three states and the possible health transitions between the two interviews are schematically presented in Figure 2. In the first example, health state one includes those who are without any activity limitation. Those who have activity limitation are included in health state two. In the second example, health state one includes those who are in good health, and those who are in poor health are included in health state two. In each of the two examples, those who die between time (t) and time ($t + 1$) end up in state three. In each example, the first two states are *transient* and persons may therefore move from one health state to the other between the two interviews. That is, transitions are assumed to take place back and forth between the following: good or better health and poor health, being free of activity limitation and being limited ($\pi 12$ and $\pi 21$), good health or poor health and death, and states with or without activity limitation and death ($\pi 13$ and $\pi 23$). Transitions from death to good or poor health or to active or inactive life ($\pi 31$ and $\pi 32$) are not possible because death is an absorbing state.

On the basis of this model, health transition probabilities are stated as follows:

$$p_{ij} = pr\,(\text{STATUS}_{t+1} = j \mid \text{STATUS}_t = i)$$
$$i = 1, 2, \text{ and } j = 1, 2, 3$$

That is, the probability of occupying a health state j at time ($t + 1$) depends on the health state i occupied at time (t). In the first example, at time (t), if STATUS equals 1, then health status is defined as being free of activity limitation, and if STATUS equals 2, health status is

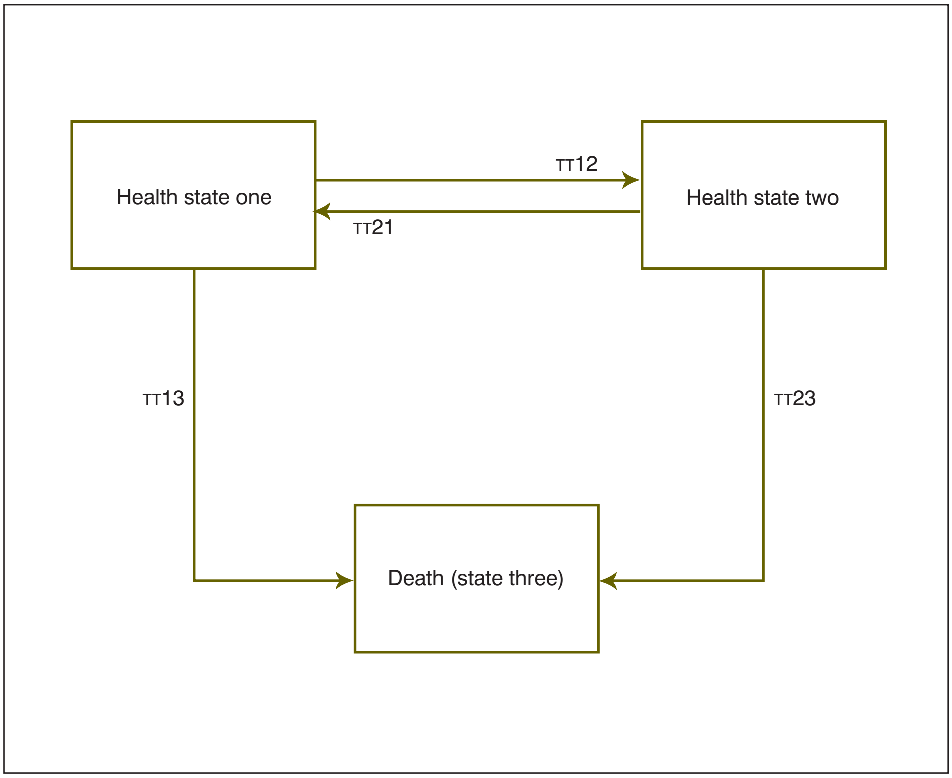


Figure 2. Health states and transitions across states: A three-state model

defined as having activity limitation. At time ($t + 1$), STATUS equals 1 and STATUS equals 2 are still defined the same way, and STATUS equals 3 stands for death.

The annual health transition probabilities for the model showing all the possible transitions can be arranged in a 3 by 3 matrix and presented as follows:

$$ {}_n p_x = \begin{bmatrix} {}_n p_x^{11} & {}_n p_x^{12} & {}_n p_x^{13} \\ {}_n p_x^{21} & {}_n p_x^{22} & {}_n p_x^{23} \\ 0 & 0 & 1 \end{bmatrix} $$

where, p is the health transition probability, x is age, and n is 1 year. The superscripts indicate the various health statuses where 1 equals without activity limitation, 2 equals with activity limitation, and 3 equals death. Because transitions from death (STATUS equals 3) to STATUS equals 1 or STATUS equals 2 are impossible, in the third row of the matrix, ${}_n p_x^{31} = {}_n p_x^{32} = 0$ and ${}_n p_x^{33} = 1$ because death is an absorbing state.

At each age, health state transition probabilities are estimated using a multinomial logistic regression of the form (52)

$$ 1\mathrm{n} \left[\frac{p_{ij}}{p_{ii}} \right] = \beta_{ij0} + \beta_{ij1}\, Age, $$

where, p is the health transition probability, i is the health state at time (t), and j is the health state at time (t + 1).

The IMaCh Program

IMaCh is a computer program that makes use of maximum likelihood and the interpolation of Markov chains. The program was developed by the Intstut National d'Etudes Démographiques (INED) in Paris (53), and it is jointly sponsored by INED and Euro-REVES (other similar computer programs have been developed by Hayward (54), Laditka and Wolf (26), and Lynch and Brown (55)).The computer program is designed to estimate parameters of transition probabilities between an initial and final health status in accordance with the health models, such as the ones presented in this report. The input data includes dates of birth and death, dates and health statuses at each interview, and covariates that are considered key to the analysis. The program outputs numerous estimates in the form of tables and figures, and all the outputs are saved in hypertext files (56).

IMaCh is a freely distributed program and runs on Linux or Windows NT or XP. The program is designed to calculate probabilities of transitions from one health state to another and vice versa. The program calculates total life expectancy and life expectancies in different health states as well as the standard errors of the estimated life and healthy life expectancies. The program provides the researcher with the option of data stratification and using or not using sample survey weights, but it does not account for some aspects of survey design such as clustering. This could be accomplished using different statistical approaches or statistical packages. Available statistical packages include SUDAAN (57) and WesVarPC (58)). The program has an option of including covariates through a multinomial logistic regression. The program adjusts for data with delays between interviews or missing values using interpolation or extrapolation, and it outputs both population and status life table estimates.

