Chapter 9
Building a Markov Cost Effectiveness Model in Excel

Abstract Within previous chapters, we have introduced you to Markov models. The objective of this and the next chapter is to provide you with an opportunity to bring together much of the material we have covered to this point by building a Markov cost effectiveness model. Working your way through this chapter, you will construct the deterministic model. We illustrate the model building process using Microsoft Excel. The model chosen is a stylised representation of a chronic disease which is characterised by periods of controlled disease with periodic disease flairs and an accumulation of long-term disability. This type of structure would suit conditions such as rheumatoid arthritis, cardiovascular disease and diabetes. We are going to consider the simplest form of cost effectiveness analysis – a comparison of two alternative treatments in terms of their costs and outcomes.

9.1 Introduction

The objective of this and the next chapter is to provide you with an opportunity to bring together much of the material we have covered to this point by building a cost effectiveness model. Working your way through this chapter, you will construct the deterministic model. Chapters 10, 11 and 12 will then provide an opportunity to modify the model to provide probabilistic results and then produce a range of outputs that reimbursement decision makers may find helpful. Section 9.2 describes the conceptual model that you are going to build. Section 9.3 provides some useful tips for building models in Excel and requires you to build the transition. Section 9.4 deals with constructing the parameter table and the fields for capturing the Markov trace data. Section 9.5 then requires you to programme your model to generate the Markov trace data. In Sect. 9.6, you will add the costs, utilities and a discount rate to your model, before programming the calculation of the deterministic incremental cost effectiveness ratio (ICER) in Sect. 9.7.

R. Edlin et al., *Cost Effectiveness Modelling for Health Technology Assessment: A Practical Course*, DOI 10.1007/978-3-319-15744-3_9

9.2 The Model

First, we need to introduce you to the conceptual model that you are going to build. We have not chosen a specific disease, but rather a stylised representation of a chronic disease which is characterised by periods of controlled disease with periodic disease flairs and an accumulation of long-term disability. This type of structure would suit conditions such as rheumatoid arthritis, cardiovascular disease and diabetes. We are going to consider the simplest form of cost effectiveness analysis – a comparison of two alternative treatments in terms of their costs and outcomes.

Whilst this chapter covers the major steps in building a deterministic model in Excel, to provide additional support, we have a series of worked exercises using Excel Workbooks available from http://medhealth.leeds.ac.uk/costeffectivenessmodelling. A total of five exercises are provided that (a) fill in the transition matrix, (b) and (c) track a cohort of patients across time, (d) add a discount rate to the model; and (e) add a half-cycle correction. These exercises track the tasks set out in this chapter but provide a more hands-on approach.

Real-world models for any of these diseases would break out the uncontrolled disease state into multiple states, as well as having multiple disability states. However, we do not require this added complexity for the purpose of demonstrating the techniques covered in this book. Figure 9.1 shows the influence diagram for the model. For convenience, we have provided labels for the transition probabilities that will be required in the model. We will use these labels throughout the exercise, and it may be useful to use these as ‘names’ in Excel as you build the model.

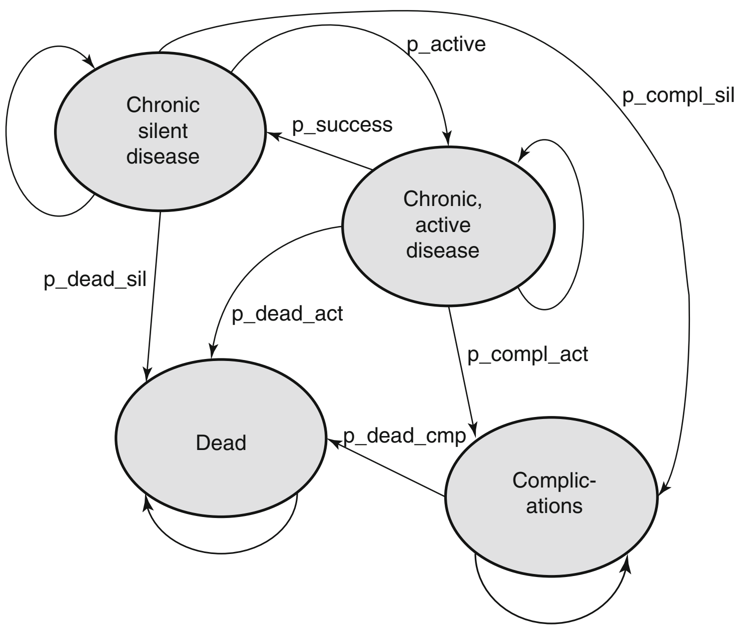

Fig. 9.1 Influence diagram for the Markov model

Table 9.1 Parameter table

Transition probabilities	**Utility parameters**
Chronic silent to chronic active (p_active)	Utility in chronic silent disease (u_silent)
Chronic silent to complications (p_compl_sil)	Utility in chronic active disease (u_active)
Chronic silent to death (p_dead-sil)	Utility in complications (u_comp)
Chronic active to chronic silent (p_success)	Utility in death (u-death)
Chronic active to chronic complications (p_compl_act)	
Chronic active to death (p_dead_act)	
Complications to death (p_dead_cmp)	
Cost parameters	
Cost in chronic silent (c_silent)	Cost in complications (c_comp)
Cost in chronic active (c_active)	Cost in death (c_death)

Table 9.2 Model transition matrix

		Outcome at end of period/start of next period			
		Silent	Active	Compl	Dead
Outcome at start of period	Silent	1-(p_silent + p_compl_sil + p_dead_sil)	p_silent	p_compl_sil	p_dead_sil
	Active	p_success	1 – (p_success + p_compl_act + p_dead_act)	p_compl_act	p_dead_act
	Compl	0	0	1-p_dead_cmp	p_dead_cmp
	Dead	0	0	0	1

Remember, in addition to transition probabilities, each state will need cost and utility data. Can you construct a full model parameter list (Step 1)? You might also want to identify a consistent naming structure for all parameters.

Does your parameter table look like Table 9.1? If it doesn't, what parameters did you miss and why? Did you have any parameters that we have not listed? What type of state is Dead in this model?

Once you have constructed the list of parameters that will be required for the model, you need to identify the transition matrix, which will be the model's engine (Step 2). Using the variable names in Table 9.1, construct the model transition matrix and then compare it to our version in Table 9.2.

If you have successfully produced a transition matrix that looks like Table 9.2 and you are confident you understand why it should look like that, then you are ready to start constructing the model in Excel. To help you construct the model efficiently, we are going to go through some 'good design' principles for modelling in Excel.

9.3 Modelling in Excel

Remarkably, the construction of a probabilistic Markov model in Excel uses relatively few functions and relies quite heavily on the use of some simple Visual Basic macros. However, those functions and macros will generate a great deal of data, and the inexperienced modeller can easily lose track of what is happening and mistakes can easily creep in and go undetected. Experience has shown that a few simple strategies can help keep control of the data and make model construction and debugging much easier.

Our first recommendation is to have a single worksheet containing all the parameters in the model. By using the naming facility within Excel and then referring to parameters by their names in the formulas, any modification of the parameter values needs only to be done once, on the parameters worksheet. The amended parameter value will then be used by all formulas in the model. This avoids the need to amend the parameter in every formula that uses it. This approach hardwires quality control into the model construction and saves huge amounts of time when changing parameter values in the light of new evidence.

Step 3 in this exercise is to open up a new Excel worksheet and construct a parameter table – reflecting the information in the parameter table you constructed in Step 1. Remember, eventually, we are going to construct a probabilistic model, so you will want to create a parameter table that has columns for the moments required to specify the parameter distribution. Think about the types of distributions we use for each type of data in order to choose which moments may need to be specified for each type of data. You will also want to name the worksheet so that any future user of the model will know that this first worksheet contains the parameter table.

Figure 9.2 shows the parameter sheet in our model. As you can see, we have rows for each of the model parameters, and for each model parameter, we have allowed space to specify the distribution parameters. Those parameters we expect to be different for the two treatments have two rows, one for each treatment-specific parameter. Did you remember to put in extra rows for the treatment-specific parameters? Our parameter table also includes a column for the deterministic parameter values and another column called 'value used'. By doing this, we will be able to compare the results of our deterministic and probabilistic analyses. How will we tell Excel which values to use for the analysis? Hint: an 'IF statement' that references a 'model-type' cell will be useful.

The final component of Fig. 9.2 that you may not have included in your parameter table is the 'Refers to' column. This is included to help the user rather than the person who built the model. It simply tells the user which interventions the parameter applies to. Hence, it takes three values: Treatment A, Treatment B and Common, the latter being for parameters that are the same for both treatments.

Figure 9.3 is also taken from our parameters worksheet, and it shows a very useful quality control check. For each parameter, we record the mean and standard deviation from the data used to construct the distribution, and next to it,

Model control				**opt_modeltype**						
opt_modeltype	Model Type	Deterministic		Deterministic						
				Stochastic						
Transition parameters								**Distribution**		
Parameter	**Description**	**Refers to**	**Value used**	**Deterministic**	**Stochastic**	**Random**	**Dist**	**Param 1**		**Param 2**
p_active	Probability of active disease from silent disease	Treatment A								
		Treatment B								
p_success	Probability of silent disease from active disease	Treatment A								
		Treatment B								
p_compl_sil	Probability of complications from silent disease	Treatment A								
		Treatment B								
p_compl_act	Probability of complications from active disease	Treatment A								
		Treatment B								
p_dead_sil	Probability of death from chronic, silent disease	Treatment A								
		Treatment B								
p_dead_act	Probability of death from chronic, active disease	Treatment A								
		Treatment B								
p_dead_cmp	Probability of death from complications state	Common								
Cost parameters										
Parameter	**Description**	**Refers to**	**Value used**	**Deterministic**	**Stochastic**	**Random**	**Dist**	**Param 1**		**Param 2**
Chronic,silent disease	Health state specific costs	Common								
Chronic,active disease	Health state specific costs	Common								
Complications	Health state specific costs	Common								
c_TxA	Cost of treament A	Treatment A								
c_TxB	Cost of teament B	Treatment B								
Utility parameters										
Parameter	**Description**	**Refers to**	**Value used**	**Deterministic**	**Stochastic**	**Random**	**Dist**	**Param 1**		**Param 2**
Chronic,silent disease	Health state specific Utility	Common								
Chronic,active disease	Health state specific Utility	Common								
Complications	Health state specific Utility	Common								

Fig. 9.2 Parameter table structure

Transition parameters	**From Data**		**From dist**	
Parameter	**Mean**	**SD**	**Mean**	**SD**
p_active				
p_success				
p_compl_sil				
p_compl_act				
p_dead_sil				
p_dead_act				
p_dead_cmp				
Cost parameters				
Parameter	**Mean**	**SD**	**Mean**	**SD**
Chronic, silent disease				
Chronic, active disease				
Complications				
c_TxA				
c_TxB				
Utility parameters				
Parameter	**Mean**	**SD**	**Mean**	**SD**
Chronic, silent disease				
Chronic, active disease				
Complications				

Fig. 9.3 Data source and prediction quality control in parameter sheet

we have cells to capture the mean and standard deviation predicted in the model. Whilst there is likely to be (very) small differences due to sampling, if these data and the predictions are substantially different, then there is a problem with the model that must be resolved before the results can be used with any confidence. We would recommend that you add this to your parameters worksheet also.

9.4 Constructing the Parameter Table

Now, having created a parameter table, we need to create a worksheet that has the transition matrices. We want this to be a separate sheet, so that users who want to examine the transition matrix do not have to search for it within a worksheet that contains other information. We suggest you call this worksheet 'transition matrix'. If you name the cells in your parameters worksheet using the same labels as you used to complete Table 9.2, then constructing your transition matrix is as simple as transposing Table 9.2 into the 'transition matrix' worksheet. Remember to put the '=' sign before the text, so that Excel recognises them as formulas.

As well as the transition matrix, you will need to create a space for the Markov traces for each treatment to be recorded over the time horizon of the model (Step 4). Call the treatments Treatment A and Treatment B. The model we are going to build will have a 5-year time horizon and a 1-month cycle.

a

	I	J	K	L	M
3		Silent	Active	Complications	Dead
4	Silent	=1-K4-L4-M4	=p_active_A A	=p_compl_sil_A	=p_dead_sil_A
5	Active	=p_success_A	=1-J5-L5-M5	=p_compl_act_A	=p_dead_act_A
6	Complications	0	0	=1-M6	=p_dead_cmp_A
7	Dead	0	0	0	1

b

Treatment A														
		Probabilities					Costs				Outcomes			
Years	Period	Silent	Active	Compl	Dead	discount	Silent	Active	Compl	Dead	Silent	Active	Compl	Dead
0.00	1													
0.01	2													
0.02	3													
0.03	4													
0.04	5													
0.05	6													

Fig. 9.4 (**a**) Transition matrix and Markov trace worksheet structure. (**b**) Markov trace structure for treatment A

- How many rows will you need to capture the Markov trace?
- How many columns will you need to capture the Markov trace for each treatment?
- What will be the heading for each column in the Markov trace?

Remember that the total costs and outcomes for each treatment are driven by the Markov trace; your structure needs to include space to capture these data.

Figure 9.4a shows our transition matrix worksheet for Treatment A. The structure for Treatment B is identical. The Markov trace table, Fig. 9.4b, consists of three sub-tables, one each for the distribution of the cohort, the costs of care and the outcomes (QALYs). In each sub-table, we have one column for each of the four states in the model.

If you look carefully, you will notice we have also included a column for the discount rate. Adding a discount rate to your worksheet is the next step in the exercise. Make sure you add it to the trace data for both treatments.

9.5 Programming Your Model

Having established the structure to capture the Markov trace data, it is time to write the functions that will actually generate that data. This is Step 5 in our exercise. At the centre of a Markov model is the application of the transition matrix to the cohort. First, you need to fill in the formulas for the transition matrix. If you have named the parameter cells in your parameters worksheet, then you can simply use the names and complete the transition matrix for Treatment A using the information in Table 9.2 above. Remember, Excel will want you to use an '=' sign at the start of each calculation so that it knows you are asking it to do a calculation. Otherwise, it will think you are simply entering text.

Table 9.3 gives you the deterministic parameter values that you will need to use in your parameters worksheet in order for the transition matrix formulas to return actual values. You will need to reference these values in the 'Value Used' cells in the parameter sheet and give them the appropriate names for use in subsequent formulas. Here, you will need to use the 'IF statement" that tells Excel where to look for the data depending upon whether the model is deterministic or probabilistic. You will need to do this for the transition matrix parameters and the cost and utility parameters. We will provide the data for these parameters shortly.

Before we can generate the Markov trace for the distribution of the cohort across the four health states in the model, we need to specify the starting distribution, i.e. how the cohort of patients is distributed across the four states at the start of the model. We are going to assume that all the patients are in the active disease state at the start of the model. Therefore, in Period 1, we put the value 1.0 in the active disease column and 0.0 in the other four columns.

Step 6 is to apply the transition matrix to the cohort to create the Markov trace for all 60 periods of the model time horizon. Remember that the transition matrix and the distribution of the cohort are arrays that can be multiplied in Excel using the matrix multiplication function =MMULT. Remember to allow the cell references in the multiplication to change with each period of the model.

Table 9.4 shows the formulas we have used to generate the trace over the first five periods. Your cell references may vary if you have positioned the Markov trace differently in your worksheet. In order to interpret the formulas, note that J22:M22 contain the Period 1 probabilities for the four states (the second row of the table); J23:M23 contain the Period 2 probabilities for the four states (third row), and so on. Here, $J$4:$M$7 contains the transition matrix, as in Fig. 9.4a. Note that all the

Table 9.3 Parameter values for the transition matrices

Probabilities	Treatment A	Treatment B
p_active	0.094	0.093
p_success	0.194	0.212
p_compl_sil	0.012	0.014
p_compl_act	0.144	0.098
p_dead_sil	0.002	0.002
p_dead_act	0.001	0.001
p_dead_cmp	0.003	0.003

Table 9.4 Functionality for Markov trace showing distribution of cohort

	Probabilities			
Period	Silent	Active	Compl	Dead
1	0	1	0	0
2	{=MMULT(J22:M22, $J$4:$M$7)}	{=MMULT(J22:M22, $J$4:$M$7)}	{=MMULT(J22:M22, $J$4:$M$7)}	{=MMULT(J22:M22, $J$4:$M$7)}
3	{=MMULT(J23:M23, $J$4:$M$7)}	{=MMULT(J23:M23, $J$4:$M$7)}	{=MMULT(J23:M23, $J$4:$M$7)}	{=MMULT(J23:M23, $J$4:$M$7)}
4	{=MMULT(J24:M24, $J$4:$M$7)}	{=MMULT(J24:M24, $J$4:$M$7)}	{=MMULT(J24:M24, $J$4:$M$7)}	{=MMULT(J24:M24, $J$4:$M$7)}
5	{=MMULT(J25:M25, $J$4:$M$7)}	{=MMULT(J25:M25, $J$4:$M$7)}	{=MMULT(J25:M25, $J$4:$M$7)}	{=MMULT(J25:M25, $J$4:$M$7)}

formulas in Table 9.4 are surrounded by curly brackets "{ }". This tells Excel that is should calculate the four probabilities in each row together as an array.

9.6 Adding a Discount Rate, Costs and Utilities

You should now have data showing the distribution of the cohort over all 60 cycles (5 years) of the model. The next task is to add in the calculation for the discount rate in the relevant column. Remember that, by convention, we apply the same discount rate to each year and the first year is not discounted. Check back to Chap. 5 (Sect. 5.6) to get the formula. The Excel symbol for raising something to a power is '^'. You may want to add the discount rate as a parameter in the parameter table. This will make it easier to examine whether varying the discount rate impacts upon your results. If you do add discount rate into the parameter table, remember to name the cell and use that in your functions. For the primary analysis, use a discount rate of 5 % per annum for both costs and benefits.

Once you have added in the discount rate, you need to go through the same process for Treatment B: construct the transition matrix, set the initial distribution for the cohort, construct the Markov trace for the distribution of the cohort across the states and then add in the discount rate.

Table 9.5 gives the monthly costs and utilities for each state in the model. You need to add these into your parameter table and then complete the Markov trace data

Table 9.5 Costs and utility parameters

	Treatment A	Treatment B
Cost		
Silent	$61.52	$926.52
Active	$122.43	$987.43
Comp	$397.00	$397.00
Dead	$0.00	$0.00
QALY parameters		
Silent	0.068	0.068
Active	0.054	0.054
Comp	–0.011	–0.011
Dead	0.000	0.000

							Silent	Active	Compl	Dead	Silent	Active	Compl	Dead
Totals	1830	1107	5.81	39.42	3.70		653.18	695.25	14461.28	0.00	0.72	0.31	–0.42	0.00
									TOTAL COST	$15,809.71			TOTAL DAYS	0.62

Fig. 9.5 Markov trace totals

Silent	Active	Compl	Dead
=0.5*O22+SUM(O23:O81) +0.5*O82	=0.5*P22+SUM(P23:P81) +0.5*P82	=0.5*Q22+SUM(Q23:Q81)+ 0.5*Q82	=0.5*R22+SUM(R23:R81) +0.5*R82
		Total Cost	=O90+P90+Q90+R90

Fig. 9.6 Formulas for half-cycle correction of costs

		Costs	Outcomes	Net benefit
Treatment A		$15,609	1.184	$19,903
Treatment B		$19,586	1.278	$18,740
Incremental		$3,977		
		for	0.094	QALYs
ICER		$42,411.02	per QALY	

Fig. 9.7 Expected incremental cost effectiveness and net benefit (deterministic)

for costs and outcomes. If you set up the formulas for calculating the costs and benefits in the first period, you should be able to simply copy these down for all periods in the model. Once you have done this, you need to calculate the total time, costs and benefits in each state over the time horizon of the model. We recommend locating these calculations at the bottom of the Markov trace columns. Figure 9.5 shows this information for Treatment A. You will need to do this for both Treatment A and B.

9.7 Adding the Calculation of the Deterministic Incremental Cost Effectiveness Ratio (ICER)

At this point, you have very nearly constructed a complete deterministic Markov cost effectiveness model. All that remains is to apply the half-cycle correction to the Markov trace data. Refer back to Chap. 5 to remind yourself why we apply a half-cycle correction and then add a 'half-cycle corrected' set of Markov trace totals below the uncorrected totals that you have just calculated. Figure 9.6 shows the formulas we used to calculate the half-cycle corrected cost totals in our model. Your cell references may differ depending upon where you have located the Markov trace data in your worksheet, but the formulas should be the same. In our worksheet, Row 22 contains the first (half) period's results and Row 82 contains the extra (half) period's results used when calculating the half-cycle correction.

Now you need to calculate the ICER and the Net Benefit. We suggest you set lambda = $50,000 per QALY. Having set our hypothetical cohort as one person, the total costs and outcomes are in fact the expected costs and outcomes, so you can use these results to calculate the ICER. Before you do that however, we suggest that you create a third worksheet: the results worksheet. Name the cells that have the expected costs and outcomes in the Markov trace worksheet, so that you can use these names in the formulas on your results worksheet. Figure 9.7 shows the results for our model.

The structure for your cost effectiveness model is now complete. In the next chapter, you will replace the deterministic parameters with probability distributions.

9.8 Summary

- We have outlined 6 steps in constructing a deterministic cost effectiveness model:
  - Step 1: Construct a full model parameter list.
  - Step 2: Construct the transition matrix.
  - Step 3: Construct a parameter table – reflecting the information in the parameter table you constructed in Step 1.
  - Step 4: In addition to the transition matrix constructed in Step 2, create a space for the Markov traces for each treatment to be recorded over the time horizon of the model. Include sub-tables, one each for the distribution of the cohort, the costs of care and the outcomes (QALYs). If applicable, add a discount rate to your worksheet. Make sure you add it to the trace data for all treatments.
  - Step 5: Write the functions that will generate that data. You need to fill in the formulas for the transition matrix; ensure you specify the starting distribution.
  - Step 6: Apply the transition matrix to the cohort to create the Markov trace for all periods of the model time horizon. Ensure you include the calculation for the discount rate in the relevant column and, if applicable, apply the half-cycle correction to the Markov trace data. Calculate the ICER and Net Benefit.