Modeling Disability-Adjusted Life Years for Policy and Decision Analysis

Abstract

This study outlines a methodological framework for joint modeling of Disability- and Quality-Adjusted Life Year outcomes. Our primary focus is on how transition matrices and state occupancy payoffs in discrete-time Markov cohort models can be structured to calculate years of life lost to disability (YLD) and years of life lost to premature death (YLL), in addition to quality-adjusted life year (QALY) outcomes. We also demonstrate how our modeling framework extends directly to microsimulation and (in part) to continuous time discrete event simulation (DES) models. In a tutorial application, we use our joint modeling framework to construct a discrete time Markov cohort natural history model for cardiovascular disease that estimates DALY and QALY outcomes for any country, region, or setting represented in the 2020 Global Burden of Disease data.

Plain Language Summary

Structuring Markov Models for Multidimensional Health Outcomes

0.1 Introduction

Source: Article Notebook

Disability-adjusted life years (DALYs) measure disease burden in a population. Conceptualized in the Global Burden of Disease (GBD) study (C. J. Murray & Lopez, 1997), DALYs quantify the total sum of years of life lost due to disability attributable to a disease (YLD), plus years of life lost to premature mortality from the disease (YLL; i.e., DALY = YLD + YLL).

In addition to their role in describing levels and trends in disease burdens worldwide, DALYs are a primary health outcome in evaluations of health interventions in low- and middle-income countries (LMICs). In these settings, resource allocation decisions are guided by modeled assessments of the incremental costs per DALY averted under alternative (often competing) strategies to improve population health.[1]

Despite the prominent role of DALYs in global health policy, scant methodological guidance is available for adapting and/or structuring decision analytic models for DALY outcomes. This methodological gap has its roots in health economics education, where textbooks and training exercises focus almost exclusively on Quality-Adjusted Life Year (QALY) outcomes—the primary health outcome used for health technology assessments (HTAs) and policy decisionmaking in high-income countries (HICs). DALYs differ from QALYs in important and model-relevant respects, including the use of reference life tables to calculate YLLs and standardized disability weights to calculate YLDs.[2] To the extent DALY-specific modeling considerations are taught, they are often considered in isolation and without a firm methodological grounding in *how* one might structure a model to measure DALY outcomes.

As a consequence, and in practice, health economic applications often resort to shortcuts and other "hacks" for calculating DALYs. For example, practitioners may simply estimate a "QALY-like" DALY that is based on a diseased state occupancy payoff of one minus the disability weight. Other approaches define a diseased-state payoff using the disability weight as an estimate of YLDs, and accumulate person-years in an absorbing death state (due to disease) as an estimate of YLLs. As this

[1] The adoption of DALYs over other common health outcomes in health economics (e.g., quality-adjusted life years, or QALYs) stems from several practical and theoretical considerations. See Feng et al. (2020) and Wilkinson et al. (2016) for futher discussion.

[2] In contrast, QALYs are calculated based on utility weights derived from general and patient surveys. See Feng et al. (2020) and Wilkinson et al. (2016) for futher discussion.

study will show, these shortcuts do not provide an accurate representation of DALY levels in a population.

This tutorial outlines a framework for direct incorporation of DALY outcomes in common decision modeling environments. Our primary focus is on discrete-time Markov cohort models—however, our framework extends directly to microsimulation and is also easily adapted for continuous time discrete event simulation (DES) models. As such, our study provides a comprehensive roadmap for incorporating DALY outcomes into common decision modeling frameworks.

To maintain consistency within the literature, this tutorial builds on an existing didactic disease progression model (Alarid-Escudero et al., 2023). The underlying discrete time Markov cohort model is time homogeneous—that is, transition probabilities do not vary as a function of age/time in model. However, the methods and code provided are easily adapted for time inhomogenous models. Finally, recognizing the wide spectrum of experience and programming comfort level among those constructing DALY-based models, we offer three approaches for modeling DALYs (beginner, intermediate and advanced) and provide replication materials for implementing our approaches in both R and Microsoft Excel.

0.2 Background

DALYs are calculated from two components. First, conditions are assigned disability weights (D) ranging from zero to one, with zero representing the absence of the condition and one representing the highest burden a condition can inflict, equivocal to death. Years lost to disability (YLD) is defined as the disability weight multiplied by the number of years a person lives with the condition (L):

$$YLD(L) = D \cdot L \tag{1}$$

The impact of disease on mortality is quantified using years of life lost to disease (YLL), which is based on remaining life expectancy $Ex(a)$ at the age of premature death from the disease (a).

$$YLL(a) = Ex(a) \tag{2}$$

.

DALYs are the sum of the two components:

$$DALY(L, a) = YLD(L) + YLL(a) \tag{3}$$

In the original GBD study, age-weighting and time discounting practices were applied to DALY calculations (C. J. Murray & Lopez, 1997). These methods respectively weighted the burden of illness more during adulthood than early childhood and old age, and valued present health over future years of illness by discounting YLD and YLL measures by 3% per year. From 2010 onwards, both practices were discontinued to make the DALY a more descriptive measure (WHO, 2013).