Another important feature of the program is its capability of calculating health transition probabilities using a new method first introduced by Laditka and Wolf (26). Because of this capability, the program can be used to calculate transition probabilities by partitioning longitudinal survey intervals into subintervals. That is, the program can be used to calculate transition probabilities for an interval longer than a year, exactly 1 year (12 months), or shorter than a year. To date, the program has been used for the following: to estimate health expectancies using data from the first LSOA (56); to estimate disability-free life expectancy of the older French population (59); to demonstrate the use of retrospective health information in longitudinal studies (60); to study the impact of obesity on active life expectancy in the older American population (45); to study the influence of education on disability-free life expectancy among the older Italian population (61); to estimate active life expectancy in people with and without diabetes (62); to estimate socioeconomic status differentials in active life expectancy among the older population in Beijing (63); to investigate educational differences in the dynamics of disability incidence, recovery, and mortality (64); and to study the effect of obesity on disability compared with mortality (65).

Running IMaCh

IMaCh is designed to run using information on two files that are saved in a user-created subdirectory. The first is a *data file* and the second is a *parameter file*. The data file includes the formatted IMaCh input data in a text format. The parameter file contains information that enables IMaCh to relate the specification of the model and the location of the input data file. A sample table with formatted IMaCh input data, including a brief explanation of the characteristics of the formatted data, is given in Appendix II. A discussion of the parameter file is given in Appendix III. A copy of the parameter file is also given in this appendix.

Because IMaCh is a PC-based program, the time needed to run IMaCh jobs depends mainly on the size of the data and the type of PC used. Most of the time, each run may take from a few minutes to about an hour. In general, the larger the size of the data set or the greater the number of covariates included in the analysis, the more time the job will take to run. Also, a PC with larger memory and a higher processing speed will run the job faster.

Once IMaCh starts running, it creates one folder and log files containing hypertext documents and outputs them to the subdirectory where the parameter file resides. The log files created at each run hold detailed information about the run and figures of various parameters plotted by the program. The folder created by the program holds several more files of tables and figures. All the tables are saved in hypertext files. IMaCh outputs provide both first- and second-moment estimates.

Expected Years With and Without Activity Limitation

Outputs of IMaCh include the estimated parameters of status transition. The estimated parameters of activity status transition using white males in the United States aged 70 years and over as an example are presented in Table A. The only covariate that is included to estimate the transition probabilities is age. All the parameters are significant at the 0.05 level of significance. At each age, transition probabilities for mortality or activity limitation are calculated using these parameters.

Total life expectancy and expected years with and without activity limitation are calculated based on the probabilities of health transitions. Activity limitation is defined as having difficulty in at least one of the following six ADLs: bathing, dressing, eating, moving in and out of a bed or a chair, walking, or using the toilet.

Table B presents total life expectancy and expected years of life with and without activity limitation using white males as an example for the years 1997–2000. The standard errors of the estimated life expectancies are given in parentheses. Total life expectancies and expected years in each health state are calculated using weighted data. The estimates indicate that in the period 1997–2000, if mortality and activity limitation-causing conditions remained unchanged, an average white male 70 years of age would expect to live a total of about 13 years. He would expect to live 8 of the 13 years free of any activity limitation, whereas he would expect to live nearly 5 of the 13 years with limitation in at least one of the six ADLs. On the other hand, under the same mortality and activity limitation-causing conditions, a white male aged 95 years would expect to live about 2.8 years, of which more than 80 percent would be with activity limitation.

Expected Years in Good and Poor Health

Expected years in good or better health and expected years in fair or poor health (hereafter referred to as expected years in good health and expected years in poor health) for females in the United States aged 70 years and over with and without college education are calculated using the MCBS data for the period 2001–2002. The parameters presented in Table C are estimated using the following model.

$$1n\left[\frac{p_{ij}}{p_{ii}}\right] = \beta_{ij\,0} + \beta_{ij1}\,\text{Age} + \beta_{ij\,2}\,\text{Education},$$

where, p is health transition probability, i is health state at time (t), and j is health state at time $(t + 1)$. Level of education is coded as 1 for those with some years of college education and 0 for those with lower than college education (including those with no education).

Some of the parameters in Table C are significant at the 95 percent level, whereas some are not. Compared with

Table A. Estimated parameters and standard errors of activity status transitions for white males aged 70 years and over: United States, 1997–2000

		Variable			
Origin state	Destination state	Constant	SE	Age	SE
Active	Inactive	[†]–8.39	1.21	[†]0.09	0.02
Active	Death	[†]–10.35	1.71	[†]0.12	0.02
Inactive	Active	[†]5.49	1.97	[†]–0.09	0.03
Inactive	Death	[†]–7.06	1.33	[†]0.08	0.02
Log likelihood	. . .	3,335.59	. . .	. . .	. . .

[†] *t* greater than 1.96.

. . .Category not applicable.

NOTES: Estimates are based on weighted data. SE is standard error.

SOURCE: CDC/NCHS, The Second Longitudinal Study of Aging.

Table B. Total life expectancies and expected years with and without activity limitation and standard errors for white males aged 70 years and over, by age: United States, 1997–2000

			Expected years			
Age	Total life expectancy	SE	Without activity limitation	SE	With activity limitation	SE
70.	13.1	0.4	8.2	0.4	5.0	0.3
75.	9.9	0.3	5.3	0.3	4.6	0.3
80.	7.2	0.3	3.2	0.2	4.0	0.3
85.	5.2	0.3	1.7	0.2	3.5	0.3
90.	3.8	0.3	0.9	0.2	2.9	0.3
95.	2.8	0.3	0.4	0.2	2.4	0.3

NOTES: Estimates are based on weighted data. SE is standard error.

SOURCE: CDC/NCHS, The Second Longitudinal Study of Aging.

