Modeling DALYs in Excel

Introduction

The objective of this notebook is to walk readers through the process of modeling Disability-Adjusted Life Years (DALYs) in Microsoft Excel. We detail both (beginner, intermediate) approaches outlined in the tutorial manuscript.

The model and results can all be found in the accompanying Excel document `sick-sicker-DALY.xlsm`.

> **Note**
>
> Outcomes and results in the Excel document are slightly different than in the tutorial manuscript. A primary reason why is that the R code used for the main results interpolates (by age) the reference life table values for YLL calculations, while the Excel-based results rely on the `LOOKUP()` function, which does not interpolate.
>
> For example, the Global Burden of Deases Reference Life Table lists remaining life expectancy of 88.0 years for a 1 year old and 84.0 for a 5 year old. Thus, our R code would assign 3 year olds a remaining life expectancy value between 84.0 and 88.0, while Excel assigns a value of 88.0 (the value at 1 years old, which is the next-closest value that does not exceed 3 years of age).
>
> A second reason why results may differ slightly is that Excel does not have the capability to perform matrix exponentiation to embed the transition rate matrix into the defined discrete time step. Consequently, as discussed below we rely on a power series expansion of the rate matrix to embed the transition probability matrix. In practice, this results in transition probabilities that are very close to, but do not exactly match, what we obtain via matrix exponentiation in R using the `expm()` command.

Step 1: Define Parameters

The worksheet **Parameters** is used to specify parameter names, values and descriptions. Note that the worksheet is divided into two (color-coded) sections: (1) parameters that are scalar inputs are shown in blue; while (2) those that are calculated as functions of other parameters are shown in red.

Parameter Name	Base Case Value	Description
n_cohort	1	Cohort size
horizon	500	Time horizon (years)
Delta_t	1	Time step (cycle length)
age0	25	Cohort starting age
r_disc_h	0.03	Annual discount rate (health outcomes)
r_disc_c	0.03	Annual discount rate (cost outcomes)
r_HS1	0.15	Rate: Healthy to sick (S1)
r_S1H	0.5	Rate: Recovery (Sick [S1] to Healthy [H])
r_S1S2	0.105	Rate: Sick (S1) to sicker (S2)
r_HD	0.002	Rate: Background mortality
hr_S1	3	Hazard Ratio: Sick (S1) to Death (D)
hr_S2	10	Hazard Ratio: Sicker (S2) to Death (D)
dw_S1	0.25	Disability weight: Sick (S1)
dw_S2	0.5	Disability weight: Sicker (S2)
c_H	2000	Annual Cost: Healthy (H)
c_S1	4000	Annual Cost: Sick (S1)
c_S2	15000	Annual Cost: Sicker (S2)
c_D	0	Annual Cost: Dead (D)
c_trtA	12000	Annual Cost: Strategy A
dw_trtA	0.05	Disbility weight: Sick (S1) under strategy A
c_trtB	13000	Annual Cost: Strategy B
hr_S1S2_trtB	0.6	Hazard Ratio: Sick (S1) to Sicker (S2) under strategy B
r_H	0.03	Cycle discount rate (health outcomes)
r_C	0.03	Cycle discount rate (cost outcomes)