While the GBD no longer uses age and time discounting, the World Health Organization's Choosing Interventions that are Cost-Effective (WHO-CHOICE) program recommends consideration of time discounting of health outcomes (Bertram et al., 2021; C. J. L. Murray et al., 2020). We therefore adopt the WHO-CHOICE recommendation and include discounting in our DALY modeling approach.[3] We do,

[3] Practitioners who do not wish to discount DALY outcomes can simply set the annual discount rate r to zero.

however, maintain the continuous-time discounting used in the original GBD DALY equations—which differs slightly from the more common use of discrete time discounting in Markov cohort models.

For an annual discount rate r, and at age a, the equation for YLDs is,

$$YLD(a, L) = D\left(\frac{1}{r}\left(1 - e^{-r(L)}\right)\right). \tag{4}$$

Similarly, YLLs are calculated as,

$$YLL(a) = \frac{1}{r}\left(1 - e^{-rEx(a)}\right). \tag{5}$$

It is important to note that the discounting shown in Equation **??** and Equation **??** yield the present value of YLD and YLL outcomes at a single point in time, when the duration of disease (L) and time of death from disease (a) are known. For a decision model where not all cohort members start off ill, that point in time very likely occurs at some point after the baseline period—and different illness durations and death times will, of course, occur across different individuals in the modeled cohort. As such, to be consistent with standard practice, we must discount YLL and YLD outcomes further—all the way back to the baseline period. This additional discounting step will become apparent in the discrete time YLD and YLL equations introduced in **?@sec-occupancy** below.

0.2.1 Reference Life Expectancy

While the creation of the DALY measure was an important step in global health research, it has received scrutiny due to its inherent assumptions and value judgements. For example, calculating YLLs requires the use of a reference life table that provides an estimate of …

0.3 Model Overview

We build on an existing progressive disease model in which healthy individuals develop a disease with two health states ("Sick" and "Sicker") (Alarid-Escudero et al., 2023). Individuals can also transition to an absorbing death state due to causes unrelated to the disease (i.e., "background" mortality), or due to disease-specific causes. In addition, the model structure is homogeneous (i.e., transition rates do not vary as a function of time). This is a simplification to distill model complexity down to only those components needed to demonstrate our DALY approach; our replication code is structured in such a way as to easily incorporate transition rates that are a function of age/time in the model.

We consider outcomes under four strategies:

- A **Standard of Care** strategy based on the baseline model parameters.
- **Strategy A**, which improves the quality of life among individuals with the disease, but does not affect disease progression.
- **Strategy B**, which reduces the rate of progression from Sick to Sicker by 40%.
- **Composite Strategy AB**, which jointly implements strategies A and B.

A state transition diagram is shown in Figure **??**. In the figure, nodes are health states and edges depict possible transitions among them. Edge labels are defined in terms of transition intensities (rates). Other key model parameters are summarized in TK….

We note that in our implementation of the Sick-Sicker model, we define the disability weight as one minus the utility weight. In general, this will not be the case as disability weights and utility weights are estimated in different ways. TK

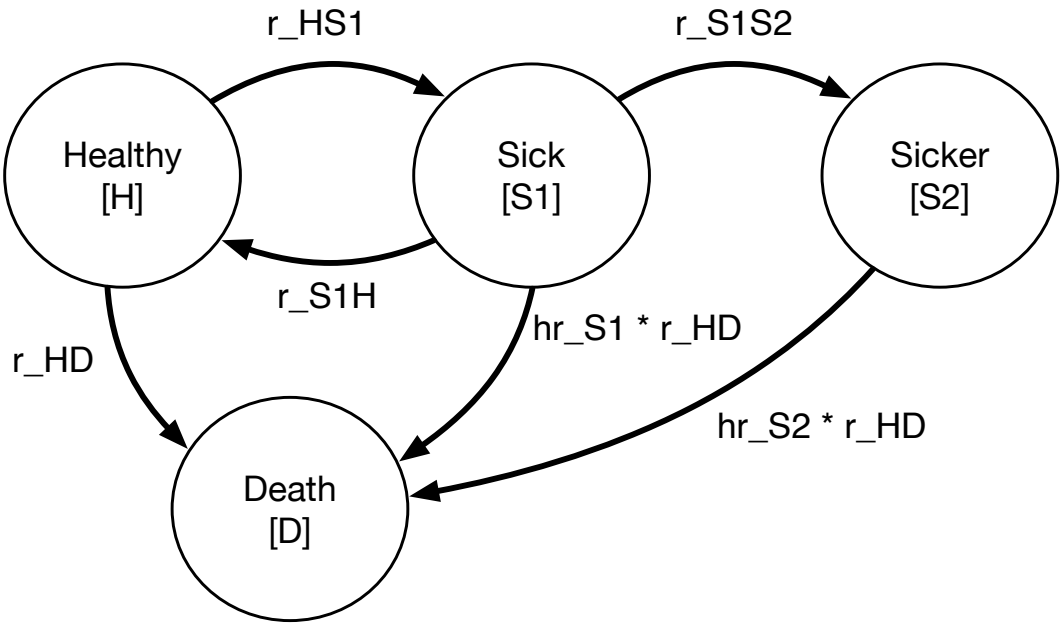

Figure 1: State transition diagram for progressive disease model

Source: Article Notebook

Source: Article Notebook

0.4 Transition Matrices

With the model parameterized, our next step is to define the matrices that govern health state transitions. The state transition diagram represented in Figure **??** is not well-suited to calculate DALY outcomes, however. A primary reason is that the death transitions reflect transitions due to all causes of death. To calculate YLLs, we need to separately track the timing and number of deaths *due to disease*.