Table C. Estimated parameters and standard errors of health status for females aged 70 years and over: United States, 2001–2002

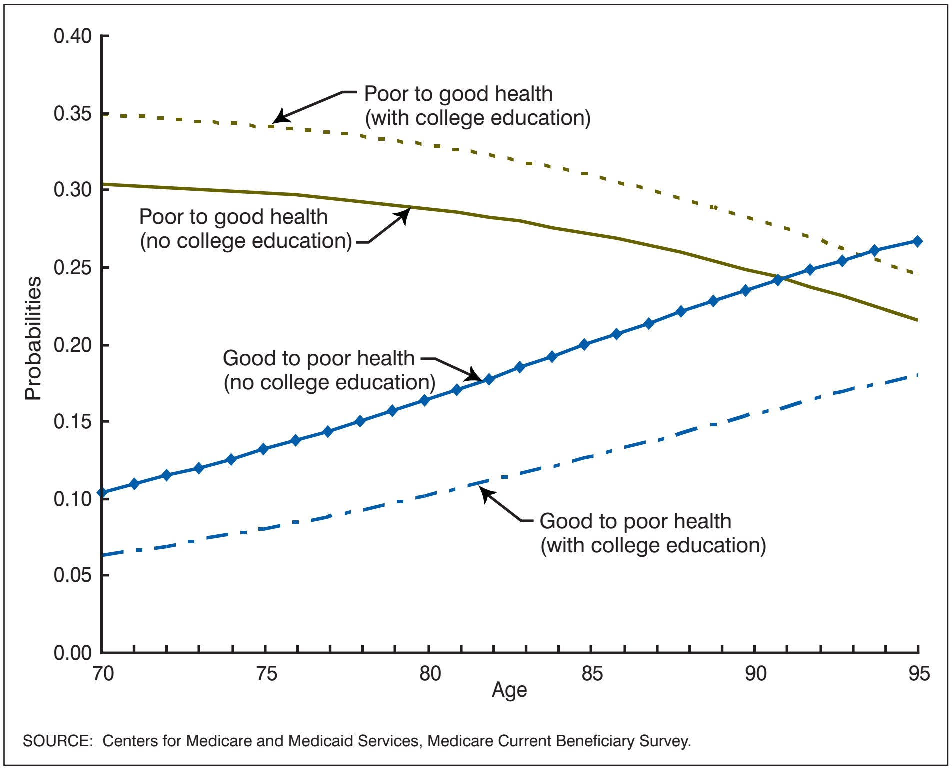
		Variable					
Origin state	Destination state	Constant	SE	Age	SE	Education	SE
Good health	Poor health	†–5.38	0.65	†0.05	0.01	†–0.55	0.12
Good health	Death	†–14.45	1.27	†0.14	0.01	–0.23	0.22
Poor health.	Good health	–0.55	0.87	†-0.03	0.01	–0.22	0.18
Poor health.	Death	†–9.06	3.66	†0.10	0.01	–0.11	0.10
Log likelihood	. . .	5,899.47	. . .	. . .	. . .	. . .	. . .

† *t* greater than 1.96.

. . . Category not applicable.

NOTES: Level of education: 0 is no college education, 1 is at least some college education. SE is standard error.

SOURCE: Centers for Medicare and Medicaid Services, Medicare Current Beneficiary Survey.



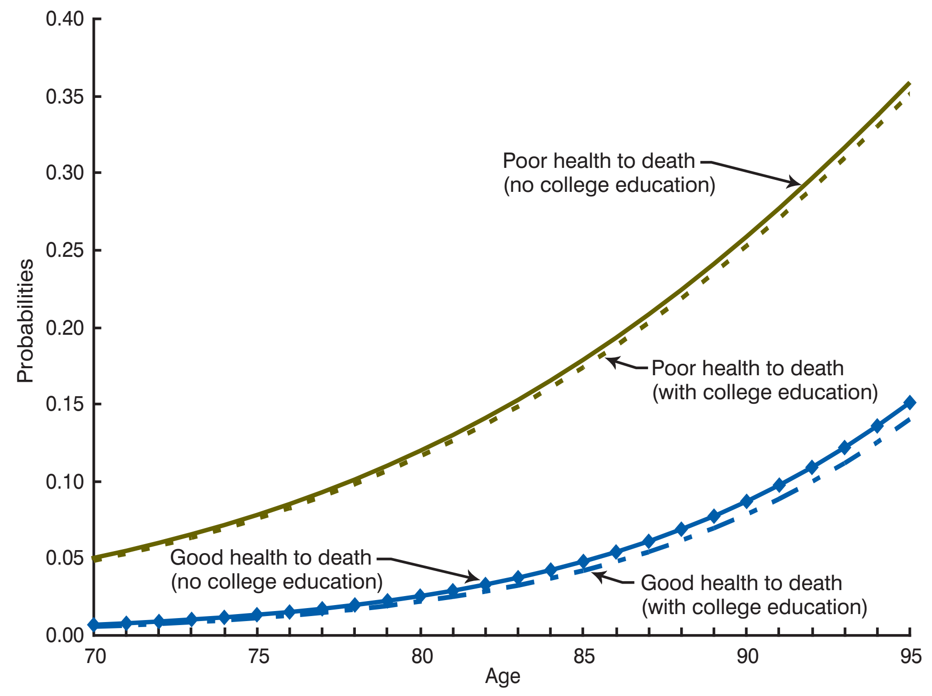
SOURCE: Centers for Medicare and Medicaid Services, Medicare Current Beneficiary Survey.

Figure 3. Probabilities of transition from good to poor health and vice versa, by level of education: Females, 2001–2002

level of education, age is a stronger determinant of transition from one health state to another. The transition probabilities from good to poor health and back by level of education using females as an example are shown in Figure 3. The transition probabilities from each health status to death for each educational group are presented in Figure 4. As would be expected, the transition probabilities from good health to poor health and from each health status to death increase with age, and the transition probabilities from poor health to good health decrease with age.

The estimated total life expectancy, expected years in good health, and expected years in poor health stratified by level of education for the same sample at selected ages are presented in Table D. As is expected, the results show health disparities by education both in terms of total life expectancy as well as expected years in good health. On average, if mortality and the way people assess their own health are assumed to remain unchanged, a female 70 years of age with at least 1 year of college education would expect to live nearly 17 years more, whereas a female of the same age but with no college education would expect to live 15.5 years more. In terms of healthy life expectancy, the average female 70 years of age with some college education would expect to spend about 13.5 years in good health, whereas an average female of the same age but without any college education would expect to spend about 11.2 years in good health. An average female aged 70 years and without any college education would also expect to spend 4.3 years in poor health, which would be about eight-tenths of a year more than the time that an average female aged 70 years with at least 1 year of college education would expect to spend in poor health. Similar health disparities between age cohorts with and without college education persist in all the other ages as well.

In addition to calculating life expectancies, expected years in good health, and expected years in poor health for the total population, the IMaCh program also calculates expected years in good health and expected years



SOURCE: Centers for Medicare and Medicaid Services, Medicare Current Beneficiary Survey.

