Conventional multistate life table accounting procedures are based on theoretical assumptions that are appropriate primarily for demographic events. Applying these approaches to the area of health care, however, may lead to serious biases given the frequent turnovers of events such as hospitalization and institutionalization. In addition, traditional approaches have been criticized for failing to capture population heterogeneity. This research introduces a new algorithm to estimate multistate life table indicators regarding health care use, taking advantage of the availability of information on average lengths of stay in hospitals and nursing homes. The survival analysis approach is used to estimate age-specific transition probabilities in order to address the issue of population heterogeneity.

Modeling Multidimensional Transitions in Health Care

XIAN LIU
JERSEY LIANG
University of Michigan
EDWARD JOW-CHING TU
Hong Kong University of Science and Technology
NANCY WHITELAW
Henry Ford Health System

1. INTRODUCTION

Recent decades have witnessed dramatic advances in modeling the demographic implications of social dynamics (Land and Rogers 1982; Schoen 1988). Sociologists and population researchers have developed a number of procedures to generalize the multidimensional transitions in sociologically relevant events such as

AUTHORS' NOTE: *This research was supported by a grant from the Michigan Health Care Education and Research Foundation, 1991-1993 (Grant No. 141 RHCA/90-01). Mark Hayward, Andrei Rogers, and Jay Teachman made several useful suggestions concerning the construction of a multistate life table model. The assistance provided by Joan Bennett, Cathy Fegan, and Albert Lau is gratefully acknowledged.*

SOCIOLOGICAL METHODS & RESEARCH, Vol. 25 No. 3, February 1997 284-317

geographical mobility, marital status patterns, occupational careers, and health status (Hayward and Grady 1990; Rogers 1975; Rogers, Rogers, and Belanger 1990; Schoen 1988; Schoen and Land 1979). These approaches have proved appropriate and useful for social processes that occur at a relatively slow pace.

However, because currently existing multistate estimating procedures are not designed to analyze events with frequent turnovers, it is highly inappropriate to apply them to more intense and rapidly unfolding social processes. Some traditional procedures allow only one transition within a 1-year period (Rogers 1975), which would yield very misleading results when applied to multiple and recurrent events. For example, a patient admitted to a hospital stays there for only a few days on average (U.S. Department of Health and Human Services [DHHS] 1987). Other sociological processes that occur rapidly over time include adolescent dating behavior, the employment experiences of marginal workers, mental health, and the like.

Several refined methods, such as Schoen's (1988) approach, relax the single-transition assumption, thereby taking into account the returning of those who have left a given state at an earlier stage. These methods, however, may still lead to substantial biases for the events with rapid processes because those who have returned to the state of origin may leave there again soon. On the other hand, scarcity of data on frequent transitions makes it extremely difficult to use shorter intervals while retaining the standard ways of estimating interval-specific transition rates. Indeed, the traditional accounting procedures are not capable of handling the issue of frequent turnovers, resulting in a characterization of life-cycle experiences at variance with the true set of experiences generated by the stochastic processes. It is thus important to adapt the standard estimation procedures to accurately characterize the social phenomena that occur intensely and rapidly over time.

The pattern of health care use typifies such processes given the frequent turnovers of hospitalization and institutionalization over time. Modeling health care is of pressing need for both planners and researchers in view of the significant impact of population aging on the demand for health services (Brock and Brody 1985; Fisher 1980). Gaining an understanding of health care perspectives also is important for sociologists in the other areas because the health care system reflects the political and economic organization of the U.S. society and is concerned with fundamental philosophical issues involving life, death, and the quality of life (Twaddle and Hessler 1987). Although single- and multiple-decrement life tables appear fairly popular in analyzing the patterns of hospital and nursing home stays (Kemper and Murtaugh 1991; Liang and Tu 1986; McConnel 1984), the practicality of these models is dubious both because of the lack of a health care system perspective (Densen 1991) and due to the misspecification of the nature of health care use (e.g., a nursing home stay is viewed as an "absorbing" rather than a "transient" state). A multistate life table, which has the analytic power of estimating both risk and duration within an integrated system and of permitting for two-way flows among multiple states, addresses both problems.

There have been few systematic attempts to gauge these issues in the fields of medical sociology and applied demography. A major obstacle may be the difficulty in finding a fitting function for estimating certain indicators such as transition probabilities and the expected durations in relevant states. This issue, however, may be solved by obtaining supplementary information such as the average length of stay in a given state for relevant episodes. Such data often are unobtainable for general demographic phenomena (e.g., migration, marital status) but usually are available in the domain of health care.

The traditional multistate accounting procedures also have been criticized for their lack of capability to capture population heterogeneity (Heckman and Singer 1982; Keyfitz 1985). Neglect of such heterogeneity would simplify models and reduce the possible discrepancies that might arise. Therefore, further generalization from a simple life table may be substantially limited. For example, using a traditional life table to predict the future pattern of health care use can be very problematic because it does not reflect the structural changes in important determinants other than age (Boult, Kane, Louis, and Ibrahim 1991; Keyfitz 1985). Because the heterogeneity always is intrinsic and would not be reduced by involving more data, it is essential to incorporate more dimensions in a multistate life table model. Although a complete solution of this issue is not yet available, employing a multivariate approach in constructing a life table model can partially overcome this limitation. A recent development is to use the hazard rate or logistic models to address this issue (Gill 1992; Guilkey and Rindfuss 1987; Guralnik, Land, Blazer, Fillenbaum, and Branch 1993; Hayward and Grady 1990; Land, Guralnik, and Blazer 1994). These approaches are appropriate for controlling some population heterogeneity in generalizing the stochastic processes of a dynamic event and have the added advantage of deriving life expectancy from data with small sample size (Land et al. 1994).

The aforementioned limitations point to the need to develop a unique system to model the multidimensional transitions in social events with rapid processes. This research introduces a new algorithm to estimate multistate life table indicators specifically on health care use. The construction of this system takes advantage of the availability of information on the average lengths of stay in both hospitals and nursing homes, which often is obtainable from hospital censuses or sample surveys. To partially overcome the issue of population heterogeneity, we employ multivariate survival analysis to estimate age-specific transition probabilities. Hence effects of some of the important determinants other than age can be captured.

In the following two sections, we briefly review the theory of multistate increment-decrement life tables and the traditional accounting procedures used to estimate life table indicators. We follow Schoen's mathematical notation given its diversity and popularity.

2. *OVERVIEW OF THEORY*

The purpose of this section is to review existing model specifications, basic concepts, and function definitions with respect to the theory of multistate life table models. Much of the presentation is based on Land and Schoen (1982), Schoen (1988), and Schoen and Land (1979).

We specify a time-inhomogeneous and continuous-time Markov process model with finite state space Ω. We then assume that the state space Ω of the process has $k + 1$ states, where k is a positive integer greater than 1. The $(k + 1)$ state is defined as an absorbing state (say, death) with all the others defined as transient states. Two-way transitions are allowed among the transient states to permit increments and decrements between given states of the model.

On the state space Ω, we define a stochastic process, $[\xi(x): x \geq 0]$, where x denotes the exact age. The transition probabilities between the $k + 1$ states of ξ, assumed absolutely continuous, are

$$\Pi_{ij}(x, n) = \text{prob}[\xi(x + n) = j | \xi(x) = i], \qquad (2.1)$$

where $\pi_{ij}(x, n)$ denotes the probability that a person in state i at exact age x will be in state j at exact age $x + n$. The corresponding transition forces, also referred to as the gross flow hazard rates, are assumed to be

$$\mu_{ij}(x) = \lim_{n \to 0} \Pi_{ij}(x, n)/n, \qquad \text{for } i \neq j \qquad (2.2)$$

where $\mu_{ij}(x)$ represents the force of decrement for transfers from state i to state j at exact age x. It is nonnegative but not necessarily smaller than 1. We also assume

$$\mu_{ii}(x) = -\lim_{n \to 0} [1 - \Pi_{ii}(x, n)/n = -\sum_{j \neq i}^{k+1} \mu_{ij}(x). \qquad (2.3)$$

where μ_{ii} always is nonpositive and often is referred to as the "force of retention" (Schoen 1988). We may arrange the transition probabilities and the forces of transition into two $(k + 1)$ by $(k + 1)$ stochastic matrices, defined as $\mathbf{\Pi}(x, n)$ and $\boldsymbol{\mu}(x + n)$, respectively, which we shall call the "matrix of transition probabilities" and "matrix of transition forces," respectively. By definition, each row in the $\mathbf{\Pi}$ matrix sums to 1, and each row in the $\boldsymbol{\mu}$ matrix sums to 0.