To accommodate this need and to accurately model DALY outcomes, several options are available. We categorize each approach based on the level of experience and skill needed to execute it (beginner, intermediate, advanced):

1. **Separate Death State (Beginner)**: Re-define the health states to include a separate cause-specific death state as depicted in Figure **??**.[4] We then draw on the resulting Markov trace and use changes in the number of cause-specific deaths in each cycle to calculate YLLs.

2. **Non-Markovian Trackers (Intermediate)**: Include a non-Markovian transition state for cause-specific deaths in the transition matrix. This approach will maintain the Markovian components captured in Figure **??**, but will allow us to add a column to the Markov trace that separately tracks the number of new deaths from the disease in each cycle. We can then apply transition state

[4] In this example, disease-specific death rates are goverened by a hazard ratio applied to the background mortality rate. Because we are operating on the rate scale, we can separate out disease-related deaths from other-cause mortality by simply taking a difference in the rates. Other applications for prevalent conditions with high death rates, however, may require us to construct a cause-deleted life table to obtain background mortality rates that net out deaths from the modeled disease.

payoffs (based on remaining life expectancy at each age/cycle) to calculate YLL outcomes.

3. **Markov Chain with Rewards (Advanced)** Define a block matrix Markov chain with rewards for occupancy (YLDs) and disease-relatd death transitions (YLLs) by adapting the methods in Caswell & van Daalen (2021). This approach draws on matrix calculus and solves for expected outcomes as well as higher order moments such as variance and skewness.

Each of these approaches facilitate the design and execution of a decision-analytic model that correctly calculates YLD, YLL, and DALY outcomes—as well as other common outcomes such as QALYs and costs. In practice, Approaches (1) and (2) will produce identical results. Approach (3) draws on slightly different cycle correction techniques, but will yield results quite similar to (1) and (2). Other shortcut approaches previously used in the literature, such as modeling a QALY-like DALY and/or accumulating time in the absorbing death state, will not in general yield similar results; we discuss reasons for this in Section TK below.

0.4.1 Approach 1 (Beginner): Cause-Specific Death State

Under this approach, we separate out deaths from disease vs. other causes by defining a separate health state for cause-specific mortality; an updated state transition diagram is shown in Figure **??**.

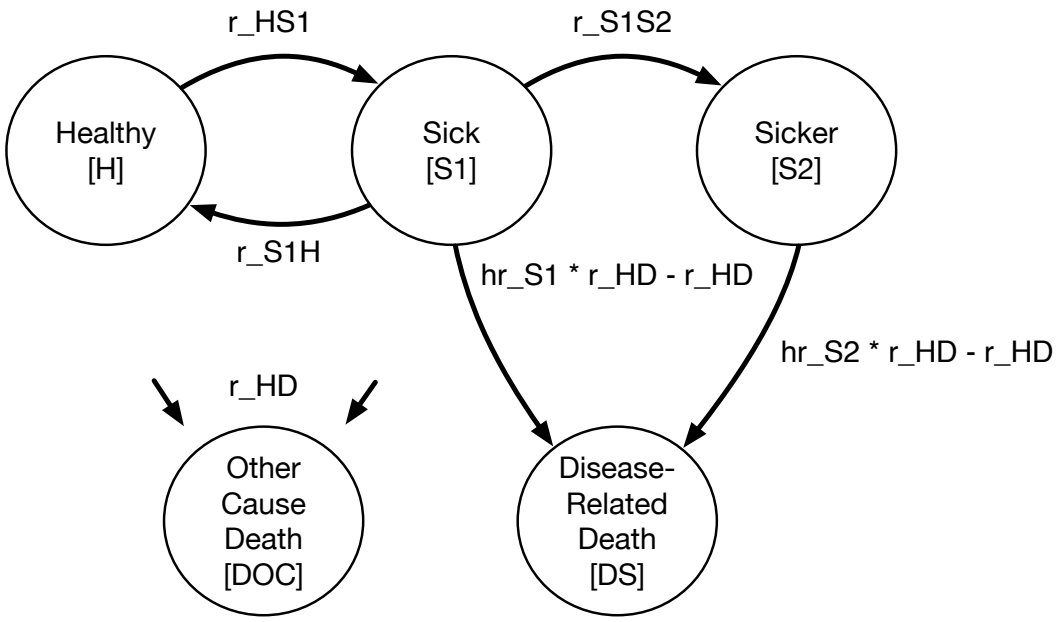

Figure 2: State transition diagram for progressive disease model with separate cause-specific death state