Figure 4. Probabilities of transition from good or poor health to death, by level of education: Females, 2001–2002

Table D. Total life expectancies, expected years in good and poor health, and standard errors, for females aged 70 years and over, by level of education and age: United States, 2001–2002

Age	Total life expectancy	SE	In good health	SE	In poor health	SE
At least 1 year of college education						
70	17.0	0.9	13.5	0.5	3.5	0.4
75	13.2	0.8	10.1	0.7	3.0	0.4
80	9.8	0.7	7.3	0.6	2.5	0.3
85	7.1	0.6	5.1	0.5	2.0	0.3
90	4.9	0.5	3.4	0.4	1.5	0.2
95	3.4	0.4	2.2	0.3	1.1	0.2
No college education						
70	15.5	0.5	11.2	0.4	4.3	0.2
75	11.9	0.4	8.3	0.3	3.6	0.2
80	8.8	0.3	5.9	0.3	2.9	0.2
85	6.3	0.3	4.0	0.3	2.3	0.2
90	4.4	0.3	2.7	0.2	1.7	0.2
95	3.0	0.3	1.7	0.2	1.3	0.2

NOTE: SE is standard error.

SOURCE: Centers for Medicare and Medicaid Services, Medicare Current Beneficiary Survey.

in poor health by initial health status. Expected years in good health and expected years in poor health by initial health status and level of education for the same sample of females at selected ages are presented in Table E. At each level of education, initial health is another important source of disparities both in expected years in good health as well as expected years in poor health. This is true at each age and at each level of education. Assuming that mortality and health conditions remain unchanged, an average female 70 years of age with college education who was in good health would expect to spend nearly 13.9 years in good health and only about 3.3 years in poor health. A female 70 years of age with the same level of education but who was in poor health would expect to spend only about 10.6 years in good health and more than 4.8 years in poor health.

Initial health status was also a cause for similar expected health disparities among the older female population without college education. If mortality and health conditions remain unchanged, an average female 70 years of age without college education and who was in good health would expect to spend about 11.7 years in good health and about 4.1 years in poor health. A female 70 years of age with the same level of education but who was in poor health would expect to spend slightly more than 9 years in good health but more than 5 years in poor health. Again, disparities clearly existed at all the ages and at both levels of education, within the group in good health and the group in poor health (as shown in Table E).

Discussion

Summary and Conclusion

The main purpose of this statistical report is to introduce the reader to a method of estimating healthy or active life expectancy using longitudinal survey data. Healthy and active life expectancies are two of the population health expectancy measures that incorporate mortality and morbidity. Population health measures such as healthy or active life expectancy can be calculated using data from longitudinal or cross-sectional surveys.

Given the choice, population health measures based on longitudinal survey data are preferable to population health measures based on cross-sectional surveys. Data from longitudinal surveys are preferred because such data provide information on changes in the health status of the population in the recent past. Such information about the population is the most appropriate information needed for forecasting the

Table E. Expected years in good and poor health and standard errors, for females aged 70 years and over, by initial health status, level of education, and age: United States, 2001–2002

Age	Initial health status							
	Good health				Poor health			
	Good health	SE	Poor health	SE	Good health	SE	Poor health	SE
With at least 1 year of college education								
70	13.9	0.8	3.3	0.9	10.6	0.4	4.8	0.5
75	10.7	0.7	2.8	0.8	7.5	0.3	4.1	0.4
80	7.9	0.6	2.3	0.6	5.0	0.3	3.3	0.3
85	5.7	0.5	1.8	0.5	3.2	0.2	2.6	0.3
90	4.0	0.4	1.3	0.4	1.9	0.2	2.0	0.3
95	2.7	0.3	1.0	0.3	1.1	0.2	1.5	0.2
With no college education								
70	11.7	0.4	4.1	0.5	9.3	0.2	5.2	0.3
75	8.9	0.3	3.4	0.3	6.5	0.2	4.3	0.2
80	6.5	0.3	2.7	0.3	4.4	0.2	3.4	0.2
85	4.6	0.2	2.1	0.2	2.8	0.2	2.7	0.2
90	3.2	0.2	1.6	0.2	1.7	0.2	2.0	0.2
95	2.2	0.2	1.1	0.2	1.1	0.1	1.5	0.1

NOTE: SE is standard error.

SOURCE: Centers for Medicare and Medicaid Services, Medicare Current Beneficiary Survey,

future health of the population.

Estimating healthy or active life expectancy using longitudinal survey data has also a model-related advantage. Healthy or active life expectancy estimated using cross-sectional survey data is partially based on a standard single-decrement life table. The model assumes that health-related events are irreversible. But in reality, during a given period of time such as a year, entering and going out of a given health state may happen frequently. Hence, any health expectancy model has to account not only for transitions from good to poor health or active life to life with activity limitation, but also for transitions from poor to good health or from life with limitation to life without activity limitation. A model that uses longitudinal survey data has the advantage of measuring health transitions in both directions. That is, the method has the advantage of accounting for possible recoveries or transitions from poor health to good health.

However, health expectancies are estimated using longitudinal survey data less frequently, despite the key advantages health expectancies estimated using longitudinal data have over health expectancies that are estimated using cross-sectional survey data. The two most often mentioned reasons for this relate to the unavailability of data and the difficulties of the technique underlying the estimation method (38,66). Cross-sectional survey data are abundantly available in all the developed and many developing countries. On the other hand, longitudinal surveys are scarce even in the United States (67). The small numbers of longitudinal surveys that are available include only the older population. The reason for the scarcity of longitudinal survey data according to some is that such data are expensive and time consuming to collect (46). However, others suggest that collecting data through a follow-up survey about the persons that participated in a baseline survey is neither more difficult nor more expensive than collecting data based on two separate cross-sectional surveys. An empirical study that has assessed the relative costs of cross-sectional and longitudinal surveys has shown that sometimes longitudinal surveys might be less expensive than repeated cross-sectional surveys (68).

The second most often mentioned reason for estimating health expectancies using cross-sectional survey data is that models that account for recovery from morbidity or illness are more complex and are based on more difficult underlying assumptions. With the advent of powerful computers, such model-related problems are solved by computer programs such as IMaCh. IMaCh is a PC-based program that is designed to calculate probabilities of transitions from one health state to another and vice versa. The program calculates total life expectancy and life expectancies in different health states with their standard errors. The program provides the researcher the options of stratification, using sample survey weights, and including covariates through a multinomial logistic regression model. The program adjusts for data with delays between interviews or missing values using interpolation or extrapolation, and it outputs both population and status-based life table estimates.

