How good is Sullivan’s method for monitoring changes in population health expectancies

Colin D Mathers, Jean-Marie Robine

Abstract

Study objective – To compare health expectancies calculated by Sullivan’s method and the multistate life table method in order to identify the magnitude of the bias in Sullivan’s method and assess how seriously this limits its use for monitoring population health expectancies.

Design – A simulation model was used to compare health expectancies calculated using Sullivan’s method and the multistate life table method under various scenarios for the evolution of disability over time in populations. The simulation model was based on abridged cohort life tables using data on French mortality from 1825–90 and disability prevalence data from the 1982 French health survey.

Main results – The Sullivan method could not detect a sudden change in disability transition rates, but the simulations suggested that it provides a good estimate of the true multistate value if there are smooth and relatively regular changes in disability prevalence over the longer term. When disability incidence rates are increasing or decreasing smoothly over time, the absolute bias in the Sullivan estimate of disability free life expectancy is relatively constant with age. The relative bias thus increases at older ages as disability free life expectancy decreases.

Conclusions – The difference between the estimates produced by the two methods was small for realistic scenarios for the evolution of population health and Sullivan’s method is thus generally acceptable for monitoring relatively smooth long term trends in health expectancies for populations.

(The Journal of Epidemiology and Community Health 1997;51:80–86)

The concept of a health indicator which combined information on mortality and morbidity was first proposed by Sanders1 and the first example of such an indicator was published in a report of the US Department of Health Education and Welfare,2 which contained preliminary estimates of “disability free life expectancy” (DFLE) calculated using a method devised by Sullivan.3 This involved using the observed prevalence of disability at each age in the current population (at a given point of time) to divide the years of life lived by a period life table cohort at different ages into years with and without disability. During the first half of the 1980s, Sullivan’s method was used to estimate DFLE for the USA,4 Canada,5 and France.6 Estimates of active life expectancy for the USA had also been made using double decrement life table methods.7

Health expectancy indices which combine mortality and morbidity into a single composite indicator are a very attractive tool for monitoring long term trends in the evolution of population health and for addressing the question of compression or expansion of morbidity in populations. During the second half of the 1980s, there was a dramatic increase in the number of health expectancy calculations carried out, almost all using the Sullivan method,4–6 which uses information on the prevalence of health states in the population. More recently, multistate life table methods have been developed which require information on transitions between health states. Various authors have claimed that Sullivan’s method produces biased or incorrect estimates and cannot be used to monitor health expectancies of populations over time.8–12 This paper compares health expectancies calculated by Sullivan’s method and the multistate life table method using a simulation model under various scenarios for the evolution of disability over time.

The problem with Sullivan’s method

Sullivan’s method uses the observed age specific prevalences of health states in a population at a given point in time (“sectional” prevalences) to calculate the years of life lived in the various health states at each age by a period life table cohort.13

Multistate life table methods were proposed by Rogers, Rogers and Belanger14 to take into account reversible transitions between good health and one or more disability (or other health) states. Their data for people aged 70 years or more from the US longitudinal survey on ageing (LSOA) show that transition rates from dependence to independence can be surprisingly high, even for the older old. In addition, the multistate life table method allows one to calculate health expectancies for population subgroups in a specific health state at a given age, eg those not disabled at age 65, whereas the Sullivan method gives only the average health expectancy for the entire population at a given age.

Problems relating to the validity of the Sullivan method were first raised in 1989 by Bebbington15 and Brouard and Robine.16 Bebbington compared the Sullivan method with the double decrement life table method using data where the disability incidence rate was rising over time. He demonstrated that the Sullivan method gives a lower estimate of
disability than the double decrement method because, in effect, the disability prevalence rate used in the Sullivan method reflects the past experience of each cohort and not the current incidence rates. Brouard and Robine similarly argued that the prevalence of disability is a stock dependent on past history, whereas the incidence of disability is a flow which can be used to compute a "pure period proportion" of disabled people, not dependent on past flows, which in turn could be used to compute a pure period indicator of disability. They noted that the question of whether to measure stocks or flows is a common dilemma in demography and in other disciplines such as economics. Brouard had previously demonstrated the large bias resulting from the use of stocks rather than flows to calculate trends in working life expectancy for French women and has reminded those debating the merits of the Sullivan and multistate methods of the very similar discussions relating to the use of prevalence versus longitudinal data in "working life tables". Rogers, Rogers and Belanger argued that the Sullivan method produces biased estimates of active life expectancy and that this bias, which is pessimistic, leads to the conclusion that active life expectancy is declining in a population whose transition probabilities between the independent and dependent states are held constant. They argued that the Sullivan index is biased in the direction of increased dependency whenever the independent population is very much larger than the corresponding dependent population. It was subsequently demonstrated that the discrepancy between estimates made using the two methods results not from bias in the prevalence index, but from the use of prevalence estimates from a non-stationary population, that is, a population in which the prevalence of dependency has not reached the equilibrium value associated with current transition rates.