Transitions among health states are defined in terms of continuous rates ("intensities") and are captured within an intensity matrix $\mathbf{Q}$,

$$\mathbf{Q} = \begin{array}{c} \\ \text{H} \\ \text{S1} \\ \text{S2} \\ \text{DOC} \\ \text{DS} \end{array} \begin{array}{c} \begin{array}{ccccc} \text{H} & \text{S1} & \text{S2} & \text{DOC} & \text{DS} \end{array} \\ \begin{pmatrix} -(\texttt{r\_HS1}+\texttt{r\_HD}) & \texttt{r\_HS1} & 0 & \texttt{r\_HD} & 0 \\ \texttt{r\_S1H} & -(\texttt{r\_S1H}+\texttt{r\_S1S2}+\texttt{hr\_S1}\cdot\texttt{r\_HD}) & \texttt{r\_S1S2} & \texttt{r\_HD} & \texttt{hr\_S1}\cdot\texttt{r\_HD}-\texttt{r\_HD} \\ 0 & 0 & -(\texttt{hr\_S2}\cdot\texttt{r\_HD}) & \texttt{r\_HD} & \texttt{hr\_S2}\cdot\texttt{r\_HD}-\texttt{r\_HD} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

Figure 3: Transition Intensity Matrix for Approach 1

Cell values in row i, column j of $\mathbf{Q}$ capture the (continuous time) transition rate from health state i to health state j. $\mathbf{Q}$ has diagonal elements defined as the negative sum of the off-diagonal row values (i.e., the row sums of $\mathbf{Q}$ are zero). This ensures that the Markov model is "closed"—that is, the total cohort size neither grows nor shrinks over time.

We next embed the continuous transition intensity matrix into a discrete time transition probability matrix by taking the matrix exponential of $\mathbf{Q}$ for a defined time step (i.e., "cycle length") Δt:[5]

$$\mathbf{P} = e^{\mathbf{Q}\Delta t} \tag{6}$$

This results in a transition probability matrix with the following probabilities defined:

$$\mathbf{P} = \begin{array}{c} \\ \text{H} \\ \text{S1} \\ \text{S2} \\ \text{DOC} \\ \text{DS} \end{array} \begin{array}{c} \begin{array}{ccccc} \text{H} & \text{S1} & \text{S2} & \text{DOC} & \text{DS} \end{array} \\ \begin{pmatrix} \texttt{p\_HH} & \texttt{p\_HS1} & \texttt{p\_HS2} & \texttt{p\_HDOC} & \texttt{p\_HDS} \\ \texttt{p\_S1H} & \texttt{p\_S1S1} & \texttt{p\_S1S2} & \texttt{p\_S1DOC} & \texttt{p\_S1DS} \\ 0 & 0 & \texttt{p\_S2S2} & \texttt{p\_S2DOC} & \texttt{p\_S2DS} \\ 0 & 0 & 0 & 1.0 & 0 \\ 0 & 0 & 0 & 0 & 1.0 \end{pmatrix} \end{array}$$

Figure 4: Transition Probability Matrix for Approach 1

Embedding the transition probability matrix using the matrix exponential ensures that the resulting transition probabilities capture the underlying continuous time disease process. In particular, $\mathbf{P}$ will capture the probability of compound ("jumpover") transitions within a single cycle.

For example, in the continuous time rate matrix $\mathbf{Q}$ above, there is a zero-valued rate defined for progressions from Healthy (H) to Disease-related death (DS), since

[5] In Markov theory, $\mathbf{P}$ is called the "discrete skeleton" of the continuous model (Iosifescu, 1980). The conversion formula used to calculate $\widetilde{\mathbf{P}}$ is the matrix analogue to the standard rate-to-probability formula commonly taught in health economics textbooks, i.e., $p = 1 - e^{r\Delta t}$, where r is the rate and Δt is the time step (i.e., "cycle length").

individuals must first become ill before they can die from disease-related causes. However, after embedding, the matrix **P** has a non-zero cycle transition probability from Healthy (H) to Disease-related death (DS) (i.e., `p_HDS`). This value captures the probability of a compound or "jumpover" transition from Healthy and through the Sick and/or Sicker state to death from disease-related causes within the same discrete time cycle; see Graves et al. (2021) for further discussion, and Iosifescu (1980) for additional theory.[6]

0.4.2 Approach 2 (Intermediate): Non-Markovian Tracking States

This section will outline an approach similar to Approach 1, but that draws on a non-Markovian "transition" state that tracks the number of disease-related deaths in each cycle; these counts will be used later to match the age of the cohort at each cycle with a reference life table to calculate YLL outcomes.

Under this approach, we maintain the overall structure as depicted in the original model (Figure **??**), but augment the transition probability matrix with non-Markovian components to facilitate accounting of disease-related deaths.[7] Approach 2 offers a more generalized method that allows practitioners to accurately account for costs and/or health payoffs (such as YLLs) that are defined by *transitions* among health states, rather than occupancy in a health state.

Figure **??** shows a state transition diagram with the tracking state added. The tracking state (shown as red nodes) simply records transitions as cohort members move from either diseased state to the absorbing death state due to causes related to the disease.

In general, tracking states can either count the total number of transitions that have occurred up to a given cycle (i.e., an "accumulator" state), or can track the total number of new transitions that occur within a single cycle (i.e., a "transition" state).[8] To calculate YLL outcomes we will add a transition state that records the total number of new disease-related deaths in each cycle.