We then introduce the initial distribution

$$l_i(0) = \text{prob}[\xi(0) = i] \qquad \text{for } i \in \Omega \qquad (2.4)$$

and define

$$l_i(x) = \text{prob}[\xi(x) = i] = \sum_{k \in \Omega} l_k(0)\Pi_{ki}(0,x). \qquad (2.5)$$

The $l_i(0)$ is the radix of the corresponding multistate life table, and the sequence of $l_i(x)$ is the stationary population corresponding to the

Markov chain. Demographers often tabulate some multiple of $l_i(x)$, such as $10^6 * l_i(x)$, for convenience (Hoem and Jensen 1982). We may define $\mathbf{l}(x)$, named the "matrix of survivors," as a $(k + 1)$ by $(k + 1)$ diagonal matrix having elements $l_i(x)$.

The gross flows of the stationary population are specified as the function

$$l_{ij}(x, n) = l_i(x)\Pi_{ij}(x, n), \tag{2.6}$$

where $l_{ij}(x, n)$ represents the number of persons in state i at exact age x who are also in state j at exact age $(x + n)$.

Because the transition probabilities satisfy the Kolmogorov forward differential equations (Schoen 1988), we may derive the following equation:

$$l_{ij}(x, n) = l_i(x) - \sum_{y \neq j}^{k+1} d_{ijy}(x, n) + \sum_{y \neq j}^{k+1} d_{iyj}(x, n), \tag{2.7}$$

where $d_{ijy}(x, n)$ represents the number of transfers from state j to state y between exact ages x and $(x + n)$ by persons in state i at exact age x, and d_{iyj} denotes those in state i at exact age x who move from state y to state j between exact ages x and $(x + n)$. We define a $(k + 1)$ by $(k + 1)$ matrix $\tilde{l}(x, n)$ containing elements $l_{ij}(x, n)$, called the "matrix of gross flows."

The actual sojourn time in state j between exact ages x and $(x + n)$ spent by persons in state i at exact age x is defined as

$$L_{ij}(x, n) = \int_0^n l_{ij}(x, t)dt, \tag{2.8}$$

where $L_{ij}(x, n)$ denotes the sojourn time, often referred to as person-years lived. We define a $(k + 1)$ by $(k + 1)$ matrix $\tilde{L}(x, n)$ in a manner analogous to $\tilde{l}(x, n)$, whose elements are the L_{ij}. Whereas the $\tilde{L}(x, n)$ matrix reflects person-years lived at the level of the gross flows, these person-years can be aggregated to the level of the net flows by

$$L_i(x, n) = \sum_{y=1}^{k} L_{yi}(x, n) = \int_0^n l_i(x + t)dt, \tag{2.9}$$

where $L_i(x, n)$ denotes the person-years lived in state i between exact ages x and $(x + n)$ without the constraint of being in state i at age x. The elements of $L_i(x, n)$ sequence constitute a diagonal matrix $\mathbf{L}(x, n)$.

These equations constitute the basic theory of multistate transitions, which has guided the constructions of various algorithms to estimate life table indicators.

3. CONVENTIONAL ESTIMATION PROCEDURE

Whereas a continuous process of redistribution of survivors accurately reflects the multidimensional flows within a well-defined state space, researchers have developed a variety of estimating algorithms to formalize the functions specified in the preceding section (Land and Schoen 1982; Namboodiri and Suchindran 1987; Rogers 1975; Schoen 1988; Schoen and Land 1979). Because these accounting procedures are based on varying assumptions on the patterns of transitions within a limited interval, there are explicit differences in the results derived from these approaches as well as a general distinction between the underlying stochastic processes of a given event and the accounting procedures (Hoem and Jensen 1982). The following is a brief review of a linear estimating procedure that has been used frequently in constructing the multistate life tables (Schoen 1988).

Demographers often start the procedure of estimating multistate life table indicators by obtaining a set of observed occurrence/exposure rates with respect to transitions from state i to state j between ages x and $(x + n)$, excluding the case where $j = i$, such that

$$M_{ij}(x, n) = D_{ij}(x, n)/P_i(x, n), \qquad (3.1)$$

where $D_{ij}(x, n)$ denotes the observed number of transfers from state i to state j of the state space Ω occurring among members of the population ages x to $(x + n)$ during the period of observation, and $P_i(x, n)$ represents the observed population in state i between ages x and $(x + n)$. Given the Markov assumption, such an observed rate applies to all persons ages x and $(x + n)$ in state i, regardless of the state they were in at exact age x. We define a $(k + 1)$ by $(k + 1)$ matrix of observed transition rates $\mathbf{M}(x, n)$, with the elements in diagonal as the summations over j, run from 1 to $(k + 1)$ excluding the case in which $j = i$, and the off-diagonal elements as $M_{ij}(x, n)$ multiplied by -1. Given the existence of an absorbing state, the $(k + 1)$ row contains all zeros.

We then define the matrix of model transition rates, $\mathbf{m}(x, n)$, in the same fashion, given by

$$\mathbf{m}(x, n) = \mathbf{M}(x, n), \tag{3.2}$$

where $m_{ij}(x, n) = d_{ij}(x, n)/L_i(x, n)$ and d_{ij} represents decrements from i to j.

We further specify the following equation:

$$\tilde{l}(x, n) = \mathbf{l}(x) - \tilde{L}(x, n)\mathbf{M}(x, n). \tag{3.3}$$

With $\tilde{l}(x, n)$ known, the matrix of transition probabilities can be found from

$$\Pi(x, n) = \mathbf{l}^{-1}(x)\tilde{l}(x, n), \tag{3.4}$$

where the superscript -1 denotes the inverse of a matrix.

Assuming linearity in the l_{ij} functions, we may write

$$\mathbf{L}(x, n) = \frac{n[\mathbf{l}(x) + \mathbf{l}(x + n)]}{2}, \tag{3.5}$$

where n represents the width of age interval.

Schoen (1975) introduces a linear scaler method to operationalize the preceding accounting procedure, allowing for return flows within a specific time period as well as the general intercommunication between states. However, multiple destination-specific transitions within a given period are restricted given the relative infrequency of demographic events.

In addition to the linear method, population researchers have specified other approaches such as mean duration, exponential, and cubic methods. When a complete life table is constructed, these methods produce very similar results (Land and Rogers 1982; Schoen 1988), and so the disparities among these methods are not discussed.

4. THREE MAJOR CRITICAL ISSUES

The following is an elaboration of the major issues regarding the application of conventional methods to multidimensional transitions in events with rapid processes. These issues are discussed on the basis of a 1-year interval, consistent with the pattern imposed by the traditional accounting procedures.

TRANSITION PROBABILITIES AND NUMBER OF SURVIVORS

As already mentioned, the traditional methods may take into account the return flows between two states within a single-year period. For example, in deriving the estimates of $\mathbf{\Pi}$ and $\mathbf{l}$ by the linear method, a given transition rate is simply multiplied by a factor (smaller than 1) to reduce the force of transition (Schoen 1975). In the domain of health care and other social events with rapid processes, however, there often are a number of transition cycles occurring during a limited interval (U.S. DHHS 1987); therefore, those who have returned to a given state would have a high risk of leaving again within the interval. Because the transition probabilities originating from a specific state, including returning to that state, always sum to unity, the misspecification of one destination-specific probability would automatically disturb all the others, in turn confounding the estimation of the number of survivors in each transient state at the beginning of the subsequent age interval.

PERSON-YEARS LIVED IN TRANSIENT STATES

Given the assumpion of linearity on the gross flows, the person-years lived in a given state by those who have moved in or out within a single age often are assumed to be around 6 months. Although there are several alternative ways in which to specify the function of transitions over time (Schoen 1988), most yield close estimates. This specification would lead to substantial biases with respect to health care use, considering its frequent flows. It is not possible that patients in a short-term hospital would stay there for 6 months on average within a single age interval. Because the total person-years lived by all survivors at the outset of a given age are restrained within a life table framework, biased estimates of the person-years lived in one state would automatically result in the misspecification of those in the others, thereby confounding further the estimation of other life table summary measures such as life expectancy. Even in terms of a general demographic event, the traditional linear system may still overestimate the person-years lived in a transient state because it does not rule out the person-years lived in other states by those who have returned to the origin, as shown in Equation 3.5.