The program can be used with longitudinal survey data at the national, state, or local level and for public health as well as clinical studies. Healthy life expectancy can be estimated using data from at least two waves of a longitudinal survey or data from a baseline and one follow-up survey. For public health studies, data could be collected through continuous longitudinal panel surveys with a rotating panel design (MCBS) or through longitudinal surveys without rotating panel design at an interval of 5 or 10 years (LSOA I and II). When resource constraints make using one of these methods difficult, data could be collected based on a cross-sectional survey followed by a single follow-up at a time, and this can be repeated as soon as more resources are available. Also, when the time interval between two consecutive waves of a longitudinal survey or the time between a baseline and a follow-up survey is longer than a year, health transition probabilities can be calculated for each month separately and annual health transition probabilities can be aggregated based on 12 monthly probabilities.

References

1. National Center for Health Statistics. Health, United States, 2006, with chartbook on trends in the health of Americans. Hyattsville, MD. 2006.
2. Schoeni RE, Liang L, Bennett J, et al. Trends in old-age functioning and disability in Japan, 1993–2002. Population Studies 60(1):39–53. 2006.
3. World Health Organization. World health report 2001: Basic indicators for all member states, statistical annex, Table A. p.136–43. Geneva. 2002.
4. Hickman JC, Estell RJ. On the use of partial life expectancies in setting health goals. Am J Public Health 59(12): 2243–50. 1969.
5. Field MJ, Gold MR, editors. Summarizing population health: Directions for the development and application of population metrics. Washington, DC: Institute of Medicine, National Academy of Sciences, National Academy Press. 1998.
6. Sanders BS. Measuring community health levels. Am J Public Health 54(7):1063–70. 1964.
7. U.S. Department of Health, Education, and Welfare. Toward a Social Report. Washington, DC. 1969.
8. Sullivan DF. Conceptual problems in developing an index of health. National Center for Health Statistics. Vital Health Stat 2(17). 1969.
9. Sullivan DF. A single index of mortality and morbidity. HSMHA Health Report 86(4):347–54. 1971.
10. Sullivan DF. Disability components for an index of health. National Center for Health Statistics. Vital Health Stat 2(42). 1971.
11. World Health Organization. Modern management methods and organization of health services. Public Health Papers No. 55. Geneva. 1974.
12. Organisation for Economic Co-operation and Development. List of social concerns common to most OECD countries. Paris. 1973.
13. Fries JF. Aging, natural death, and the compression of morbidity. N Engl J Med 303(3):130–5. 1980.
14. Fries JF. The compression of morbidity, near or far? Milbank Q 67(2):208–32. 1989.
15. Gruenberg EM. The failure of success. Milbank Memorial Fund Quarterly: Health and Society 55(1):3–24. 1977.
16. Manton KG. Changing concepts of morbidity and mortality in the elderly population. Milbank Memorial Fund Quarterly: Health and Society 60(2):183–244. 1982.
17. Hayward MD, Heron M. Racial inequality in active life among adult Americans. Demography 36(1):77–91. 1999.
18. Mathers CD. Health expectancies: An overview and critical appraisal. In: Murray CJL, et al., editors. Summary Measures of Population Health: Concept, Ethics, Measurement and Application. Geneva: World Health Organization. p. 177–204. 2002.
19. Murray CJL, Mathers CD, Salomon JA, et al. Health gaps: An overview and critical appraisal. In: Murray CJL, et al., editors. Summary Measures of Population Health: Concept, Ethics, Measurement and Application. Geneva: World Health Organization. p. 233–44. 2002.
20. Murray CJL, Lopez AD. On the comparable quantification of health risks: Lessons from the global burden of disease study. Epidemiology 10(5):594–605. 1999.
21. Hyder AA, Rotlant G, Morrow RH. Measuring the burden of disease: Healthy life years (HeaLYs). Am J Public Health 88(2):196–202. 1998.
22. Hyder AA, Morrow RH. Healthy life years (HeaLYs). In: Murray CJL, et al., editors. Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications. Geneva: World Health Organization. p. 245–57. 2002.
23. Robine JM, Romieu I. Healthy active aging: Health expectancies at age 65 in the different parts of the world. REVES paper no 318. A paper contributed to the World Health Organization, Geneva. 1998.
24. Mathers CD, Robine JM, Wilkins R. Health expectancy indicators: Recommendations for terminology. In: Mathers CD, McCallum J, Robines JM, editors. Advances in Health Expectancies. Canberra: Australian Institute of Health. p. 18–33. 1994.
25. Katz S, Laurence MD, Branch G, et al. Active life expectancy. N Engl J Med 17:1218–24. 1983.
26. Laditka SB, Wolf DA. New methods for analyzing active life expectancy. J Aging Health 10(2):214–41. 1998.
27. Rogers A. Introduction to multi-regional mathematical demography. New York: John Wiley. 1975.
28. U.S. Bureau of Labor Statistics. Tables of working life: Length of working life for men. Bulletin 1001. Washington, DC: U.S. Government Printing Office. 1950.
29. Molla MT, Wagener DK, Madans JH. Summary measures of population health: Methods for calculating healthy life expectancy. Healthy People 2010 Statistical Notes, no 21. Hyattsville, MD: National Center for Health Statistics. August 2001.
30. Laditka SB, Hayward MD. The evaluation of demographic methods to calculate health expectancy. In: Rubine JM, Jagger C, Mathers CD, et al., editors. Determinants of Health Expectancy. West Sussex, UK: John Wiley. p. 221–34. 2003.
31. Bartholomew DJ. Stochastic Models for Social Processes. London: John Wiley. 1973.
32. Schoen R. Constructing increment-decrement life table. Demography 12:313–24. 1975.
33. Hoem JM, Fong MS. A Markovian chain model of working life tables, WP–2. Laboratory of Actuarial Mathematics. Copenhagen: University of Copenhagen. 1977.
34. Schoen R, Woodrow K. Labor force status life tables for the United States, 1972. Demography 17:297–322. 1980.
35. Willekens F. A multi-state analysis: Tables of working life. Environment and Planning A12:563–88. 1980.
36. Rogers A, Rogers RG, Branch LG. A multi-state analysis of active life expectancy. Public Health Rep 104:222–6. 1989.
37. Rogers A, Rogers RG, Belanger A. Longer life but worse health? Measurement and dynamics. Gerontologist 30(5):640–9. 1990.
38. Land KC, Guralnik JM, Blazer DG. Estimating increment-decrement life tables with multiple covariates from panel data: The case of active life expectancy. Demography 31(2):297–319. 1994.
39. Crimmins EM, et al. Changing mortality and morbidity rates and the health status and life expectancy of the older population. Demography 31:159–75. 1994.
40. Nusselder WA. A multi-state life table analysis of health expectancy for the Netherlands. In: Nusselder WA, editor. Compression or Expansion of Morbidity: A Life Table Approach. Amsterdam: Thesis Publishers. p. 85–128. 1998.
41. Anderson PK. Multi-state models in survival analysis: A study of nephropathy and mortality in diabetes. Stat Med 7:661–70. 1988.