It is now well understood that Sullivan's method, unlike the standard life table method for calculating period life expectancy, does not produce a pure cross sectional indicator derived from the current health conditions of the population. This is because the prevalence rates are partly dependent on earlier health conditions of each age cohort, that is, incidence, recovery, and state specific mortality rates applying at earlier times (or ages). To construct a purely cross sectional indicator, one would have to use the equilibrium prevalences observed in a fictitious cohort which had always been exposed to the observed cross sectional transition rates between health states. In an equilibrium or stationary population, where all transition rates are constant over time, Sullivan's method gives the same health expectancies as the multistate methods. The problems with Sullivan's method arise not because it uses prevalence and mortality data averaged over all health states, but because the data it uses are dependent on past conditions in the population. This problem had been illustrated in a scenario simulating large changes in cardiovascular case fatality rates over a short time period, although the authors incorrectly generalise from this extreme case to conclude that Sullivan's method is always inaccurate.

Theoretically, the multistate life table method is to be preferred for calculating health expectancies, but its use requires longitudinal data which are expensive and time consuming to collect and are rarely available. Hence we have developed a simulation model using French data, which allows us to compare the Sullivan estimate with the pure period estimate from the multistate life table for a non-stationary population which has experienced realistic changes in transition rates over time. Our aim is to identify the magnitude of the bias in Sullivan's method and assess how seriously it limits the use of the method for monitoring trends in population health.

### The multistate life table method

For these simulations, we work with abridged life tables using five year age groups throughout the lifespan, and terminating in a final open ended age interval of 85 years and over. It has been shown that abridged life tables provide estimates of health expectancies that are of acceptable accuracy compared to estimates based on full life tables. The starting point for the Sullivan method is the standard abridged life table, and particularly the life table functions \( l_x \) (the number of persons surviving to exact age \( x \)) and \( L_x \) (the total number of years lived in the age interval \( (x, x + 5) \)). The observed prevalence of disability, \( d_x \), in the age interval \( (x, x + 5) \) is used to divide the years lived in the age interval into years spent with disability, \( d_xL_x \), and years without disability \((1 - d_x)L_x\).

The average expectation of life free of disability at age \( x \), or DFLE at age \( x \) (DFLEX), is the total years of life lived free of disability from age \( x \) onwards divided by the number of persons alive at age \( x \):

\[
DFLEX_x = \sum_{i=1}^{n} (1 - d_x)L_x/l_x
\]

where the summation on index \( i \) is across the five year age intervals of the abridged life table \((x, x + 5, \ldots, w)\), where \( w = 85 \) denotes the final open ended age interval.

The multistate life table generalises the single state life table (which analyses the transition from a single "alive" state to the "absorbing" death state) to include reversible transitions between two or more non-absorbing "alive" states. In order to calculate DFLEX, we consider a two state life table with a non-disabled state (denoted by subscript 1) and a disabled state (denoted by subscript 2).

The number of survivors in state \( k = 1, 2 \) and age \( x \) is denoted by \( l_{xs} \) (see figure 1). The transition probability \( q_x \) of the single state life table becomes a transition matrix giving the probabilities of transition between states \( j \) and \( k \) in the age interval \( (x, x + 5) \). We ignore the complexities of multiple transitions and consider only the probabilities of transitions between an initial state at age \( x \) and a final state
Figure 1 Transitions in the multistate life table.

![Diagram of multistate life table]

at age \(x + 5\). The transition probability \(l_x\), the probability of a person not disabled at exact age \(x\) being disabled at exact age \(x + 5\), is closely related to the incidence rate of disability for the age interval \((x, x + 5)\). The transition probability \(r_x\), the probability of a person disabled at exact age \(x\) being free of disability at exact age \(x + 5\), is closely related to the recovery rate from disability for the age interval \((x, x + 5)\).

The transition rates \(q_x\) and \(q_{x+5}\) are the probabilities of dying within the interval \((x, x + 5)\) for a non-disabled and disabled person respectively.

