Chapter 2

Key aspects of decision modelling for economic evaluation

This chapter considers the basic elements of decision modelling for economic evaluation. It considers the key stages in developing a decision analytic model and describes the cohort model, the main type of decision model used in the field. The decision tree and Markov model are described in detail and examples provided of their use in economic evaluation.

2.1. The stages of developing a decision model

It is possible to identify a series of stages in developing a decision model for economic evaluation. In part, this will involve some general choices concerning the nature of the evaluation. This will include the measure of effect and the time horizon, but also the perspective of the analysis; that is, whose costs and effects we are interested in? Below is a list of the stages in the development process which relate specifically to the decision modelling.

2.1.1. Specifying the decision problem

This involves clearly identifying the question to be addressed in the analysis. This requires a definition of the recipient population and subpopulations. This will typically be the relevant patients, but may include nonpatients (e.g. in the case of screening and prevention programmes). This requires specific details about the characteristics of individuals, but should also include information about the locations (e.g. the UK NHS) and setting (e.g. secondary care) in which the options are being delivered. The specific options being evaluated also need to be detailed. These will usually be programmes or interventions, but could include sequences of treatments with particular starting and stopping rules.

Part of the definition of the decision problem relates to which institution(s) is/are (assumed to be) making the relevant decision. In some cases this will be explicitly stated – for example, in the case of a submission to a reimbursement agency – but it will often have to be implied by the characteristics of the evaluation, such as the sources of data used.

2.1.2. Defining the boundaries of the model

All models are simplifications of reality and it will never be possible for a model to include all the possible ramifications of the particular option being considered. Choices need to be taken, therefore, about which of the possible consequences of the options under evaluation will be formally modelled. For example, should the possible implications of antibiotic resistance be assessed in all economic evaluations of interventions for infectious diseases? Another example relates to whether or not to include changes in disease transmission resulting from screening programmes for HIV. It has been shown that including reductions in the horizontal transmission of HIV in such models has a marked impact on the cost-effectiveness of screening (Sanders *et al.* 2005).

2.1.3. Structuring a decision model

Given a stated decision problem and set of model boundaries, choices have to be made about how to structure the possible consequences of the options being evaluated. In part, this will be based on the nature of the interventions themselves. For example, for an economic evaluation of alternative diagnostic strategies for urinary tract infection in children, it was necessary to use a complex decision tree to reflect the prior probability (prevalence) and diagnostic accuracy (sensitivity and specificity) of the various single and sequential screening tests (Downs 1999). In part, model structure will reflect what is known about the natural history of a particular condition and the impact of the options on that process; for example, the future risks faced by patients surviving a myocardial infarction and the impact of options for secondary prevention on those risks.

As a general approach to structuring a decision model, there is value in having some sort of underlying biological or clinical process driving the model. Examples of the former include the use of CD4 counts or viral load in HIV models (Sanders *et al.* 2005). The latter approach is more common and examples include the use of the Kurtzke Expanded Disability Status Scale in multiple sclerosis (Chilcott *et al.* 2003), the Mini Mental State Examination in Alzheimer's disease (Neumann *et al.* 1999) and clinical events, such as myocardial infarction and revascularization in coronary heart disease (Palmer *et al.* 2005). The cost-effectiveness of the relevant interventions can then be assessed by attaching health-related quality-of-life weights and costs to states or pathways defined in this way. The advantage of using these biologically- or clinically-defined states is that they should be well-understood. In particular, there should be good evidence about the natural history of a disease in terms of

these definitions. This is particularly important when modelling a baseline (e.g. disease progression without treatment) and in extrapolating beyond the data from randomized trials.

There are no general rules about appropriate model structure in a given situation. However, some of the features of a disease/technology that are likely to influence choices about structure include:

- Whether the disease is acute or chronic and, if the latter, the number of possible health-related events which could occur over time.
- Whether the risks of events change over time or can reasonably be assumed to be constant.
- Whether the effectiveness of the intervention(s) (relative to some usual care baseline) can be assumed constant over time or time-limited in some way.
- If and when treatment is stopped, the appropriate assumptions about the future profile of those changes in health that were achieved during treatment. For example, would there be some sort of 'rebound' effect or would the gains, relative to a comparator group, be maintained over time (Drummond *et al.* 2005).
- Whether the probability of health-related events over time is dependent on what has happened to 'a patient' in the past.

2.1.4. Identifying and synthesizing evidence

The process of populating a model involves bringing together all relevant evidence, given a selected structure, and synthesizing it appropriately in terms of input parameters in the model. Consistent with the general principles of evidence-based medicine (Sackett *et al.* 1996), there needs to be a systematic approach to identifying relevant evidence. Evidence synthesis is a key area of clinical evaluation in its own right (Sutton *et al.* 2000) which is of importance outside the requirements of economic evaluation. However, the requirements of decision analytic models for economic evaluation have placed some important demands on the methods of evidence synthesis. These include:

- The need to estimate the effectiveness of interventions despite the absence of head-to-head randomized trials. This involves the use of indirect and mixed treatment comparisons to create a network of evidence between trials.
- The need to obtain probabilities of clinical events for models over a standardized period of follow-up despite the fact that clinical reports present these over varying follow-up times.

- The need for estimates of treatment effectiveness in terms of a common endpoint although trials report various measures.
- The need to assess heterogeneity in measures between different types of patients. Ideally this would be undertaken using individual patient data, but metaregression can be used with summary data in some situations.

These issues in evidence synthesis are being tackled by statisticians, often within a Bayesian framework (Sutton and Abrams 2001; Ades 2003; Spiegelhalter *et al.* 2004), and these are increasingly being used in decision models for economic evaluation (Ades *et al.* 2006). An important area of methodological research in the field relates to incorporating evidence synthesis and decision modelling into the same analytic framework – 'comprehensive decision modelling' (Parmigiani 2002; Cooper *et al.* 2004). This has the advantage of facilitating a fuller expression of the uncertainty in the evidence base in the economic evaluation.

2.1.5. Dealing with uncertainty and heterogeneity

Uncertainty and heterogeneity exist in all economic evaluations. This is an area of economic evaluation methodology that has developed rapidly in recent years (Briggs 2001), and its implications for decision modelling represent an important element of this book. Chapters 4 and 5 provide more detail about appropriate methods to handle uncertainty and heterogeneity. Box 2.1 summarizes the key concepts, and these are further developed in Chapter 4.

2.1.6. Assessing the value of additional research

The purpose of evaluative research, such as randomized control trials, is to reduce uncertainty in decision making by measuring one or more parameters (which may be specific to particular subgroups) with greater precision. This is generally true in clinical research, but also in assessing cost-effectiveness. Given limited resources, it is just as appropriate to use decision analytic models to assess the value for money of additional research projects as to assess alternative approaches to patient management. In quantifying the decision uncertainty associated with a particular comparison, decision models can provide a framework within which it is possible to begin an assessment of the cost-effectiveness of additional research. This can be undertaken informally using simple sensitivity analysis by assessing the extent to which a model's conclusions are sensitive to the uncertainty in one (or a small number)

Box 2.1. Key concept in understanding uncertainty and heterogeneity in decision models for cost-effectiveness analysis

Variability: Individual patients will inevitably differ from one another in terms, for example, of the clinical events that they experience and the associated health-related quality of life. This variability cannot be reduced through the collection of additional data.

Parameter uncertainty: The precision with which an input parameter is estimated (e.g. the probability of an event, a mean cost or a mean utility). The imprecision is a result of the fact that input parameters are estimated for *populations* on the basis of limited available information. Hence uncertainty can, in principle, be reduced through the acquisition of additional evidence.