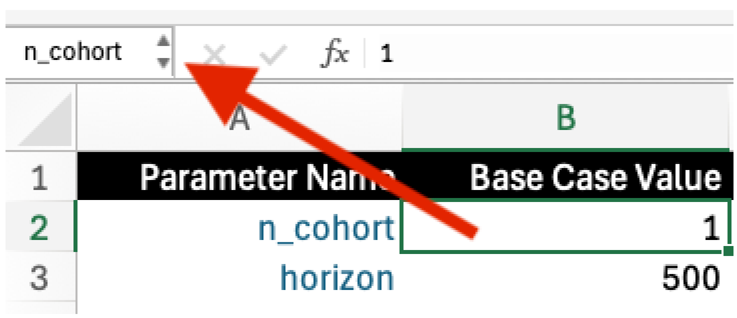
Scalar parameters in blue

Calculated parameters in red

Parameters Tab

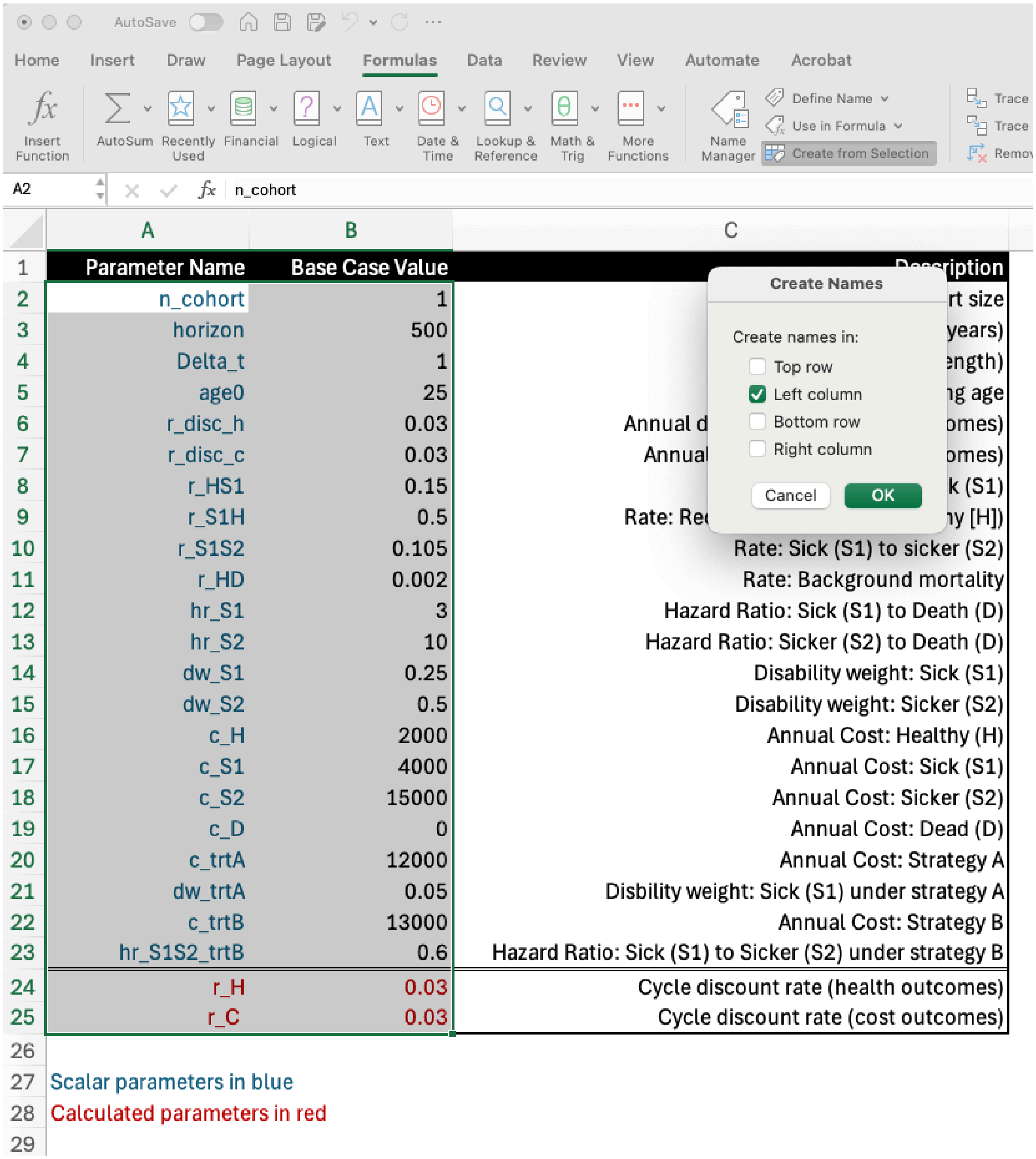
Naming Variables

Throughout the Excel document we rely on parameter and matrix names rather than cell definitions (e.g., “r_HD” is used in lieu of “Parameters$B11”). To name variables, one can simply type in the preferred name into the upper left box (shown below for the parameter “n_cohort”).



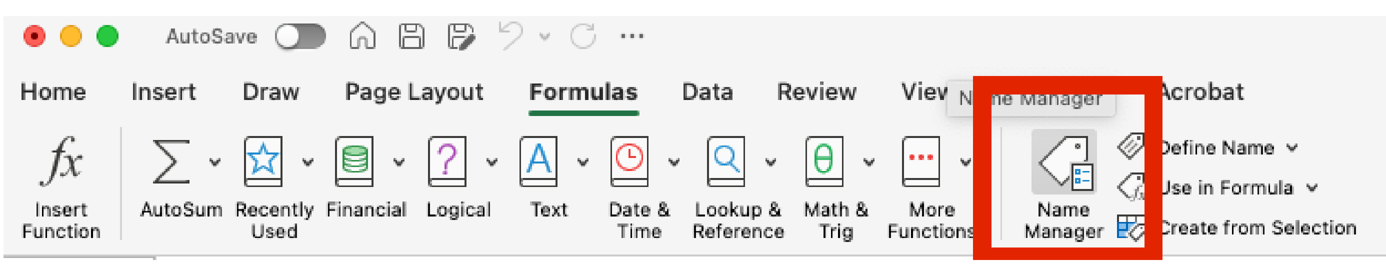
A similar approach will be used later to name matrices—but instead of selecting a single cell to name, we will select a region of cells to assign a single name to .

An alternative (more efficient) approach may be to batch name parameters all at once. To do this, select the region with parameters and values to be named, and then click “Create from Selection” in the ***Formulas*** tab. Make sure that the parameter names are to the left of the value for that parameter. Select “Left column” in the resulting popup box, and Excel will automatically assign parameter names to each value.



Naming parameters

Names can be managed (e.g., edited, deleted, etc.) by clicking on the “Name Manager” in the *Formulas* tab:



Step 2: Construct Transition Matrices

The Excel document also contains two worksheets for constructing transition probability matrices—one for each approach.

Beginner Approach

Transition probability matrices for Approach 1 (Beginner) are constructed in the Worksheet **Transition Matrices (App. 1).** Transition rates are inserted into a separate rate matrix for each strategy under the section titled “Transition Rate Matrices”:

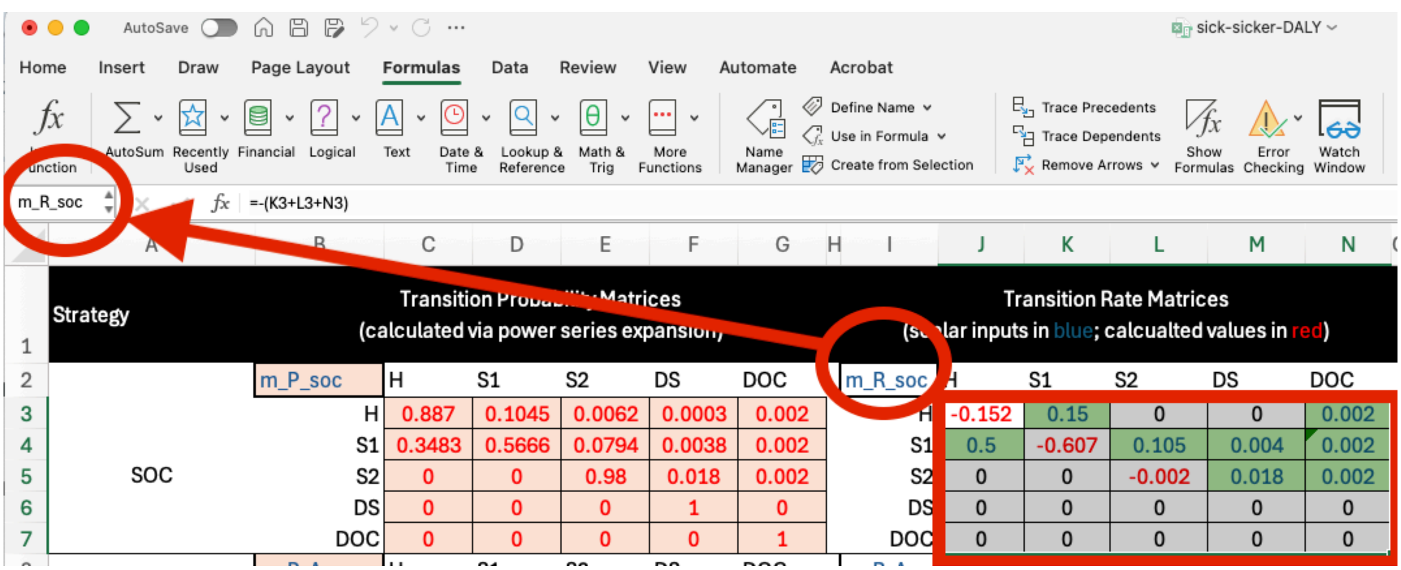
Transition Rate Matrices (scalar inputs in blue; calcualted values in red)

m_R_soc	H	S1	S2	DS	DOC
H	-0.152	0.15	0	0	0.002
S1	0.5	-0.607	0.105	0.004	0.002
S2	0	0	-0.002	0.018	0.002
DS	0	0	0	0	0
DOC	0	0	0	0	0

Transition rates

Again, direct inputs are shown in Blue text while calculated values are shown in Red. In the table above, the rates are inserted using their parameter names, while the diagonal elements are calculated as the negative sum of the off-diagonal elements in the same row.

Above and to the left of each matrix you will also find a matrix name. To define matrices, we select the cell values of the matrix and then input the name into the cell directly to the left of the formula bar:



To the right of the transition rate matrix is a series of matrices that are used in the power series expansion that helps create the transition probability matrix (shown in red in the image above). The formula used for the power series expansion is

$$e^{\mathbf{R}} = \sum_{k=0}^{4} \frac{\mathbf{R}^k}{k!}$$

Note that we only go up to $k = 1, \ldots, 4$ in the Excel document. Higher-order (e.g., $k = 1, \ldots, 10$) approaches do not materially change the transition probabilities—though practitioners may find additional expansions are necessary for their applications.

The far-most left panel shows the final transition probability matrices calculated via the power series expansion. Again, one can see the defined matrix names (one for each strategy) just above and to the left of each matrix.

Approach 2 (Intermediate)

Our intermediate approach relies on non-Markovian tracking states added to the "edges" of each matrix. The approach to embedding the transition probability matrix is essentially the same, with a few minor wrinkles which we discuss below.

The worksheet **Transition Matrices (App. 2)** details the transition matrices and calculations for Approach 2. As above, transition rates are entered directly (as parameter names) in the section titled "Transition Rate Matrices.

A key difference, however, is that the transition rate matrix is split into two parts: a Markovian submatrix and the Non-Markovian tracking states (which are highlighted in light blue):

Transition Rate Matrices
(scalar inputs in blue; calcualted values in red)

m2_R_soc	H	S1	S2	D	trDS
H	-0.152	0.15	0	0.002	0
S1	0.5	-0.611	0.105	0.006	0.004
S2	0	0	-0.02	0.02	0.018
D	0	0	0	0	0
trDS	0	0	0	0	0

It is critically important that the diagonal elements are set to sum to zero *only for the cells in the Markovian submatrix* (i.e., these values do not include the cells highlighted in blue). In our formulation, all transitions to death are recorded in the "D" column, while only transitions to death due to disease-related causes are captured in the "trDS" tracking state column. This is a slight difference from Approach 1, which treats disease-related and other-cause-related death transitions as separate (mutually exclusive) health states.

As under Approach 1, the transition probability matrix is calculated via a power series expansion. However, there is one manual change that must be made to the matrix. After embedding, the tracking state "trDS" will be treated as an absorbing state (i.e., the probability of transition from trDS to trDS is 1.0). We must manually change this value to 0 to make "trDs" a transition tracking state:

	B	C	D	E	F	G
	Transition Probability Matrices (calculated via power series expansion)					
	m2_P_soc	H	S1	S2	D	trDS
	H	0.8872	0.1043	0.0062	0.0023	0.0003
	S1	0.3476	0.5682	0.0784	0.0058	0.0038
	S2	0	0	0.9802	0.0198	0.0178
	DS	0	0	0	1	0
	DOC	0	0	0	0	0

Step 3: Outcomes

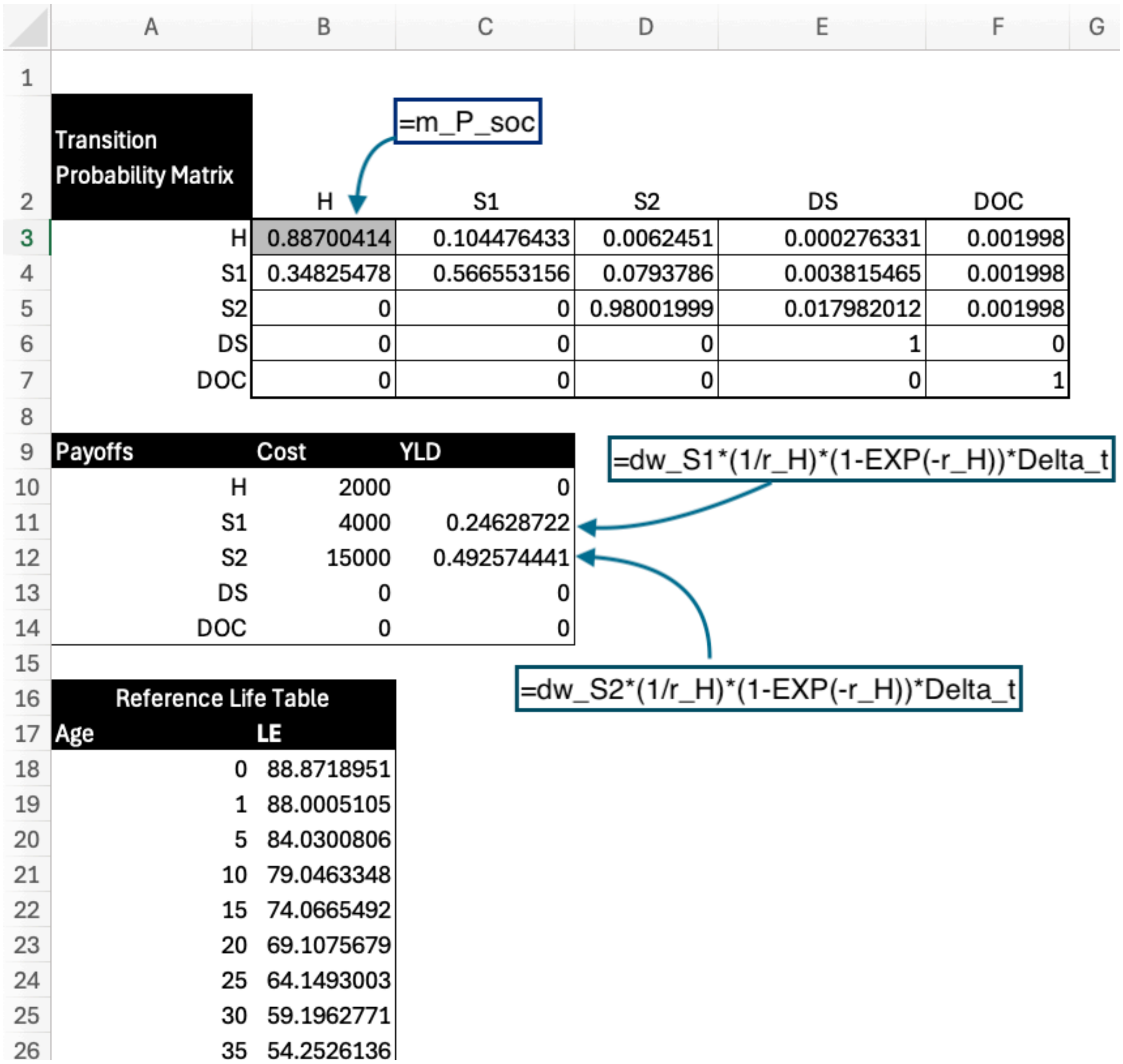
The Excel document contains a series of worksheets that work through outcome calculations under each strategy and methodological approach. For a given approach (beginner, intermediate), the worksheets for each strategy are essentially identical—the only differences are in the transition probability matrices and payoff vectors, which vary by strategy.

We will walk readers through a single strategy (“Standard of Care,” or “SOC”) as an example.

Input Objects

The leftmost side of the worksheet has a variety of input objects; these are merely summaries that refer to parameter and matrix names defined above. This area also includes the reference life table from the GBD to facilitate calculation of YLL outcomes.

Payoffs are defined using parameters and the formulas as outlined in the main manuscript. A few are highlighted in the image below:



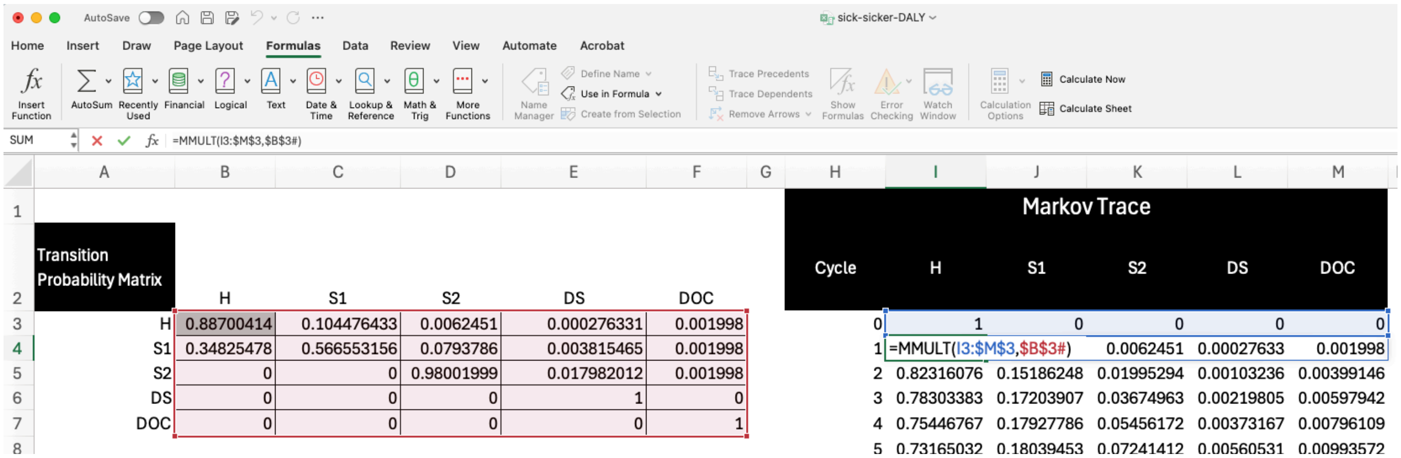
Transition Probability Matrix	H	S1	S2	DS	DOC
H	0.88700414	0.104476433	0.0062451	0.000276331	0.001998
S1	0.34825478	0.566553156	0.0793786	0.003815465	0.001998
S2	0	0	0.98001999	0.017982012	0.001998
DS	0	0	0	1	0
DOC	0	0	0	0	1

Payoffs	Cost	YLD
H	2000	0
S1	4000	0.24628722
S2	15000	0.492574441
DS	0	0
DOC	0	0

Reference Life Table	
Age	LE
0	88.8718951
1	88.0005105
5	84.0300806
10	79.0463348
15	74.0665492
20	69.1075679
25	64.1493003
30	59.1962771
35	54.2526136

Markov Trace

The next section over calculates the Markov trace for the modeled time horizon (500 years in this case). In cycle zero, we start with a healthy cohort (with size defined by the `n_cohort` parameter). The next row (cycle 1) is then calculated via matrix multiplication of the current state occupancy (i.e., the row for cycle=0) and the transition probability matrix):



=MMULT(I3:$M$3,$B$3#)

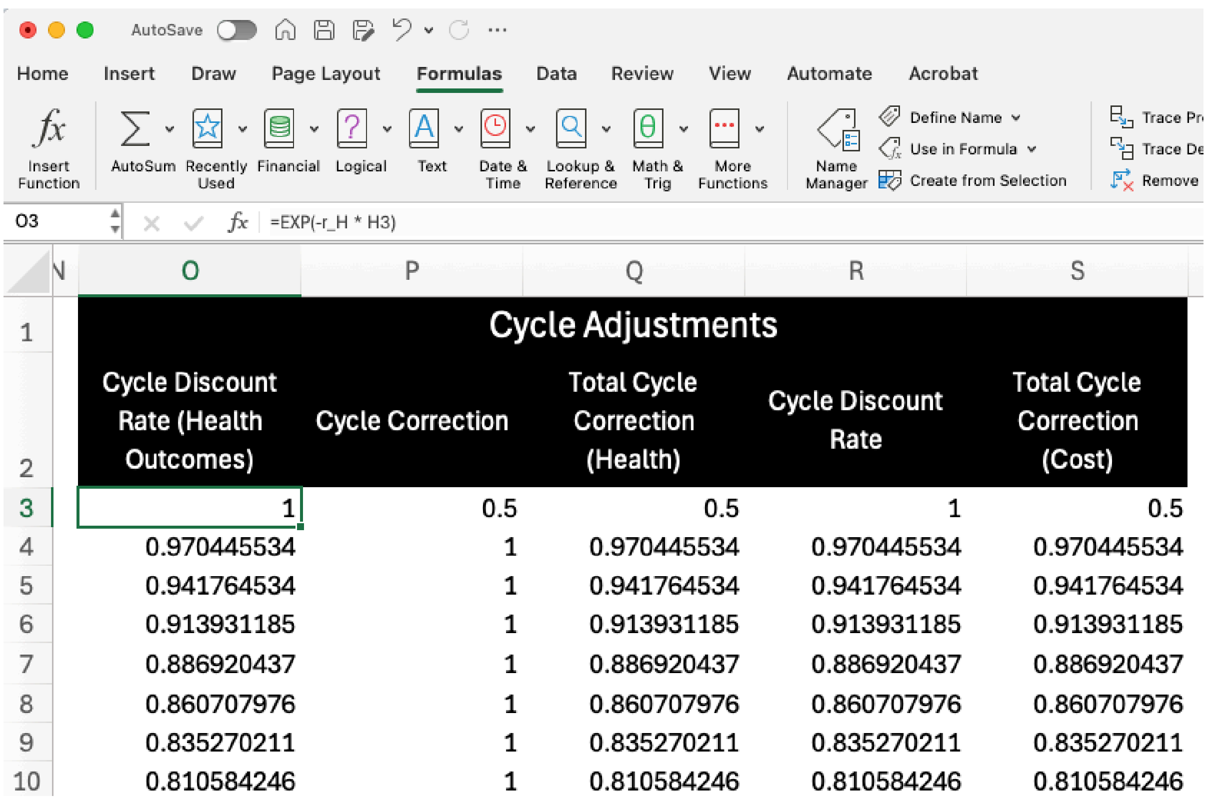
Transition Probability Matrix	H	S1	S2	DS	DOC
H	0.88700414	0.104476433	0.0062451	0.000276331	0.001998
S1	0.34825478	0.566553156	0.0793786	0.003815465	0.001998
S2	0	0	0.98001999	0.017982012	0.001998
DS	0	0	0	1	0
DOC	0	0	0	0	1

Markov Trace					
Cycle	H	S1	S2	DS	DOC
0	1	0	0	0	0
1	=MMULT(I3:$M$3,$B$3#)		0.0062451	0.00027633	0.001998
2	0.82316076	0.15186248	0.01995294	0.00103236	0.00399146
3	0.78303383	0.17203907	0.03674963	0.00219805	0.00597942
4	0.75446767	0.17927786	0.05456172	0.00373167	0.00796109
5	0.73165032	0.18039453	0.07241412	0.00560531	0.00993572

Each subsequent row is calculated using a similar process.

Cycle Adjustments

The next panel contains columns for cycle adjustments based on discounting and a half-cycle correction. We combine these together to construct a cycle adjustment factor (c_t in the manuscript) that will be used later.



=EXP(-r_H * H3)

Cycle Discount Rate (Health Outcomes)	Cycle Correction	Total Cycle Correction (Health)	Cycle Discount Rate	Total Cycle Correction (Cost)
1	0.5	0.5	1	0.5
0.970445534	1	0.970445534	0.970445534	0.970445534
0.941764534	1	0.941764534	0.941764534	0.941764534
0.913931185	1	0.913931185	0.913931185	0.913931185
0.886920437	1	0.886920437	0.886920437	0.886920437
0.860707976	1	0.860707976	0.860707976	0.860707976
0.835270211	1	0.835270211	0.835270211	0.835270211
0.810584246	1	0.810584246	0.810584246	0.810584246

Years of Life Lost to Premature Mortality (YLLs)

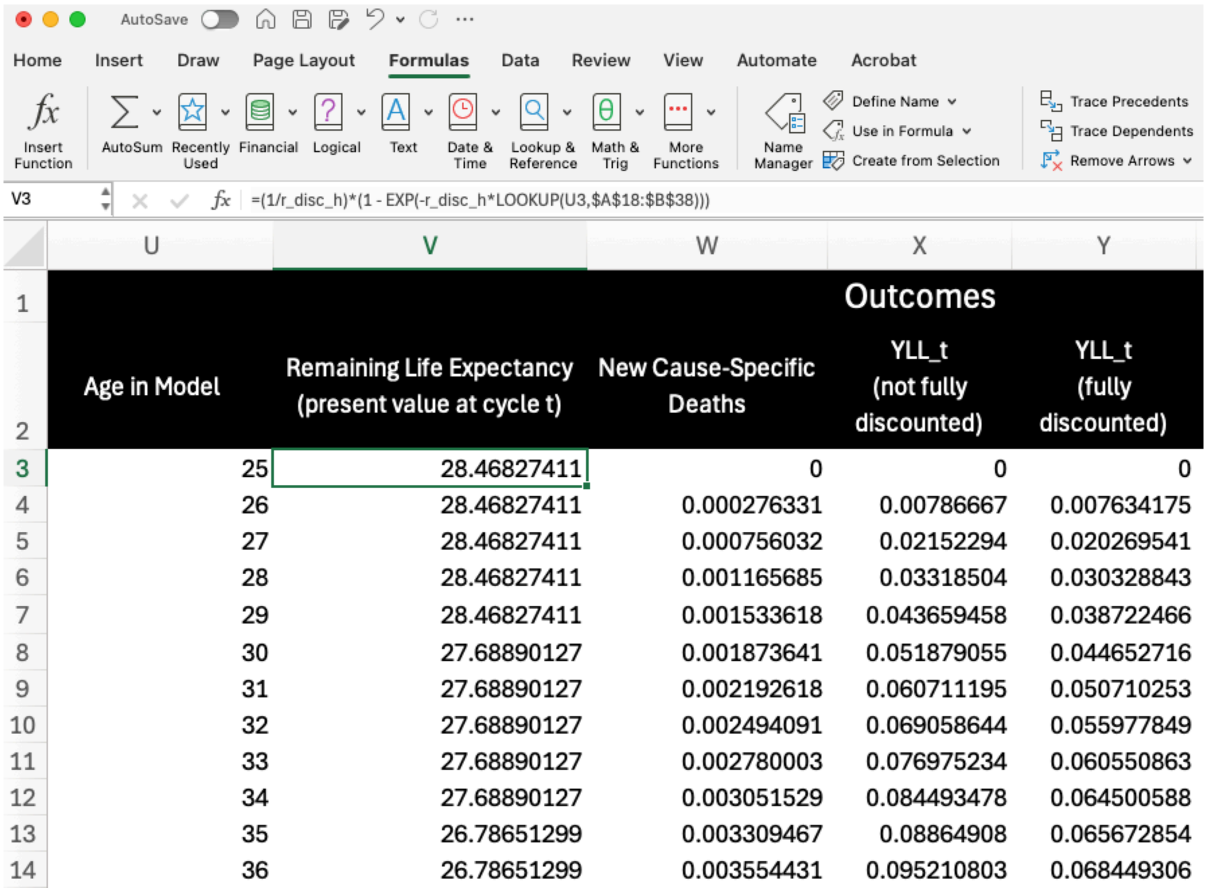
Columns U through Y provide the necessary numbers for YLL outcomes. We start by determining the age of the cohort in the model (a_t; Column U):

$$a_t = a_0 + t \cdot \Delta_t$$

where a_0 is the age of the cohort at $t = 0$.

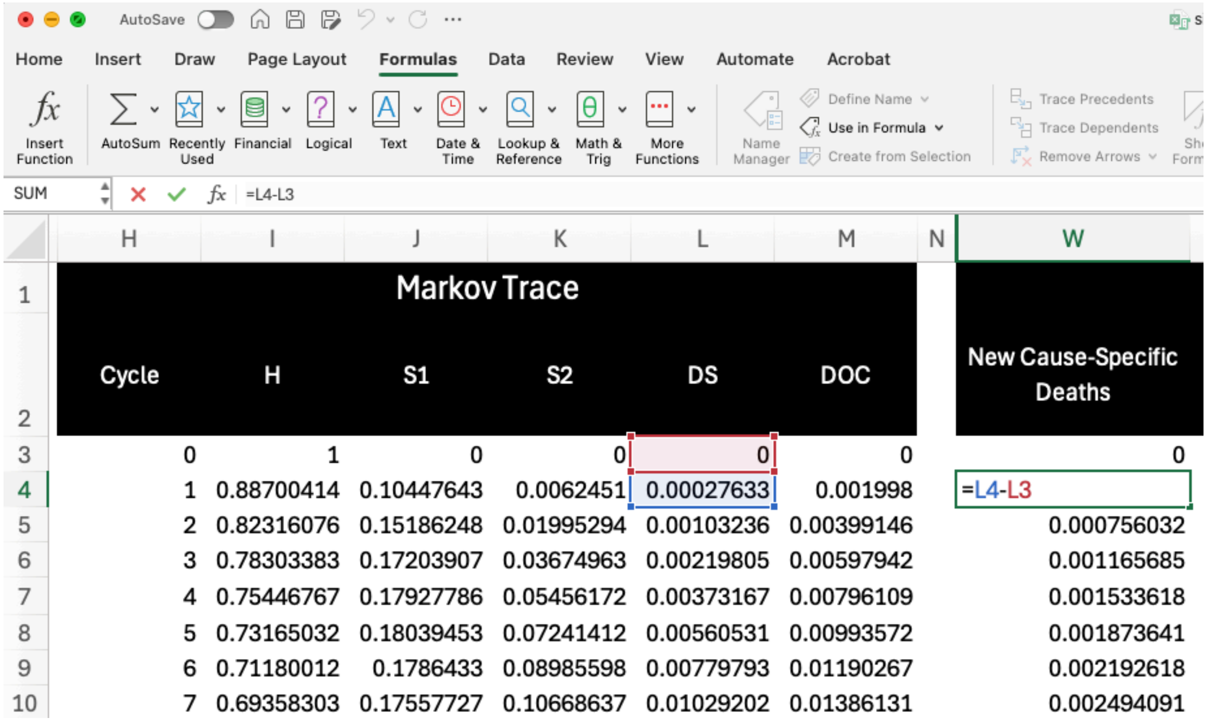
Column V then plugs this calculated age into a `LOOKUP()` function to obtain the value for remaining life expectancy from the reference life table ($Ex(a_t)$). As noted at the beginning of this document, `LOOKUP()` finds the value associated with closest age. The formulas then discount this value to the present value at the cycle using the following formula:

$$Ex_t = \frac{1}{r}\left(1 - e^{-rEx(a_t)}\right)$$



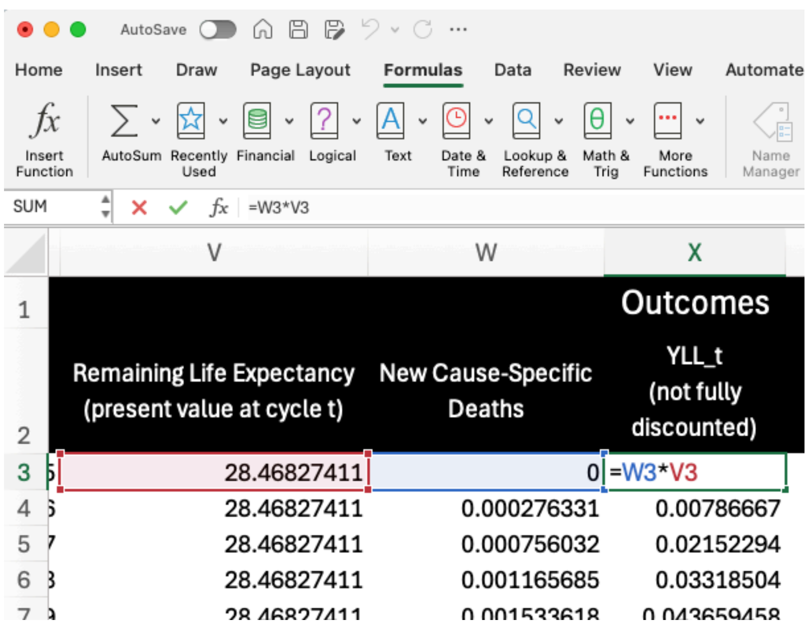
	U	V	W	X	Y
1				Outcomes	
2	Age in Model	Remaining Life Expectancy (present value at cycle t)	New Cause-Specific Deaths	YLL_t (not fully discounted)	YLL_t (fully discounted)
3	25	28.46827411	0	0	0
4	26	28.46827411	0.000276331	0.00786667	0.007634175
5	27	28.46827411	0.000756032	0.02152294	0.020269541
6	28	28.46827411	0.001165685	0.03318504	0.030328843
7	29	28.46827411	0.001533618	0.043659458	0.038722466
8	30	27.68890127	0.001873641	0.051879055	0.044652716
9	31	27.68890127	0.002192618	0.060711195	0.050710253
10	32	27.68890127	0.002494091	0.069058644	0.055977849
11	33	27.68890127	0.002780003	0.076975234	0.060550863
12	34	27.68890127	0.003051529	0.084493478	0.064500588
13	35	26.78651299	0.003309467	0.08864908	0.065672854
14	36	26.78651299	0.003554431	0.095210803	0.068449306

Next, Column W calculates the number of cause-specific deaths in the cycle based on the “DS” column in the Markov Trace:



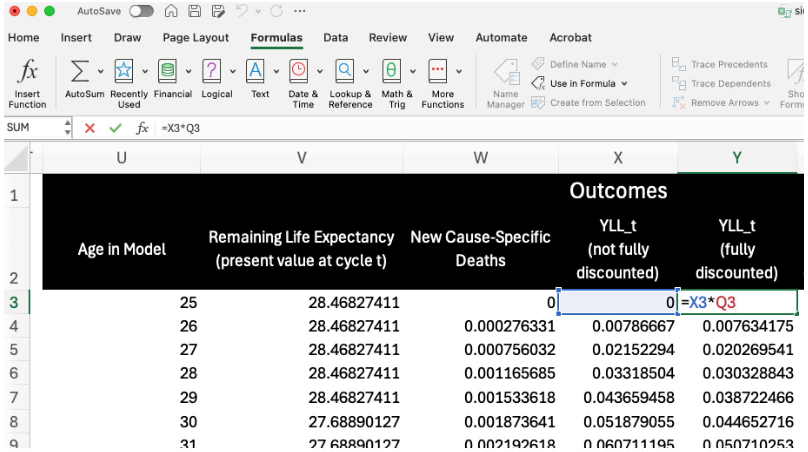
	Cycle (H)	H (I)	S1 (J)	S2 (K)	DS (L)	DOC (M)	N	New Cause-Specific Deaths (W)
1	Markov Trace							
2	Cycle	H	S1	S2	DS	DOC		New Cause-Specific Deaths
3	0	1	0	0	0	0		0
4	1	0.88700414	0.10447643	0.0062451	0.00027633	0.001998		=L4-L3
5	2	0.82316076	0.15186248	0.01995294	0.00103236	0.00399146		0.000756032
6	3	0.78303383	0.17203907	0.03674963	0.00219805	0.00597942		0.001165685
7	4	0.75446767	0.17927786	0.05456172	0.00373167	0.00796109		0.001533618
8	5	0.73165032	0.18039453	0.07241412	0.00560531	0.00993572		0.001873641
9	6	0.71180012	0.1786433	0.08985598	0.00779793	0.01190267		0.002192618
10	7	0.69358303	0.17557727	0.10668637	0.01029202	0.01386131		0.002494091

In Column L, we calculate the initial cycle-specific values for YLLs (`YLL_t`) by multiplying remaining life expectancy by the number of cause-specific deaths in the cycle.



	V	W	X
1			Outcomes
2	Remaining Life Expectancy (present value at cycle t)	New Cause-Specific Deaths	YLL_t (not fully discounted)
3	28.46827411	0	=W3*V3
4	28.46827411	0.000276331	0.00786667
5	28.46827411	0.000756032	0.02152294
6	28.46827411	0.001165685	0.03318504
7	28.46827411	0.001533618	0.043659458

Column Y then calculates the final cycle-specific YLL values by multiplying by the cycle adjustment factor for health outcomes (Column Q). Again, this cycle-adjustment reflects both discounting and the half-cycle correction:



SUM =X3*Q3

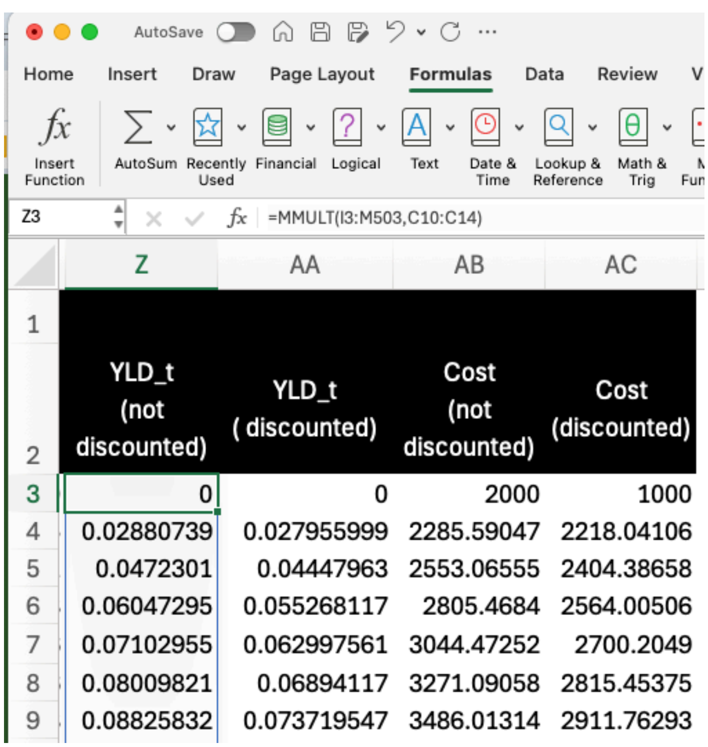
	U	V	W	X	Y
1				Outcomes	
2	Age in Model	Remaining Life Expectancy (present value at cycle t)	New Cause-Specific Deaths	YLL_t (not fully discounted)	YLL_t (fully discounted)
3	25	28.46827411	0	0	=X3*Q3
4	26	28.46827411	0.000276331	0.00786667	0.007634175
5	27	28.46827411	0.000756032	0.02152294	0.020269541
6	28	28.46827411	0.001165685	0.03318504	0.030328843
7	29	28.46827411	0.001533618	0.043659458	0.038722466
8	30	27.68890127	0.001873641	0.051879055	0.044652716
9	31	27.68890127	0.002192618	0.060711195	0.050710253

We then sum up the cycle specific YLL values to obtain expected YLLs over the model time horizon. This value is stored in the "Outcomes" table on the left-hand side of the worksheet. We also assign this value a name ("YLL_SOC") using the parameter naming methods outlined above; this is useful later for creating summary tables.