42. Silverstein MD, et al. A prognosis in SLE: Comparison of Markov model to life table analysis. J Clin Epidemiol 41(7):623–33. 1988.
43. Land KC, Hough GC. New method tables of school life, with applications to U.S. data from recent school years. Journal of the American Statistical Association 84(405):63–75. 1989.
44. Xie Y. A demographic approach to studying the process of becoming a scientist/engineer. In: Career in Science and Technology: An International Perspective. Washington, DC: National Academy Press. p. 43–57. 1995.
45. Reynolds SL, Saito Y, Crimmins EM. The impact of obesity on active life expectancy in older American men and women. Gerontologist 45(4): 438–44. 2005.
46. Freedman D, Thornton A, Camburn D, et al. The life history calendar: A technique for collecting retrospective data. Sociological Methodology 18:37–68. 1988.
47. Gordis L. Epidemiology. 2d ed. Philadelphia: W. B. Sanders. 2000.
48. Selvin S. Statistical Analysis of Epidemiologic Data. 2d ed. New York: Oxford University Press. 1996.
49. National Center for Health Statistics. The Second Longitudinal Study of Aging 1994–2000 [online]. Available from: http://www.cdc.gov/nchs/about/otheract/aging/lsoa2.htm. 2007.
50. Adler JS. A profile of the Medicare Current Beneficiary Survey. Health Care Financ Rev 15(4):153–63. 1994.
51. Centers for Medicare and Medicaid Services, Office of Strategic Planning, Information, and Methods Group. CY 1997 cost and use: Public use file documentation. Baltimore, MD: Center for Medicare and Medicaid Services. 2000.
52. Cox DR. Regression models and life tables. Journal of the Royal Statistical Society, Series B (Methodology) 34:187–220. 1972.
53. Brouard N, Liévre A. IMaCh: Maximum Likelihood Computer Program using Interpolation of Markov Chains. Paris: Institut National d'Etudes Démographiques. With a contribution to the mathematical part by C. R. Heathcote of the Australian National University, Canberra. Available from: http://euroreves.ined.fr/imach/. Accessed June 2007.
54. Hayward MD, Grady WR. Work and retirement among a cohort of older men in the United States, 1966–1983. Demography 27(3):337–56. 1990.
55. Lynch SM, Brown JS. A new approach to estimating life tables with covariates and constructing interval estimates of life table quantities. Sociological Methodology 35(1):189–238. 2005.
56. Liévre A, Brouard N, Heathcote CR. The estimation of health expectancies from cross-longitudinal surveys. Mathematical Population Studies 10(4):211–48. 2003.
57. Shah BV, BG Barnwell, Bieler GS. SUDAAN: Software for the statistical analysis of correlated data. Research Triangle Park, NC: Research Triangle Institute. 2001.
58. Westat. WesVar: Software for analysis of data from complex samples. Rockville, MD: Westat, Inc. Available from: http://www.westat.com/wesvar/. Accessed June 2007.
59. Pérés K, Jagger C, Liévre A, et al. Disability-free life expectancy of older French people: Gender and education differentials from the PAQUID cohort. European Journal of Aging 2:225–33. 2005.
60. Molla MT, Lubitz J. Longitudinal surveys, retrospective information and population health measures. Proceedings of Statistics Canada Symposium 2005: Methodological Challenges for Future Information Needs. Ottawa, Canada. 2005.
61. Minicuci N, Noale N. Influence of level of education on disability free life expectancy by sex: The ILSA study. Exp Gerontol 40:997–1003. 2005.
62. Jagger C, Goydere E, Clark M, et al. Active life expectancy in people with and without diabetes. Journal of Public Health Medicine 25:42–6. 2003.
63. Kaneda T, Zimmer Z, Tang Z, et al. Socioeconomic status differentials in life and active life expectancy among older adults in Beijing. Disabil Rehabil 27(5):241–51. 2005.
64. Jagger C, Matthews R, Melzer D, Matthews F, Brayne C, and the Medical Research Council Cognitive Function and Ageing Study. Educational differences in the dynamics of disability incidence, recovery and mortality: Findings from the MRC Cognitive Function and Ageing Study (MRC CFAS). Int J Epidemiol 36:358–65. 2007.
65. Al Snih S, Ottenbacher KJ, Markides KS, Kuo Y-F, Eschbech K, Goodwin JS. The effect of obesity on disability vs mortality in older Americans. Arch Intern Med 167(8):774–80. 2007.
66. Stone RA. Markovian education model and other examples linking social behavior to the economy. Journal of the Royal Statistical Society: Series A 135:511–44. 1972.
67. Sondik E. Summary measures of population health: Application and issues in the United States. In Murray CJL, et al., editors. Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications. Geneva: World Health Organization. p. 75–81. 2002.
68. Pearson RW. The advantages and disadvantages of longitudinal surveys. In: Federal Committee on Statistical Methodology, Office of Information and Regulatory Affairs. Seminar on Quality of Federal Data: Statistical Policy Working Paper 20. Washington, DC: Office of Management and Budget. p. 407–24. 1991.

Appendix I

Questions About Activities of Daily Living From the Second Longitudinal Study of Aging (LSOA II) and About Self-Assessed Health from the Medicare Current Beneficiary Survey (MCBS) Questionnaires

Questions about activities of daily living from LSOA II

[Respondents were asked to circle 1 or 2.]

18a. Because of a health or physical problem, do you have **any** difficulty bathing or showering?

Yes 1
No 2

18b. Because of a health or physical problem, do you have **any** difficulty dressing?

Yes 1
No 2

18c. Because of a health or physical problem, do you have **any** difficulty eating?

Yes 1
No 2

18d. Because of a health or physical problem, do you have **any** difficulty going in and out of bed or chairs?

Yes 1
No 2

18e. Because of a health or physical problem, do you have **any** difficulty walking?

Yes 1
No 2

18f. Because of a health or physical problem, do you have **any** difficulty using the toilet, including getting to the toilet?