Decision uncertainty: The joint implications of parameter uncertainty in a model result in a distribution of possible cost-effectiveness relating to the options under comparison. There is a strong normative argument for basing decisions, given available evidence, on the expectation of this distribution. But the distribution can be used to indicate the probability that the correct decision has been taken.

Heterogeneity: Heterogeneity relates to the extent to which it is possible to explain a proportion of the interpatient variability in a particular measurement on the basis of one or more patient characteristics. For example, a particular clinical event may be more likely in men and in individuals aged over 60 years. It will then be possible to estimate input parameters (and cost-effectiveness and decision uncertainty) conditional on a patient's characteristics (subgroup estimates), although uncertainty in those parameters will remain.

of parameters. Formal value-of-information methods are considered fully in Chapters 6 and 7. These methods have the strength of reflecting the joint uncertainty in all parameters. They also assess the extent to which reduction in uncertainty through additional research would result in a change in decision about the use of a technology and, if there is a change, its value in terms of improved health and/or reduced costs.

Each of these stages is crucial to the development of a decision model that is fit for the purpose of informing real policy decisions.

2.2. Some introductory concepts in decision analysis

Decision analysis is based on some key 'building blocks' which are common to all models. These are covered more fully in introductory texts (Weinstein and Fineberg 1980; Hunink *et al.* 2001; Drummond *et al.* 2005), and are only summarized here.

2.2.1. Probabilities

In decision analysis, a probability is taken as a number indicating the likelihood of an event taking place in the future. As such, decision analysis shares the same perspective as Bayesian statistics (O'Hagan and Luce 2003). This concept of probability can be generalized to represent a strength of belief which, for a given individual, is based on their previous knowledge and experience. This more 'subjective' conceptualization of probabilities is consistent with the philosophy of decision analysis, which recognizes that decisions cannot be avoided just because data are unavailable to inform them, and 'expert judgement' will frequently be necessary.

Specific probability concepts frequently used in decision analysis are:

- *Joint probability*. The probability of two events occurring concomitantly. In terms of notation, the joint probability of events A and B occurring is P(A and B).
- *Conditional probability*. The probability of an event A given that an event B is known to have occurred. The notation is P(A|B).
- *Independence*. Events A and B are independent if the probability of event A, P(A), is the same as the probability of P(A|B). When the events are independent $P(A \text{ and } B) = P(A) \times P(B)$.
- Joint and conditional probabilities are related in the following equation: $P(A \text{ and } B) = P(A|B) \times P(B)$. Sometimes information is available on the joint probability, and the above expression can be manipulated to 'condition out' the probabilities.

2.2.2. Expected values

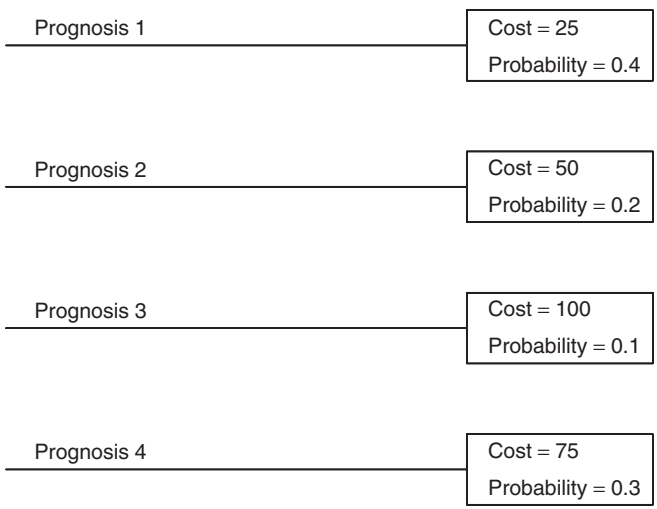
Central to the decision analytic approach to identifying a 'preferred' option from those being compared under conditions of uncertainty is the concept of expected value. If the options under comparison relate to alternative treatments for a given patient (or an apparently homogeneous group of patients), then the structure of the decision model will reflect the variability between patients in the events that may occur with each of the treatments. The probabilities will show the likelihood of those events for a given patient. On this basis, the

model will indicate a number of mutually exclusive 'prognoses' for a given patient and option (more generally, these are alternative 'states of the world' that could possibly occur with a given option). Depending on the type of model, these prognoses may be characterized, for example, as alternative pathways or sequences of states. For a given option, the likelihood of each possible prognosis can be quantified in terms of a probability, and their implications in terms of cost and/or some measure of outcome. The calculation of an expected value is shown in Box 2.2 using the example of costs. It is derived by adding together the cost of each of the possible prognoses weighted by the probability of it occurring. This is analogous to a sample mean calculated on the basis of patient-level data.

2.2.3. Payoffs

As described in the previous section, each possible 'prognosis' or 'state of the world' can be given some sort of cost or outcome. These can be termed 'payoffs', and expected values of these measures are calculated. The origins of

Box 2.2. An illustration of the concept of expected values using costs



Expected cost = $(25 \times 0.4) + (50 \times 0.2) + (100 \times 0.1) + (75 \times 0.3) = 52.50$

decision analysis are closely tied to those of expected utility theory (Raiffa 1968), so the standard payoff would have been a 'utility' as defined by von Neumann-Morgenstern (von Neumann and Morgenstern 1944). In practice, this would equate to a utility based on the standard gamble method of preference elicitation (Torrance 1986). As used for economic evaluation in health care, the payoffs in decision models have been more broadly defined. Costs would typically be one form of payoff but, on the effects side, a range of outcomes may be defined depending on the type of study (see Chapter 1). Increasingly, quality-adjusted life-years would be one of the payoffs in a decision model for cost-effectiveness analysis, which may or may not be based on utilities elicited using the standard gamble.

The principle of identifying a preferred option on the basis of a decision analytic model is on the basis of expected values. When payoffs are defined in terms of 'von Neumann-Morgenstern utilities', this would equate with a preferred option having the highest expected utility; this is consistent with expected utility theory as a normative framework for decision making under uncertainty. Although a wider set of payoffs are used in decision models for economic evaluation, the focus on expected values as a basis for decision making remains. This follows the normative theory presented by Arrow and Lind (1970) arguing that public resource allocation decisions should exhibit risk neutrality. For example, in cost-effectiveness analysis, the common incremental cost-effectiveness ratio would be based on the differences between options in terms of their expected costs and expected effects. However, the uncertainty around expected values is also important for establishing the value and design of future research, and this should also be quantified as part of a decision analytic model. The methods for quantifying and presenting uncertainty in models are described in Chapters 4 and 5, respectively; and the uses of information on uncertainty for research prioritization are considered in Chapters 6 and 7.

2.3. **Cohort models**

The overall purpose of a model structure is to characterize the consequences of alternative options in a way that is appropriate for the stated decision problem and the boundaries of the model. The structure should also be consistent with the key features of the economic evaluation, such as the perspective, time horizon and measure of outcome. There are several mathematical approaches to decision modelling from which the analyst can choose. One important consideration is whether the model should seek to characterize the experience of the 'average' patient from a population sharing the same characteristics, or

should explicitly consider the individual patient and allow for variability between patients. As described previously, the focus of economic evaluation is on expected costs and effects, and uncertainty in those expected values. This has resulted in most decision models focusing on the average patient experience – these are referred to as cohort models. In certain circumstances, a more appropriate way of estimating expected values may be to move away from the cohort model, to models focused on characterizing variability between patients. These 'micro simulation' models are discussed in Chapter 3, but the focus of the remainder of this chapter is on cohort models.

The two most common forms of cohort model used in decision analysis for economic evaluation are the decision tree and the Markov model. These are considered below.