To implement Approach 2, we add a transition state row and column to the transition intensity matrix. This transition state, called `trDS`, is included in the augmented intensity matrix **Q** below:

[6] Because we embed the transition probability matrix using matrix exponentiation, rather than through pairwise application of rate-to-probability formulas to each transition type, our results will differ from those in Alarid-Escudero et al. (2023)—even though we use identical input parameters. Application of standard rate-to-probability formulas in health states with competing risks (i.e., the possibility of transitioning to more than one other health state in a given cycle) will ignore the possibility of compound transitions within a single cycle. Though not (yet) widely used in health economics, embedding the transition probability matrix using the matrix exponential is the technically correct way to construct a transition probability matrix from underlying transition rates.

[7] Tracking states also allow for accurate bookeeping for other outcomes such as costs. For example, if developing the disease incurs a one-time diagnosis or treatment cost, the compound transitions implied by the embedded transition probability matrix indicate that some individuals will transiently enter (and then exit) the Sick state in a single cycle. When calculating costs, practitioners may want to include a tracking state for the Sick state to be sure to capture these one-time costs, which would be masked if cost payoffs are simply multiplied by state occupancy at the end of each cycle (e.g., costs for individuals with a sojourn through the Sick state in a single cycle would not be accounted for).

[8] More generally, accumulator and transition states can be defined for any number of transition types, as they are useful for capturing one-time costs in the model, or for for calculating other decision-relevant outcomes such as the total number of people who developed the disease or died from the disease as secondary outcomes.

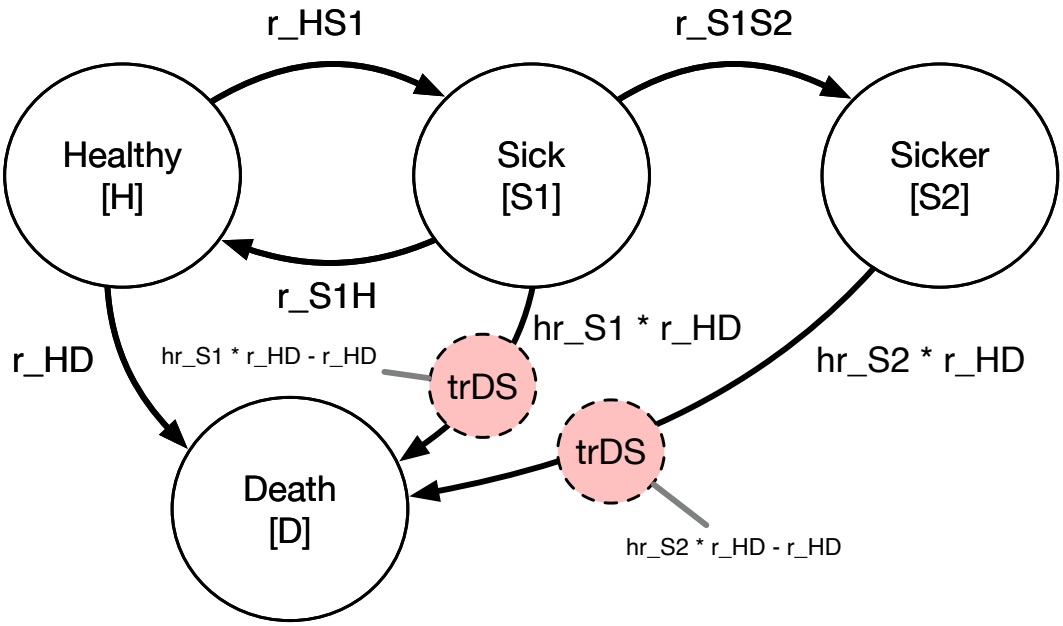

Figure 5: State Transition Diagram with Transition State (Red)

$$\mathbf{Q} = \begin{array}{c} \\ \text{H} \\ \text{S1} \\ \text{S2} \\ \text{D} \\ \text{trDS} \end{array} \begin{array}{c} \begin{array}{ccccc} \text{H} & \text{S1} & \text{S2} & \text{D} & \text{trDS} \end{array} \\ \left(\begin{array}{ccccc:c} -(\texttt{r\_HS1}+\texttt{r\_HD}) & \texttt{r\_HS1} & 0 & \texttt{r\_HD} & 0 \\ \texttt{r\_S1H} & -(\texttt{r\_S1H}+\texttt{r\_S1S2}+\texttt{hr\_S1}\cdot\texttt{r\_HD}) & \texttt{r\_S1S2} & \texttt{hr\_S1}\cdot\texttt{r\_HD} & \texttt{hr\_S1}\cdot\texttt{r\_HD}-\texttt{r\_HD} \\ 0 & 0 & -(\texttt{hr\_S2}\cdot\texttt{r\_HD}) & \texttt{hr\_S2}\cdot\texttt{r\_HD} & \texttt{hr\_S2}\cdot\texttt{r\_HD}-\texttt{r\_HD} \\ 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 \end{array}\right) \end{array}$$

Figure 6: Transition intensity matrix with transition state added

Two aspects of **Q** are worth highlighting. First, **Q** is divided into a Markovian submatrix and the non-Markovian tracking row and column. This division is made apparent using dotted vertical and horizontal lines. Critically, the Markovian submatrix remains closed—that is, the diagonal elements remain unchanged so that the row sums of the submatrix remain zero, even after the addition of the tracking column along the "edges" of **Q**. This ensures that the Markovian submatrix can be used to calculate state occupancy for a closed cohort that neither gains nor loses cohort members over the modeled time horizon.