Yes 1
No 2

Question about self-assessed health from MCBS

In general, compared to other people (your/sample person's) age, would you say that (your/his/her) health is

[1] Excellent
[2] Very good
[3] Good
[4] Fair or
[5] Poor?

Appendix II

Data Format

The data needed to calculate life expectancies in the various health states using IMaCh include the identification number of a sample member, date of birth, date of death (for decedents), date of first interview, heath status at the time of first interview, date of the follow-up survey, health status at the time of the follow-up survey, sample weight, and covariates such as sex, race, and level of education. Although the sample weights and the covariates are optional, all the other variables are required. Because the minimum length of an interval the program uses to calculate transition probabilities is 1 month, only month and year are required for the various dates. Regardless of the inclusion of the optional variables, the program requires the input data to be presented in a specific format. The following data with 10 records will be used to explain the format of the IMaCh input data.

The data in Figure I include 10 records and 10 fields. The fields included in the data are personal identification number, sex, level of education, sample weight, date of birth, date of death (for decedents), date of first interview, health status at first interview, date of second interview, and health status at the time of the second interview. The data indicate that of the 10 survey participants, 4 died between the first and the second interviews and 2 respondents did not participate in the second interview. Note that all the dates in these data are given in the SAS date format ‘‘date9.’’

The data presented in Figure II are the same as in Figure I. However, the data in Figure II are formatted as IMaCh input data. The formatting changed the way the various dates are presented. In Figure I, the dates included the day, month, and year. In Figure II, the dates included only the month and year separated by a slash. The months, which were given in three-letter character format in Figure I, are changed to months in numeric format in Figure II. The other highly visible change due to formatting is the way missing values are presented. In the IMaCh input data, a month with a missing value must be recoded as ‘‘99,’’ a year with a missing value as ‘‘9999,’’ and health status with a missing value as ‘‘-1.’’

The data in Figure II are also formatted to fit the model. In the input data, the sequence of the fields is very important. Each field has to be separated from the next by a blank space. The first field is for the personal identification number. The next two fields are for covariates such as sex and level of education. These are followed by the field for the sample weight, which is followed by the field for date of birth. The next two fields stand for the date of death and the date of the first interview, which are followed by the field for health status at the time of the first interview. The last two fields stand for the date of the second interview and health status at the time of the second interview.

The fields that are required to run the model include the personal identification number, date of birth, date of death, date of the first interview, health status at the time of the first interview, date of the second interview, and health status at the time of the second interview. Each respondent’s age is calculated based on date of birth and the date of the first interview or date of death (for decedents). Because no extra covariates are included in the model and the sample weight is not used, values of the two additional covariates (the second and third fields in the input data) and the sample weight (the fourth field) are recoded as ‘‘1.’’

00695525	1	1	1	5/1892	12/2001	8/2001	2	99/9999	3
00509374	1	1	1	1/1896	12/2001	99/9999	2	99/9999	3
00570976	1	1	1	4/1896	99/9999	10/2001	1	11/2002	2
00705029	1	1	1	9/1897	9/2002	12/2001	2	3/2002	3
00567926	1	1	1	11/1898	99/9999	9/2001	1	10/2002	2
00547312	1	1	1	4/1899	99/9999	10/2001	1	9/2002	2
00559692	1	1	1	2/1899	99/9999	11/2001	1	99/9999	-1
00636532	1	1	1	8/1899	99/9999	9/2001	2	9/2002	1
00699311	1	1	1	8/1899	7/2002	11/2001	1	3/2002	3
00492980	1	1	1	10/1900	99/9999	10/2001	1	99/9999	-1

Figure II. Illustrative example of data in IMaCh input data format

00695525	1	2	1378	13MAY1892	02DEC2001	22AUG2001	2	.	3
00509374	2	1	1654	12JAN1896	24DEC2001	.	2	.	3
00570976	2	3	1824	06APR1896	.	03OCT2001	1	29NOV2002	2
00705029	1	2	1697	07SEP1897	11SEP2002	11DEC2001	2	04MAR2002	3
00567926	1	1	2094	03NOV1898	.	18SEP2001	1	20OCT2002	2
00547312	2	1	1900	17APR1899	.	15OCT2001	1	20SEP2002	2
00559692	2	3	1672	21FEB1899	.	01NOV2001	1	.	.
00636532	2	3	1493	14AUG1899	.	17SEP2001	2	11SEP2002	1
00699311	2	2	1613	20AUG1899	14JUL2002	05NOV2001	1	14MAR2002	3
00492980	1	1	1356	04OCT1900	.	10OCT2001	1	.	.

Figure I. Illustrative example of health data from the baseline and follow-up surveys

Appendix III

The Parameter File Format

IMaCh is designed to run using information on two files that need to be included in one of the subdirectories that are already created for the purpose. The first is a *data file* and the second is a *parameter file*. The data file includes the formatted IMaCh input data shown in Figure II in a text format. The parameter file shown in Figure III contains information that enables IMaCh to relate the specification of the model and the location of the input data file.

The parameter file includes the title, which in this case is “MLE,” or maximum likelihood estimates. This is followed by the name of the subdirectory where the input data reside. Note that the subdirectory name ends with the dataset name. This is followed by “lastobs,” or last observation, which indicates the total number of records in the input data (2,444). The word “firstpass” stands for the first interview, and “lastpass” stands for the last interview. In this example, “lastpass” equals 2. But, “lastpass” will equal 3 if the data used were from three interviews, 4 if the data used were from four interviews, and so on. The term “stepm” stands for the length of the interval used to calculate the transition probabilities and is measured in months. The program calculates transition probabilities by intervals as short as a month and also by intervals longer than a year.

The number of covariate columns is indicated by “ncovcol” and equals 2. The term “nlstate” stands for the number of health states that are transients and equals 2, and “ndeath” stands for the number of absorbing states and equals 1. The maximum number of waves is indicated by “maxwav” and equals 2. The term “mle” stands for maximum likelihood estimates, and its value equals 1 to indicate that maximum likelihood is used in the process of calculating health expectancies.