If we assume that persons who make a transition between states \((1, 2 \text{ or } 3)\) spend on average one half of the interval \((x, x + 5)\) in the originating state (this could be generalised to different fractions for each pair of states if such information were available), then it is straightforward to calculate the survivors \(l_{x+5,1}\) and \(l_{x+5,2}\) in each state at age \(x + 5\) and the years of life \(L_{x+5,1}\) and \(L_{x+5,2}\) lived in states 1 and 2 during the interval \((x, x + 5)\).

The “disability free” life expectancy for the entire cohort at age \(x\) is

\[
DFLEX = \sum_{i=4}^{8} L_{x}/(l_{x1} + l_{x2})
\]

Experimentation with different values for the proportion of the interval \((x, x + 5)\) lived in the originating state has shown that the simulation results are not highly sensitive to the value of this parameter, and that the assumption that it is 0.5 for all transitions does not significantly affect the validity of the comparisons between the Sullivan and multistate methods for the scenarios examined below.

### The multistate simulation model

**ESTIMATION OF TRANSITION RATES FOR FRENCH MALES IN 1982**

Robine et al calculated the DFLEX for French males using data on the prevalence of disability (and institutionalisation) by five-year age group in 1982. We have chosen a set of transition rates to and from disability which give a good fit of the calculated prevalence of disability to the observed prevalence at each age using the following assumptions:

1. 1.05% of the life table cohort of 100 000 are born disabled (so \(l_{01} = 98 950\) and \(l_{02} = 1050\)).

2. The relative risk of death for disabled compared to non-disabled persons is assumed to be \(R = 3.66\) at each age. This allows the computation of the state specific mortality risks \(q_i\) and \(q_{i+5}\) using \(l_{i+5,1}\) and \(l_{i+5,2}\), the proportion of the cohort in each state and the total life table mortality risk, as follows:

\[
q_i = l_{i+5,1} q_i / (l_{i+5,1} + R l_{i+5,2})
\]

\[
q_{i+5} = R l_{i+5,2} q_i.
\]

Under this assumption, the state specific mortality risks will vary not only in accordance with trends in the overall mortality risk, but also with the relative prevalence of disability \((l_{i+5,2}/l_{i+5,1})\).

The value 3.66 was taken from LSOA data for persons aged 70 years or more. The simulations are not highly sensitive to the value assumed over a range of \(R = 1.0\) to 10.0, or to the assumption that the relative risk is the same at all ages.

Figure 2 illustrates the transition rates chosen for the stationary population scenario (incidence and recovery transition rates constant over time). The “recovery” rates were initially set at \(r_x = 0.2\) for ages 65 and over, based on estimated two-year recovery rates from the LSOA, and then incidence rates chosen to result in life table prevalence rates \(L_{i+5,2}/L_i\) close to the observed prevalence rates. It was found necessary to vary the recovery rates also to obtain reasonable fits. Note that there is not a unique solution for the incidence and recovery rates, although they are constrained severely by the observed prevalences. The rates chosen are purely for the purpose of carrying out the simulations described below.

### THE SIMULATION MODEL

In order to compare the Sullivan and multistate methods, we set up a model containing 34 fictitious cohorts of 100 000 males born at five
Assessment of Sullivan’s method in France

Figure 3 Comparison of Sullivan and multistate life table methods for calculating disability free life expectancy (DFLE) under scenario 1: constant mortality rates, disability incidence rates increase by 50% from 1945.

Figure 4 Comparison of the Sullivan and multistate life table methods for calculating disability free life expectancy (DFLE) under scenario 2: age specific disability incidence and recovery rates are constant over time.

year intervals from 1825 to 1990. The abridged period life tables for French males for each of these calendar years was obtained from published sources. For some of these years, where life tables were not available for the specific year, a life table was constructed using log-linear interpolation of the mortality indices $q$, from the life tables for the nearest available year on either side. The five year interval allowed us to ignore the effects of the two world wars during which life expectancy decreased substantially. These period life tables were then used to construct cohort life tables for each of the 34 birth cohorts, representing the actual survival experience of French male birth cohorts.

For each scenario for the evolution of disability transition rates over time (described below), a set of assumed transition rates were specified for each period from 1825 to 1990. These transition rates were chosen to give a good fit of the calculated period prevalence of disability by age in 1982 to the observed prevalence in 1982 as described above. These period transition rates together with the period life tables were used to construct multistate life tables for calendar years from 1920 to 1990 at five year intervals. The multistate life tables were used to calculate the total years lived in each state and hence the DFLE. This is a pure period expectancy based on the assumed period transition rates. The life tables for each period and birth cohort were all constructed using the three assumptions outlined in the previous section.