Second, two transition intensities—from the S1 (Sick) and S2 (Sicker) states to Death—appear in the tracking column. This ensures that `trDeadDisease` will track all relevant transitions to death due to the disease. Because we are operating on the rate scale, we can net out non-disease related deaths as captured by the background mortality rate among healthy individuals (i.e., `r_HD`).

As above, we obtain the transition probability matrix by embedding **Q** into the discrete time step (Equation **??**). However, the resulting transition probability matrix treats `trDS` as an absorbing state (i.e., individuals are retained in the `trDS` with

probability one). Using the terminology introduced above, this absorbing state could serve as an **accumulator** state that (in the constructed Markov trace) records the total number of people who have died from the disease up to any given cycle. This may be a decision-relevant health outcome to consider on its own; indeed, so long as the Markovian submatrix remains closed, there is no limit to the number of accumulator and/or transition states one might add along the "edges" of a model.[9]

To change `trDS` to a **transition** state, we simply replace the absorbing probability of one in the cell [`trDS`, `trDS`] with a zero. This cell-level change is highlighted in grey in the bottom right cell of **P** below:

$$\mathbf{P} = \begin{array}{c} \\ \text{H} \\ \text{S1} \\ \text{S2} \\ \text{D} \\ \text{trDS} \end{array} \begin{array}{c} \begin{array}{ccccc} \text{H} & \text{S1} & \text{S2} & \text{D} & \text{trDS} \end{array} \\ \left(\begin{array}{ccccc} \texttt{p\_HH} & \texttt{p\_HS1} & \texttt{p\_HS2} & \texttt{p\_HD} & \texttt{p\_HDS} \\ \texttt{p\_S1H} & \texttt{p\_S1S1} & \texttt{p\_S1S2} & \texttt{p\_S1D} & \texttt{p\_S1DS} \\ 0 & 0 & \texttt{p\_S2S2} & \texttt{p\_S2D} & \texttt{p\_S2DS} \\ 0 & 0 & 0 & 1.0 & 0 \\ 0 & 0 & 0 & 0 & 0.0 \end{array}\right) \end{array}$$

Figure 7: Transition Probability Matrix for Approach 2

Source: Article Notebook

0.4.3 Approach 3 (Advanced): Markov Chain With Rewards

Define

$$\begin{aligned} \tau &= \text{Number of transient (non-absorbing) states} \\ \alpha &= \text{Number of absorbing states} \\ \omega &= \text{Number of cycles} \\ s &= \text{Total number of states; } s = \tau\omega + \alpha \\ \mathbf{U}_x &= \text{Transition matrix for age } x, \text{for } x = 1, \ldots, \omega \\ \mathbf{D}_j &= \text{Age advancement matrix for stage } j, \text{for } j = 1, \ldots, \tau \\ \mathbf{M}_i &= \text{Mortality matrix for age class } x, \text{for } x = 1, \ldots \omega \\ \mathbf{K} &= \text{vec-permutation matrix; parameters } \tau, \omega \end{aligned}$$

In the above notation, the matrix $\mathbf{U}_x$ captures transition probabilities among transient (i.e., non-absorbing) health states, while $\mathbf{M}_x$ captures transitions from transient health states to the absorbing death states (non-disease mortality and disease-related mortality). Indexing by age class x indicates that separate matrices are defined for each age in the Markov model.

[9] To build on the example of compound "jumpover" transitions above, suppose an individual starts off healthy in a cycle, then rapidly transitions through the Sick and Sicker state and dies due to disease-related causes within the same cycle. If there is some treatment cost associated with being in the Sicker state, a traditional approach that applies cost payoffs to state occupancy at the (beginning) end of the cycle would miss treatment costs for this individual because they *transition* through the Sicker state, but never occupy it at the beginning or end of a cycle. Adding a non-Markovian transition state to the model facilitates more accurate bookkeeping because the transition state would pick up on this transition through the Sicker state.

To construct $\mathbf{U}_x$ and $\mathbf{M}_x$ we define transition rate ("intensity") matrices as in Approaches 1 and 2 above.[10] The overall intensity matrix $\mathbf{Q_x}$ is given by

$$\mathbf{Q}_x = \left(\begin{array}{c|c} \mathbf{V}_x & \mathbf{0} \\ \hline \mathbf{S}_x & \mathbf{0} \end{array}\right)$$

where $\mathbf{V}_x$ is the rate matrix for the transitory (i.e., non-absorbing) states and $\mathbf{S}_x$ is the rate matrix for the absorbing states. The diagonal elements of $\mathbf{Q}_x$ are the negative sum of the non-diagonal column elements, thus making the column sums of $\mathbf{Q}_x$ zero.

