Stochastic trees and the *StoTree* modeling software

StoTree is software tool for the formulation and solution of continuous-time Markov models of medical interventions. Markov chain models were introduced to the medical literature by Beck and Pauker [1] and provide a convenient means to account for medical treatment options and risks that occur not only in the present but also in the near and distant future. For a more recent introduction to Markov models in medicine, see [2] or [3]. *StoTree* represents Markov models using the graphic technique of *stochastic trees*, introduced by Hazen [4-6].

StoTree is implemented in the Excel spreadsheet environment. It is a Microsoft Excel addin for Office 2007, 2010 and 2013 or later that occupies its own tab on the Office ribbon. It is written in VBA for Office and therefore requires an Office installation on a PC. I do not know whether it will work on a MacIntosh computer running one of these versions of Office – in principle it should as long as VBA for Excel is available.

This document contains three major sections. The first section *Stochastic trees* gives an introduction to stochastic trees. This is the place to start if you are a beginner. Those familiar with stochastic trees may skip to the second section *Using StoTree to construct a stochastic tree*. The third major section *Factored stochastic trees* discusses factoring, a powerful tool that is one of the unique features of *StoTree*. In its final subsection *The benefits of factoring,* we describe the advantages of using factoring in model construction.

Installing StoTree

You can download a copy of *StoTree* from this website. It should appear as the file StoTree7?.xlam on your computer, where ? is a single letter indicating the version of StoTree7 (currently version *v*). Store it in any directory you like, but note where you have stored it. To activate it in Excel, follow the path File\Options\Addins, then click on Go… and Browse… to find the location at which you stored *StoTree*. (Alternately, you may follow the path Developer\Excel Addins if you have made the Developer tab visible on the Excel ribbon.) Click on the appropriate buttons to complete the installation.

The default path for Office addins is c:\Users\[your identifier]\Appdata\Roaming\Microsoft\AddIns. Although it is not required, you may wish to store StoTree7?.xlam at this location, as it is the default place Excel looks for addins.

Contents

StoTree v.7

June 2023

Contents

Stochastic trees

For those familiar with stochastic or decision-analytic modeling, a stochastic tree can be characterized in several equivalent ways:

- as a continuous-time Markov chain transition diagram with chance and decision nodes added,
- as a decision tree with stochastic transitions added,
- as a continuous-time version of a Markov cycle tree.

Persistent versus instantaneous states

A simple stochastic tree is shown in Figure 1[1], which depicts the risk of pulmonary embolism over time. It consists of *nodes* (*No Embolism, Embolism, Dead*) and arrows connecting the nodes. Nodes depict health states. The arrows are of two types, either *wavy* or *straight*. A node or health state can be either *persistent* (occupied continuously for some positive duration) or *instantaneous* (occupied for vanishingly small duration), depending on the type of arrows emanating from it. Wavy arrows emanate from an persistent state (such as *No Embolism*) and indicate that the persistent state is occupied for a duration that is uncertain but dependent on the *rate* at which exit from that state occurs. In Figure 1, the exit rate from *No Embolism* is *rPE*, and it labels the wavy arrow emanating from that health state. The units of the rate *rPE* is events per year (or per month or day or whatever the unit of time is). In *StoTree*, persistent states are known as *stochastic nodes*.
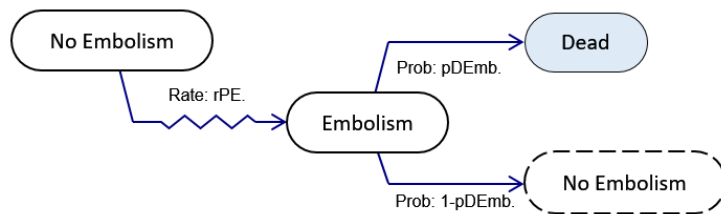

Figure 1. A simple stochastic tree, depicting the risk of pulmonary embolism over time.

It is possible for a node to have *no* arrows emanating from it. *StoTree* interprets such nodes to be persistent states – they are occupied from the time they are entered until the end of the time horizon. For instance, in Figure 2 below, the nodes *Cured, Not Cured,* and *Survive Rs* are persistent states.

Straight arrows emanate from an instantaneous state (such as *Embolism*), and point to the possible new states immediately following. Each straight arrow is labeled with a *probability* representing the chance that exit from that state occurs along that arrow. Exit from that state occurs immediately – the state is occupied for zero duration. (This is obviously an idealization – any health state must be occupied for a positive amount of time. However, zero duration can be a convenient modeling assumption for short-duration states.) Probabilities exiting instantaneous states must sum to 1. In *StoTree*, instantaneous states are known as *chance nodes*.

In Figure 1, exit from the instantaneous state *Embolism* leads to *Dead* with probability *pDEmb*, and back to *No Embolism* with probability 1−*pDEmb*. The dashed border around the second *No Embolism* node means that it is a *copy* of the first node and that transition occurs back to the first node. Copy nodes are a stochastic tree convention meant to facilitate tree-like display of a Markov model. They are optional – one could equally well draw a (straight) arrow in this case from *Embolism* back to *No Embolism.*

---

[1] All stochastic tree diagrams in this document are screen captures from *StoTree*.

Overall, this stochastic tree depicts a patient subject to repeated pulmonary embolisms with a chance of death after each embolism.

A second stochastic tree is depicted in Figure 2. Here the tree begins with a *square* node 0 depicting a decision that must be made initially. Choice must be made between *Liver-Directed Therapy, Living Donor Transplant,* and *Resection* (i.e. surgical removal of a tumor). Straight arrows emanate from node 0 with probabilities 1,0,0 summing to one like with chance nodes – but here each probability must be 1 or 0 to correspond to a decision chosen (or declined) with certainty. In Figure 2, the decision chosen is *Liver-Directed Therapy*. Decision nodes are also instantaneous, are therefore occupied for zero duration, and are exited immediately.
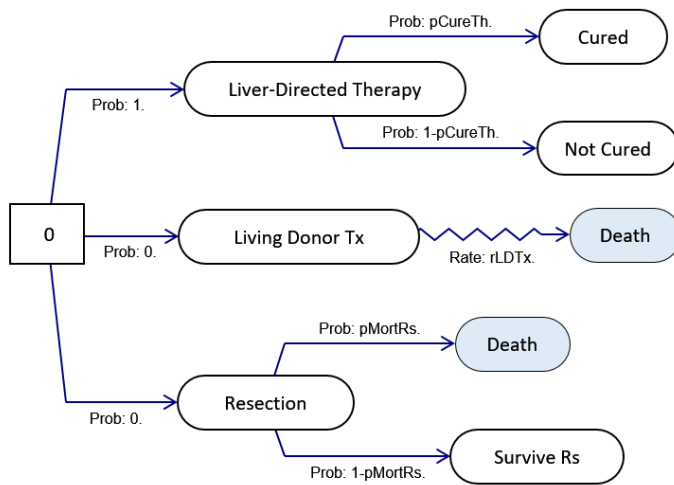
Figure 2. A stochastic tree with initial decision node.

Figure 3 shows a third stochastic tree, modeling HIV progression. Here all nodes are stochastic, and two stochastic transitions compete at both node A and node B.
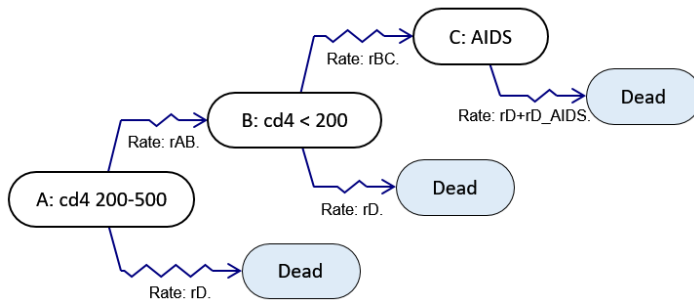
Figure 3. A stochastic tree with competing stochastic transitions.

Continuous time versus discrete time

Stochastic trees are formulated in *continuous time*. Alternate graphical depictions for Markov models arise and are commonly used in the medical literature. In these, time is *discretized* into steps 0, Δs, $2\Delta s$, $3\Delta s$, … of length Δs. Here, Δs is termed the *cycle length*. Rates r are converted into probabilities p via the formula $p = 1 - \exp(-r{\cdot}\Delta s)$. (For more on this conversion, see the next section.)

For instance, Figure 4 portrays the Markov model of Figure 1 converted to discrete time. The fact that state *No Embolism* can persist over time is indicated by the self-loop with label $1 - pPE$. The fact that the state *Embolism* is instantaneous is indicated by the lack of any such self-loop. In particular, this state must be departed for either *No Embolism* or *Dead* after only one cycle.

Figure 4. A discrete-time portrayal of the Markov model of *Figure 1*. The rate *rPE* of *Figure 1* has been converted into a transition probability $pPE = 1 - \exp(-rPE \cdot \Delta s)$, where Δs is the cycle length. The loop with label $1 - pPE$ on the *No Embolism* state means that the state can be occupied over multiple successive cycles. The lack of any such loop at the *Embolism* state means that once entered it must be departed immediately, that is, after only one cycle.

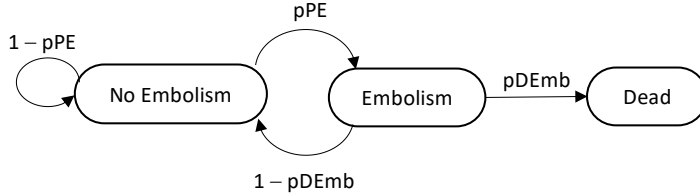We provide Figure 4 in case the reader may be more familiar with graphical conventions traditional in the medical literature. However, the remainder of this documentation uses the stochastic tree conventions of Figure 1.

Top

Rates versus probabilities

The distinction between a rate and a probability may be made as follows. Consider an event (such as *pulmonary embolism* in Figure 1, or *death* for living donor transplant recipients in Figure 2) that can occur to an individual once or repeatedly as long as he or she is at risk for that event. A *rate* satisfies

> *rate = event count per person per unit time at risk.*

So if we observe k events among multiple subjects who are at risk for that event for a total of $n\Delta t$ person-years[2], then the event rate r is estimated by

$$r = \frac{k}{n\Delta t}.$$

Rates have units of time^{-1}. For example, we might observe $k = 250$ events in a total of $n\Delta t = 200$ person-years at risk. This would yield an estimated rate of $250/200 = 1.25$ events per person per year at risk. As you can see, a rate must be nonnegative, but can exceed one.

Rates are sometimes referred to as *hazard rates*, and under this name assume a key role in survival analysis and in standard statistical models, such as the proportional hazards model [7]. Stochastic trees have the conceptual advantage of speaking directly in the language of rates.

On the other hand, when time is not an issue, and we are only interested in whether or not a given event occurs (e.g, death given pulmonary embolism in Figure 1, or cure given liver-directed therapy in Figure 2), then the probability of that event satisfies

---

[2] If there are n individuals, each at risk for Δt years, then total person-years at risk would be $n \cdot \Delta t$. If individuals are at risk for different amounts of time, then person-years at risk would be $n \cdot \overline{\Delta t}$, where $\overline{\Delta t}$ is the average time at risk across all n individuals.

probability = event count per person.

So if we observe *k* events among *n* individuals, then the event probability *p* is estimated by

$$p = k/n.$$

The notions of rate and probability are related whenever we are counting events that occur over time but we are interested in a particular duration Δs. If events occur at rate *r*, and *p* is the probability of at least one event in duration Δs, then

$$p = 1 - \exp(-r \cdot \Delta s). \tag{1}$$

For small Δs, probability is approximately rate times duration:

$$p \approx r\Delta s.$$

For example, if the estimated rate is 1.25 events per person-year at risk, as above, and we are interested in a one-month duration, then $\Delta s = 1/12$ yr. and the probability of at least one event in a one-month duration is

$$p = 1 - \exp(-(1.25)(1/12)) = 0.0989.$$

Note that the approximation $p \approx r\Delta s$ gives $p = (1.25)(1/12) = 0.1042$, reasonably close since Δs is small, but not exactly correct.

Most software for Markov modeling is inherently discrete-time, and allows only probability parameters as input. Users are then required to convert rates to probabilities via (1). *StoTree* models *continuous-time* Markov chains, and allows both rate and probability inputs. *StoTree* converts rates to probabilities automatically as needed for calculations using (1), relieving the user of this task.

The relation (1) assumes that the event rate *r* is constant over time. Some rates – for example, mortality rates – are time-dependent. In this case the relation between rates and particular-duration probabilities is more complex. See, for example [3, 7].

Top

Cohort analysis: Quality, cost and discounting

The traditional method in the medical literature of solving Markov models is to perform a *cohort analysis*. This method assumes that a cohort of identical individuals begins in some initial health state x_0 at time $t = 0$, and progresses forward in time through the health states in the model, according to the assumed rates and probabilities in the model. The analyst selects a small time increment *dt*, called the *cycle length*. The rates and probabilities in the model determine the *transition probabilities*

$$\lambda_{xy}(dt,t) = \text{probability the state at time } t+dt \text{ is } y \text{ given that the state at time } t \text{ is } x.$$

The analyst then sets the cohort size $n_{x_0}(0)$ in state x_0 at time $t = 0$, and calculates the values of

$$n_y(t) = \text{expected count of individuals in state } y \text{ at time } t$$

for all time instants $t = 0, dt, 2dt, 3dt, \ldots$ using the Kolmogorov forward equations

$$n_y(t+dt) = n_y(t) + \sum_{x \neq y} n_x(t)\lambda_{xy}(dt,t) - n_y(t)\sum_{z \neq y} \lambda_{yz}(dt,t).$$

The purpose of computing the expected counts $n_y(t)$ is to use them to calculate expected *quality adjusted life years* (*QALY*s) and expected costs. The analyst uses the inputs

Q_y = expected quality accrued per unit time in state y (the so-called *utility* of state y)

C_y = expected cost accrued per unit time in state y

to compute expected *QALY*s by forming the sum of products $\sum_y Q_y n_y(t)$, possibly weighting the result by a time-discount factor, and then summing the all resulting quantities over all time instants $t = 0$, dt, $2dt$, $3dt$, …. The analogous operations are performed to calculate expected costs. It is also possible to include *quality tolls* Q_{xy} and *cost tolls* C_{xy} that are incurred whenever the cohort or a portion of it moves from state x to state y.

Note that if the utilities Q_y are all equal to 1 and there are no quality tolls, then expected *QALY*s is simply expected lifetime.

Top

8

Using *StoTree* to construct a stochastic tree

Overview of *StoTree* capabilities

StoTree is a Microsoft Excel addin possessing its own tab on Excel's ribbon (see Figure 5). You can use *StoTree* to formulate a continuous-time Markov model of health interventions of interest. *StoTree* allows you to graphically characterize the consequences of interventions in one or more compact diagrams. You can next specify rates for events that occur over time, and probabilities for one-time events. You can also input quality and cost parameters. You can then instruct *StoTree* to perform a cohort analysis in order to compute comparative effectiveness and/or cost-effectiveness. Finally, if you are concerned about the impact of particular parameter inputs, *StoTree* can perform a sensitivity analysis on such parameters.
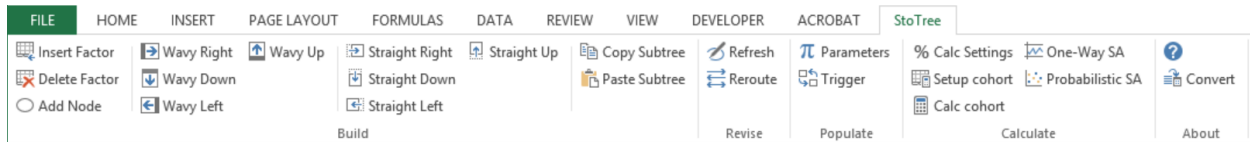

Figure 5. The StoTree tab on the Microsoft Excel ribbon. [Note: The identity of the remaining tabs will depend on what other software may be installed on your computer.]

Constructing a basic stochastic tree – in short

Once you install *StoTree* as an addin, you can use the commands on the *StoTree* tab to construct, populate and solve a stochastic tree. In overview, the steps are:

1. Insert a new factor.

2. Add one or more nodes.

3. Connect the nodes with arrows as desired.

4. Populate arrows with rate or probability parameters.

5. Instruct *StoTree* to setup a cohort analysis worksheet, and have *StoTree* calculate that sheet to compute expected *QALY*s and costs.

We discuss these steps in detail in the following sections.

Top

Constructing a stochastic tree – philosophy

StoTree uses a *node-focused* philosophy in constructing stochastic trees. By this we mean that most of the important graphical construction techniques you will use begin with creating, selecting or deleting a node. You should keep the following points in mind.

- Node selection is *sticky* in *StoTree* – if you click on two or more nodes consecutively, they will all be selected (in contrast to Excel's normal behavior, where clicking on a second graphic object will deselect the first one).

- If you want to change the layout of a tree you have constructed, simply drag one or more nodes to new locations, and then click on Refresh on the *StoTree* toolbar – the arrows will follow.

    StoTree does not arrange the layout of a tree for you. It places nodes you insert at the cell you have selected, and does not move or rearrange nodes – that is your responsibility.

- If you want to delete a node, simply select it and press Delete on your keyboard.

Constructing a basic stochastic tree – detail

To construct a basic stochastic tree using *StoTree*, you should open an empty Excel workbook, and proceed more or less from left to right on the *StoTree* tab. You should execute steps in roughly the following order:

1. *Insert a new factor.*

   Click on Insert Factor to insert a blank worksheet into your workbook, and set up your workbook to host stochastic tree construction. (You will see a dialog asking whether you wish to insert a stochastic tree into this workbook, and you should click Yes.) The resulting worksheet will have cells but no cell borders, and there will be an Active/Inactive toggle button in its upper left corner.

2. *Add one or more nodes.*

   a. Select a cell in the newly created worksheet where you would like *StoTree* to place a new node, and click on Add Node. A userform will appear that allows you to name the node and choose its type. *StoTree* will insert the name you provide as a caption in the node it draws. Click on Insert and a node will appear at your selected cell.

   b. Resize the node if it is too large or too small by clicking on it and dragging the shape handles that appear.

   c. Add other nodes to other locations in the worksheet by selecting other cells and clicking on Add Node again.

   d. You can move a node around by dragging it with your mouse. If needed, you can type in a new caption as well.

3. *Connect the nodes with arrows as desired.*

   a. To draw an arrow from one node to another:

      i. Click first on an empty cell to deselect any node you may have currently selected.

      ii. Click on the node where you want your new arrow to start.

      iii. Click on the node where you want your new arrow to end. (Note: You do not have to press the Ctrl key. Node selection in *StoTree* is *sticky*, and both nodes will remain selected when you click on the second one.)

      iv. Click on one of the arrow button on the *StoTree* tab.

   StoTree will draw an arrow of the type you chose from the first node to the second.

   b. You can drag to move nodes around your worksheet, but doing this will leave the arrows behind. *Do not attempt to drag the arrows*. Instead, click on the Refresh command on the *StoTree* tab, and arrows will re-attach to all nodes you have selected (or will reattach to all nodes if none are selected when you click Refresh).

4. *Revise your construction as needed*.

    a. You may copy/paste entire subtrees as follows. (However, using factoring is potentially more simple and elegant – see the section below.)

        i. Select the root node of any subtree you wish to copy.

        ii. Click the Copy Subtree button on the *StoTree* tab. *StoTree* will copy the subtree to a new blank worksheet.

        iii. Click on a cell in your original worksheet where you would like the copied subtree to be pasted.

        iv. Click on the Paste Subtree node. *StoTree* will paste the copied subtree into your worksheet with its root node at the cell you selected.

    (Please note that you *cannot* use Excel's Copy and Paste commands to duplicate existing nodes or arrows – or rather you can, but *StoTree* will not recognize the duplicates as part of your stochastic tree.)

    b. If needed, you can reroute arrows you have mistakenly drawn. Select any part of the arrow, and click on Reroute on the *StoTree* tab.

5. *Populate arrows with rate or probability parameters*.

You will notice that the arrows that *StoTree* draws are initially labeled "Needs populating". Click on any such arrow and then click the Parameters button on the *StoTree* tab. *StoTree* will open a user form where you may enter a desired rate (for a wavy arrow) or probability (for a straight arrow). You can ignore the Quality Toll and Cost Toll entries for now. You can enter a numeric value, or any numeric expression that will successfully evaluate in your Excel spreadsheet. You can include spreadsheet cell references in your expression if you provide defined names for the relevant cells (see below).

6. *Instruct StoTree to setup a cohort analysis worksheet*.

    a. Click on Setup Cohort Sheet. *StoTree* will open a new worksheet called the *cohort sheet*, and will insert formulas that will perform a cohort analysis when the sheet is repeatedly calculated.

    b. You can manually perform a cohort analysis by clicking on the button Step ahead one cycle. Or you can simply click on the Calculate Sheet button on the cohort sheet or the Calc Cohort button on the *StoTree* tab to complete the cohort analysis in one step.

Top

Calculation settings

StoTree provides default calculation settings for the cohort sheet it creates. However, you may modify these settings. Click on the Calc Settings button on the *StoTree* tab to enter your desired settings. The settings you can choose are

- *Initial Cohort*. The number of individuals in the cohort. *StoTree* calculates fractional values of the expected counts $n_y(t)$ as needed, so it is adequate to set the initial cohort size to 1.

- *Initial Time*. The clock time t at which the cohort begins. Usually this would be $t = 0$.

- *Cycle Time.* The increment *dt* by which time is incremented in each cycle of the cohort analysis. Ideally this should be a small value, but small *dt* increases calculation time. Larger values of *dt* will calculate quickly but results will be less accurate. Use large values of *dt* when debugging your model, but switch to small values after debugging is complete and you want accurate results.

- *Time Horizon.* The clock time at which you wish cohort analysis to stop.

- *Discount Rate.* If you choose discount rate to have value *r*, then *QALY*s and costs incurred at time *t* will be weighted by the discount factor e^{-rt}. Set the discount rate to zero if you want no discounting.

- *Calculate Expected QALYs* (Yes/No).

- *Calculate Expected Costs* (Yes/No).

For the first five settings above, you would usually enter numeric constants (e.g. 1 for *Initial Cohort*, 0 for *Initial Time*, and so on). However, you may also enter cell references if you provide defined names (see below) for the relevant cells. This could be useful if you wish to perform a sensitivity analysis on any of these settings.

Top

The cohort sheet

We illustrate here the cohort sheet output that *StoTree* produces for the HIV-AIDS model of Figure 3. The user has supplied values for the parameters *rAB, rBC, rD, rD_AIDS* (these are defined names in Excel – see below), and in addition has entered cost rates in each state. When the user clicks on the Setup cohort sheet button on the *StoTree* tab, *StoTree* produces the cohort sheet in Figure 6. In that figure, *StoTree* has performed one cycle of cohort analysis: The cycle time is 0.1, and in the *Time* column, only the time-instants $t = 0$, $t = 0.1$ appear. The initial cohort size is 1000 in state A, and during this cycle, the state counts have moved from (1000,0,0,0) to (948.749, 34.3946, 0, 16.8563) in states A,B,C and *Terminal Node* (death). Accrued *QALY*s for the cohort are 99.157 yr., and accrued costs are $225,880. The user can click the plus (+) signs on the left to inspect details in the calculations for counts, *QALY*s and costs that may be useful for debugging.

If the user presses the Calculate Sheet button, or the Calc Cohort button on the *StoTree* tab, *StoTree* will perform all additional cycles of cohort analysis, using a cycle length of 0.1 yr., until the time horizon $t = 20$ is reached. The results are presented in Figure 7. Discounted cumulative *QALY*s for this cohort of 1000 is 3588.62, and discounted cumulative cost is 10^7.

	A	B	C	D	E	F	G	H	I
1									
2	**Cycle Time =**	0.1		HIV_AIDS:	A: cd4 200	B: cd4 < 20	C: AIDS	Terminal Node	
3	**Starting Time =**	0		Chance/Decision?	0	0	0		
4	**Time Horizon =**	20							
5	**Discount Rate =**	0.03		**Initial Cohort =**	1000				
6									
7	Calculate Sheet								
8				**Time**	**Counts**				
9	Reset to Start Time			0	1000	0	0	0	FALSE
16				0.1	948.749	34.3946	0	16.8563	
17	Step ahead one cycle								
18				Quality Coefficients:	1	1	1		
19				Cost Coefficients:	2278	2278	7276		
20									
25				Incremental QALY Tolls:	0	0	0		
26				Incremental QALYs:	97.4375	1.71973	0		
27				Incremental Total:	97.4375	1.71973	0		
28				Cum Discounted QALYs:	97.4375	1.71973	0		
29				**Overall Cum QALYs =**	**99.157**				
30									
35				Incremental Cost Tolls:	0	0	0		
36				Incremental Costs:	221963	3917.54	0		
37				Incremental Total:	221963	3917.54	0		
38				Cum Discounted Costs:	221963	3917.54	0		
39				**Overall Cum Costs =**	**225880**				

Figure 6. The cohort sheet StoTree produces for the stochastic tree model of Figure 3 when the user clicks the Setup cohort sheet *button on the StoTree tab.*

1 2		A	B	C	D	E	F	G	H	I
	1									
	2	Cycle Time =	0.1		HIV_AIDS:	A: cd4 200	B: cd4 < 20	C: AIDS	Terminal Node	
	3	Starting Time =	0		Chance/Decision?	0	0	0		
	4	Time Horizon =	20							
	5	Discount Rate =	0.03		Initial Cohort =	1000				
	6									
	7	Calculate Sheet								
	8					Time	Counts			
	9	Reset to Start Time				19.9	0.02839	0.06056	0.57583	999.335	FALSE
+	16					20	0.02693	0.05749	0.55077	999.365	
	17	Step ahead one cycle								
	18				Quality Coefficients:	1	1	1		
	19				Cost Coefficients:	2278	2278	7276		
	20									
+	25				Incremental QALY Tolls:	0	0	0		
	26				Incremental QALYs:	0.00277	0.0059	0.05633		
	27				Incremental Total:	0.00277	0.0059	0.05633		
	28				Cum Discounted QALYs:	1798.72	911.259	877.883		
	29				Overall Cum QALYs =	3587.9				
	30									
+	35				Incremental Cost Tolls:	0	0	0		
	36				Incremental Costs:	6.30092	13.4461	409.855		
	37				Incremental Total:	6.30092	13.4461	409.855		
	38				Cum Discounted Costs:	4097481	2075849	6387477		
	39				Overall Cum Costs =	1E+07				
	40									

Figure 7. The revised cohort sheet produced when the user clicks Calculate sheet.

As suggested by the other two buttons on the cohort worksheet, the user also has the option of resetting calculations to the start time, which will produce once again the worksheet displayed in Figure 6; and then stepping the calculations ahead one cycle at a time. This procedure could be useful for debugging.

Once you have instructed *StoTree* to create a cohort sheet of this type, you should save your workbook as a macro-enabled workbook, that is, with a ".xlsm" suffix. The workbook now contains Visual Basic macros in a hidden code sheet, and must be saved with this suffix in order for the macros to function.

Although you need a licensed copy of *StoTree* to create a cohort sheet, the resulting workbook may be viewed and manipulated by anyone who has even an unlicensed version of *StoTree* installed. You may therefore share your workbook (see below) with colleagues possessing only unlicensed copies of the *StoTree* addin. Your colleagues may wish to perform what-if analyses on your model, and may do so without acquiring a license. See the section *Defined names for documenting and sensitivity* below for more information on how to do *what-if* analyses.

Top

Tips and shortcuts for constructing a basic tree

There are a number of shortcuts you can use to construct basic trees more quickly and elegantly.

1. You can instruct *StoTree* to draw multiple arrows in one click by first selecting node you want to serve as origin (call it node A), and next selecting all destination nodes (call them nodes B,C and D). Then a single click of an arrow button on the *StoTree* tab will draw arrows from A to B, A to C and A to D.

14

2. As already mentioned, node selection in *StoTree* is *sticky* – you can click several nodes in succession to select them all, without using the Ctrl key in Excel. In fact, if you wish to take advantage of the shortcut in point 1, you *must* select nodes *without* using the Ctrl key – otherwise *StoTree* does not recognize the multiple selection. You should also be careful, if you want to select only one node, that other nodes are not already selected. To achieve this, simply click on any cell in the worksheet and all nodes will be de-selected.

3. If you determine that an arrow or a node is no longer needed, you can simply click on the node or any part of the arrow and press the Delete key on your keyboard. *StoTree* will notice the deletion the next time you click one of the buttons on its tab. At this point, if you have deleted a node, then *StoTree* will also delete any arrows connected to that node. If you delete just part of an arrow, then *StoTree* will delete the entire arrow when it notices the partial deletion.

4. If you mistakenly draw an arrow of the wrong type or orientation, you can correct your mistake without deleting the arrow and redrawing it. Instead select the arrow and click the desired arrow button on the *StoTree* tab. *StoTree* will replace the mistaken arrow with the corrected arrow.

5. You can move labels on arrows from one side of the arrow to the other, or slide them along the arrow. Clicking the Refresh button will leave the label on the side of the arrow on which *StoTree* finds it.

6. You may wish to move or delete a group of nodes and arrows together. You can do so by activating Excel's Select Objects pointer ↕ (Home/Editing/Find & Select). You can use this pointer to select multiple graphics at once by dragging it around the objects. Once a group of nodes and arrows are selected, you can drag them to any desired location or delete them. If you wish to access Excel's Select Objects pointer more quickly, you can add it to the Quick Access Toolbar (File/Options/Quick Access Toolbar).

7. Do *not* attempt to use Excel's Copy/Paste tools to create new nodes or arrows. This will only result in new worksheet graphics that *StoTree* will not recognize as being part of your model. The reason is that *StoTree* maintains hidden databases of nodes and arcs, and these databases will *not* be updated by Excel's Copy/Paste tools. You must use the *StoTree* tab on the Office ribbon to create new nodes or arrows. For similar reasons, do not attempt to create a new factor (see below for a discussion of factors) by having Excel copy a worksheet.

Top

QALYs and costs

You can enter quality accrual rates (*quality coefficients* or so-called *utilities*) and cost accrual rates for any stochastic node in your tree. (Note chance and decision nodes are instantaneous, so quality and cost accrual does not apply to them.) Simply select the node and then click on the Parameters button on the *StoTree* tab. A userform appears into which you may enter these values for the selected node. You may enter quality tolls and cost tolls for any arrow in your model – again simply select part of the arrow and then click on the Parameters button.

Top

Sharing your constructed workbook

Once you have used the *StoTree* addin to construct a Markov model in your workbook, and have saved your workbook as a macro-enabled workbook (.xlsm), you may wish to share your workbook with colleagues. Because it is an Excel workbook, anyone with Microsoft Office can view it. However the

graphic objects and cohort-sheet formulas in your workbook contain references to macros in *your* copy of the *StoTree* addin. When a user on another computer recalculates the cohort sheet, or clicks on a graphic, error messages will result – because your workbook is on her computer, the macro references there will not be able to find your copy of the *StoTree* addin.

However, if the user has a copy of the *StoTree* addin installed, she can convert her copy of your workbook to access *her* copy of *StoTree*. She does this by opening her copy of your workbook and clicking on the Convert button in the About section of the *StoTree* tab. A successful conversion will result in a confirmation message similar to the one illustrated in Figure 8.
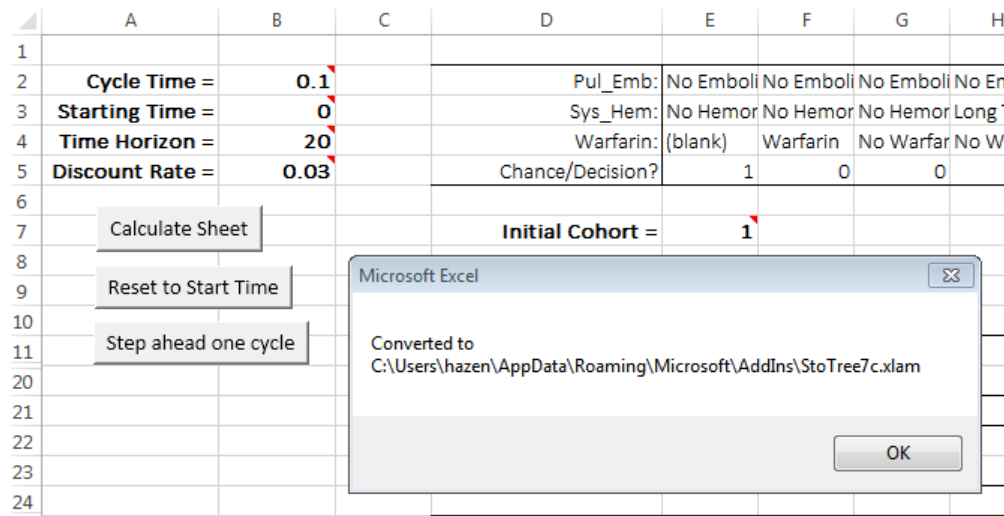

Figure 8. To convert another user's workbook to a form compatible with your copy of Microsoft Office, you should clicking on the Convert button in the About section of the StoTree tab. This figure illustrates the confirmation message displayed when conversion is complete.

Top

Defined names for documenting and sensitivity

You can enter any numeric value or expression (e.g., 22.3, 1/10, 12+3) into the Parameters userform. However, to properly document your model and make it more transparent and flexible, you are strongly advised to avoid numeric expressions, and instead use Excel's *defined names* feature (on the Formulas tab) to provide names for cells containing your input data. Then use these names, or expressions containing these names, in *StoTree*'s Parameter userform when you populate parameters for your model. *StoTree* displays rate and probability expressions next to the appropriate arrows, so using defined names makes it easy to see to what arrows these inputs apply. See for instance, Figure 9. This is a method you can use to document your model.

A second advantage of using Excel's defined names in the Parameter userform is that you may change the values in the named cells, and then Calculate the cohort sheet to obtain new expected *QALY*s and costs, without having to request *StoTree* to setup the cohort sheet again. You can thereby quickly perform a manual sensitivity analysis by repeatedly changing values in named cells and clicking on Calculate Sheet. Note, however, that *StoTree* has automated this process via the One-Way Sensitivity button on the *StoTree* tab.
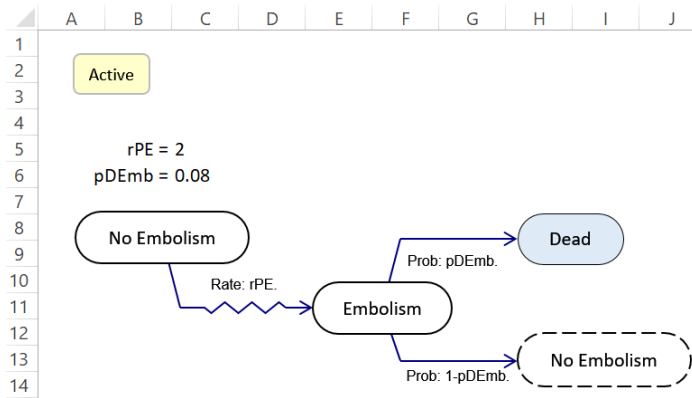
16

Figure 9. The use of Excel's defined names in StoTree. The user has given cell C5 the name rPE, and cell C6 the name pDEmb.

One-way sensitivity analysis

By clicking on the One-Way SA button on the *StoTree* tab, you can obtain instructions in how to set up a one-way sensitivity analysis on any input parameter that you have stored in a cell of your workbook. You proceed by setting up one of Excel's *column data tables*, then selecting its corner cell, and finally pressing Ctrl+Shift+s. Details are provided by clicking on the One-Way SA button on the Excel ribbon's *StoTree* tab.

Probabilistic sensitivity analysis

Probabilistic sensitivity analysis is a type of sensitivity analysis in which the uncertainty in parameters is represented by probability distributions. The primary advantage of probabilistic sensitivity analysis is that it allows analysts to perform a sensitivity analysis on all parameters jointly, instead of only one or two at a time.

If you have set up your workbook for probabilistic sensitivity analysis, then you can obtain sensitivity estimates via Monte Carlo simulation, by having StoTree repeatedly recalculate your workbook, once for each Monte Carlo iteration. These recalculations are done in a structured way that is not equivalent to using the F9 key on your keyboard.

For information on how to set up your workbook for probabilistic sensitivity analysis, click on the Probabilistic SA button on the Excel ribbon's *StoTree* tab.

Time-dependent parameters

As mentioned above, you may enter any valid Excel formula or expression containing defined names into *StoTree*'s Parameter userform. There is an additional defined name _t that *StoTree* automatically creates when you instruct it to setup a cohort sheet. The name _t represents the *elapsed time* in the cohort sheet, and you may use this name in any of the formulas or expressions you enter into the Parameter userform. You would most commonly use _t in a rate formula when you desire the rate to depend on elapsed time in your cohort analysis. Figure 10 shows an example. However, _t may be used in any parameter expression, for example for quality or cost coefficients. Figure 11 illustrates how you could use _t in a quality coefficient formula.
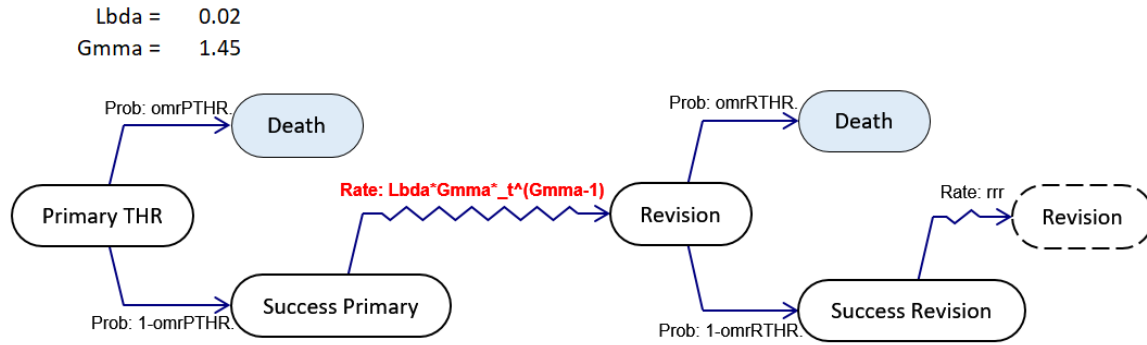
17

Figure 10. You can use the defined name _t *to represent elapsed time in parameter expressions. Here, in a model of the outcome of hip replacement, the rate out of* Success Primary *is the Weibull hazard rate* $h(t) = \lambda \gamma t^{\gamma-1.}$
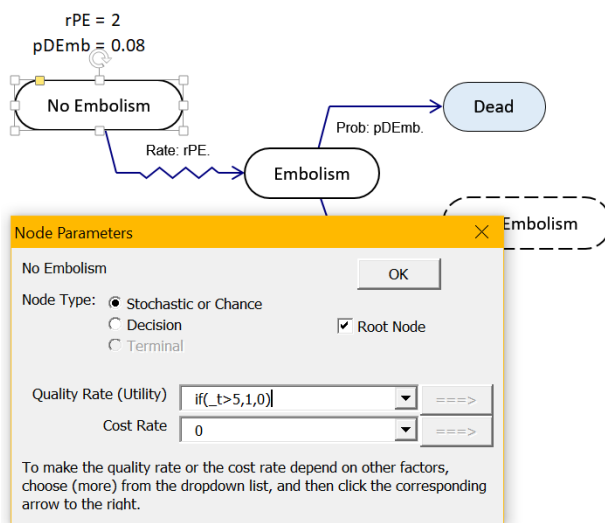

Figure 11. You can use the elapsed-time defined name _t *in a quality coefficient. Here the node* No Embolism *has zero quality until 5 years have passed, and quality 1 after that. The resulting cohort analysis (with initial cohort set to 1.0) will calculate expected survival in excess of 5 years..*

[Top](#)

Human mortality rates

StoTree includes a function *rMortUS* that returns annual human mortality rates in the United States for any age, gender and race. You can employ this function in conjunction with the current time variable _t to supply the correct human time-dependent mortality at any clock time in a cohort analysis – see for example Figure 12. See the section *Built-in Functions* for documentation on this function.
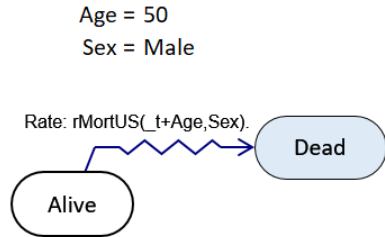
Age = 50
Sex = Male

Rate: rMortUS(_t+Age,Sex).
Dead
Alive

Figure 12. Use of StoTree's function rMortUS *to model human mortality.*

Constant intervals of clock time

Occasionally you may wish a state transition to occur at a specific clock time t_0, or repeatedly at time intervals of length t_{Inc}. In the first case this means that the associated transition rate at t_0 is essentially infinite, and in the second case essentially infinite every t_{Inc} units of clock time; but zero for other clock times. The *StoTree* jump functions JmpAt_() and JmpEvery_() produce these respective effects.

Consider a stochastic (wavy) arrow connecting node A to node B in a stochastic tree. When the user inserts a rate expression JmpAt_(t_0) on that arrow, then whatever cohort is at A will jump to B at time t_0 (but at no other times). There are also optional parameters. The expression JmpAt_(t_0,t_{Inc}) forces whatever cohort is at A at time t_0 to jump to B, and also whatever cohorts are at A at subsequent times $t_0 + t_{Inc}$, $t_0 + 2t_{Inc}$, $t_0 + 3t_{Inc}$, … to jump to B. The expression JmpAt_(t_0,t_{Inc},t_F) stops jumps after time t_F.

The expression JmpEvery_(t_{Inc}) forces whatever cohort is at A at times 0, t_{Inc}, $2t_{Inc}$, … (but at no other times) to jump to B. JmpEvery_(t_{Inc}, t_0) produces a jump from A to B at times t_0, $t_0 + t_{Inc}$, $t_0 + 2t_{Inc}$, …. (and is therefore equivalent to JmpAt_(t_0,t_{Inc})). And JmpEvery_(t_{Inc}, t_0, t_F) stops jumps after time t_F.

Figure 13 illustrates the use of JmpAt_() and JmpEvery_() in a model of treating atopic dermatitis (AD) in infants. Because JmpAt_() and JmpEvery_() produce zero or infinite *rates*, they can only be used on stochastic (wavy) arrows.

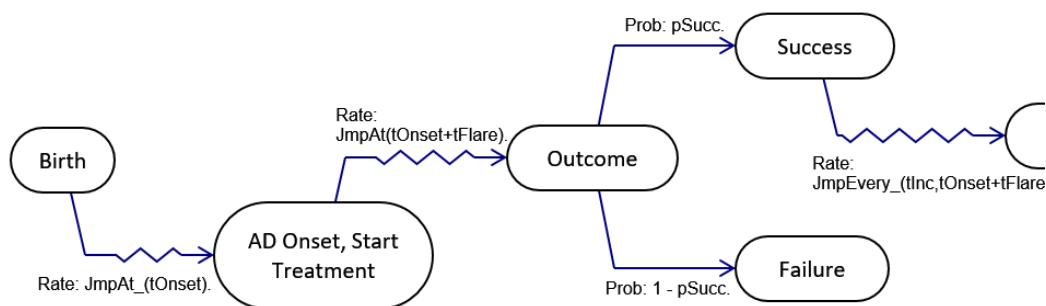Prob: pSucc.
Success

Rate:
JmpAt(tOnset+tFlare).
Outcome
Rate:
JmpEvery_(tInc,tOnset+tFlare

Birth

Rate: JmpAt_(tOnset).
AD Onset, Start
Treatment

Prob: 1 - pSucc.
Failure

Figure 13. A stochastic tree fragment illustrating the use of the StoTree functions JmpAt_() and JmpEvery_() in a model of treating atopic dermatitis (AD) in infants. Here the duration tOnset from birth to AD onset, and the duration tFlare of a flare are taken as known deterministic values. The function JmpEvery_() on the arrow out of Treatment Success *is used to model repeated flares at intervals of tInc.*

Figure 14 shows the use of JmpEvery_() to model repeated testing.

19

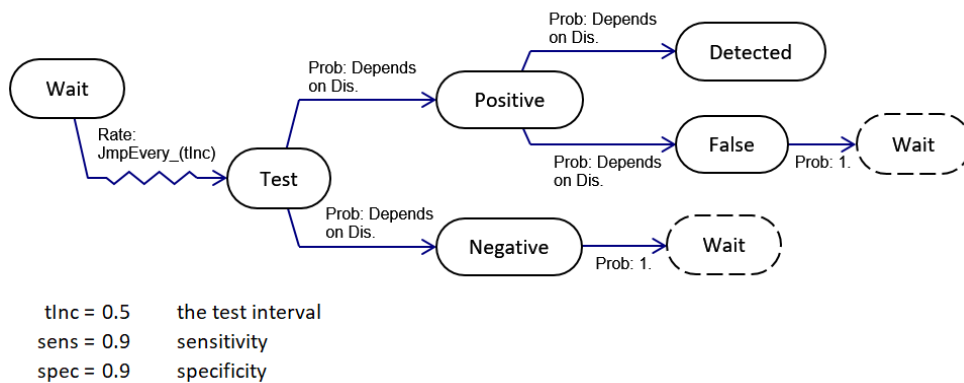*Figure 14. A stochastic tree illustrating the use of the StoTree function JmpEvery_() to model repeated testing at intervals of tInc = 0.5 yr. Negative tests or false positive tests result in return to the Wait state, where testing is repeated due to the rate JmpEvery_(tInc).*

It is possible you may wish transitions to occur only at predetermined times $t_1, t_2, \ldots$ that are not equally spaced. For example, you might want testing to occur more frequently initially than later. The multiple-jump function JmpAtM_ () accomplishes this. If you enter the times $t_1, t_2, \ldots$ into a spreadsheet region and assign that region an Excel defined name, then using the region name in a call to JmpAtM_ () will result in jumps at those times and no others. For example, if the name of the region containing times $t_1, t_2, \ldots$ is *tTimes*, then the call should be JmpAtM_(*tTimes*).

Top

Clock time versus sojourn time

You may at times wish to have a rate out of some state depend not on the clock time _t but on the so-called *sojourn time* in a state – the time elapsed since the state was entered. Unfortunately, Markov models as usually formulated do not explicitly include a sojourn-time variable, except in the special case in which sojourn time and clock time coincide, which happens only for states occupied initially at time 0.

The distinction can be seen in Figure 10, where the failure rate of the initially implanted artificial hip depends on its age, which is equal to the sojourn time in the *Success Primary* state. In that state, age is equal to clock time _t because a new hip is implanted at time 0. Therefore the age-dependent failure rate can be made a Weibull function of _t. On the other hand, in the state *Success Revision*, age is never equal to clock time because the replacement hip is never implanted (if it is at all) at time 0. The modeler would like to make the failure rate of the replacement hip depend on its age, but cannot because there is no sojourn time variable available. As an approximation, the modeler has in this case entered a constant failure rate *rrr* for the replacement hip.

This approximation may be adequate in some cases, but if it is not, then the modeler must resort to using what are known as *tunnel states*, a sequence of states that mimics elapsed sojourn time. *StoTree* can handle tunnel states – for a discussion of the methodology, see [8].

Top

Decision nodes

In a stochastic tree, a decision node is meant to indicate that an individual has a choice among the branches leading from that node. However, in *StoTree*, decision nodes are treated the same as chance nodes. You the user would assign probability one to the branch corresponding to whatever decision is to

be analyzed, and probability zero to other branches. *StoTree* does not perform any kind of optimization among decision branches – it merely evaluates *QALY*s and/or costs associated with the desired decision branch. To evaluate a different decision, you would have to change its branch probability to one. This can be automated using Excel's defined names, as is illustrated in Figure 15.
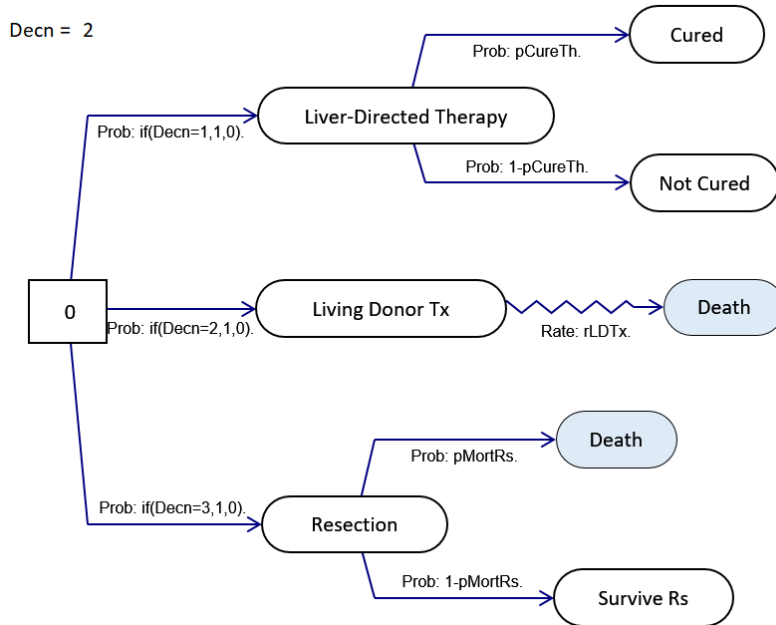

Figure 15. You can use a defined name (here, Decn*) to specify which decision branch is to be selected. Change the value of the* Decn *cell and press* Calculate Sheet *to evaluate a different decision. You can also use StoTree's One-Way Sensitivity command to perform a sensitivity analysis on* Decn, *and thereby evaluate all decision alternatives in one step.*

Top

Terminal nodes

Death states are common in medical decision analyses. They can be represented in *StoTree* as stochastic nodes with zero quality and no emanating arrows. However, it is computationally more efficient to use the *StoTree* node type *terminal* to represent death states. When you create a terminal node, it appears shaded in your spreadsheet (see several of the preceding figures). You can change an existing stochastic node to terminal in the Parameters userform if the node has no successors.

Top

Root nodes

StoTree begins its cohort analysis by assuming all cohort members occupy the *root node* in the stochastic tree. You can set which node in your tree is the root node using a checkbox in the Parameters userform for that node. If you neglect to do so, *StoTree* will make an educated guess about which node is the root node in your tree. However, for some types of trees, there may be more than one plausible root node, and you should make sure *StoTree* has selected the one you desire.

Top

Factored stochastic trees

Often in constructing models of medical prognosis, one encounters processes that evolve *in parallel*. Consider for example, the pulmonary embolism model depicted in Figure 1, which we duplicate below in Figure 16 for convenience.

Tsevat et al.[9] consider the use of Warfarin to reduce the incidence of pulmonary embolism. However, Warfarin can increase the incidence of systemic hemorrhage, a process depicted in Figure 17. The pulmonary embolism and systemic hemorrhage processes are proceeding *in parallel* – both processes evolve *simultaneously* and *in parallel*. It is possible, but cumbersome, to formulate a single Markov model or stochastic tree that depicts both processes. Fortunately, this is not necessary in *StoTree*. One of the most important features of *StoTree* is that it allows the modeler to *formulate parallel processes as inputs*. This is done by formulating the two processes on *separate worksheets* in the Excel workbook.

StoTree also allows the modeler to specify links between the processes, so that what happens in one process (worksheet) can influence what happens in others, and can in turn be influenced by others. Parallel processes are known as *factors*, and the overall combination of processes is called a *factored stochastic tree.*

For emphasis, we repeat the key properties of factors in *StoTree*.

1. The user composes different factors on different worksheets in Excel.

2. *StoTree* interprets different factors (different worksheets) as proceding *simultaneously* and *in parallel* along the *same* timeline. At all times, one state in each factor is occupied.
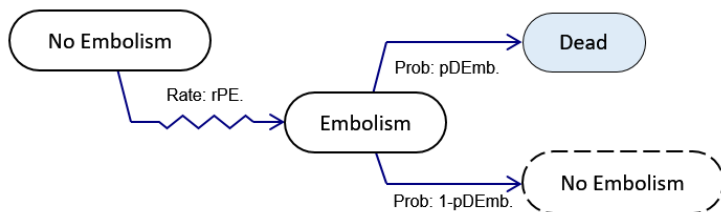

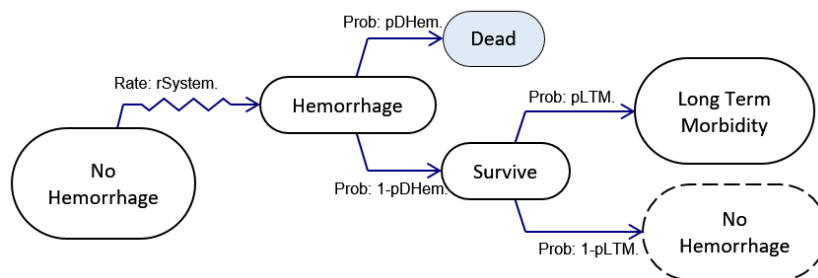*Figure 16. A model of repeated pulmonary embolism, duplicating Figure 1.*

Figure 17. A model of systemic hemorrhage. This process evolves in parallel *with the pulmonary embolism process depicted in Figure 16.*

Using *StoTree* to construct a factored stochastic tree

How do you construct a factored stochastic tree? The procedure in *StoTree* is quite simple: Click on the Insert Factor button on the *StoTree* tab for each factor you wish to create. *StoTree* will insert new worksheets into your workbook in which you can construct new stochastic factors in the same way as described in the sections above. Once you have created multiple worksheets, each containing its own factor, *StoTree* allows you to link what happens in one factor to what happens in others using *dependencies* and *triggers*. We discuss each of these in turn.

Top

Dependencies

In *StoTree*, a *dependency* indicates how parameters in one factor depend on the state or states of other factors. For example, in the *StoTree* model with factors depicted in Figure 16 and Figure 17, both the rate of embolism and the rate of hemorrhage should depend on whether Warfarin is being administered. The decision on whether to do so is shown in Figure 18, which depicts a *third* factor with an initial decision node *Choose* selecting between *Warfarin* and *No Warfarin*.
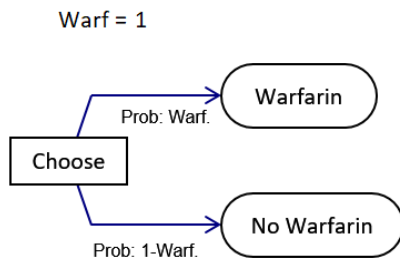

Figure 18. The Therapy *factor, depicting the decision of whether to administer Warfarin.*

Once the user has created this *Therapy* factor, s/he can make the hemorrhage rate in the *Systemic Hemorrhage* factor depend on whether Warfarin is being administered. The user activates the *Systemic Hemorrhage* factor, clicks on the arrow out of *No Hemorrhage*, and proceeds as in Figure 19, Figure 20, Figure 21, and Figure 22 below.
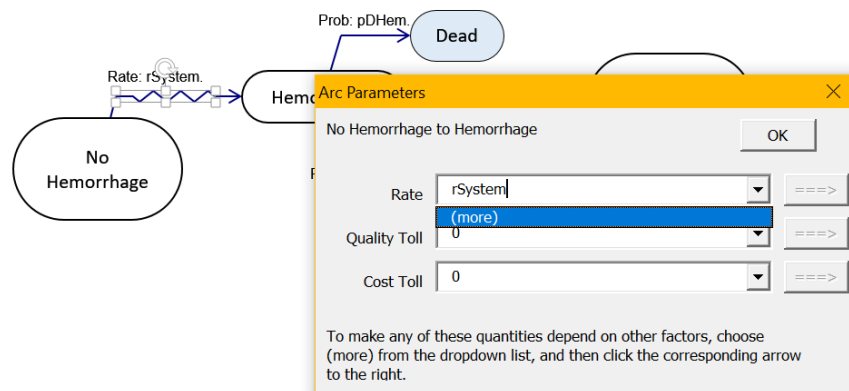

Figure 19. Creating a dependency in the Systemic Hemorrhage *factor. The user has selected the arrow from* No Hemorrhage *and clicked on the* Parameters *button to bring up the userform pictured here. Next the user has clicked on the dropdown arrow in the* Rate *combobox to reveal the option* (more).
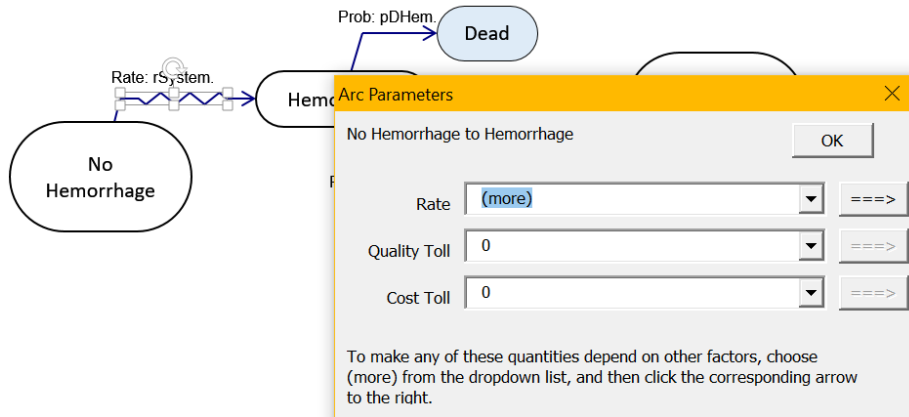
Figure 20. The user has clicked on the (more) *entry in the Rate combobox dropdown. As a result, the ===> button next to the combo box has been activated. The user will next click on this button to bring up the Dependency userform.*
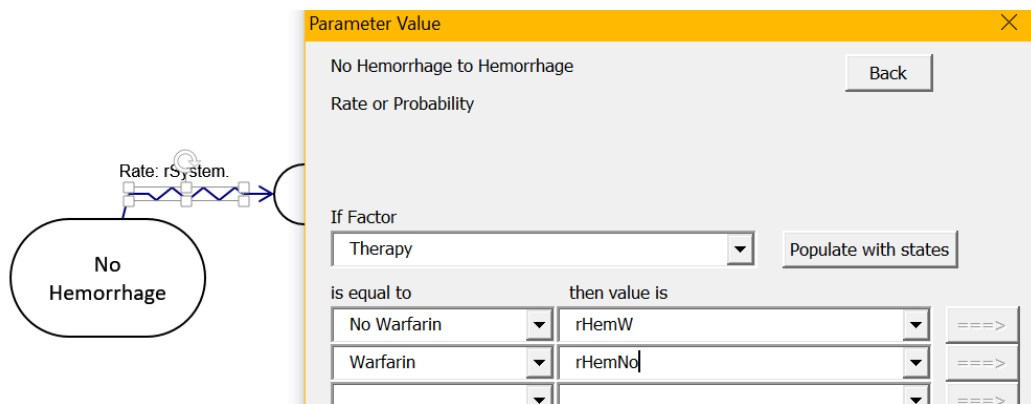

Figure 21. The Dependency userform. The user has selected Therapy *from the* If Factor *dropdown, and indicated that the chosen rate will equal rHemW if the* Therapy *factor is in state* Warfarin, *and will equal rHemNo if the* Therapy *factor is in state* No Warfarin. *The user will next click on* Back *to return to the* Parameters *userform and then click* Ok..
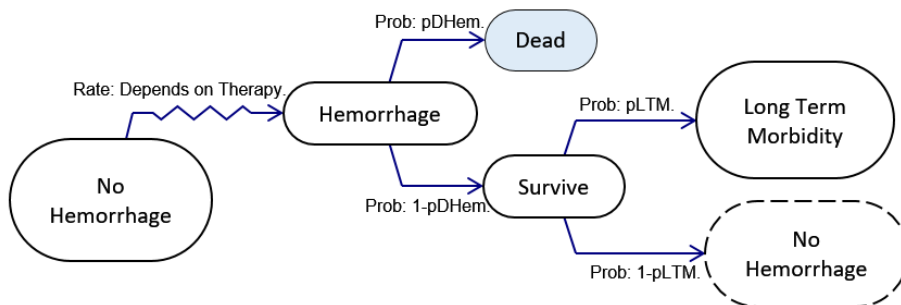

Figure 22. The process is complete. The arrow out of No Hemorrhage *is labeled to indicate that its rate depends on the status of the* Therapy *factor.*

Top

Triggers

Sometimes it is desirable to have a transition in one factor produce a state change in another factor. The user can arrange this by using the *Trigger* button on the *StoTree* tab.

Consider once more the Tsevat model of Warfarin treatment (*Embolism* in Figure 16, *Hemorrhage* in Figure 17, and *Therapy* in Figure 18). If a hemorrhage occurs while the patient is on Warfarin, it would seem natural to discontinue Warfarin for some period of time. In other words, the occurrence of a hemorrhage should *trigger* transition from *Warfarin* to *No Warfarin* in the *Therapy* factor. The user needs to add a trigger that sends the *Therapy* factor from *Warfarin* to *No Warfarin*. The method is shown in Figure 23, Figure 24, and Figure 25.

Triggers may also be attached to nodes. *StoTree* executes such a trigger *just after* the node is entered. This could be useful if there is a node with multiple entering arrows, on each of which the same trigger is needed. Then it would be more efficient to simply put that trigger on the node.
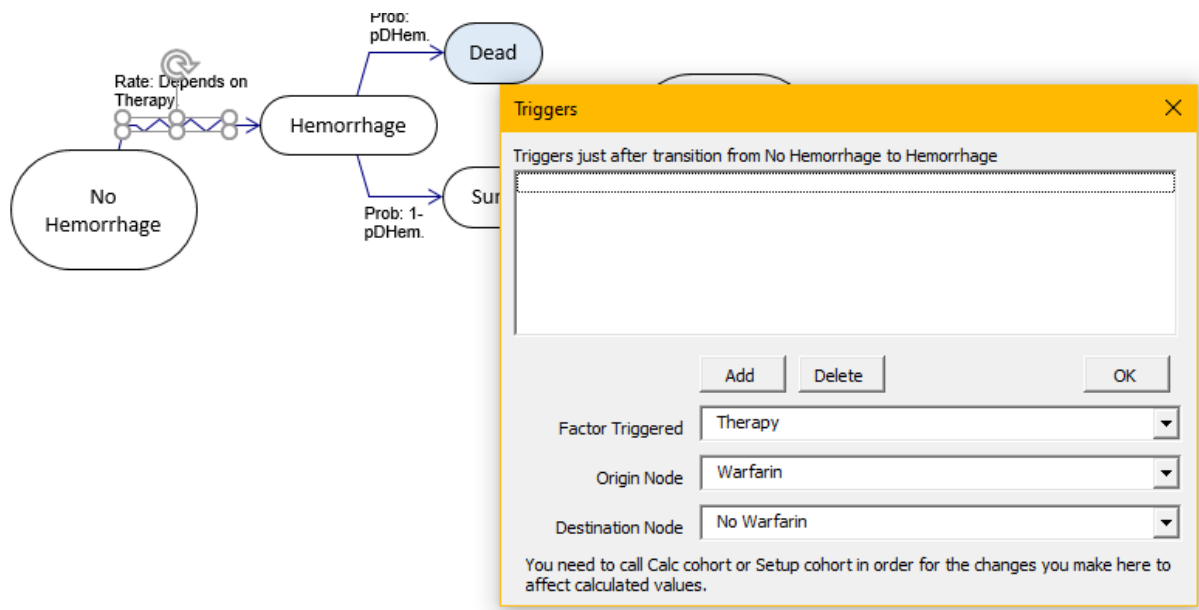

Figure 23. Inserting a trigger. The user has selected the arrow into Hemorrhage *and clicked on the* Trigger *button on the StoTree tab to show the dialog above. Then s/he has selected* Therapy *from the* Factor Triggered *dropdown*, Warfarin *from the* Origin Node *dropdown, and* No Warfarin *from the* Destination Node *dropdown, reflecting the desired trigger.*
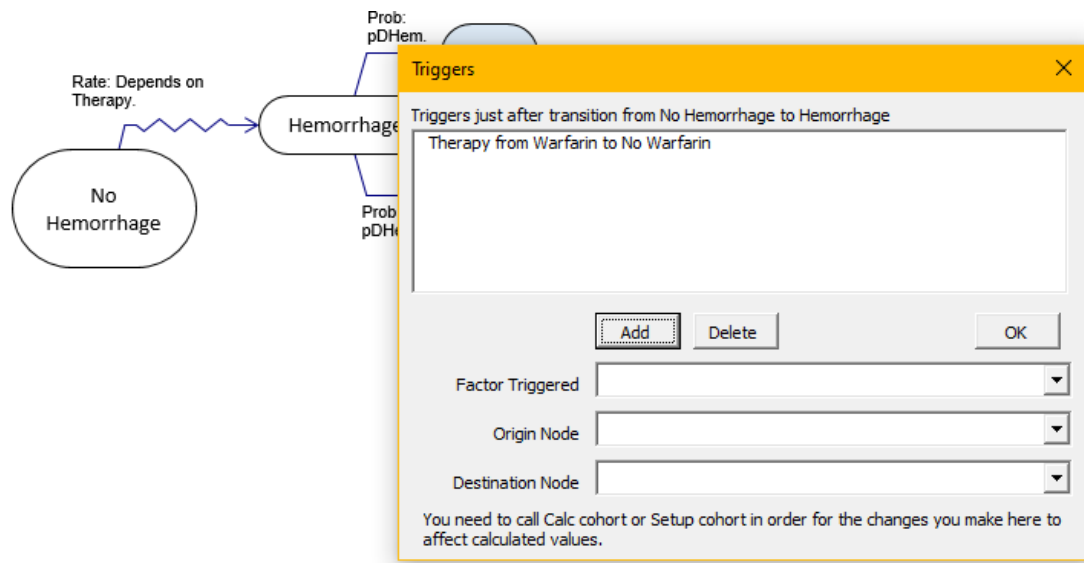
Figure 24. Inserting a trigger (continued). The user has next clicked on the Add *button to add the chosen trigger to the list above. The user will click on* OK *to complete the process.*
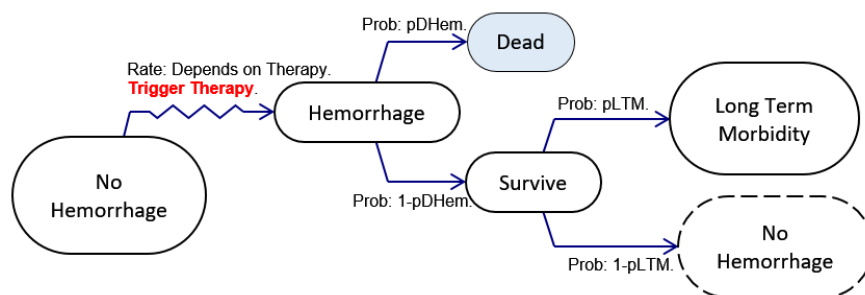

Figure 25. After the user enters the trigger shown in Figure 24, StoTree displays a comment (depicted in red here for emphasis) indicating that the Therapy *factor is triggered by transition along the relevant arrow.*

Top

Priorities among triggers*

A user may enter several triggers on an arrow. *StoTree* executes these in the order entered, which is the order listed in the trigger userform. Triggers are executed *after* any tolls on an arrow are incurred. For example, suppose there is a toll on an arrow $A_1 \rightarrow B_1$ in factor 1 that depends on whether factor 2 is in state *High* or state *Low*. But suppose also that the transition $A_1 \rightarrow B_1$ triggers factor 2 from *High* to *Low*. When the transition $A_1 \rightarrow B_1$ occurs and its toll is incurred, which state will factor 2 be in – *High* or *Low*? The answer is state *High* – the toll is incurred *before* the trigger is executed.

Top

Terminal nodes

Terminal nodes are the shaded nodes in diagrams above that usually represent death. They have a special status in a factored stochastic tree, in that a transition to a terminal node automatically triggers the entire process to stop. For example, consider the factored stochastic tree with an embolism factor depicted in Figure 16, and a hemorrhage factor depicted in Figure 17. Should a transition to Dead occur in the hemorrhage factor, the embolism factor would not continue, but rather terminate also. This would happen

26

even if the embolism factor did not have any terminal nodes. Transition to a terminal node in any factor stops everything.

This is another reason to use terminal nodes to represent death in a stochastic tree model. One could instead use a stochastic node with zero quality, but transition to this node would not stop other factors from proceeding through their states unless the user deliberately set up triggers to force this to happen.

Top

Priorities among arrows*

In a multifactor tree, there can be many combinations of health states. Consider a particular health-state combination. There may be multiple ways to exit this combination health state, corresponding to multiple exit arrows across different factors.

Recall that *StoTree* treats decision nodes the same as chance nodes. You the user would assign probability one to the branch corresponding to whatever decision is to be analyzed, and probability zero to other branches. For the purpose of assigning priorities, there are therefore only two types of nodes that can appear in a combination health state – decision/ chance nodes, and stochastic nodes.

The following precedence rules are enforced:

1. If a combination state contains only one decision/ chance node, *StoTree only* executes arrows out of this node, thereby reaching a new combination state. No other arrows are executed.

   For instance, consider the combination state (*Choose*, *No Embolism, No Hemorrhage*) in the model composed of the *Therapy* factor (Figure 18), the Embolism factor (Figure 16), and the Hemorrhage factor (Figure 17). The combination state contains only one decision/ chance node *Choose* (a decision node). It takes priority, and the only possible transitions out of this combination state are those in the *Therapy* factor, namely the transitions to *Warfarin* and to *No Warfarin*. *StoTree* ignores the transitions from *No Embolism* in the Embolism factor, and from *No Hemorrhage* in the Hemorrhage factor. (Note, however, that these transitions will be considered subsequently because they will be present at resulting combination states.)

   As another example, consider the combination state (*Warfarin, Embolism, No Hemorrhage*) in this Warfarin model. *Warfarin* and *No Hemorrhage* are stochastic nodes, and there is one decision/ chance node *Embolism* (a chance node). Therefore *StoTree* considers only transitions out of *Embolism* and ignores transitions out of *No Hemorrhage* (and would ignore transitions out of *Warfarin* if there were any). Again, these transitions will be subsequently be executed because they will be present at the resulting combination states.

2. If the combination state contains decision/ chance nodes in two or more factors, *StoTree* executes only the decision/ chance arrows from the *leftmost* of these factors, that is, the factor contained in the *leftmost worksheet* in the Excel workbook containing the model, or in other words, the worksheet having the leftmost tab in the Excel display. No other arrows are executed. (But again, decision/ chance arrows in other factors would subsequently be executed because they would be present at the new combination states.)

   [This leftmost priority was established in version *7T* of *StoTree*. For earlier versions *7S* and before, priority was given to the *alphanumerically first* factor, that is, the factor whose worksheet name is lexicographically first in alphanumeric order.]

3. If only stochastic nodes are present, they are all executed across all factors, that is transitions *compete* across factors.

For instance, consider the combination state (*Warfarin, No Embolism, No Hemorrhage*) in the Warfarin model discussed in point 1. This state consists of three stochastic nodes. *StoTree* considers all transitions from these three nodes, which in this case consist of a transition to *Embolism* at rate *rPE* and a transition to *Hemorrhage* at rate *rSysHem*. These two transitions compete, and *StoTree* accounts for the possibility that either one may occur first in time.

The cohort sheet for a factored tree

It is important to remember that in a factored stochastic tree, the factors depict processes that are proceeding in parallel. So *any combination of states*, one from each factor, may in principle occur. *StoTree* accounts for this in constructing the cohort sheet, and lists many such state combinations, as long as they can be reached from the combination root node. Figure 26 illustrates this for the example tree of this section, a combination of the trees in Figure 16, Figure 23, Figure 25.

	A	B	C	D	E	F	G	H	I	J	K
1											
2	Cycle Time =	0.1									
3	Start Time =	0									
4	Time Horizon =	20									
5	Discount Rate =	0.03									
6					Pul Emb:	No Emboli	No Emboli	No Emboli	No Emboli	No Emboli	Terminal Node
7	Calculate Sheet				Sys Hem:	No Hemor	No Hemor	No Hemor	Long Term	Long Term	Morbidity
8					Therapy:	Choose	Warfarin	No Warfar	Warfarin	No Warfarin	
9	Reset to Start Time				Chance/Decision?	1	0	0	0	0	
10											
11	Step ahead one cycle				Initial Cohort =	1					
12											

Figure 26. The cohort sheet StoTree generates when there are multiple factors. In this case the stochastic tree is the one with factors Pul_Emb, Sys_Hem, *and* Therapy *depicted in Figure 16, Figure 23, Figure 25, respectively.*

You may note that the three factor trees in this model contain respectively 3, 5, and 4 states. Yet there are nowhere near $3 \times 5 \times 4 = 60$ combination states in the cohort sheet of Figure 26. *StoTree* exploits known properties of stochastic processes to bypass the portrayal of terminal (death) nodes, chance nodes, and decision nodes (unless they are root nodes). Only stochastic nodes and root chance/decision nodes need be portrayed on the cohort sheet. This is the primary reason you should use a terminal node rather than a stochastic node to model the death state – a stochastic-node death state will add to the number of combinations that need display, whereas a terminal node will not. The three factors in this model contain respectively 1, 2, and 4 stochastic or root nodes, producing $1 \times 2 \times 4 = 8$ combinations, the ones displayed in the cohort sheet.

Top

Active and Inactive factors

StoTree will form the cohort sheet using only combinations of states from *active* factors. You can use the active/ inactive toggle button in each factor worksheet to change a factor from active to inactive and vice versa. This toggle button lies in the upper left corner of each factor worksheet (for example, see Figure 9).

You might, for example, wish to deactivate a factor rather than deleting it if you think you might use it later and do not wish to have to reconstruct it.

Quality and cost for factored trees

In order to perform cohort analysis for a factored stochastic tree, *StoTree* must assign quality coefficients and cost rates to *combinations* of stochastic nodes. The rule *StoTree* uses to accomplish this is that the cost rate for a combination is the *sum* of the cost rates in each factor, and the quality coefficient for a combination is the *product* of the quality coefficients in each factor.

Tips and tricks for constructing factored stochastic trees

Use an additional factor instead of duplicate subtrees

As mentioned above, *StoTree* has *Copy Subtree/ Paste subtree* features for easily duplicating subtrees in any particular factor. A disadvantage of doing so, however, is that copy/paste can produce large bushy trees. For example, in Figure 27, the user has duplicated the *Test* subtree, because periodic screening applies to both the healthy and diseased populations. A more elegant approach is to create a separate factor depicting the outcomes of testing, and use dependencies to link its parameters back to health and disease. This approach is illustrated in Figure 28.
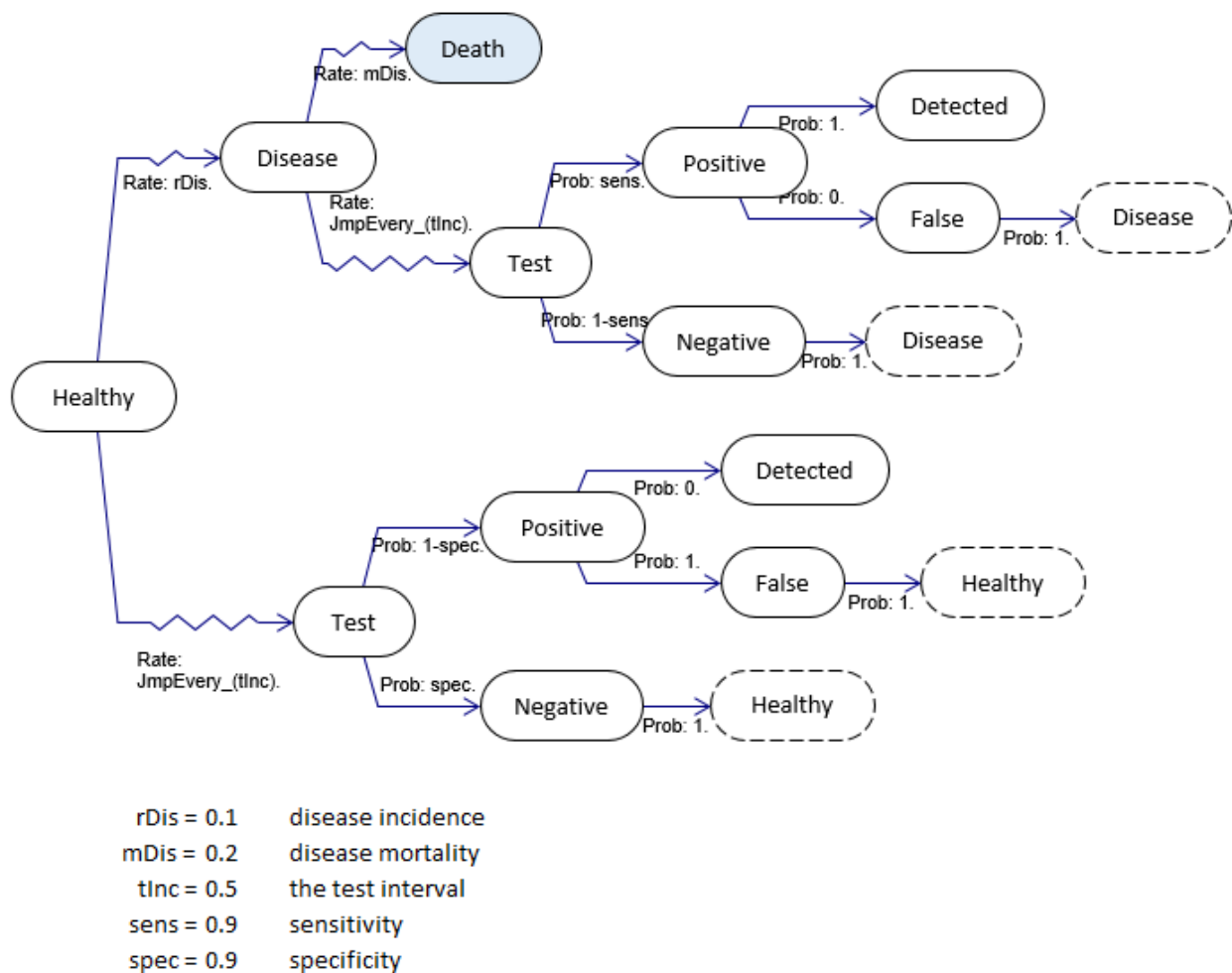

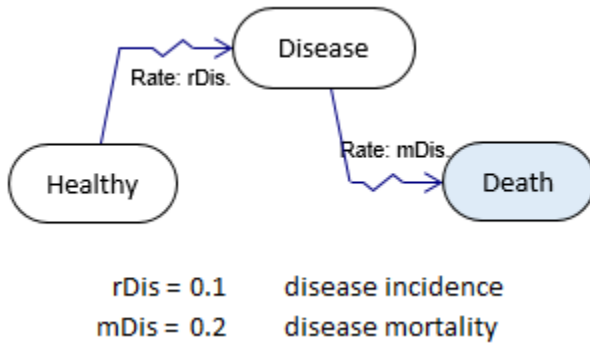|	
rDis = 0.1	disease incidence
mDis = 0.2	disease mortality
tInc = 0.5	the test interval
sens = 0.9	sensitivity
spec = 0.9	specificity

Figure 27. A generic stochastic tree for periodic screening. Note the Test *subtree is repeated.*

The advantage is a simpler and more elegant model formulation. However, the factored approach does require the use of dependencies at the *Test* branches and the *Positive* branches (see Figure 28). The resulting model details are thereby obscured.

Whether to factor in this way is largely a matter of taste – computationally there are no advantages. Had the parameters on the *Test* and *Positive* branches been independent of health or disease, dependencies would not be necessary and the factored approach would be superior. The general rule is to *consider factoring when you find yourself duplicating subtrees*.

Factor *Dis*

rDis = 0.1   disease incidence
mDis = 0.2   disease mortality

Factor *Test*

tInc = 0.5   the test interval
sens = 0.9   sensitivity
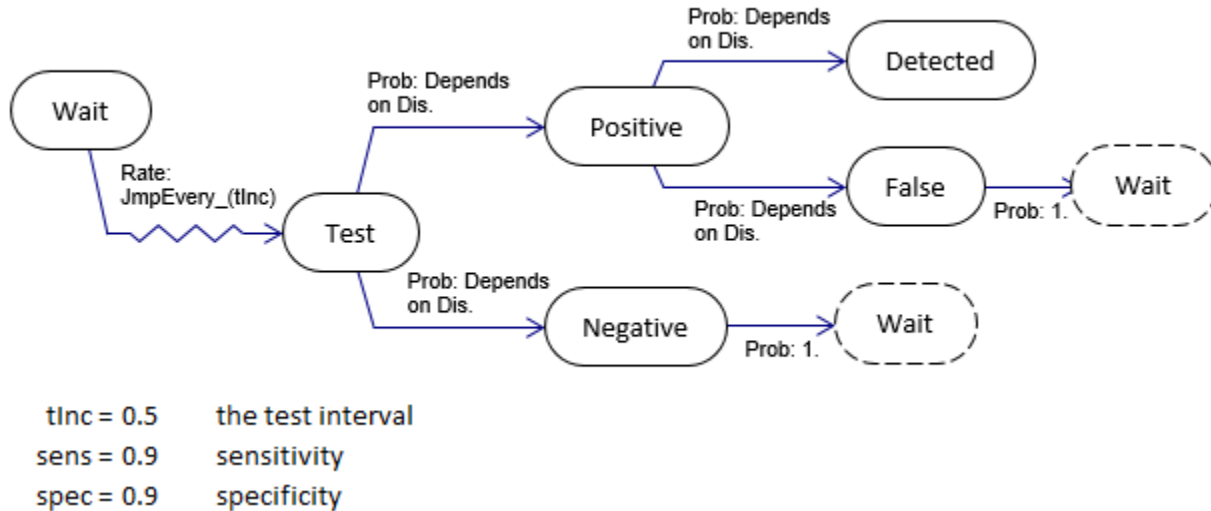spec = 0.9   specificity

Figure 28. A factored approach to model periodic screening. Compare to Figure 27. The Test *subtree has been split off into its own factor.*

Dependencies may only reference stochastic nodes

You may notice that the dependency userform only allows you to access stochastic nodes in each factor. Recall that chance nodes, being instantaneous, have zero duration. They therefore cannot influence rates, probabilities, *QALY*s or costs in other factors.

Top

The benefits of factoring

The ability to handle multiple factors is one of the primary strengths of *StoTree*. Formulating factored stochastic trees has several advantages:

- It simplifies the model formulation process. The analyst can focus attention on one component (factor) of the model at a time, refining it as needed, and only later worry about other components and interactions between components.

- It allows the easy formulation of large Markov models. The models shown in this document are small examples. In our experience, models containing tens or hundreds of states may easily arise. These are readily formulated in factored form because most factors contain on the order of 2–5 states. Formulating large models such as these is impractical without factoring.

- Greater *transparency* in the presentation of the model to others. Even those not versed in the theory of continuous-time Markov decision processes can gain an intuitive graphical understanding of simple 2- to 5-state factors of most models.

- It allows the easy swapping or adding of model components. The *StoTree* software allows the analyst to add, remove, or substitute model components (factors) as needed. This opens up the possibility that model components formulated and populated by other analysts may be imported, or that repositories of model components may be created for the community of modelers.

Top

Built-in functions

StoTree contains several built-in functions you may find useful. You can use these functions in any Excel formula you type into a cell, or in any dialog box that arises after you click on the Parameters button on the *StoTree* ribbon.

Human mortality rates

As mentioned above, *StoTree* includes a function *rMortUS* that returns annual human mortality rates in the United States for any age, sex, and race. As of this writing, these are for the year 2017, and were obtained from "United States Life Tables, 2017", *National Vital Statistics Reports* Vol. 68, No.7, June 24, 2019. See https://www.cdc.gov/nchs/data/nvsr/nvsr68/nvsr68_07-508.pdf.

The syntax of the function is:

rMortUS(Age [,Sex] [,Race])

Part	Description
Age	Any numeric value or numeric variable.
Sex	If present, must be either "Male" or "Female" (quotes included) or a variable whose contents is one of these strings.
Race	If present, must be either "White" or "Black" (quotes included) or a variable whose contents is one of these strings.

The expression returns a numeric value indicating the average annual mortality rate per person for the age, sex, and race specified. When Sex or Race or both are omitted, the mortality rate returned is for the combination of all relevant populations. For example, rMortUS(35,,"White") returns the average annual mortality per person for a 35 year-old whites, either male or female.

Weibull hazard rate function

The Weibull is a commonly used distribution for modeling survival. With parameters $\lambda > 0$ and $\gamma > 0$, it's hazard rate function is

$$h(t) = \lambda \gamma t^{\gamma - 1} \qquad t > 0.$$

The corresponding survival function is

$$S(t) = P(T > t) = \exp(-\lambda t^{\gamma}) \qquad t > 0.$$

The syntax of the function is:

WeiHR(t, Lambda, Gamma)

Part	Description
t	Any numeric value or numeric variable.
Lambda	Any numeric value or numeric variable.
Gamma	Any numeric value or numeric variable.

The expression returns a numeric variable or the string "Error". "Error" is returned if t is not a number or is a number with a nonpositive value, or if t is equal to zero and the value of Gamma is less than 1.

Top

References

1. Beck, J., Pauker SG, *The Markov process in medical prognosis.* Med Decis Making, 1983. **3**: p. 419-453.
2. Beck, J., Sonnenberg FA, *Markov models in medical decision making - a practical guide.* Med Decis Making, 1993. **13**: p. 322-338.
3. Briggs, A., M. Sculpher, and K.C. Claxton, *Decision Modeling for Health Economic Evaluation*. Handbooks in Health Economic Evaluation, ed. A.G.a.A. Briggs. 2006, Oxford: Oxford.
4. Hazen, G.B., *Stochastic trees: a new technique for temporal medical decision modeling.* Med Decis Making, 1992. **12**(3): p. 163-78.
5. Hazen, G.B., *Factored stochastic trees: a tool for solving complex temporal medical decision models.* Med Decis Making, 1993. **13**(3): p. 227-36.
6. Hazen, G.B., *Stochastic trees and the StoTree modeling environment: models and software for medical decision analysis.* J Med Syst, 2002. **26**(5): p. 399-413.
7. Collett, D., *Modelling Survival Data in Medical Research*. Second ed. 2003, Boca Raton, Florida: Chapman & Hall/ CRC.
8. van Rosmalen, J., M. Toy, and J.F. O'Mahony, *A Mathematical Approach for Evaluating Markov Models in Continuous Time without Discrete-Event Simulation.* Medical Decision Making, 2013. **33**: p. 767-779.
9. Tsevat, J., et al., *Warfarin for dilated cardiomyopathy: a bloody tough pill to swallow?* Med Decis Making, 1989. **9**(3): p. 162-9.