2.3.1. The decision tree

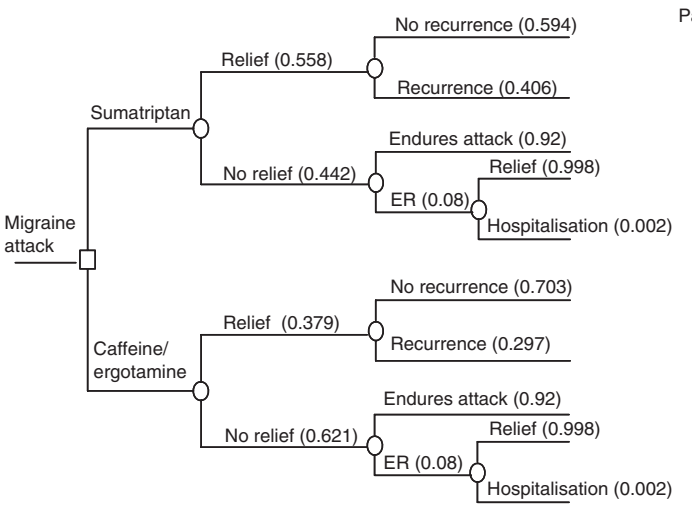
The decision tree is probably the simplest form of decision model. Box 2.3 provides a brief revision of the key concepts using a simple example from the management of migraine (Evans *et al.* 1997); the decision tree has been described in more detail elsewhere (Hunink *et al.* 2001; Drummond *et al.* 2005). The key features of a decision tree approach are:

- A square decision node – typically at the start of a tree – indicates a decision point between alternative options.
- A circular chance node shows a point where two or more alternative events for a patient are possible; these are shown as branches coming out of the node. For an individual patient, which event they experience is uncertain.
- Pathways are mutually exclusive sequences of events and are the routes through the tree.
- Probabilities show the likelihood of a particular event occurring at a chance node (or the proportion of a cohort of apparently homogeneous patients expected to experience the event). Moving left to right, the first probabilities in the tree show the probability of an event. Subsequent probabilities are conditional; that is, the probability of an event given that an earlier event has or has not occurred. Multiplying probabilities along pathways estimates the pathway probability which is a joint probability (as discussed previously).

Expected costs and outcomes (utilities in Box 2.3) are based on the principles in Box 2.2. Expected values are based on the summation of the pathway values weighted by the pathway probabilities.

A somewhat more complicated decision tree model comes from a cost-effectiveness analysis of alternative pharmaceutical therapies for

Box 2.3. Example of a decision tree based on Evans *et al*. (1997)



Pathway	Probability	Cost	Expected cost	Utility	Expected utility
Sumatriptan					
A	0.331	16.10	5.34	1.00	0.33
B	0.227	32.20	7.29	0.90	0.20
C	0.407	16.10	6.55	–0.30	–0.12
D	0.035	79.26	2.77	0.10	0.0035
E	0.0001	1172.00	0.11	–0.30	–0.00003
Total	**1.0000**		**22.06**		**0.41**
Caffeine/erogotamine					
F	0.266	1.32	0.35	1.00	0.27
G	0.113	2.64	0.30	0.90	0.10
H	0.571	1.32	0.75	–0.30	–0.17
I	0.050	64.45	3.22	0.10	0.0050
J	0.0001	1157.00	0.11	–0.30	–0.00003
Total	**1.0000**		**4.73**		**0.20**

gastro-oesophageal reflux disease (GORD) (Goeree *et al.* 1999). This model is described in some detail here, both to ensure an understanding of the decision tree, and to set up the case study used in Chapter 5 where the same model is used to demonstrate the appropriate analysis of a probabilistic model.

As shown in Fig. 2.1, six treatment options are considered in the form of strategies, as they define sequences of treatments rather than individual therapies:

- *A: Intermittent proton-pump inhibitor (PPI).* Patients would be given PPI and, if this heals the GORD, they would taken off therapy. If they experience

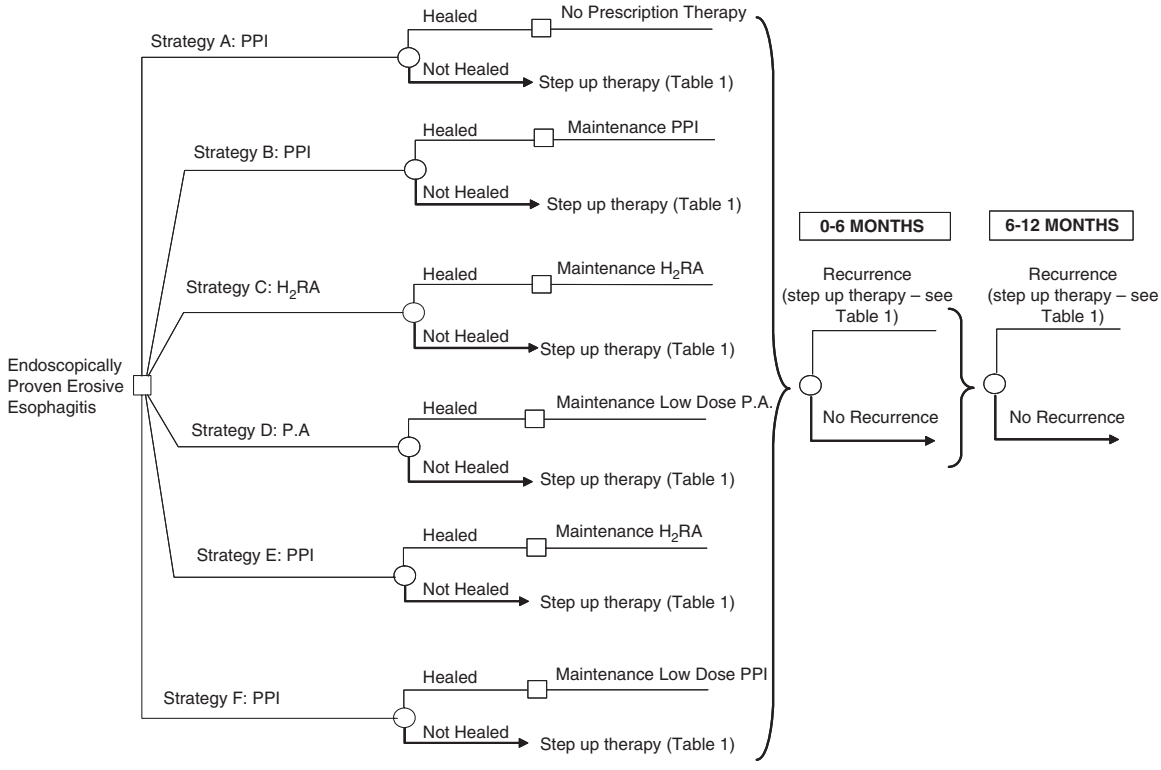


Fig. 2.1 Decision tree used to evaluate six treatment strategies for gastro-oesophageal reflux disease (adapted from Goeree *et al*. (1999)). PPP, proton pump inhibitor; H_2RA: H_2 receptor antagonists; PA: prokinetic agent. Table 1 refers to the table from the original article.

a recurrence of the condition, they would be returned to the PPI regimen. If patients fail to heal with their initial dose of PPI, the dose of the therapy would be doubled (DD PPI) and, once healed, patients would be maintained on standard dose PPI. If the GORD recurs, they would be given DD PPI.