For the defined time step Δ_t, the discrete time transition probability matrix $\mathbf{P}_x$ is obtained by taking the matrix exponential of the intensity matrix ($\mathbf{Q}_x$) multipled by the time step (Δ_t):

$$\mathbf{P}_x = e^{\mathbf{Q}_x \Delta t}$$

We can then obtain $\mathbf{U}_x$ and $\mathbf{M}_x$ based on:

$$\mathbf{P}_x = \left(\begin{array}{c|c} \mathbf{U}_x & \mathbf{0} \\ \hline \mathbf{M}_x & \mathbf{0} \end{array}\right)$$

In addition, the matrix $\mathbf{D}_j$ defines age advancement in the Markov chain. Using the example from Caswell & van Daalen (2021), if $\omega = 3$ then

$$\mathbf{D}_j = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & [1] \end{pmatrix} \quad j = 1, \ldots, \tau$$

In our implementation, we include the (optional) 1 value in the lower right corner; this assumes that after the last specified age, the cohort continues to experience transitions among health states according to the transition probabilities defined for the last age class. If this value is 0, the model will assume that everyone dies after the last cycle.

We next combine the transition matrices (for all age classes) together into a series of block-structured matrices as follows:

$$\mathbb{U} = \left(\begin{array}{c|c|c} \mathbf{U}_1 & \cdots & \mathbf{0} \\ \hline & \ddots & \\ \hline \mathbf{0} & \cdots & \mathbf{U}_\omega \end{array}\right)$$

$$\mathbb{D} = \left(\begin{array}{c|c|c} \mathbf{D}_1 & \cdots & \mathbf{0} \\ \hline & \ddots & \\ \hline \mathbf{0} & \cdots & \mathbf{D}_\tau \end{array}\right)$$

$$\tilde{\mathbf{U}} = \mathbf{K}^\top \mathbb{D} \mathbf{K} \mathbb{U} \quad \tau\omega \times \tau\omega$$

[10] The only difference with the two approaches above is that the rows in these rate matrices correspond to the final state, while the columns correspond to the starting state; this is the opposite of the rate matrices defined above, where the rows corresponded to the starting health state and the columns to the ending health state in a given cycle.

where $\mathbf{K}$ is a permutation matrix known as the vec-permutation matrix.[11] See Henderson & Searle (1981) and Appendix B in Caswell & van Daalen (2021) for further information.

We also define

$$\tilde{\mathbf{M}} = \begin{pmatrix} \mathbf{M}_1 & \cdots & \mathbf{M}_\omega \end{pmatrix} \quad \alpha \times \tau\omega$$

and

$$\tilde{\mathbf{P}} = \left(\begin{array}{c|c} \tilde{\mathbf{U}} & \mathbf{0}_{\tau\omega\times\alpha} \\ \hline \tilde{\mathbf{M}} & \mathbf{I}_{\alpha\times\alpha} \end{array}\right) \quad (\tau\omega+\alpha)\times(\tau\omega+\alpha)$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{0}$ is a matrix of zeros.

We now (nearly) have the components needed to calculate outcomes. A key difference in the Healthy Longevity approach, relative to the approaches above, is that we do not calculate outcomes separately in each cycle, and then sum them. Rather, the method utilizes matrix calculus to solve for *expected outcomes* and other moments in the outcome distribution (e.g., variance, skewness, etc.).[12]

Source: Article Notebook

0.5 Outcomes

With $\mathbf{P}$ defined under either Approach 1 or Approach 2, we now have the necessary ingredients to construct a Markov trace. Define $\mathbf{s}_0$ as the initial state occupancy vector at time $t = 0$. The vector $\mathbf{s}_0$ has size k, where k is the total number of states captured in the $k \times k$ matrix $\mathbf{P}$ (including transition states, if using Approach 2). This vector summarizes the number or fraction of the population in each health state at baseline. Often, this vector will be set such that the entire cohort starts off healthy—though this need not always be the case.

For a time-homogenous model such as considered here, health state occupancy at cycle t is is calculated as:

$$\mathbf{s}_t' = \mathbf{s}_0'\mathbf{P}^t \tag{7}$$

For a time-inhomongenous model in which transition probabilities change over time (e.g., death rates increase due to aging), we must construct separate transition probability matrices for each cycle (i.e., $\mathbf{P}(t)$). State occupancy at cycle t is calculated by sequentially applying the transition matrices corresponding to each time step leading up to cycle t, i.e.,

$$\mathbf{s}_t' = \mathbf{s}_0'\mathbf{P}(1)\mathbf{P}(2)\ldots\mathbf{P}(t) \tag{8}$$

0.5.1 Years of Life Lived with Disability (YLD)

To calculate YLDs, we define a disability weight payoff vector $\mathbf{d}_{YLD}$,

[11] A function to construct a vec-permultation matrix is provided within the code snippet below.