OUTCOMES	
YLL_SOC	2.8651

Years of Life Lost to Disability and Costs

YLDs and costs are more straightforward to calculate as they require only matrix multiplication of the markov trace by the respective payoff vectors. We first calculate undiscounted and unadjusted cycle-specific values using matrix multiplication of the trace and the payoff vector. We then apply the health- and cost-specific cycle adjustment values (Columns Q and S) to obtain the final adjusted cycle-specific values for YLDs and Costs, and then sum up each column for the outcomes table.



Z3 =MMULT(I3:M503,C10:C14)

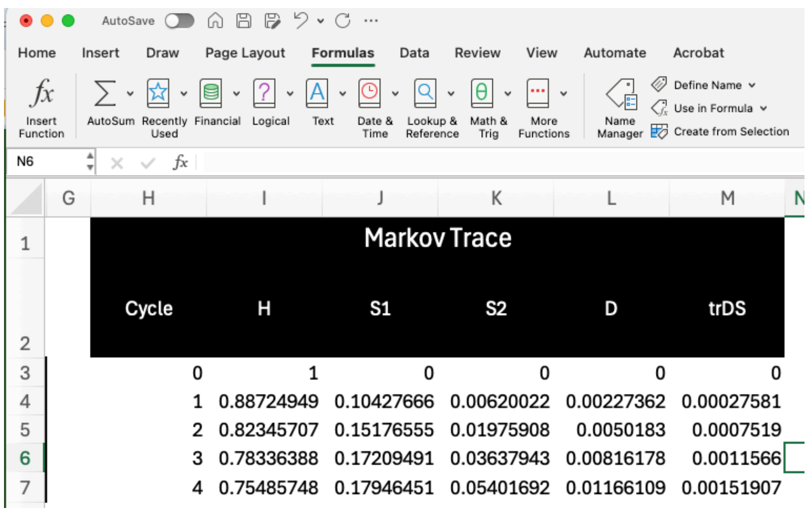
	Z	AA	AB	AC
1				
2	YLD_t (not discounted)	YLD_t (discounted)	Cost (not discounted)	Cost (discounted)
3	0	0	2000	1000
4	0.02880739	0.027955999	2285.59047	2218.04106
5	0.0472301	0.04447963	2553.06555	2404.38658
6	0.06047295	0.055268117	2805.4684	2564.00506
7	0.07102955	0.062997561	3044.47252	2700.2049
8	0.08009821	0.06894117	3271.09058	2815.45375
9	0.08825832	0.073719547	3486.01314	2911.76293

OUTCOMES		
YLL_SOC		2.8651
YLD_SOC		4.6128
DALY_SOC		7.4780
COST_SOC	$	155,458

Finally, expected DALYs are calculated as the sum of YLL and YLD outcomes.

Outcomes Under Approach 2

Expected outcomes are calculated for Approach 2 using a nearly identical process as above. The only difference is that the Markov trace itself already contains a count of new disease-related deaths in each cycle (in the “trDS” column); it does not need to be separately calculated, as we did in Column W above.



Markov Trace					
Cycle	H	S1	S2	D	trDS
0	1	0	0	0	0
1	0.88724949	0.10427666	0.00620022	0.00227362	0.00027581
2	0.82345707	0.15176555	0.01975908	0.0050183	0.0007519
3	0.78336388	0.17209491	0.03637943	0.00816178	0.0011566
4	0.75485748	0.17946451	0.05401692	0.01166109	0.00151907

Summary and Cost-Effectiveness Analysis

Finally, the worksheet **Summary and CEA** draws on the named expected outcome objects to summarize YLLs, YLDs, and DALYs under each approach and strategy. These values, along with the expected cost outcome, are then used to construct the CEA table.

Our results here closely match those in the tutorial manuscript, though are not identical due to the reasons outlined at the beginning of this document.

Table 1. Expected Health Outcomes, by Strategy and Approach

Strategy	Years Living with Disease (YLDs)	Years of Life Lost to Premature Mortality (YLLs)	Disability-Adjusted Life Years (DALYs)
Approach 1 (beginner)			
Standard of Care	4.613	2.865	7.478
Strategy A: Quality of Life Improvement	3.910	2.865	6.775
Strategy B: Reduce Disease Progression	3.826	2.180	6.006
Composite: Strategy A + Strategy B	2.963	2.180	5.143
Approach 2 (intermediate)			
Standard of Care	4.608	2.833	7.440
Strategy A: Quality of Life Improvement	3.901	2.833	6.734
Strategy B: Reduce Disease Progression	3.819	2.155	5.974
Composite: Strategy A + Strategy B	2.953	2.155	5.108

Table 2. Cost-Effectiveness Analysis

Outcome and Strategy	Approach 1 (beginner)	Approach 2 (intermediate)
Cost		
Standard of Care	$ 155,458	$ 155,349
Strategy B: Reduce Disease Progression	$ 261,186	$ 260,968
Composite: Strategy A + Strategy B	$ 379,387	$ 379,109
Strategy A: Quality of Life Improvement	$ 287,961	$ 287,837
DALY		
Standard of Care	7.478	7.440
Strategy B: Reduce Disease Progression	6.006	5.974
Composite: Strategy A + Strategy B	5.143	5.108
Strategy A: Quality of Life Improvement	6.775	6.734
ICER ($/DALY Averted)		
Standard of Care	ref.	ref.
Strategy B: Reduce Disease Progression	$ 71,812	$ 72,049
Composite: Strategy A + Strategy B	$ 136,953	$ 136,398
Strategy A: Quality of Life Improvement	Dominated	Dominated