CALCULATING THE DFLE USING THE SULLIVAN METHOD
From the age specific transition rates assumed for the various periods in the model, we can obtain the age specific transition rates experienced by each birth cohort as it ages with time. Together with the cohort life tables, we can thus construct the multistate life table for each birth cohort and calculate the prevalence of disability at each age in the cohort. These prevalence rates at given ages are influenced by all the transition rates experienced by the cohort at earlier ages in earlier years. These were then used to estimate the period prevalence of disability by age for each calendar year. Note that these prevalences come from a different birth cohort for each age group and hence are different functions of transition rates for prior periods through which the cohorts have lived. Finally, these simulated “observed” period prevalences of disability by age are used together with the period life table to calculate DFLE using the Sullivan method.

CONSTRUCTION OF SCENARIOS
The following section examines a number of illustrative scenarios in which age specific transition rates from non-disabled to disabled change over time but, for simplicity, recovery rates are assumed to be constant over time for each age. Note that simple but reasonably realistic scenarios have been chosen which illustrate relatively regular long term trends in population health and also some short term and drastic changes in population health. Separate simulations were carried out to test the sensitivity of the results reported here to the assumption that recovery rates are constant over time. It was found that, while time trends in recovery rates contributed to differences between Sullivan and multistate estimates of DFLE, the Sullivan method is less sensitive to variations in recovery rates than disability rates for the types of scenarios examined here. This is because the recovery rates are generally higher in magnitude than incidence rates and so their time variations are more rapidly reflected in the resulting prevalence rates.
prevalence for French males in 1982. This is because the computed prevalences of disability for 1982 in the simulation model are dependent on the changes in transition rates assumed under the various scenarios.

After examining a first scenario which graphically illustrates the limitations of the Sullivan method, we examine an equilibrium scenario in which disability transition rates are constant over time while mortality rates decline as observed historically in the French male population. We then examine two extreme scenarios in which disability incidence rates increase and decrease at twice the absolute magnitude of the change in mortality rate at each age. These scenarios illustrate strong versions of the two major hypotheses for the evolution of health: compression of morbidity and expansion of morbidity. It is very likely that the real evolution of morbidity in the populations of developed countries falls somewhere between these extremes.

Results

Scenario 1
Mortality rates are assumed constant (at 1982 values) and in 1945 the incidence rate of disability is assumed to increase by 50% at each age.

Up to 1945, Sullivan's method and the multistate method give the same value for DFLE at birth (see figure 3). In 1945, the multistate estimate faithfully follows the drop in incidence rates, but the Sullivan estimate takes 30 years to reach the same value. This graphically illustrates how Sullivan's method has averaged the change in incidence rates in 1945 over 30 years through the use of period prevalence rates which take decades to reach their equilibrium values for the new transition rates.

Scenario 2
Age specific incidence and recovery rates are assumed constant with time. Mortality rates decline as observed historically in the French male population. As shown in figure 4, Sullivan's method gives almost the same results as the pure period estimate from the multistate life tables.

Scenario 3
Starting in 1945, at each age from 5 onwards, the disability incidence rate is assumed to decrease by twice the amount which the death rate of non-disabled persons decreases. Mortality rates decline as observed historically in the French male population. In this scenario, it is assumed that the "prevention of disease" reduces equally both the incidence of death from a healthy state and the incidence of non-fatal disability.

Sullivan's method and the multistate method give results that are reasonably close (see figure 5). The value given by Sullivan's method is lower than the multistate estimate, but gives a reasonably good indication of the trend over time.
**Assessment of Sullivan’s method in France**

**Figure 7** Postulated changes in disability rates over time for scenario 5: the disability incidence rate decreases except during the two world wars and the 1960s.

**Figure 8** Comparison of the Sullivan and multistate table methods for calculating disability free life expectancy (DFLE) under scenario 5: the disability incidence rate decreases except during the two world wars and the 1960s.

**Scenario 4**

At each age from 5 onwards, the disability incidence rate is assumed to increase by twice the amount which the death rate of non-disabled persons decreases. Mortality rates decline as observed historically in the French male population. In this scenario, it is assumed that all people “saved” from death are left disabled and that an equal extra number of non-disabled persons become disabled. This expansion of morbidity is assumed to start in 1945 (any earlier and disability incidence rates would go negative).

**Scenario 5**

Sullivan’s method and the multistate method give results that are still reasonably close (see figure 6), but in this case, the value given by Sullivan’s method is higher than the multistate estimate. For scenarios 3 and 4 (disability incidence rates decreasing and increasing by twice the mortality trend respectively), the Sullivan DFLE in 1990 are approximately lower and higher respectively than the multistate DFLE for ages 0 to around 65. As the absolute bias in the Sullivan estimate of DFLE is relatively constant with age, the relative bias increases from around 1.5% at birth to around 7% at older ages where DFLE is much lower.

**Discussion**

Data requirements for multistate methods are considerable and there are very few countries where national data is available or likely to be available for some time. Sullivan’s method is being used more and more widely. The Sullivan method is not capable of detecting a sudden change in disability transition rates, but our simulations suggest it provides a good estimate of the true period value if there are smooth and relatively regular changes over the longer term, as postulated by the principal scenarios for evolution of population health. We illustrate both these conclusions using a more complex but not unrealistic scenario for French males (see figure 7). In scenario 5, disability incidence rates jump substantially during the two world wars, but otherwise decline at a constant annual percentage except during the 1960s when the epidemic of cardiovascular disease was peaking. In order to illustrate clearly the differences between the methods, we ignore any effects of the two world wars on mortality. The resulting estimates of DFLE obtained by the Sullivan method and the multistate method are shown in figure 8. We conclude that Sullivan’s method seems to be a quite acceptable method for monitoring long term trends in health expectancies at the population level.

Barendregt et al also carried out a simulation comparing the Sullivan method with the multistate method. Their hypothetical example was based on a sudden and very large change in survival rates, stable incidence rates and varying prevalence rates and shows that the Sullivan method is quite incapable for monitoring the resulting changes in health expectancy. This confirms the conclusion reached here using scenario 5, that Sullivan’s method is not appropriate for detecting sudden changes in population health. However, they generalise their conclusion to state that Sullivan’s method is not appropriate for the analysis of changes in health expectancy over time in any circumstances. We disagree with the conclusion: the simulations presented here have shown that Sullivan’s method provides acceptable estimates of the true period value of health expectancy if there are smooth and relatively regular changes in transition rates over a reasonably long term period, and in particular, provides good estimates of trends in health expectancy.

We fully agree that the multistate life table method is to be technically preferred for calculating period health expectancies, since it is based on period transition rates only and will
detect sudden or gradual changes in disability transition rates. However, its use requires longitudinal data which are expensive and time consuming to collect and are rarely available. The LSOA has opened new prospects for calculation of truly cross sectional health expectancies, as evidenced by recent papers, but it should be noted that in order to avoid bias, institutionalised persons must be interviewed beginning with the first round of the longitudinal survey – which was not the case with the LSOA.

It may be possible to develop alternative methods to correct for the conservative bias inherent in Sullivan’s method. One possibility is to include a retrospective question on health state one or two years ago in cross sectional health surveys to obtain information about flows into and out of health states. There are two problems with this, long term recall of health states may be biased and the question can only be asked of survivors. Pollard, Golini, and Miledia have developed a method for calculating transition rates from such cross sectional survey data containing information on the time of onset of disease. It may also be possible to obtain information on differential survival through record linkage or use of a national death index.

Methods of analysis and calculation have proliferated at a far greater speed than the availability of good quality population health data, and we feel that the simulations reported here should lend confidence to the continued use of Sullivan’s method to calculate population health expectancy time series and assist in identifying the circumstances in which Sullivan’s method provides an acceptable alternative to the technically exact multistate method for monitoring trends in population health.

CONCLUSIONS

The Sullivan method is not capable of detecting a sudden change in disability transition rates, but it provides a very good estimate of the multistate model if there are smooth and relatively regular changes over the longer term and we conclude that Sullivan’s method seems to be a quite acceptable method for monitoring relatively smooth long term trends in health expectancies at the population level in non-stationary populations.

Data requirements for multistate methods are considerable and there are very few countries where national data is available or likely to be available for some time. Sullivan’s method is being used more and more widely. Sullivan’s method provides a useful indicator which can be used with confidence for monitoring trends, as long as its limitations are understood.

This work was assisted by the National Health and Medical Research Council (Australia) which provided a travelling Fellowship to Colin Mathers, Naturalia et Biologica (France), and a grant from CNAMTS (France).