[12] Again, for our purposes here we will focus on expected outcomes—though note that formulas for higher-order moments are provided in Caswell & van Daalen (2021) and Caswell & Zarulli (2018).

$$\mathbf{d}_{YLD} = \begin{array}{r} \text{H} \\ \text{S1} \\ \text{S2} \\ \text{DOC} \\ \text{DS} \end{array} \begin{pmatrix} 0 \\ \texttt{dwS1}\frac{1}{r_{\Delta_t}}(1-e^{-r_{\Delta_t}})\Delta_t \\ \texttt{dwS2}\frac{1}{r_{\Delta_t}}(1-e^{-r_{\Delta_t}})\Delta_t \\ 0 \\ 0 \end{pmatrix}$$

Figure 8: YLD Payoff Vector

where `dwS1` and `dwS2` are the disability weights for the Sick and Sicker states, respectively. In addition, r_{Δ_t} is the cycle discount rate, which is calculated as,

$$r_{\Delta_t} = r\Delta_t \tag{9}$$

where r is the annual discount rate and Δ_t is the time step.

In the YLD payoff vector, the term $\frac{1}{r_{\Delta_t}}(1 - e^{-r_{\Delta_t}})$ is included as a continuous-time discounting adjustment factor for the defined cycle length Δ_t. This term is included to discount time *within* each cycle in order to maintain the continuous-time discounting approach used in the original GBD equations.[13]

To fully discount outcomes, we still must discount all future outcome values back to baseline ($t = 0$). For a time-homogeneous model, discounted years of life lost to disability (YLD) at cycle t is given by

$$YLD(t) = \mathbf{s}_0'\mathbf{P}^t\mathbf{d}_{YLD} \times e^{-r_{\Delta_t}t} \tag{10}$$

For a time-inhomogenous model, YLDs are calculated as

$$YLD(t) = \mathbf{s}_0'\mathbf{P}(1)\mathbf{P}(2)\ldots\mathbf{P}(t)\mathbf{d}_{YLD} \times e^{-r_{\Delta_t}t} \tag{11}$$

0.5.2 Years of Life Lost to Disease (YLLs)

As noted in Section **??** and in Equation **??**, YLLs are based on the present value of remaining life expectancy among deaths that occur in each cycle t. Define $a(t)$ as the age of the cohort at cycle t, i.e., $a(t) = t \cdot \Delta t + a0$, where a_0 is the age of the cohort at $t = 0$. Define $e(t) = e(a(t))$ as the present value of remaining life expectancy at cycle t. Following the GBD continuous time discounting approach, $e(a(t))$ is given by

$$e(a(t)) = \frac{1}{r}(1 - e^{-rEx(a(t))}) \tag{12}$$

where $Ex(a)$ is the remaining life expectancy at age a. $Ex(a)$ is drawn from either an exogenous (reference) or an endogenous life table, depending on the objectives of the modeling exercise (Anand & Reddy, 2019).

We next define a remaining life expectancy payoff vector at cycle t:

[13] For example, if $r = 0.03$ and $\Delta_t = 1$ (i.e., annual cycle length), this adjustment factor will be $0.985 = \frac{1}{r_{\Delta_t}}(1 - e^{-r_{\Delta_t}})$. If $\Delta_t = 1/12$ (monthly cycle) the cycle discounting adjustment factor is 0.99875.

$$\mathbf{d}_{YLL} = \begin{matrix} \text{H} \\ \text{S1} \\ \text{S2} \\ \text{DOC} \\ \text{DS} \end{matrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{r}\left(1 - e^{-rEx(a(t))}\right) \end{pmatrix}$$

Figure 9: YLL Payoff Vector

This payoff vector reflects discounting, but only in terms of the present value of remaining life expectancy *at time t*.

$$YLL(t) = \mathbf{s}_t' \mathbf{e}(t) \times e^{-r\Delta_t t} = \mathbf{s}_0' \mathbf{P}^t \mathbf{e}(t) \times e^{-r\Delta_t t} \tag{13}$$

Total discounted YLLs at time $t = 0$ is given by:

$$YLL = \sum_{t=0}^{N-1} YLL(t) = \sum_{t=0}^{N-1} \left(\mathbf{s}_0' \mathbf{P}^t \mathbf{e}(t) \times e^{-r\Delta_t t}\right) \tag{14}$$

Source: Article Notebook

Source: Article Notebook

To calculate outcomes, we must next define "reward" matrices $\mathbf{R}_m$, where m indexes the moment of interest (e.g., expected value, variance, etc.). The structure and values of $\mathbf{R}_m$ will differ, however, depending on the outcome.

To facilitate how we define rewards (i.e., payoffs), we briefly classify each of our outcomes (LE, YLL, YLD, DALYs, QALYs) into broad classes corresponding to whether the payoff or "reward" applies to occupancy, or to transitions, within or to a given health state:

Table 1: Classification of Health Outcomes

Outcome	Reward Class
Life Expectancy	Occupancy (1.0 for each alive health state)
Years of life lived with disability (YLD)	Occupancy (disability weight applied to time in CVD state)
Yearls of life lost to disease (YLL)	Transition (remaining life expectancy applied to CVD death transitions)
Disability-Adjusted Life Years (DALYs)	Occupancy (YLD) + Transition (YLL)
Quality-Adjusted Life Years	Occupancy (utility weights applied to living health states)

0.5.3 Occupancy-Based Outcomes

To calculate occupancy-based outcomes, we start with a reward matrix $\mathbf{H}$, which has dimensions $\tau \times \omega$ and is structured as shown in Figure **??**: