SUPPLEMENTARY MATERIAL

Technical Appendix to Forecasting the Global Burden of Alzheimer's Disease

by

Ron Brookmeyer, Elizabeth Johnson, Katie Ziegler and Michael Arrighi

The Model

In this supplementary appendix, we describe the technical details of the methods in "Forecasting the Global Burden of Alzheimer's Disease." The underlying model is a multi-state discrete time Markov model. Individuals may transition from the healthy state (S_0), to early stage Alzheimer's disease (S_1) and then to late stage disease (S_2). Persons are at risk of death in each state. The model is illustrated by the figure below:

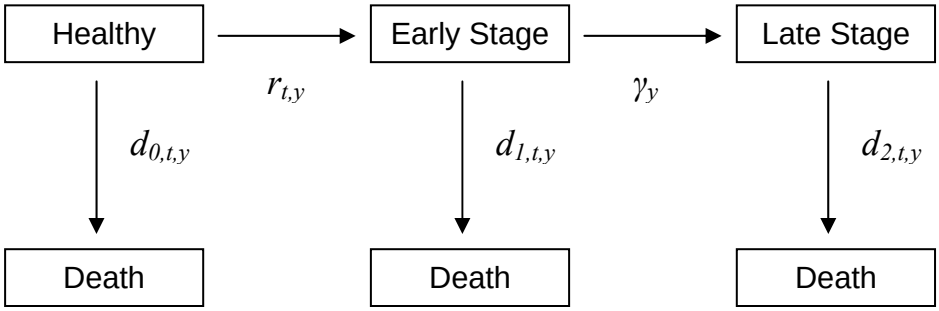

Figure: Schematic illustration of multi-state model for Alzheimer's disease showing transition probabilities between states.

The figure shows the transition rates which are the probabilities of moving between states in a year. The transition rates may depend on age (t) and calendar year (y). We define $r_{t,y}$ to be the conditional probability that a healthy individual who is age t in year y has onset of early stage disease during year y. We define γ_y to be the probability that a person with early stage disease in year y progresses to late stage disease during the year. The death rates are the probabilities that a person who is age t in year y dies during the year , which are allowed to depend on the disease state; specifically, the death rates are called $d_{0,t,y}$ for persons in the healthy state, $d_{1,t,y}$ for persons with early stage disease, and $d_{2,t,y}$ for persons with late stage disease.

We model the effect of interventions on disease onset that begin in year Y through a multiplicative model: $r_{ty}=r_t$ if $y<Y$, and $r_{ty}=\theta r_t$ if $y\geq Y$, where θ is the relative risk associated with the efficacy of the intervention, and r_t is the age specific incidence curve of onset of Alzheimer's disease in the absence of interventions. Similarly, we model the effect of interventions on disease progression rates that begin in year Y through the multiplicative model: $\gamma_y=\gamma$ if $y<Y$, and $\gamma_y=\varphi\gamma$ if $y\geq Y$ where φ is the relative risk associated with the treatment to delay disease progression, and γ is the progression rate from early to advanced disease in the absence of any intervention.

We model the effect of Alzheimer's disease on mortality through an additive model:

$$d_{1,t,y} = d_{0,t,y} + k_1$$

$$d_{2,t,y} = d_{0,t,y} + k_2$$

The parameters k_1 and k_2 represent the additional effect of Alzheimer's disease on background mortality rates for persons with early and late stage disease respectively.

Estimating and Forecasting Prevalence

We now describe the equations for estimating and forecasting disease prevalence.

The probability that a person who was born in year y-t is alive at age t in year y and living in state S_j is called $p_{j,t,y}$ We give expressions for the probabilities $p_{j,t,y}$ in terms of the transition probabilities at the end of this appendix. The age-specific disease prevalence rate at age t in year y, is the fraction of persons alive at age t in year y who have disease; that is we define the age specific prevalence to be the conditional probability of disease given that the person is alive. Thus, the age specific prevalence rate for early stage Alzheimer's disease at age t in year y is :

$$\frac{p_{1,t,y}}{p_{0,t,y} + p_{1,t,y} + p_{2,t,y}} \qquad (1)$$

The age specific prevalence rate for late stage Alzheimer's disease at age t in year y is:

$$\frac{p_{2,t,y}}{p_{0,t,y} + p_{1,t,y} + p_{2,t,y}} \qquad (2)$$

Then, the age- specific prevalence rate for Alzheimer's disease (either early or late stage disease) at age t in year y is obtained by summing equations 1 and 2:

$$R_{t,y} = \frac{p_{1,t,y} + p_{2,t,y}}{p_{0,t,y} + p_{1,t,y} + p_{2,t,y}}$$

The age specific prevalence rates are then combined with population projections to obtain forecasts of numbers of current and future cases of disease. Let $N_{t,y}$ be the population forecast of the numbers of persons alive at age t in year y. Then, the estimated numbers of persons living with Alzheimer's disease at age t in year y is $N_{t,y}$ x $R_{t,y}$. Similarly, the numbers of persons living with early and late stage Alzheimer's are obtained by multiplying $N_{t,y}$ by equations (1) and (2), respectively.

Probability Expressions for $p_{j,t,y}$

We now derive the equations for $p_{j,t,y}$ which are the probabilities that a person born in year y-t is alive at age t and in state S_j in year y. We express these equations in terms of the underlying transition probabilities(see figure). To avoid ambiguity, we assume all transitions occur in the beginning of each year, with transitions to early stage disease occurring first, followed by transitions to late stage disease, and finally death. Thus, an individual could experience multiple transitions in a year. Our equations for disease prevalence refer to the probability of being in a particular state after all transitions for the year have occurred.

The probability that an individual who was born in year y-t is living in the healthy state (S_0) at age t in year y is:

$$p_{0,t,y} = \prod_{j=1}^{t}\left(1-r_{j,y-t+j}\right)\left(1-d_{0,j,y-t+j}\right)$$

The probability that an individual born in year y-t is alive at age t in year y and living with early stage Alzheimer's disease (S_1) is:

$$p_{1,t,y} = \sum_{i=1}^{t}\left[\prod_{1\le j\le i-1}\left(1-r_{j,y-t+j}\right)\left(1-d_{0,j,y-t+j}\right)\right]\left[r_{i,y-t+i}\right] \; x \left[\prod_{k\ge i}^{t}\left(1-\gamma_{y}\right)\left(1-d_{1,k,y-t+k}\right)\right]$$

This expression for $p_{1,t,y}$ is derived from the following considerations. In order to have early stage disease at age t, an individual must have had disease onset at some prior age i. The first term in brackets refers to the probability of remaining in the healthy state until onset at age i; the second term in brackets refers to the probability of disease onset at age i; and the third term refers to the probability of remaining in the early stage of disease through age t.

The probability that an individual who was born in year y-t is alive at age t in year y and living with late stage disease (S_2) is:

$$p_{2,t,y} = \sum_{i=1}^{t} \sum_{l=i}^{t} \left[\prod_{1 \le j \le i-1} \left(1 - r_{j,y-t+j}\right)\left(1 - d_{0,j,y-t+j}\right) \right] \left[r_{i,y-t+i} \right]$$
$$x \left[\prod_{k=i}^{l-1} \left(1 - \gamma_{y-t+k}\right)\left(1 - d_{1,k,y-t+k}\right) \right] \left[\gamma_{y-t+l} \right] \left[\prod_{k=l}^{t} \left(1 - d_{2,k,y-t+k}\right) \right]$$

This expression above for $p_{2,t,y}$ is derived from the following considerations. In order to have late stage disease at age *t,* an individual must have had onset of early stage disease at some prior age *i* (where *i* varies between 1 and *t*), and then subsequently progressed to late stage disease at age *l*, (where *l* varies between *i* and *t*). The first term in brackets refers to the probability of remaining in the healthy state until onset at age *i*; the second term refers to the probability of onset at age *i*; the third term in brackets refers to the probability of remaining in early stage disease until age *l* (we set this term to 1 if *l*=*i*); the fourth term refers to the probability of progressing to late stage disease; and finally the fifth term in brackets refers to the probability of remaining in late stage until age *t.*

Input parameters and Computing Software

A computer program to implement the model is available from the authors. The program incorporates U.N. world population forecasts through 2050, and background gender and age specific mortality rates from vital statistics data from 1959 to the present. The program allows the users to enter user defined input parameters including population projections for specific countries or regions, and disease incidence and progression rates, as well as the flexibility to modify the background mortality rates and the efficacy and start date of potential interventions. As such, the computer program has the flexibility to forecast the prevalence of other diseases in elderly populations in addition to Alzheimer's disease.