- *B: Maintenance PPI.* Patients would initially be treated with PPI and, once healed, they would be maintained on PPI. If they fail to heal with their initial dose, or if the GORD recurs subsequent to healing, they would be given DD PPI.
- *C: Maintenance H_2* receptor antagonists (H_2RA). Patients are initially treated with H_2RA and, if they heal, they are maintained on that drug. If their GORD subsequently recurs, they are placed on double-dose H_2RA (DD H_2RA). If patients fail to heal on the initial dose, they are given PPI and, if they then heal, are maintained with H_2RA. If the GORD subsequently recurs, they are given PPI to heal. If patients fail to heal initially with PPI, they are moved to DD PPI and, if they then heal, are maintained with PPI. If the GORD subsequently recurs on maintenance PPI, they are given DD PPI to heal.
- *D: Step-down maintenance prokinetic agent (PA).* Patients would be given PA for initial healing and, if this is successful, maintained on low dose (LD) PA; if their GORD subsequently recurs, they would be put on PA to heal again. If patients fail their initial healing dose of PA, they would be moved to PPI to heal and, if successful, maintained on LD PA. If their GORD subsequently recurs, they would be treated with PPI for healing. If patients fail their initial healing dose of PPI, they would be moved to DD PPI and, if successful, maintained on PPI. If the GORD subsequently recurs, they would receive DD PPI to heal.
- *E: Step-down maintenance H_2*RA. Patients would initially receive PPI to heal and, if this is successful, they would be maintained on H_2RA. If they subsequently recur, they would be given PPI to heal. Patients who initially fail on PPI would be given DD PPI and, if this heals the GORD, they would be maintained on PPI. If their GORD subsequently recurs, healing would be attempted with DD PPI.
- *F: Step-down maintenance PPI.* Patients would initially be treated with PPI and, if this heals the GORD, would move to LD PPI. In the case of a subsequent recurrence, patients would be given PPI to heal. Patients who fail on their initial dose of PPI would be given DD PPI and, if successful, maintained on PPI. If their GORD recurs, they would be given DD PPI to heal.

The structure of the decision tree used in the study is shown in Fig. 2.1. For each strategy, the initial pathway shows whether their GORD initially heals and, if so, it indicates the maintenance therapy a patient will move to. If they do not heal, they move to step up therapy as defined for each of the five strategies. The figure shows that, for each pathway on the tree, there is a probability of GORD recurrence during two periods: 0–6 months and 6–12 months. Should this happen, step-up therapy is used as defined above. It should be noted that the tree contains decision nodes to the right of chance nodes. However, this indicates a treatment decision defined by the strategies rather than a point in the tree where alternative courses of action are being compared.

To populate the model, the authors undertook a meta-analysis of randomized trials to estimate, for each drug, the proportion of patients healed at different time points. They also used available trial data to calculate the proportions of patients who recur with GORD over the two time periods. Routine evidence sources and clinical opinion were used to estimate the cost of therapies and of recurrence.

The decision tree was evaluated over a time horizon of 12 months. Costs were considered from the perspective of the health system and outcomes were expressed in terms of the expected number of weeks (out of 52) during which a patient was free of GORD. Table 2.1 shows the base-case results of the analysis. For each strategy over 1 year, it shows the expected costs and time with (and without) GORD symptoms. The table shows the options that are dominated or subject to extended dominance (Johannesson and Weinstein 1993), and the incremental cost per week of GORD symptoms avoided are shown for the remaining options. Figure 2.2 shows the base-case cost-effectiveness results on the cost-effectiveness plane (Black 1990; Johannesson and Weinstein 1993). It shows that Option D is dominated as it is more costly and less effective than Options C, A and E. It also shows that Option F is subject to extended dominance. That is, it lies to the left of the efficiency frontier defined by non-dominated options. This means that it would be possible to give a proportion of patients Option E and a proportion Option B and the combined costs and effects of this mixed option would dominate Option F (see the discussion of cost-effectiveness decision rules in Chapter 1).

2.3.2. Markov models

The advantages of Markov models

Although various aspects of the GORD case study, such as the choice of outcome measure, can be criticized, the study provides a good example of a cost-effectiveness model based around a decision tree structure. Some of the

Table 2.1 Base-case results from the gastro-oesophageal reflux disease case study (Goeree *et al*. 1999)

Strategy	Expected 1-year cost per patient $	Expected weeks with (without) GORD per patient in 1 year	Incremental costs (DC) $	Incremental effects (DE) GORD weeks averted	DC/DE ($/GORD week averted)
C: Maintenance H_2RA	657	10.41 (41.59)	–	–	–
A: Intermittent PPI	678	7.78 (44.22)	21	2.63	8
E: Step-down maintenance H_2RA	748	6.17 (45.83)	70*	1.61*	44*
B: Maintenance PPI	1093	4.82 (47.18)	345†	1.35†	256†
D: Step-down maintenance PA	805	12.60 (39.40)			Dominated by A,C,E
F: Step-down maintenance PPI	955	5.54 (46.46)			Dominated by extended dominance

* Relative to strategy A

†Relative to strategy E

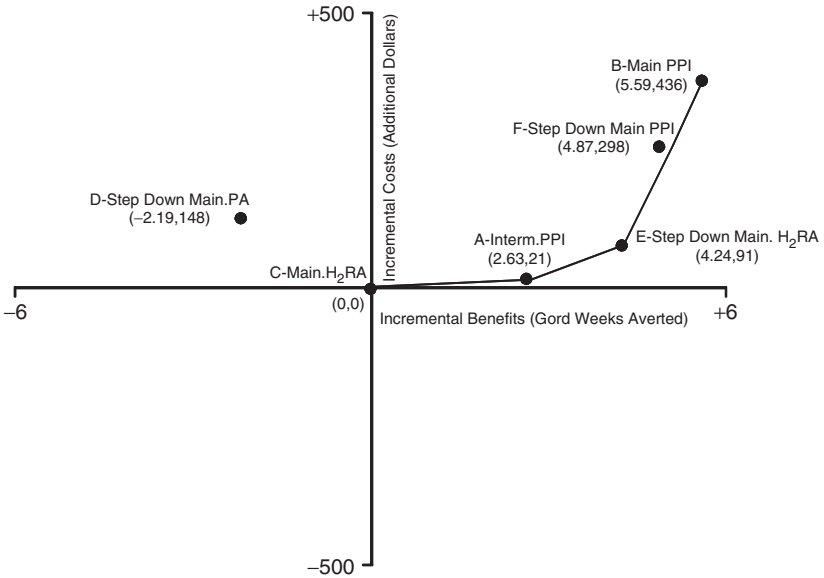


Fig. 2.2 Base-case results of the gastro-oesophageal reflux disease case study shown on the cost-effectiveness plane (taken from Goeree *et al*. (1999)). The line joining Strategies C, A, E and B is the 'efficiency frontier'.

potential limitations of the decision tree are also evident in the case study. The first is that the elapse of time is not explicit in decision trees. GORD relapse between 0 and 6 months and between 6 and 12 months has to be separately built into the model as no element of the structure explicitly relates to failure rate over time.

A second limitation of decision trees made evident in the GORD example is the speed with which the tree format can become unwieldy. In the GORD case study, only three consequences of interventions are directly modelled: initial healing, relapse between 0 and 6 months and relapse between 6 and 12 months. GORD is a chronic condition and it can be argued that for this analysis, a lifetime time horizon may have been more appropriate than one of 12 months. If a longer time horizon had been adopted, several further features of the model structure would have been necessary. The first is the need to reflect the continuing risk of GORD recurrence (and hence the need for step-up therapy) over time. The second is the requirement to allow for the competing risk of death as the cohort ages. The third is the consideration of other clinical developments, such as the possible occurrence of oesophageal cancer in patients experiencing recurrent GORD over a period of years. This pattern of recurring-remitting disease over a period of many years and of competing

clinical risks is characteristic of many chronic diseases such as diabetes, ischaemic heart disease and some forms of cancer. In such situations, the need to reflect a large number of possible consequences over time would result in the decision tree becoming very 'bushy' and, therefore, difficult to program and to present. As such, a Markov framework was used to further develop the GORD model described above (Goeree *et al.* 2002).