Schoen (1988) specifies a mean duration method to estimate the actual duration in a given state. For a 1-year interval, this approach can be expressed as

$$L_i(x, 1) = l_i(x + 1) + \sum_{y \neq i}^{k+1} a_{iy}(x,1)d_{iy}(x,1) - \sum_{y \neq i}^{k+1} a_{yi}(x,1)d_{yi}(x,1), \qquad (4.1)$$

where $a_{ij}(x, 1)$ is the mean duration into age interval $(x, 1)$ for transfers from state i to state j, and $d_{ij}(x, 1)$ denotes the number of transfers from i to j between exact ages x and $(x + n)$. This state survival function does include the occurrences of multiple transitions within a limited interval. However, several problems arise if we make use of this equation to generalize the pattern of health care use. First, the number of transfers between two states of health care cannot be derived from the traditional approaches, as noted subsequently. Second, multiple transitions always occur within a truncated time period; therefore, the mean duration in a given state within a time bound is a function of the order of transfers. The traditional methods may estimate the mean duration for the first transfer but not the subsequent spacings within a 1-year interval. Third, the length of stay in a hospital or nursing home has been observed to be positively linked to age. The traditional mean duration method does not have the capability to integrate such a relationship in the multistate model. These limitations will greatly hamper the application of this approach in the estimation process.

NUMBER OF TRANSFERS

Given the bias in the estimates of number of survivors and transition probabilities, the total number of transitions from one state to another also would be erroneous because it is partly derived from them. Even if the estimates of the number of survivors and transition probabilities are valid, the number of transfers still would be misspecified because the conventional accounting procedures do not reflect multiple and recurring moves.

Theoretically, the preceding three concerns can be resolved by using shorter intervals. If the interval width is sufficiently short (say, weekly) to suit the circumstances of a given process, then it may be reasonable to assume that the rate of transition from state *i* to state *j* is the same across all subintervals within a single age interval. Such a strategy, however, is not realistic. First, the estimating procedure would become extremely tedious. Second, it would be very difficult, if not impossible, to obtain data of weekly based populations, which should serve as the proxy of exposure for calculating transition rates. Third, the use of data from several sample surveys, which often is the case for multistate analyses, would be very limited given an insufficient sample size for each subinterval.

Indeed, with the guidance of the multistate life table theory, it is critical to develop a new algorithm reflecting the unique stochastic processes of social events with rapid turnovers.

5. MODELING A MULTISTATE LIFE TABLE FOR HEALTH CARE

In this section, we specify an age-inhomogeneous, finite-space, and continuous-time process model to formalize the functions specified in the Overview of Theory section with respect to health care use. It introduces a set of accounting rules that take into consideration the multiple sequences of transfers per single age.

To make the model operational to the data usually available, we specify four mutually exclusive states—community, short-term hospital, nursing home, and death—which are denoted as c, h, n, and d, respectively, with c as the initial state and d as the only absorbing state, as displayed in Figure 1. The long-stay hospital sector is contained in the state of community due to the data scarcity and an effort to simplify the state space. We also assume that all hospitalized patients in short-term hospitals would be discharged in the same year as they are

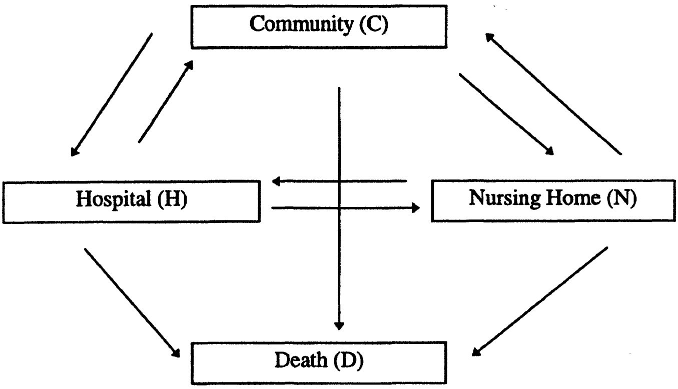

Figure 1: Four-State Increment-Decrement Life Table Model

admitted given the rapid turnovers in this regard. (Data of the National Hospital Discharge Survey have shown that about 95% of admitted patients would be discharged within 15 days and that the median length of stay for admitted patients was only 5 to 7 days.)

We propose to start the estimation procedure by deriving a series of age-specific transfer probabilities (Land and Schoen 1982) among the four designated states. Given a Markov process, a transfer probability, which is defined subsequently, is assumed to be constant throughout an age interval. We attempt to use such probabilities to reflect the occurrences of multiple and recurring events in health care. Whereas using the hazard rate models in this regard is well rationalized (Gill 1992), we employ multivariate survival analysis, an effective statistical approach, to estimate these probabilities so as to address the issue of population heterogeneity. Therefore, constructing the life table model within this context is divided into two steps: (1) estimating transfer probabilities and (2) deriving life table indicators.

ESTIMATING TRANSFER PROBABILITIES

Initially, we aim at obtaining, first, the probabilities of making at least one destination-specific movement within a 1-year period (rather than those that a person in state i at exact age x would be in state j at exact age $[x + n]$). This is because the conventional indicator cannot capture frequent transitions between two exact ages. We denote such a probability, termed "transfer probability" within this context, as p_{ij} to differentiate from the conventional transition probability π_{ij}. There are basically nine origin-and-destination transfer probabilities given $i \neq j$. The value of p_{ii}, the probability of staying in state i throughout the whole period, may be estimated residually given that origin-specific transition probabilities, including that to the origin state itself, always sum to 1. Note that Schoen and Land (1979) used p_{ii} to define another life table concept, whereas Schoen (1988) defined p_i as the ratios of survivorship functions.

Provided that relevant data are available, we construct three sets of hazard rate models, dealing respectively with the transfers from three transient states: community, hospital, and nursing home. Because there are multiple destinations with respect to each origin state, we employ hazard rate models with competing risks for each set. Although our research concentrates on constructing a complete multistate life table, we intend to derive various hazard rates with respect to a specific calendar year, and so those who have yet to move out by the end of the year would be treated as censored cases. In generating an origin-and destination-specific transfer rate, all transitions to other destinations also are considered right censored. Because it presumably involves all person-years lived in a given state, a hazard model deals with the number of destination-specific transfers versus all person-years lived in the origin state before making the first transfer; therefore, existence of multiple counts in an individual's record would not affect the estimation of the hazard rates under study. Note that the unit of analysis for generating a transfer probability is *each interval between events for each person* rather than an individual, and so it represents the likelihood of having a single run of the chain.

Age, as well as other causally related individual and societal factors, is assumed to influence the hazard rates according to a conceptual framework. The transition forces $\mu_{ij}(t)$, where j denotes the state of destination (j = c, h, n, d) and i represents the origin state for a specific transition type (i = c, h, n), is defined by Equations 5.1 and 5.2.

$$\mu_{id}(t) = \mu_{0,id}(t)\exp[a_{id} + b_{id}\log(A) + X'C_{id}], \qquad (5.1)$$

where d represents death, A is age, and $\mu_{0,id}$ denotes a specified (such as Weibull or exponential) or an unspecified baseline hazard function for continuous time T, the duration from the onset of a given observation period to the occurrence of a specific transfer. X is a vector of explanatory variables other than age, and C is a vector of coefficients. The increase of mortality with age commonly is modeled as an exponential function of age (see, e.g., Rogers et al. 1990). The specification of time function in fitting such a hazard rate model relies on the primary assumption regarding the shape of a given type of transfers over time. And the function for discharges to the states other than death is specified as

$$\mu_{ig}(t) = \mu_{0,ig}(t)\exp(a_{ig} + b_{ig}A + X'C_{ig}), \tag{5.2}$$

where g = c, h, n. In constructing a complete multistate life table, these two equations typically are used to operate within a 1-year interval assuming a specific time function.

The selection of explanatory variables should rely on an underlying conceptual framework or specific research interest; they may incorporate both categorical and continuous variables, interaction terms, and time-varying covariates. Theoretically, these equations also can be used to model a non-Markovian process. The estimation of the parameters in these two equations may be based on either a maximum likelihood approach for a fully parametric model or a partial likelihood method for a semiparametric model (Lawless 1982). Inserting relevant values into the preceding formulas, we may obtain the estimates of hazard rates for a given population subgroup.

The use of age as a continuous variable needs to be discussed. A piecewise application in this regard, which has been widely used in sociology, often is impractical in the life table analysis because of the restriction of sample size. Even if the sample size is large enough for each single-year age group, exclusion of the age factor from the multivariate framework is inconsistent with our model specification. Additionally, how to statistically evaluate the difference in a life table indicator between two age groups would be a critical issue. On the other hand, by treating age as a continuous variable with the specification of a given time function, we may readily assess the quality of relevant parameter estimates by reading standard errors.

Generally, the survival function with respect to a specific transfer is intimately related to the hazard rates given by

$$S_{ij}(t) = \exp[-\int_0^t \mu_{ij}(u)du], \tag{5.3}$$

where $S_{ij}(t)$ represents the proportion surviving from a transfer from state i to state j at time t. Note that this survival function refers not to the gross flows but rather to a single cycle of transitions. A specific survival function for an observation interval can be derived readily from a set of hazard rate estimates given a specific time function for specific transfers. Usually, very small differences between survival functions can be associated with large discrepancies in hazard functions (Teachman 1983).