When the sample weights are not included in the analysis, then weight is assigned a value of 0 (but 1 in the data file). The value of weight equals 1 when the sample weights are included in the analysis of the data. If none of the covariates in fields two and three are included in the analysis, then “model” equals dot. If a covariate in field two of the input data is included, then “model” equals “V1.” If the covariate in field three is included instead of the covariate in field two, then “model” equals “V2.” When covariates in fields two and three are both included, then “model” equals “V1+V2.” The model indicator is followed by the parameters, gradient, and covariance matrices. The two terms “agemin” and “agemax” stand for the calculation of the period prevalence. The minimum and maximum ages for the calculation of health expectancies are denoted by “bage” (beginning age) and “fage” (final age).

```
# title=Expected years with and without activity limitations datafile=. . . \dataname.txt lastobs=2444 firstpass=1
lastpass=2
ftol=1.000000e-008 stepm=24 ncovcol=2 nlstate=2 ndeath=1 maxwav=2 mle=1 weight=0
model=.
#Parameters
12          0.000    0.000
13          0.000    0.000
21          0.000    0.000
23          0.000    0.000
#scales
12          0.000    0.000
13          0.000    0.000
21          0.000    0.000
23          0.000    0.000
#covariance
121         0.000
122         0.000    0.000
131         0.000    0.000    0.000
132         0.000    0.000    0.000    0.000
211         0.000    0.000    0.000    0.000    0.000
212         0.000    0.000    0.000    0.000    0.000    0.000
231         0.000    0.000    0.000    0.000    0.000    0.000    0.000
232         0.000    0.000    0.000    0.000    0.000    0.000    0.000  0.000
# agemin=minimum age, agemax=maximum age for life expectancy, bage=beginning age, fage=final age.
agemin=70 agemax=95 bage=65 fage=95
```

NOTE: lastobs is last observation or record, firstpass is first wave, lastpass is last wave, stepm is number of months in each interval, ncovcol is number of covariate columns, nlstate is number of health states, ndeath is number of death sates, maxwav is maximum number of waves, and mle is maximum likelihood estimates.

Figure III. Illustrative example of the parameter file

Vital and Health Statistics series descriptions

SERIES 1. **Programs and Collection Procedures**—These reports describe the data collection programs of the National Center for Health Statistics. They include descriptions of the methods used to collect and process the data, definitions, and other material necessary for understanding the data.

SERIES 2. **Data Evaluation and Methods Research**—These reports are studies of new statistical methods and include analytical techniques, objective evaluations of reliability of collected data, and contributions to statistical theory. These studies also include experimental tests of new survey methods and comparisons of U.S. methodology with those of other countries.

SERIES 3. **Analytical and Epidemiological Studies**—These reports present analytical or interpretive studies based on vital and health statistics. These reports carry the analyses further than the expository types of reports in the other series.

SERIES 4. **Documents and Committee Reports**—These are 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.

SERIES 5. **International Vital and Health Statistics Reports**—These reports are analytical or descriptive reports that compare U.S. vital and health statistics with those of other countries or present other international data of relevance to the health statistics system of the United States.

SERIES 6. **Cognition and Survey Measurement**—These reports are from the National Laboratory for Collaborative Research in Cognition and Survey Measurement. They use methods of cognitive science to design, evaluate, and test survey instruments.

SERIES 10. **Data From the National Health Interview Survey**—These reports contain statistics on illness; unintentional injuries; disability; use of hospital, medical, and other health services; and a wide range of special current health topics covering many aspects of health behaviors, health status, and health care utilization. They are based on data collected in a continuing national household interview survey.

SERIES 11. **Data From the National Health Examination Survey, the National Health and Nutrition Examination Surveys, and the Hispanic Health and Nutrition Examination Survey**—Data from direct examination, testing, and measurement on representative samples of the civilian noninstitutionalized population provide the basis for (1) medically defined total prevalence of specific diseases or conditions in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics, and (2) analyses of trends and relationships among various measurements and between survey periods.

SERIES 12. **Data From the Institutionalized Population Surveys**—Discontinued in 1975. Reports from these surveys are included in Series 13.

SERIES 13. **Data From the National Health Care Survey**—These reports contain statistics on health resources and the public's use of health care resources including ambulatory, hospital, and long-term care services based on data collected directly from health care providers and provider records.

SERIES 14. **Data on Health Resources: Manpower and Facilities**—Discontinued in 1990. Reports on the numbers, geographic distribution, and characteristics of health resources are now included in Series 13.

SERIES 15. **Data From Special Surveys**—These reports contain statistics on health and health-related topics collected in special surveys that are not part of the continuing data systems of the National Center for Health Statistics.

SERIES 16. **Compilations of Advance Data From Vital and Health Statistics**—Advance Data Reports provide early release of information from the National Center for Health Statistics' health and demographic surveys. They are compiled in the order in which they are published. Some of these releases may be followed by detailed reports in Series 10–13.

SERIES 20. **Data on Mortality**—These reports contain statistics on mortality that are not included in regular, annual, or monthly reports. Special analyses by cause of death, age, other demographic variables, and geographic and trend analyses are included.

SERIES 21. **Data on Natality, Marriage, and Divorce**—These reports contain statistics on natality, marriage, and divorce that are not included in regular, annual, or monthly reports. Special analyses by health and demographic variables and geographic and trend analyses are included.

SERIES 22. **Data From the National Mortality and Natality Surveys**—Discontinued in 1975. Reports from these sample surveys, based on vital records, are now published in Series 20 or 21.

SERIES 23. **Data From the National Survey of Family Growth**—These reports contain statistics on factors that affect birth rates, including contraception, infertility, cohabitation, marriage, divorce, and remarriage; adoption; use of medical care for family planning and infertility; and related maternal and infant health topics. These statistics are based on national surveys of women of childbearing age.

SERIES 24. **Compilations of Data on Natality, Mortality, Marriage, and Divorce**—These include advance reports of births, deaths, marriages, and divorces based on final data from the National Vital Statistics System that were published as *National Vital Statistics Reports* (NVSR), formerly *Monthly Vital Statistics Report*. These reports provide highlights and summaries of detailed data subsequently published in *Vital Statistics of the United States*. Other special reports published here provide selected findings based on final data from the National Vital Statistics System and may be followed by detailed reports in Series 20 or 21.

For answers to questions about this report or for a list of reports published in these series, contact:

Information Dissemination Staff
National Center for Health Statistics
Centers for Disease Control and Prevention
3311 Toledo Road, Room 5412
Hyattsville, MD 20782

1-866-441-NCHS (6247)
E-mail: nchsquery@cdc.gov
Internet: www.cdc.gov/nchs

U.S. DEPARTMENT OF
HEALTH & HUMAN SERVICES

Centers for Disease Control and Prevention
National Center for Health Statistics
3311 Toledo Road
Hyattsville, MD 20782

CS117909
T31107 (05/2008)
DHHS Publication No. (PHS) 2008–1346, Series 2, No. 146