The Markov model is a commonly used approach in decision analysis to handle the added complexity of modelling options with a multiplicity of possible consequences. Such models have been used in the evaluation of screening programmes (Sanders *et al.* 2005), diagnostic technologies (Kuntz *et al.* 1999) and therapeutic interventions (Sculpher *et al.* 1996). The added flexibility of the Markov model relates to the fact that it is structured around mutually exclusive disease states, representing the possible consequences of the options under evaluation. Instead of possible consequences over time being modelled as a large number of possible pathways as in a decision tree, a more complex prognosis is reflected as a set of possible transitions between the disease states over a series of discrete time periods (cycles). Costs and effects are typically incorporated into these models as a mean value per state per cycle, and expected values are calculated by adding the costs and outcomes across the states and weighting according to the time the patient is expected to be in each state.

A case study in HIV

The details of the Markov model can be illustrated using a case study. This is a cost-effectiveness analysis of zidovudine monotherapy compared with zidovudine plus lamivudine (combination) therapy in patients with HIV infection (Chancellor *et al.* 1997). This example has been used for didactic purposes before (Drummond *et al.* 2005), but is further developed here, and in Chapter 4 for purposes of probabilistic analysis.

The structure of the Markov model is shown in Fig. 2.3. This model characterizes a patient's prognosis in terms of four states. Two of these are based on CD4 count: 200–500 cells/mm^3 (the least severe disease state – State A) and less than 200 cells/mm^3 (State B). The third state is AIDS (State C) and the final state is death (State D). The arrows on the Markov diagram indicate the transitions patients can make in the model. The key structural assumption in this early HIV model (now clinically doubtful, at least in developed countries) is that patients can only remain in the same state or progress; it is not feasible for them to move back to a less severe state. More recent models have allowed patients to move back from an AIDS state to non-AIDS states and, through therapy, to experience an increase in CD4 count. These models have also

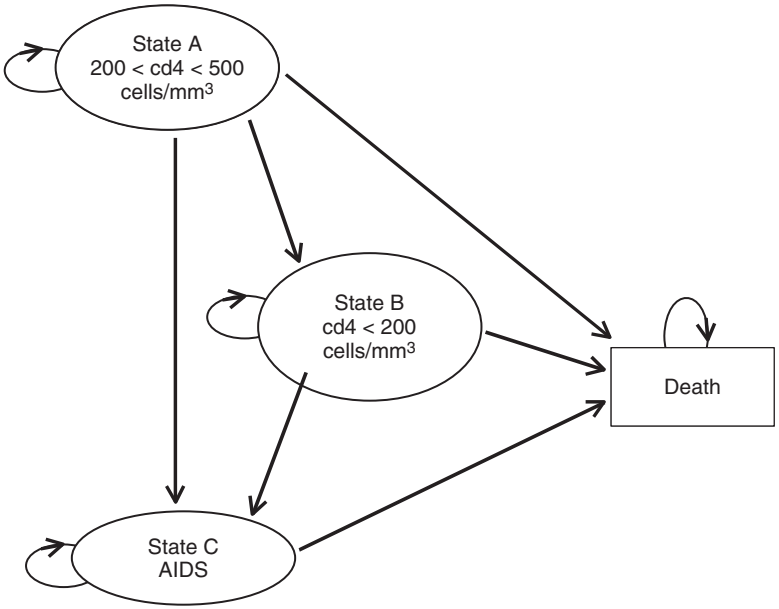


Fig. 2.3 The structure of the Markov model used in the case study (Chancellor *et al*. 1997).

allowed for the fact that prognosis for these patients is now understood in terms of viral load as well as CD4 count (Sanders *et al*. 2005).

The transition probabilities governing the direction and speed of transitions between disease states in the model are shown in Table 2.2 where a cycle is taken as 1 year. For monotherapy, these 'baseline' (i.e. control group) probabilities are taken from a longitudinal cohort study where data were collected prior to any use of combination therapy. The zeros indicate that backwards transitions are assumed not to be feasible. The transition probabilities for combination therapy were based on an adjustment to the baseline values according to the treatment effect of combination therapy relative to monotherapy. This treatment effect took the form of a relative risk (0.509) which was derived from a meta-analysis of trials. Although the treatment effect in the trials was something rather different, it was assumed that the relative risk worked to reduce the transition probabilities from one state to any worse state. The calculation of these revised (combination) transition probabilities is shown in Table 2.2. Any probability relating to the movement to a worse state is multiplied by 0.509, and the probability of remaining in a state is correspondingly increased. The separation of baseline probabilities from a relative

Table 2.2 Transition probabilities and costs for the HIV Markov model used in the case study (Chancellor *et al*. 1997)

State at start of cycle	**State at end of cycle**			
1. Annual transition probabilities				
(a) Monotherapy				
	State A	State B	State C	State D
State A	0.721	0.202	0.067	0.010
State B	0.000	0.581	0.407	0.012
State C	0.000	0.000	0.750	0.250
State D	0.000	0.000	0.000	0.000
(b) Combination therapy				
	State A	State B	State C	State D
State A	0.858 (1 – sum)	0.103 (0.202 × RR)	0.034 (0.067 × RR)	0.005 (0.010 × RR)
State B	0.000	0.787 (1 – sum)	0.207 (0.407 × RR)	0.006 (0.012 × RR)
State C	0.000	0.000	0.873 (1 – sum)	0.127 (0.25 × RR)
State D	0.000	0.000	0.000	1.000
2. Annual costs				
Direct medical	£1701	£1774	£6948	–
Community	£1055	£1278	£2059	–
Total	£2756	£3052	£9007	–

RR, relative risk of disease progression. Estimated as 0.509 in a meta-analysis.

The drug costs were £2278 (zidovudine) and £2086 (lamivudine).

treatment effect is a common feature of many decision models used for cost-effectiveness. One advantage of this approach relates to the important task of estimating cost-effectiveness for a particular location and population subgroup. Often it is assumed that baseline event probabilities should be as specific as possible to the location(s) and subgroup(s) of interest, but that the relative treatment effect is assumed fixed.

It can be seen that all the transition probabilities are fixed with respect to time. That is, the baseline annual probability of progressing from, for example, State A to State B is 0.202, and this is the case 1 year after start of therapy and it is also the case, for those remaining in State A, after 10 years. When these time invariant probabilities are used, this is sometimes referred to as a Markov Chain.

Table 2.2 also shows the annual costs associated with the different states. These are assumed identical for both treatment options being compared – excluding the costs of the drugs being evaluated. The drug costs were

£2278 (zidovudine) and £2086 (lamivudine). Outcomes were assessed in terms of changes in mean survival duration so no health-related quality-of-life weights (utilities) were included.

Cohort simulation

As explained previously, the calculation of expected costs and outcomes for all cohort models involves summing the costs and outcomes of all possible consequences weighted by the probability of the consequence. In the case of Markov models, this involves calculating how long patients would spend in a given disease state. This can be achieved using matrix algebra, but is more commonly undertaken using a process known as cohort simulation. This is illustrated, for the first two cycles for monotherapy, in Fig. 2.4. Exactly the same process would be used for combination therapy based on the adjusted transition probabilities in Table 2.2. This example uses a cohort size of 1000, but this number is arbitrary and, for a cohort model, the same answer will emerge for any size of starting cohort. Cohort simulation simply involves multiplying the proportion of the cohort ending in one state by the relevant transition probability to derive the proportion starting in another state. In a spreadsheet, this is achieved by setting up the formulae for one cycle and then copying down for subsequent cycles.

Calculating expected costs for a cohort model simply involves, for each cycle, adding the costs of each state weighted by the proportion in the state, and then adding across cycles. This is shown in Table 2.3 for monotherapy based on the costs shown in Table 2.2 and annual drug costs for zidovudine of £2278. The process of discounting to a present value is very straightforward in a cohort simulation and this is also shown in Table 2.3.

Table 2.4 shows the calculation of expected survival duration (life expectancy) for the monotherapy group. At its simplest, this involves adding the proportion of the living cohort for each cycle, and adding across the cycles. Of course, the units in which survival duration is estimated will depend on the cycle length. In the case study, all transitions are assumed to take place at the start of the cycle, so the simple approach to calculating life expectancy assumes those dying during a cycle do not survive for any proportion of the cycle. Strictly, an unbiased estimate of life expectancy would assume that deaths occur halfway through a cycle. It is possible, therefore, to employ a half-cycle correction as shown in the final column of Table 2.3. The half-cycle correction is shown here just for life expectancy calculation. In principle, however, it also applies to the calculation of expected costs. In the context of a cost-effectiveness analysis where the focus is on the incremental costs and outcomes of alternative options, it is unlikely that the half-cycle correction

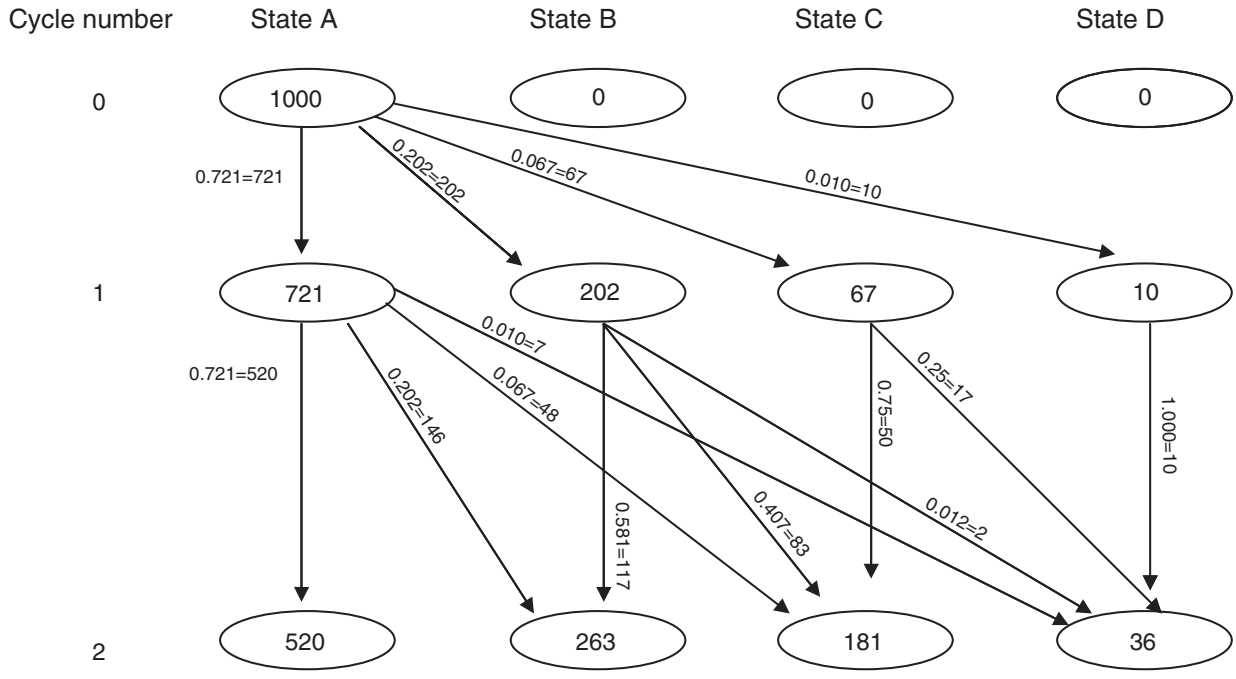


Fig. 2.4 Illustration of the first two cycles of a cohort simulation for monotherapy for the Markov model used in the case study.

Table 2.3 Calculation of expected costs for monotherapy in the case study Markov model based on the costs shown in Table 2.2, annual drug costs of £2278 and an annual discount rate of 6%

Cycle (year)	Proportion of cohort in each state				Costs (£)	
	A	B	C	D	Undiscounted	Discounted
0	1000					
1	721	202	67	10	5463*	5153†
2	520	263	181	36	6060	5393
3	376	258	277	89	6394	5368
4	271	226	338	165	6381	5055
5	195	186	364	255	6077	4541
6	141	147	361	350	5574	3929
7	102	114	341	444	4963	3301
8	73	87	309	531	4316	2708†
9	53	65	272	610	3682	2179
10	38	49	234	679	3092	1727
11	28	36	198	739	2564	1350
12	20	26	165	789	2102	1045
13	14	19	136	830	1708	801
14	10	14	111	865	1377	609
15	7	10	90	893	1103	460
16	5	8	72	915	878	346
17	4	5	57	933	695	258
18	3	4	45	948	548	192
19	2	3	36	959	431	142
20	1	2	28	968	337	105
					63 745	44 663

*{[721 × (2756 + 2278)] + [202 × (3052 + 2278)] + [67 × (9007 + 2278)] + [10 × 0]} / 1000

†5463 / [(1 + 0.06)^{1}]

will make a lot of difference to results unless the cycle length is long as a proportion of the model's time horizon.

Of course, the cohort simulation and calculation of expected costs and expected survival duration, shown in Fig. 2.4 and Tables 2.3 and 2.4 for the monotherapy baseline, would also need to be undertaken for combination therapy. The process would be exactly the same but would involve different transition probabilities and costs.

Table 2.4 Calculation of life expectancy over 20 cycles for monotherapy for the case study Markov model. Calculation with and without a half-cycle correction is shown

Cycle (year)	Proportion of cohort in each state				Life years	
	A	B	C	D	Die at start	Die in middle
0	1000					
1	721	202	67	10	0.990	0.995
2	520	263	181	36	0.964	0.977
3	376	258	277	89	0.911	0.937
4	271	226	338	165	0.835*	0.873
5	195	186	364	255	0.745	0.790
6	141	147	361	350	0.650	0.697
7	102	114	341	444	0.556	0.603†
8	73	87	309	531	0.469	0.513
9	53	65	272	610	0.390	0.429
10	38	49	234	679	0.321	0.355
11	28	36	198	739	0.261	0.291
12	20	26	165	789	0.211	0.236
13	14	19	136	830	0.170	0.190
14	10	14	111	865	0.135	0.152
15	7	10	90	893	0.107	0.121
16	5	8	72	915	0.085	0.096
17	4	5	57	933	0.067	0.076
18	3	4	45	948	0.052	0.059
19	2	3	36	959	0.041	0.047
20	1	2	28	968	0.032	0.036
					7.996	8.475

*(271 + 226 + 338) / 1000

†{102 + 114 + 341 + [0.5 × (444 – 350)]}/1000

The Markov assumption

Although the Markov version of a cohort model provides greater flexibility than a decision tree, it also has some important restrictions in the context of structuring complex prognoses. The restriction relates to what is known as the Markov assumption, or 'memoryless' feature of Markov models. This assumption means that once a notional patient has moved from one state to another,

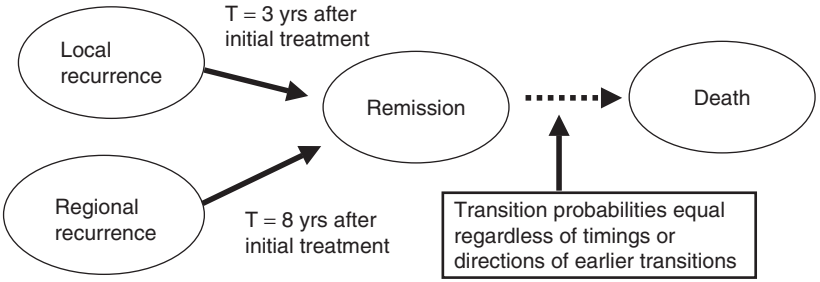


Fig. 2.5 Example of a simple Markov model to illustrate the Markov assumption.

the Markov model will have 'no memory' regarding where the patient has come from or the timing of that transition. This is illustrated in Fig. 2.5 using a simple Markov model to characterize prognosis following treatment for cancer. It shows that patients can have two forms of cancer recurrence: local and regional. Once treated for their recurrence, patients then move into a remission state but, once in that state, all patients are considered homogeneous regardless of where they have come from (what type of recurrence) or the timing of the recurrence after initial treatment for the cancer. The implication of this is that it is not possible to have different probabilities of future transitions (e.g. the probability of death in Fig. 2.5) according to the nature or timing of the recurrence.

In general, the Markov assumption means that it may not be straightforward to build 'history' into this type of model. That is, to model a process where future events depend on past events. This may be possible by adding additional states to the model and incorporating time dependency into transition probabilities. Both of these extensions to the Markov model are considered in Chapter 3.

2.4. **Summary**

This chapter has described some of the key concepts for the use of decision analysis for economic evaluation. The majority of cost-effectiveness studies are based on cohort modelling. This provides a flexible framework which can be programmed relatively easily (often in spreadsheets) and evaluated rapidly. However, the cohort framework can be restricting. In the case of Markov models, for example, the memoryless feature of these models can cause difficulties in modelling a complicated prognosis.

2.5. Exercise: replicating a Markov model of HIV/AIDS

2.5.1. Overview

The aim of this exercise is to get you to replicate a previously published Markov model by Chancellor and colleagues (1997). The model itself is somewhat out of date, in that it compares combination therapy (lamivudine and AZT) with monotherapy (AZT alone). Nevertheless, the model is straightforward enough to serve as a simple example that it is possible to replicate in a short period of time.

The basic structure of the model was given in Fig. 2.3, which shows that the disease process is structured as chronic, such that patients can move to successively more serious disease states, but cannot recover.

The cycle length of the model is 1 year and is evaluated over 20 years (after which more than 95 per cent of patients are expected to have died).

Use the data given below to populate the model and calculate the incremental cost-effectiveness ratio for AZT monotherapy. You should be able to replicate a figure for the ICER of £6276.

1. *Transition probabilities.* These were calculated from the counts of individuals that were observed to move between the four health states each year in a longitudinal data set from the Chelsea and Westminster Hospital in London. These counts are given in Table 2.5 and you should be able to verify that these counts give the transition probabilities presented in Table 2.2 part 1(a).
2. *State costs.* These are given in the original article separately by state for 'direct medical' and 'community care' costs and these were reproduced in Table 2.2 part 2.
3. *Costs of drugs.* The yearly cost of AZT monotheray is given as £2278 and lamivudine is given as £2086.50.
4. *Treatment effect.* A meta-analysis of four trials is reported in the paper giving a pooled relative risk of disease progression as 0.509. Note from Table 2.2 part 1(b) how this treatment effect is applied to the monotherapy transitions in the model.

Table 2.5 Counts of transition between the four model states per year

	A	B	C	D	Total
A	1251	350	116	17	1734
B	0	731	512	15	1258
C	0	0	1312	437	
D					

5. *Discounting*. The original analysis was based on discounting the costs at 6 per cent but not discounting the estimates of life years.

2.5.2. Step-by-step guide to constructing the model

Open the file '*Exercise 2.5 – Template*' and start at the <*Parameters*> worksheet. You will see that the cells to be completed are coloured yellow.

Using the figures from Table 2.5, complete columns C and D on the <*Parameters*> worksheet for the transition probabilities (rows 9–17). Column C should contain the number of events of interest and column D should contain the complement, such that columns C plus D should equal the appropriate row total from Table 2.5 above. Now calculate the respective transition probabilities in column B using the counts in columns C and D (the reason for structuring the spreadsheet like this will become apparent in the exercise in Chapter 4).

Enter the other information for the state costs, drug costs, treatment effect and discounting given above in the appropriate place in the template.

Having entered the input parameters of the model, the task now is to construct the Markov models for the treatment alternatives: combination and monotherapy. Note that the parameters you have just entered are in named cells, such that they can be referred to by the shortened name in column A of the <*Parameters*> worksheet.

If you move to the Markov worksheet, you will see the structure of the Markov model laid out for you for each treatment alternative. Start with the monotherapy model.

1. Generating the Markov trace

The initial concern is with generating the Markov trace, that is, showing the proportions of patients that are in any one state at any one time. This is to be entered in columns C to F representing the four model states, and G provides a check (as the sum across C to F must always equal the size of the original cohort – which is set to 1 in cell C7 such that the Markov trace represents proportions of patients).

The first step in building a Markov model is to define the transition matrix. This is a matrix that shows the probability of transition from states represented in the rows to states represented in the columns. To save you time, a copy of the transition matrix is reproduced in Table 2.6 for the monotherapy arm.

i. Start by making sure that you understand the transition matrix. In particular, make sure you can replicate it from the information given in the diagram of the model in Fig. 2.3.

Table 2.6 Transition matrix for the monotherapy arm

Transition matrix	A	B	C	D
A	tpA2A	tpA2B	tpA2C	tpA2D
B	0	tpB2B	tpB2C	tpB2D
C	0	0	tpC2C	tpC2D
D	0	0	0	1

ii. Use the transition matrix to populate the Markov model. This will involve representing the transitions between the different states represented in columns C to F.
Hints: Start with just one row at a time (e.g. the first – row 8). Once you have this row correct you can simply copy it down to the other rows. Although it is tempting to look across the rows of the transition matrix – consider looking down the columns as this indicates, for a given state, where the people entering that state come from. Indeed, the transition matrix describes the proportion of patients in the other states in the previous cycle arriving in the given state for this cycle. For example, looking down for 'State B' we find that tpA2B *patients in 'State A' last period arrive in State B this period and that* tpB2B *patients in 'State B' last period stay in 'State B' this period. The formula for cell D8 should therefore read: =C7*tpA2B+D7*tpB2B*

iii. When you get to the dead state, do not be tempted to make this the remainder of the other cells in the row – the use of remainders in this way will mean that any errors could go unnoticed. Instead, it is good practice to complete all states as you think they should be then sum across the states to make sure the total sums to one. Do this check in column G and make sure it does!

iv. When you think you have the first row correct, copy this row down to the 19 rows below. If your check in column G is still looking good then most likely you have done it correctly.

By far the most tricky bit of building the Markov model in Excel is now complete. Now that you have the Markov trace you can calculate the cost and effects.

2. Estimating life years

In column H calculate the proportion of the cohort that is in one of the 'alive' states for each of the years in the model. In column I apply the standard discount formula[1] (although we have set this to zero in the base case). Finally,

in row 29, sum the columns to give the expected life years both discounted and undiscounted.

3. Estimating costs

In column K, calculate the costs for each time period by applying the state costs, not forgetting that AZT is given for the whole time period. In column L, apply the standard rate of discount for costs. Again, in Row 29, sum the columns to give costs both undiscounted and discounted.