Because statistically the joint density for several independent states is the product of the marginal densities (Greene 1993), the destination-independent survival function, $S_{i.}(t)$, is given by

$$S_{i.}(t) = \prod_{\substack{j=1 \\ j \neq i}}^{k+1} S_{ij}(t). \tag{5.4}$$

As indicated earlier, in the present research we are interested in deriving a set of discrete survival functions with respect to a 1-year period.

Given the intimate relationship between the survival and probability functions (Teachman 1983), the age-specific probability for a specific type of transfers can be calculated as follows:

$$\mathrm{p}_{ij}(t, t+n) = \frac{S_{ij}(t) - S_{ij}(t+n)}{S_{ij}(t)}, \tag{5.5}$$

where $\mathrm{p}_{ij}(t, t + n)$ is the transfer probability from state i to state j between time t and time $(t + n)$. When $t = 0$, this equation becomes

$$\mathrm{p}_{ij}(0, n) = 1 - S_{ij}(n). \tag{5.6}$$

For a state survival function, the timing for a specific transfer can be viewed as that of the initial transfer from the origin to the destination states (Allison 1984), and so we also consider p_{ij} the initial transfer probability from state i to state j between exact ages x and $(x + 1)$. Similarly, given a Markov process, p_{ij} also can be viewed as the one-step transition probability defined by statisticians, constant throughout a given age interval (Cox and Miller 1978; Karlin and Taylor 1975). When a specific time function is defined, we may readily identify the exact timing for the first transfer from state i to state j and the remaining time in the interval for the occurrence of another transfer.

For the transfers from hospital to other states, the annual age-specific transfer probabilities also can be viewed as those for each transfer cycle, under the assumptions that all hospital patients would be discharged within the same year as they are admitted and that such transfers are linearly distributed. Theoretically, when a multiple destination-specific transfer occurs and another cycle can be completed within the truncated time interval, $p^*_{ij}(x) = p_{ij}(x)$ where $p^*_{ij}(x)$ represents the probability of multiple occurrences within a shortened time period. Specifically, we let all $p^*_{hj}(x)$ be $p_{hj}(x)$.

When the remaining time within a given age interval is not long enough for the completion of a cycle, the value of a given probability of multiple occurrences would be smaller than the corresponding p_{ij}. Assuming a constant probability density function within a specific age, the probability of multiple occurrences would simply change proportionally with the truncation of the age interval such that

$$p^*_{ij}(x) = p_{ij}(x)\,\frac{V_i(x)}{Q_i(x)}, \tag{5.7}$$

where V_i is the days remaining in the age interval, exposed to experiencing a multiple transition from state i to state j, and Q_i refers to the number of days required for the completion of an annual cycle. (It is operationally defined as 365 days for transfers from community and nursing home because the average lengths of stay in these two states usually are beyond a year.) On the other hand, if the function of constant forces of transfers (i.e., exponential function) are posited, then such a conditional probability would change exponentially with the linearly truncated central rate such that

$$p^*_{ij}(x) = p_{ij}(x) | V_i < Q_i = M^*(x, 1)\exp[0.5M^*(x, 1)], \tag{5.8}$$

where M^* denotes the force of multiple occurrences within a truncated period, defined as

$$M^*_{ij} = \begin{Bmatrix} M_{ij} & \text{if } Q_i \leq V_i \\ \\ \dfrac{M_{ij}V_i}{Q_i} & \text{if } Q_i > V_i \end{Bmatrix}. \tag{5.9}$$

Because the occurrence of a multiple destination-specific transition involves more than two consecutive moves within a single age, fitting a linear change in p_{ij} with the truncation of the central rate would not pose much threat to the accuracy of estimating a discrete life table model in this context. For ease of computation, we may use the following approximate equation given constant forces:

$$p^*_{ij}(x) \approx p_{ij}(x)\frac{V_i(x)}{Q_i(x)}. \tag{5.10}$$

This approximate equation also can be used for other time functions when the accurate derivation of the probabilities of multiple occurrences is unduly complicated.

DERIVING LIFE TABLE INDICATORS

In this subsection, we consider parallel independent runs of the Markov chain and the related spacings between realizations within a 1-year interval (Smith and Roberts 1993). With the availability of various transfer probabilities, we may estimate the values of l_i and d_{ij}, the number of transfers from state i to state j. To generate these indicators, we make three additional assumptions. First, we allow three return flows to hospital within a single age. Second, between the community and nursing home, one return flow is allowed within a single age because the transitions initiating from these two states are relatively infrequent. Third, only one transition is allowed before death within each age interval. We start with $l_c(0) = 100{,}000$ (community is the initial state in this context). The estimation of the number of survivors at the onset of each age and the number of transfers from a given state to another within a 1-year interval are based on the principle of Chapman-Kolmogorov relation (Cox and Miller 1978), given by

$$p_{ij}^{(n)} = \sum_{k=1}^{K+1} p_{ik}^{(n-1)} p_{ki}^{(I)} \qquad (n = 1, 2, \ldots), \tag{5.11}$$

where $p_{ij}^{(n)}$ is defined as the probability of being at state j at time n for those who are at state i at time 0, termed "n-step transition probability" (Cox and Miller 1978). Similarly, $p_{ik}^{(n-1)}$ is an $(n-1)$-step transition probability and $p_{kj}^{(I)}$ is a one-step transition probability. This Markov process equation indicates that, with the presence of repeated transitions within a limited time interval, the n-step transition probability is virtually the outcome of a series of one-step transition probabilities within period n. The transfer probability p_{ij} reflects only a single move from states i to j and may not be relevant to the conditional probability distribution on k states (Karlin and Taylor 1975).

Given the aforementioned three assumptions, Equation (5.11) will give rise to the pattern of the multidimensional random walks in health care and the following equations to estimate $l_i(x)$ sequence, assuming a constant probability density function.

$$\begin{aligned}
l_c(x+1) = {} & l_c(x)\{1 - [p_{ch}(x) - p_{ch}(x)p_{hc}(x) \\
& - p_{ch}(x)p_{hn}(x)Z_1(x)p_{nc}(x) - p_{ch}(x)p_{hn}(x)Z_1(x)p_{nh}(x)p_{hc}(x)] \\
& - p_{cd}(x) - [p_{cn}(x) - p_{cn}(x)0.5p_{nc}(x) - p_{cn}(x)0.5p_{nh}(x)p_{hc}(x)]\} \\
& + l_n(x)\{p_{nc}(x)[1 - 0.5p_{cn}(x) - 0.5p_{ch}(x)p_{hn}(x) - 0.5p_{cd}(x)] \\
& + [p_{nh}(x)p_{hc}(x)] - [p_{nc}(x)0.5p_{ch}(x)]/[Y_h(x)/2]\} + l_h(x)\{[p_{hc}(x) \\
& + p_{hn}(x)Z_2(x)p_{nc}(x) - p_{hc}(x)Z_2(x)p_{cd}(x)] \\
& - [p_{hc}(x)Z_2(x)p_{ch}(x)]/[Y_h(x) - 0.5]\} - [l_c(x)p_{ch}(x)]/Y_h(x)
\end{aligned} \tag{5.12}$$

$$\begin{aligned}
l_n(x+1) = {} & l_n(x)\{1 - [p_{nh}(x) - p_{nh}(x)p_{hn}(x) \\
& - p_{nh}(x)p_{hc}(x)Z_1(x)p_{cn}(x) - p_{nh}(x)p_{hc}(x)Z_1(x)p_{cn}(x)p_{hn}(x)] \\
& - p_{nd}(x) - [p_{nc}(x) - p_{nc}(x)0.5p_{cn}(x) - p_{nc}(x)0.5p_{ch}(x)p_{hn}(x)]\} \\
& + l_c(x)\{p_{cn}(x)[1 - 0.5p_{nc}(x) - 0.5p_{nh}(x)p_{hc}(x) - 0.5p_{nd}(x)] \\
& + [p_{ch}(x)p_{hn}(x)] - [p_{cn}(x)0.5p_{nh}(x)]/[Y_h(x)/2]\} + l_h(x)\{[p_{hn}(x) \\
& + p_{hc}(x)Z_2(x)p_{cn}(x) - p_{hn}(x)Z_2(x)p_{nd}(x)] - [p_{hn}(x)Z_2(x)p_{nh}(x)]/[Y_h(x) \\
& - 0.5]\} - [l_n(x)p_{nh}(x)]/Y_h(x)
\end{aligned} \tag{5.13}$$

$$
\begin{aligned}
l_h(x+1) = {} & [l_c(x)p_{ch}(x)]/Y_h(x) + [l_n(x)p_{nh}(x)]/Y_h(x) \\
& + [l_h(x)p_{hc}(x)Z_2(x)p_{ch}(x)]/[Y_h(x) - 0.5] \\
& + [l_h(x)p_{hn}(x)Z_2(x)p_{nh}(x)]/[Y_h(x) - 0.5] \\
& + [l_c(x)p_{cn}(x)0.5p_{nh}(x)]/[Y_h(x)/2] \\
& + [l_n(x)p_{nc}(x)0.5p_{ch}(x)]/[Y_h(x)/2]
\end{aligned}
\tag{5.14}
$$

In these equations, Y_h denotes the number of cycles of hospital admission in a year, which can be estimated by 365 divided by the average length of stay in a hospital, Q_h. In addition,

$$Z_1(x) = [0.5Y_h(x) - 1]/Y_h(x)$$

and

$$Z_2(x) = [Y_h(x) - 0.5]/Y_h(x).$$

It is clear that a multiple transition probability from one state to another within a truncated time period is viewed as proportional to the initial transfer probability, with Zs and other numerals as the multiplicative factors to derive these recurring probabilities. Because a sequence of possible transitions is included within each brace in Equations 5.12 and 5.13, a set of n-step transition probabilities also can be calculated in the form of these equations. For example, the term within the first brace in Equation 5.12 yields the estimate of the conventional n-step transition probability from C to C (π_{cc}), that is, the probability of being in the community at exact age (x + 1) for those who are in the community at exact age x. Similarly, the terms within the second and third braces of Equation 5.12 calculate the n-step transition probabilities from the nursing home and the hospital, respectively, to the community between exact ages x and (x + 1) (π_{cn} and π_{ch}). The final term estimates the number of survivors who are in the community at age x and in the hospital at age (x + 1), serving as a correction factor.

In these equations, only the multiple transitions are relevant to the linear function given that p_{ij}'s are derived directly from multivariate survival analysis. Hence these equations also can be used for other time functions given Equation 5.10. Our empirical data analysis shows that the estimates of a given $l_i(x)$ derived from two different time functions are identical within the framework of a complete life table. There are some other combinations of transition probabilities, such as $p_{ch}p_{hn}p_{nh}p_{hc}$, which are not included in the first two equations. This is because the chance of making so many transitions within a 1-year period would be negligible and the exclusion of them would not disturb the construction of a multistate life table.

Next, we estimate the number of transfers from one state to another so as to calculate the person-years lived in each transient state beyond a given age, using the empirical data of the average length of stay in that state. Because the average length of stay in community is unknown and sometimes is meaningless, we calculate the number of transfers to the other two states first according to the Chapman-Kolmogorov principle and the assumptions stated earlier. Specifically, the distribution of survivors at initial time and the pattern of the "multidimensional random walks" (Karlin and Taylor 1975), as seen through a set of one-step transition probabilities, would lead to the following formulas.

$$\begin{aligned} d_{c.h}(x, 1) &= l_c(x)p_{ch}(x)\{1 + [(Y_h(x) - 1)/2Y_h(x)]p_{hc}(x)p_{ch}(x) \\ &+ [(Y_h(x) - 1)(Y_h(x) - 10)/8Y_h(x)^2]p_{hc}(x)^2p_{ch}(x)^2 \ldots\} \\ &+ l_c(x)[p_{cn}(x)0.5p_{nh}(x)+p_{ch}(x)p_{hn}(x)Z_1(x)p_{nh}(x)] \ldots \end{aligned} \tag{5.15}$$

$$\begin{aligned} d_{n.h}(x, 1) &= l_n(x)p_{nh}(x)[1 + p_{hn}(x)Z_1(x)p_{nh}(x) \\ &+ Z_3(x)Z_1(x)p_{hn}(x)^2p_{nh}(x)^2 \ldots] \\ &+ l_n(x)[0.5p_{nc}(x)p_{ch}(x) + Z_1(x)p_{nh}(x)p_{hc}(x)p_{ch}(x)] \ldots \end{aligned} \tag{5.16}$$

$$\begin{aligned} d_{h.h}(x, 1) &= l_h(x)\{Z_2(x)p_{hc}(x)p_{ch}(x)[1 + Z_4(x)p_{hc}(x)p_{ch}(x) \\ &+ Z_4(x)p_{hn}(x)p_{nh}(x) \ldots]\} + l_h(x)Z_2(x)p_{hn}(x)p_{nh}(x) \\ &\{1 + Z_4(x)p_{hn}(x)p_{nh}(x) + Z_4(x)p_{hc}(x)p_{ch}(x) \ldots\} \end{aligned} \tag{5.17}$$

$$\begin{aligned} d_{c.n}(x, 1) &= l_c(x)p_{cn}(x)[1 + 0.5p_{nh}(x)p_{hn}(x) \ldots] \\ &+ l_c(x)p_{ch}(x)p_{hn}(x)[1 + Z_1(x)p_{nh}(x)p_{hn}(x) \ldots] \end{aligned} \tag{5.18}$$

$$\begin{aligned} d_{h.n}(x, 1) &= l_h(x)p_{hn}(x)[1 + Z_2(x)p_{nh}(x)p_{hn}(x) \\ &+ Z_2(x)Z_5(x)p_{nh}(x)^2p_{hn}(x)^2 + Z_2(x)Z_6(x)p_{nc}p_{cn} \ldots] \\ &+ l_h(x)Z_2(x)p_{hc}(x)[p_{cn}(x) + p_{ch}(x)p_{hn}(x) \ldots] \end{aligned} \tag{5.19}$$

$$d_{\mathrm{n.n}}(x, 1) = l_{\mathrm{n}}(x)\mathrm{p}_{\mathrm{nc}}(x)0.5[\mathrm{p}_{\mathrm{cn}}(x)+\mathrm{p}_{\mathrm{ch}}(x)\mathrm{p}_{\mathrm{hn}}(x)] + l_{\mathrm{n}}(x)\mathrm{p}_{\mathrm{nh}}(x)[\mathrm{p}_{\mathrm{hn}}(x) + Z_1(x)\mathrm{p}_{\mathrm{hc}}(x)\mathrm{p}_{\mathrm{cn}}(x) \ldots\}, \quad (5.20)$$

where $d_{i.j}(x, 1)$ denotes the number of transfers to state j between exact ages x and $(x + 1)$ for those who are in state i at exact age x. The ellipsis dots represent possible but negligible terms. In addition,

$$Z_3(x) = [0.25Y_{\mathrm{h}}(x) - 1]/Y_{\mathrm{h}}(x),$$

$$Z_4(x) = [0.5Y_{\mathrm{h}}(x) - 1.25]/Y_{\mathrm{h}}(x),$$

$$Z_5(x) = [Y_{\mathrm{h}}(x) - 2.5]/2Y_{\mathrm{h}}(x), \text{ and}$$

$$Z_6(x) = [0.5Y_{\mathrm{h}}(x) - 0.25]/Y_{\mathrm{h}}(x).$$

The Zs are the multiplicative factors to calculate the number of recurring events. These equations illustrate that the estimation of the number of transfers within a 1-year period takes into consideration the multiple transfers within a 1-year interval according to the imposed conditions. Knowing these numbers, we may further calculate the person-years lived in hospitals and nursing homes by those who have made the transfers during the age interval as follows:

$$^{*}L_{\mathrm{h}}(x) = [d_{\mathrm{c.h}}(x, 1) + d_{\mathrm{n.h}}(x, 1) + d_{\mathrm{h.h}}(x, 1)] * [Q_{\mathrm{h}}(x)/365] \quad (5.21)$$

and

$$^{*}L_{\mathrm{n}}(x) = [d_{\mathrm{c.n}}(x, 1) + d_{\mathrm{h.n}}(x, 1) + d_{\mathrm{n.n}}(x, 1)] * [Q_{\mathrm{n}}(x)/365], \quad (5.22)$$

where $^{*}L_{\mathrm{h}}(x)$ and $^{*}L_{\mathrm{n}}(x)$ are the person-years lived in hospitals and nursing homes, respectively, by those who have moved to hospitals or nursing homes between exact ages x and $(x + 1)$. We put a superscript asterisk to this indicator because it differs intuitively from the conventional $L(x)$, given that the person-years lived by those who have made relevant transfers are not necessarily positioned between exact ages x and $(x + 1)$. Although this age-crossing phenomenon may influence the implication of this indicator, it would not affect the derivation of the cumulative person-years lived beyond exact age x, often referred to as $T_i(x)$.

The estimation of such cumulative person-years lived is based primarily on the $^*L_i(x)$'s. Because $^*L_i(x)$ covers only the years spent by the new entries, we must take into account the years spent by those staying in state i at the beginning of the age. Assuming that the entries into a state are evenly distributed, we have

$$T_h(x) = \sum_{i=x}^{\omega} {}^*L_h(i) + l_h(x)^*[Q_h(x)/(2^*365)] \tag{5.23}$$

and

$$T_n(x) = \sum_{i=x}^{\omega} {}^*L_n(i) + l_n(x)^*[Q_n(x)/(2^*365)], \tag{5.24}$$

where ω denotes the upper limit of a life span.

The life expectancies in the hospital and the nursing home beyond exact age x, denoted by $e_h(x)$ and $e_n(x)$, respectively, are calculated as follows:

$$e_h(x) = T_h(x)/l(x) \tag{5.25}$$

and

$$e_n(x) = T_n(x)/l(x), \tag{5.26}$$

where $l(x)$ represents the number of survivors at exact age x, usually obtained from the complete single decrement life table.

The corresponding indicators with respect to the community state are calculated as

$$T_c(x) = T(x) - T_h(x) - T_n(x) \tag{5.27}$$

and

$$e_c(x) = T_c(x)/l(x) = e(x) - e_h(x) - e_n(x), \tag{5.28}$$

where $T(x)$ represents the aggregate cumulative person-years lived beyond exact age x.

6. APPLICATION

In this section, we contrast two sets of multistate life table indicators, derived from the traditional linear method and our new estimating procedure, using data of health care use among U.S. civilians during the mid-1980s. The purpose of the comparison is to highlight the need for a unique algorithm for events with rapid processes. To make the application conform to the state space specified previously, we distinguish four states in this context: community, short-stay hospital, nursing home, and death, with death as an absorbing state and the others as transient states.

DATA AND OPERATIONALIZATION

Data for the application came primarily from several national probability sample surveys including the 1985 National Nursing Home Survey (NNHS), the 1985 National Hospital Discharge Survey (NHDS), and the 1987 National Medical Care Expenditure Survey. The data from the first two surveys are used to capture the pattern of the risk and duration of nursing home and short-term hospital stays during the mid-1980s. The 1985 NNHS covered all types of nursing homes that provided some level of nursing care in the conterminous United States. The sampling was basically a stratified two-stage probability design, with the selection of facilities as the first stage and the selection of residents, discharges, and registered nurses as the second stage. Whereas this survey included two subsamples dealing with current residents and discharges in the previous year, a merged data file was constructed involving all nursing home residents in the year prior to the survey. The merged file involved 11,129 respondents, both current residents and discharges, and incorporated a number of variables such as date of birth, date of interview, sex, marital status, whether the respondent was a current resident, destination of discharge if it occurred, and dates of admission and discharge. Such information permits dynamic analysis on the transitions from nursing home to other states.

The 1985 NHDS encompassed patients discharged from noninstitutional hospitals located in the 50 states and the District of Columbia, with the National Master Facility Inventory (NMFI) constituting the sampling frame. (NMFI is a census of all inpatient health facilities conducted periodically by mail by the National Center for Health Statistics.) It is assumed that all hospitalized patients in short-stay hospitals would be discharged in the same year as they are admitted. The data file contained about 200,000 cases and a number of variables including dates of admission and discharge, date of birth, sex, race, marital status, destination of discharge, and length of stay. Given the large sample size, a 5% random sample of respondents was selected for data analysis. Due to the execution of the multivariate survival analysis, an effective parameter estimating approach, the use of a large subsample would not pose serious threats to the quality of the life table construction. Like the NNHS data set, the NHDS provided sufficient information to execute an event history analysis for the transfer probabilities from hospital to other states.

For the transfers from the community to other states, we used the 1987 National Medical Care Expenditure Survey as the major data source. This survey covered the entire civilian noninstitutionalized population of the United States. The sampling used a stratified multi-stage area probability sample design. The data in the person file included age, sex, race, marital status, employment status, health status, and the like. Although the file consisted of information on hospitalization (whether the respondents had been hospitalized in the prior year), the status of institutionalization was unknown. However, the number of occurrences from the community to the nursing home can be derived from the NNHS using weights and information on residence before institutionalization. Similarly, the transitions from the community to death may be estimated as a residual after capturing the risk and number of deaths in hospitals and nursing homes.

First, we employed the multivariate survival analysis to estimate the transfer probabilities from short-stay hospital or nursing home to other states, using the equations specified earlier. Cox proportional hazard rate models with competing risks were used because specific time functions of transfers within a 1-year interval were not of concern in this application. In estimating the survival functions, we used Kalbfleisch and Prentice's (1980) method to derive the baseline hazard rate for operationalizing Equation 5.3. To construct a period life table, the duration in the origin state was defined as the length of stay during the previous 1-year period. Therefore, for those who were in the origin state at the beginning of an age, the starting date of the period, rather than the date of admission, served as the onset of the observation period given the assumption of a Markov process.

Because each destination state represents an independent event type, we employed a separate model for the transfers from a given origin state to each destination state, treating the other types of transfers as censored cases. Age, sex, race, and marital status were included as covariates in these hazard models. Age, a continuous variable, was defined as the age at the beginning of the observation period. Sex and race were indexed by two dummy variables, with "female" and "White" coded 1. Because there was a large number of missing values on marital status in both data sets, we created two dummies, "currently married" and "not currently married," for this variable, with "others" coded 0. Therefore, the "missing" cases were treated as a reference group. In deriving the transfer probabilities, exact ages and age-specific sample means of other covariates were inserted into Equations 5.1 and 5.2.

With respect to transfers from the community to other states, a binary logit model was executed to estimate the transfer probabilities from the community to short-stay hospitals because the exact timing for making such transfers is not available in the data. The probability of death, as indicated earlier, was derived by substracting the observed deaths in hospitals and nursing homes from the total number of deaths in 1985, which was obtained from the *1986 Demographic Yearbook* (United Nations 1988), and then calculating the observed death rates that were finally transferred into probabilities. The results of the multivariate analysis, not presented here, are available on request.

For the transfer probabilities from the community to nursing homes, first the age-specific transfer rates were calculated using data of the 1985 NNHS and 1987 National Medical Care Expenditure Survey, as indicated previously, and then transferred into probabilities. Because the latter survey, which provided estimates of mid-year population, involved only a portion of deaths, the transfer probabilities thus derived may be overestimated to some extent. Such a bias, however, would not be serious in terms of the construction of a multistate life table because most deaths, among older adults in particular, tend to occur in hospitals and nursing homes and also because the community involves a huge base population.

We used the transfer probabilities as inputs into a computer program that operationalized the newly developed algorithm, as specified by Equations 5.12 through 5.28. The age-specific average lengths of stay in short-stay hospitals and nursing homes were derived from the 1985 NHDS and NNHS. A complete multistate life table, both by single year of age and by the birth of a synthetic cohort, was then constructed, reflecting the patterns of health care use for U.S. civilians during the mid-1980s.

The application of the conventional linear method, on the other hand, was based on traditional perspectives. We first created the observed age-specific occurrence/exposure rates from the three data sets and then estimated the life table summary measures using Schoen's (1988) computer program. In calculating the occurrence/exposure rates, all events of a particular transition type within an age interval were counted, and the mid-period population was used as the proxy for exposure. Because all hospital admissions were assumed to be discharged in the same year as they were admitted, the exposure of the transitions from hospital to other states within a 1-year interval was estimated as about half of all discharges. Because the life expectancy at birth thus derived tends to be overestimated due to the inadequate assumption on the frequency of turnovers in short-stay hospitals, relevant life table indicators in the three states were adjusted according to the 1985 national life table, recognizing the pattern of distribution.

RESULTS OF TWO MULTISTATE LIFE TABLES

Although the traditional linear method is essentially a bivariate procedure, we hold that the two aforementioned schedules are highly comparable. First, we used age-specific sample means of the four explanatory variables to generate the transfer probabilities so that they would be fairly close to bivariate estimates for large samples. Second, although Keyfitz (1985) revealed the biases of ignoring population heterogeneity in estimating central rates, they were not sizable when only a few covariates were considered. Therefore, the differences in the estimates of various life table indicators, as shown subsequently, would arise mainly from the discord of the two estimating procedures.

Table 1 compares the age-specific transition or transfer probabilities derived from the two approaches. The differences are striking. Although there is no solid base to judge which method is valid simply from the differentials in the probabilities originating from the community and nursing homes, a comparison of the pattern of transitions from short-stay hospitals to other states provides instructive insights. When applying the traditional method, which does not permit multiple destination-specific transfers within a single age interval, a sizable percentage of hospitalized patients presumably would remain in short-stay hospitals 1 year later (at almost all ages, it is more than 20%). Even at the youngest several ages shown in the table, only about three fourths of these people would return to the community by the end of the 1-year period. Both phenomena obviously are unrealistic. Because the origin-specific transition probabilities always sum to unit, the estimation of the probabilities to the other two states also would be influenced. By contrast, the results from our new procedure illustrate that none of the hospitalized patients would remain in a short-stay hospital throughout a 1-year period and that the chance of returning to the community remains greater than 90% before reaching age 50 years. These contrasts on the transitions from hospital to other states are perhaps the strongest evidence that a traditional accounting procedure does not possess the capability to model the events with frequent turnovers.

It must be noted that the two sets of probabilities reflect different concepts, as indicated earlier. However, our algorithm is capable of deriving the traditional transition probabilities, also as indicated earlier. It would be helpful to compare some transition probabilities (π_{ij}) using both the proposed method and the traditional approach. In our system, for example, $\pi_{cc}(80)$, $\pi_{ch}(80)$, $\pi_{cn}(80)$, and $\pi_{hh}(80)$ are estimated as .8779, .0049, .0988, and .0056, respectively, in contrast to .7438, .1822, .0398, and .2705 as derived from the traditional method. It is evident that our program provides more reasonable estimates.

Table 2, which compares three life table summary measures by age, further shows the differences. Let us look at the $l_i(x)$ sequence first. The traditional method generates many more survivors in short-stay

TABLE 1: Transition Probabilities Among Four States at Selected Ages, Derived From Two Approaches: U.S. Civilians, 1985

Age	From Community to:				From hospital to:				From nursing home to:			
	Community	*Hospital*	*Nursing Home*	*Death*	*Community*	*Hospital*	*Nursing Home*	*Death*	*Community*	*Hospital*	*Nursing Home*	*Death*
Estimates from Schoen's approach (π_{ij})												
0	.9632	.0265	.0000	.0103	.7730	.2137	.0000	.0133	—	—	—	—
10	.9887	.0112	.0000	.0002	.8033	.1936	.0000	.0031	—	—	—	—
20	.9528	.0462	.0002	.0008	.7708	.2229	.0035	.0029	.6022	.0264	.3712	.0003
30	.9487	.0498	.0006	.0010	.7646	.2254	.0063	.0036	.3081	.0866	.5891	.0163
40	.9572	.0405	.0007	.0016	.7686	.2149	.0073	.0092	.3354	.1308	.4373	.0965
50	.9445	.0506	.0010	.0039	.7521	.2201	.0102	.0175	.3017	.1429	.4533	.1021
60	.9189	.0689	.0025	.0097	.7224	.2291	.0152	.0334	.2808	.1228	.4994	.0970
70	.8579	.1139	.0088	.0195	.6559	.2515	.0394	.0532	.2473	.1153	.4764	.1610
80	.7438	.1822	.0398	.0343	.5400	.2705	.1078	.0818	.1934	.1331	.4821	.1914
90	.5618	.2444	.1331	.0607	.3758	.2427	.2414	.1400	.1003	.1124	.5364	.2508
Estimates from new approach (p_{ij})												
0	.8210	.1715	.0000	.0075	.9848	.0000	.0000	.0152	—	—	—	—
10	.9479	.0520	.0000	.0001	.9977	.0000	.0000	.0023	—	—	—	—
20	.9337	.0659	.0000	.0004	.9685	.0000	.0227	.0088	.3266	.3709	.2948	.0076
30	.9192	.0804	.0001	.0003	.9537	.0000	.0354	.0109	.2807	.3606	.3472	.0115
40	.9044	.0951	.0002	.0003	.9259	.0000	.0578	.0163	.2508	.3474	.3828	.0190
50	.8849	.1137	.0004	.0010	.8778	.0000	.0928	.0294	.2160	.3460	.3993	.0387
60	.8616	.1343	.0009	.0032	.7960	.0000	.1529	.0511	.1904	.3312	.4070	.0714
70	.8353	.1547	.0035	.0065	.6560	.0000	.2621	.0819	.1676	.3247	.3871	.1206
80	.7868	.1813	.0239	.0080	.4292	.0000	.4440	.1268	.1393	.2920	.3835	.1852
90	.6835	.2074	.1090	.0001	.1448	.0000	.6725	.1827	.1113	.2654	.3619	.2614

hospitals at each of the selected ages than does the new model; accordingly, the proportions of survivors in the community and in nursing homes appear substantially lower given an identical number of total survivors. Such differences also are reflected in the estimates of cumulative person-years lived beyond the selected ages; each of the $T_h(x)$ values from the traditional method appears substantially higher than that from the new approach. Not surprisingly, given this pattern, a U.S. civilian would be expected to spend about 67.54 years in the community, 5.56 years in a short-stay hospital, and 1.68 years in a nursing home. Obviously, such a distribution is totally unrealistic, particularly for the life expectancy in a short-stay hospital, which in turn reduces the life expectancies in the community and nursing homes given a fixed average remaining years of life. By contrast, our model produces much more reasonable results. The life expectancies at birth in the three states are 72.35 years in the community, 0.16 year (59.5 days) in a short-stay hospital, and 2.28 years in a nursing home. This is because our estimation procedure takes into account the frequent turnovers of hospitalization and institutionalization and uses the information regarding the average length of stay in a hospital and a nursing home for each episode; hence the frequent multidimensional transitions in health care use are captured.

7. *DISCUSSION*

We have introduced a new algorithm that takes into account the multiple sequences of transfers per single age. The system is an adaptation of the standard multistate accounting procedures to characterize social events involving intense and rapid processes. We make specific theoretical assumptions about the frequencies of turnovers and use information on the average lengths of stay in relevant states to estimate the risk and duration of health care use. If this stochastic process is assumed correctly, then the estimating system described in this article provides fairly reasonable life table indicators.

We also have made an effort to employ multivariate survival analysis generating transfer probabilities as the primary inputs, as have other researchers (Gill 1992; Guilkey and Rindfuss 1987; Guralnik et al. 1993; Land et al. 1994). In addition to the statistical implications

TABLE 2: Life Table Indicators in Community, Hospital, and Nursing Home, by Selected Ages, Derived From Two Approaches: U.S. Civilians, 1985

	Community			*Hospital*			*Nursing Home*		
Age (x)	l_x	T_x	e_x	l_x	T_x	e_x	l_x	T_x	e_x
Estimates from Schoen's approach									
0	100,000	6,753,888	67.5389	0	556,340	5.5634	0	168,247	1.6825
10	97,176	5,791,575	58.7333	1,432	531,630	5.3913	0	167,583	1.6995
20	92,661	4,838,776	49.3369	5,415	501,053	5.1089	0	166,722	1.6999
30	90,524	3,931,384	4.5343	6,282	435,072	4.4858	183	164,686	1.6980
40	90,268	3,029,370	31.7178	5,114	377,259	3.9499	128	161,404	1.6899
50	86,172	2,150,984	23.2634	6,054	318,233	3.4417	236	156,776	1.6956
60	77,259	1,336,833	15.7105	7,348	247,634	2.9102	485	148,719	1.7477
70	59,845	655,126	9.3463	8,909	162,586	2.3195	1,341	132,248	1.8867
80	32,871	200,269	4.4589	8,513	72,049	1.6041	3,530	94,644	2.1072
90	7,320	25,756	1.7022	3,592	12,767	0.8437	4,218	32,303	2.1349
Estimates from new approach									
0	100,000	7,234,559	72.3456	0	16,299	0.1630	0	227,618	2.2762
10	98,583	6,247,548	63.3574	25	15,621	0.1584	0	227,618	2.3083
20	98,000	5,264,039	53.6730	76	14,894	0.1519	0	227,618	2.3208
30	96,485	4,291,827	44.2506	94	13,991	0.1443	410	225,323	2.3232
40	94,557	3,334,623	34.9137	140	12,623	0.1322	814	220,786	2.3116
50	90,700	2,403,160	25.9907	188	10,756	0.1163	1,574	212,077	2.2937
60	82,022	1,527,951	17.9565	235	8,337	0.0980	2,835	196,898	2.3139
70	64,992	773,535	11.0355	265	5,494	0.0784	4,838	170,930	2.4386
80	37,532	244,482	5.4433	243	2,575	0.0573	7,139	119,906	2.6697
90	9,616	20,490	1.3542	117	568	0.0376	5,398	49,767	3.2891

of taking some population heterogeneity into account, multivariate survival analysis may have potential for developing a non-Markovian model because the effects of previous health care events can be addressed. However, this will require a much more complicated procedure to model probabilities of multiple transfers, which will be assumed as a function of previous transitions within a non-Markovian system.

The construction of multistate life tables to generalize events with rapid and recurring processes could have extensive applications in sociology, generating important policy implications and new research orientations. For example, sociologists studying mental health can gain insight from integrating the transitions in mental health status into a life table framework. Researchers in employment and occupation may increase their understanding of risks for various transitions and durations in working states, including short-term employment. Our method can be readily adapted to modeling these phenomena provided that the average length of stay in the short-term state is known.

This approach also can be used to model transitions among more states given its flexibility in using average length of stay as the primary variable in calculating life table summary measures. For example, if data were available, we might incorporate long-term hospital stays as an independent transient state. The pattern of flows both to and from this state would be examined through various transfer probabilities and age-specific life expectancies in conjunction with the dynamic patterns of other states.

Questions may be raised about the estimation of the ${}^{*}L_i(x)$ sequence. Because some of those who have moved into a given state within an age interval may stay there beyond this age, this type of indicator does not possess the primary nature of the conventional $L_i(x)$. However, such an age-crossing phenomenon would not disturb the construction of $T_i(x)$, a cumulative indicator, because it counts all person-years lived beyond a given age. In our new system, ${}^{*}L_i(x)$ lacks intuitive meaning and is used only for estimating summary indicators such as $T_i(x)$ and $e_i(x)$. If needed, however, the conventional $L_i(x)$ sequence can be readily estimated from two adjacent $T_i(x)$ values.

It may be argued that repeated events and multiple spacings can be readily addressed with the application of the traditional mean duration method, but this is not the case. As we have pointed out, multiple transitions always occur within a truncated time interval, and the mean duration for transfers in a limited time interval is a function of their orders within the context of health care. An adaptation of the mean duration method, taking into account the order of transfers within a single age interval, may yield better estimates. However, the length of stay in certain states (e.g., nursing home), a function of age in health care use, still would be misspecified because of the restriction of a limited time interval. In addition, traditional approaches, when applied to social phenomena with rapid processes, cannot give the valid estimates of the number of transfers that are required to execute the mean duration method.

REFERENCES

Allison, Paul D. 1984. *Event History Analysis: Regression for Longitudinal Event Data*. Beverly Hills, CA: Sage.

Boult, Chad, Robert L. Kane, Thomas A. Louis, and Joseph G. Ibrahim. 1991. "Forecasting the Number of Future Disabled Elderly Using Markovian and Mathematical Models." *Journal of Clinical Epidemiology* 44:973-80.

Brock, Dwight B. and Jacob A. Brody. 1985. "Statistical and Epidemiological Characteristics." Pp. 53-71 in *Principles of Geriatric Medicine*, edited by R. Andres, E. L. Biermand, and W. R. Hazzard. New York: McGraw-Hill.

Cox, David R. and Hilton D. Miller. 1978. *The Theory of Stochastic Processes* (2nd ed.). New York: Chapman & Hall.

Densen, Paul M. 1991. *Tracing the Elderly Through the Health Care System: An Update*. Washington, DC: Agency for Health Care Policy and Research.

Fisher, Charles R. 1980. "Differences by Age Group in Health Care Spending." *Health Care Financing Review* 1:65-90.

Gill, Richard D. 1992. "Multistate Life-Tables and Regression Models." *Mathematical Population Studies* 3:259-76.

Greene, William H. 1993. *Econometric Analysis*. 2nd ed. New York: Macmillan.

Guilkey, David K. and Ronald R. Rindfuss. 1987. "Logistic Regression Multivariate Life Tables: A Communicable Approach." *Sociological Methods & Research* 16:276-300.

Guralnik, Jack M., Kenneth C. Land, Dan Blazer, Gerda G. Fillenbaum, and Laurence G. Branch. 1993. "Educational Status and Active Life Expectancy Among Older Blacks and Whites." *New England Journal of Medicine* 329(2):110-16.

Hayward, Mark D. and William R. Grady. 1990. "Work and Retirement Among a Cohort of Older Men in the United States: 1966-1983." *Demography* 27:337-56.

Heckman, James J. and Burton Singer. 1982. "Population Heterogeneity in Demographic Models." Pp. 567-99 in *Multidimensional Mathematical Demography*, edited by K. C. Land and A. Rogers. New York: Academic Press.

Hoem, Jan M. and Ulla Funck Jensen. 1982. "Multistate Life Table Methodology: A Probabilist Critique." Pp. 155-264 in *Multidimensional Mathematical Demography*, edited by K. C. Land and A. Rogers. New York: Academic Press.

Kalbfleisch, John D. and Ross L. Prentice. 1980. *The Statistical Analysis of Failure Time Data*. New York: John Wiley.

Karlin, Samuel and Howard M. Taylor. 1975. *A First Course in Stochastic Processes* (2nd ed.). New York: Academic Press.

Kemper, Peter and Christopher M. Murtaugh. 1991. "Lifetime Use of Nursing Home Care." *New England Journal of Medicine* 324(9):595-600.

Keyfitz, Nathan. 1985. *Applied Mathematical Demography* (2nd ed.). New York: Springer-Verlag.

Land, Kenneth C., Jack M. Guralnik, and Dan G. Blazer. 1994. "Estimating Increment-Decrement Life Tables With Multiple Covariates From Panel Data: The Case of Active Life Expectancy." *Demography* 31:297-319.

Land, Kenneth C. and Andrei Rogers. 1982. "Multidimensional Mathematical Demography: An Overview." Pp. 1-41 in *Multidimensional Mathematical Demography*, edited by K. C. Land and A. Rogers. New York: Academic Press.

Land, Kenneth C. and Robert Schoen. 1982. "Statistical Methods for Markov-Generated Increment-Decrement Life Tables With Polynomial Gross Flow Functions." Pp. 265-346 in *Multidimensional Mathematical Demography*, edited by K. C. Land and A. Rogers. New York: Academic Press.

Lawless, Jerald F. 1982. *Statistical Models and Methods for Lifetime Data*. New York: John Wiley.

Liang, Jersey and Edward J. C. Tu. 1986. "Estimating Lifetime Risk of Nursing Home Residency: A Further Note." *The Gerontologist* 26:560-63.

McConnel, Charles E. 1984. "A Note on the Lifetime Risk of Nursing Home Residence." *The Gerontologist* 24:193-98.

Namboodiri, Krishnan and C. M. Suchindran. 1987. *Life Table Techniques and Their Applications*. New York: Academic Press.

Rogers, Andrei. 1975. *Introduction to Multiregional Mathematical Demography*. New York: John Wiley.

Rogers, Andrei, R. G. Rogers, and A. Belanger. 1990. "Longer Life But Worse Health? Measurement and Dynamics." *The Gerontologist* 30:640-49.

Schoen, Robert. 1975. "Constructing Increment-Decrement Life Tables." *Demography* 12:313-24.

———. 1988. *Modeling Multigroup Populations*. New York: Plenum.

Schoen, Robert and Kenneth C. Land. 1979. "A General Algorithm for Estimating a Markov-Generated Increment-Decrement Life Table With Application to Marital-Status Patterns." *Journal of the American Statistical Association* 74:761-76.

Schoen, Robert and Verne E. Nelson. 1974. "Marriage, Divorce, and Mortality: A Life Table Analysis." *Demography* 11:267-90.

Smith, A.F.M. and G. O. Roberts. 1993. "Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods." *Journal of the Royal Statistical Society B* 55:3-23.

Teachman, Jay D. 1983. "Analyzing Social Processes: Life Tables and Proportional Hazards Models." *Social Science Research* 12:263-301.

Twaddle, Andrew C. and Richard M. Hessler. 1987. *A Sociology of Health* (2nd ed.). New York: Macmillan.

United Nations. 1988. *1986 Demographic Yearbook.* New York: United Nations.

U.S. Department of Health and Human Services, Public Health Service, and National Center for Health Statistics. 1987. "Health Statistics on Older Persons, United States, 1986." *Vital and Health Statistics*, Series 3, No. 25.

Xian Liu is an assistant research scientist in the Institute of Gerontology and a research affiliate in the Population Studies Center at the University of Michigan. His recent research includes the investigation of transitions in health status and the development of multistate active life table models. His articles have appeared in Population and Development Review, Journal of Gerontology: Social Sciences, *and* Social Science and Medicine.

Jersey Liang is a professor in the School of Public Health and a research scientist in the Institute of Gerontology at the University of Michigan. He currently is involved in issues related to the quality of life in the elderly, comparative aging, health, and health care.

Edward Jow-Ching Tu is a professor of sociology at Hong Kong University of Science and Technology. His current research interests include comparative study of population change and its implications, population aging, and migration.

Nancy Whitelaw is associate director of the Center for Health System Studies at the Henry Ford Health System in Detroit, Michigan, and a health science specialist in Health Services Research and Development at the Ann Arbor Department of Veterans Affairs Medical Center. Her research focuses on the integration of health services for the elderly and the measurement of health status in older populations.