4. Adapting the model for combination therapy

You now need to repeat the steps above, but this time for combination therapy.

i. Start by copying the whole monotherapy model to the corresponding combination therapy model. This will minimize repetition as tasks now relate to adjusting the model for combination treatment. (Note: remember to anchor the year (column A) in the discounting formula before copying).

ii. Firstly, the treatment effect needs to be added in. In the original article, the relative risk parameter was applied to all transitions. The corresponding transition matrix for combination therapy is, therefore, given in Table 2.7.

iii. Use this transition matrix to adjust the trace in rows 8 and 9 only (as the base case assumption in the originally reported model was that treatment effect was limited to 2 years).

Now add in the cost of lamivudine for these 2 years only in column V (as the drug is assumed to be given for only 2 years)

[1] Recall that the standard discount rate is given by $1/(1 + r)^t$ where r is the discount rate and t represents time (in years).

Table 2.7 Transition matrix for the combination therapy arm

Transition matrix	A	B	C	D
A	1-tpA2B*RR-tpA2C*RR-tpA2D*RR	tpA2B*RR	tpA2C*RR	tpA2D*RR
B	0	1-tpB2C*RR-tpB2D*RR	tpB2C*RR	tpB2D*RR
C	0	0	1-tpC2D*RR	tpC2D*RR
D	0	0	0	1

5. Cost-effectiveness estimates

The final task is simply to copy the corresponding results from the sums at the bottom of the Markov model sheet (row 29) to the <*Analysis*> worksheet (row 5) and to calculate the appropriate increments and ICER.

Congratulations, you have now replicated the Markov model. Compare your result for the ICER to that given in the solution (£6276). Is any debugging required? If it is, then you may want to compare your Markov trace and stage costs for monotherapy against those reported in Table 2.3.

References

Ades, A. E. (2003) A chain of evidence with mixed comparisons: models for multi-parameter evidence synthesis and consistency of evidence, *Statistics in Medicine*, 22: 2995–3016.

Ades, A. E., Sculpher, M. J., Sutton, A., Abrams, K., Cooper, N., Welton, N., *et al.* (2006) 'Bayesian methods for evidence synthesis in cost-effectiveness analysis', *PharmacoEconomics* 24: 1–19.

Arrow, K. J. and Lind, R. C. (1970) 'Risk and uncertainty: uncertainty and the evaluation of public investment decisions', *American Economic Review*, 60: 364–78.

Black, W. C. (1990) 'The CE plane: a graphic representation of cost-effectiveness', *Medical Decision Making*, 10: 212–214.

Briggs, A. (2001) 'Handling uncertainty in economic evaluation and presenting the results' in M. Drummond and A. J. McGuire (eds) *Economic evaluation in health care: merging theory with practice*. Oxford, Oxford University Press; pp. 172–214.

Chancellor, J. V., Hill, A. M., Sabin, C. A., Simpson, K. N. and Youle, M. (1997) 'Modelling the cost effectiveness of lamivudine/zidovudine combination therapy in HIV infection', *PharmacoEconomics*, 12: 1–13.

Chilcott, J., McCabe, C., Tappenden, P., O'Hagan, A., Cooper, N. J., Abrams, K., *et al.* (2003) 'Modelling the cost effectiveness of interferon beta and glatiramer acete in the management of multiple sclerosis', *British Medical Journal*, 326: 522.

Cooper, N. J., Sutton, A. J., Abrams, K. R., Turner, D. and Wailoo, A. (2004) 'Comprehensive decision analytical modelling in economic evaluation: A Bayesian approach', *Health Economics*, 13: 203–226.

Downs, S. (1999) 'Technical report: urinary tract infections in febrile infants and young children', *Pediatrics*, 103: 54.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. and Stoddart, G. L. (2005). *Methods for the economic evaluation of health care programmes*. Oxford, Oxford University Press.

Evans, K. W., Boan, J. A., Evans, J. L. and Shuaib, A. (1997) 'Economic evaluation of oral sumatriptan compared with oral caffeine/ergotamine for migraine', *PharmacoEconomics*, 12: 565–577.

Goeree, R., O'Brien, B. J., Hunt, R., Blackhouse, G., Willan, A. and Watson, J. (1999) 'Economic evaluation of long term management strategies for erosive oesphagitis', *PharmacoEconomics*, 16: 679–697.

Goeree, R., O'Brien, B. J., Blackhouse, G., Marshall, J., Briggs, A. H. and Lad, R. (2002) 'Cost-effectiveness and cost-utility of long-term management strategies for heartburn', *Value in Health*, 5: 312–328.

Hunink, M., Glaziou, P., Siegel, J., Weeks, J., Pliskin, J., Elstein, A., *et al.* (2001) *Decision making in health and medicine. Integrating evidence and values.* Cambridge, Cambridge University Press.

Johannesson, M. and Weinstein, S. (1993) 'On the decision rules of cost-effectiveness analysis', *Journal of Health Economics*, 12: 459–467.

Kuntz, K. M., Fleishmann, K. E., Hunink, M. G. M. and Douglas, P. S. (1999) 'Cost-effectiveness of diagnostic strategies for patients with chest pain', *Annals of Internal Medicine*, 130: 709–718.

Neumann, P. J., Hermann, R. C. and Kuntz, K. M. (1999) 'Cost-effectiveness of donepezil in the treatment of mild or moderate Alzheimer's disease', *Neurology*, 52: 1138–45.

O'Hagan, A. and Luce, B. (2003) *A primer on Bayesian statistics in health economics and outcomes research.* Bethesda, MD, Medtap International.

Palmer, S., Sculpher, M., Philips, Z., Robinson, M., Ginnelly, L., Bakhai, A., *et al.* (2005) 'Management of non-ST-elevation acute coronary syndromes: how cost-effective are glycoprotein IIb/IIIa antagonists in the UK National Health Service?', *International Journal of Cardiology*, 100: 229–240.

Parmigiani, G. (2002) *Modeling in medical decision making: a Bayesian approach.* Chichester, Wiley.

Raiffa, H. (1968) *Decision analysis: introductory lectures on choices under uncertainty.* Reading, MA, Addison-Wesley.

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B. and Richardson, W. S. (1996) 'Evidence-based medicine: what it is and what it isn't', *British Medical Journal*, 312: 71–72.

Sanders, G. D., Bayoumi, A. M., Sundaram, V., Bilir, S. P., Neukermans, C. P. and *et al.* (2005) 'Cost-effectiveness of screening for HIV in the era of highly active antiretroviral therapy', *New England Journal of Medicine*, 352: 570–585.

Sculpher, M., Michaels, J., McKenna, M. and Minor, J. (1996) 'A cost-utility analysis of laser-assisted angioplasty for peripheral arterial occlusions', International Journal of Technology Assessment in Health Care 12: 104–125.

Spiegelhalter, D. J., Abrams, K. R. and Myles, J. P. (2004). *Bayesian approaches to clinical trials and health-care evaluation.* Chichester, Wiley.

Sutton, A. J. and Abrams, K. R. (2001) 'Bayesian methods in meta-analysis and evidence synthesis', *Statistical Methods in Medical Research*, 10: 277–303.

Sutton, A. J., Abrams, K. R., Jones, D. R., Sheldon, T. A. and Song, T. A. (2000) *Methods for meta-analysis in medical research.* Chichester, Wiley.

Torrance, G. W. (1986) 'Measurement of health state utilities for economic appraisal – a review', *Journal of Health Economics*, 5: 1–30.

von Neumann, J. and Morgenstern, O. (1944) *Theory of games and economic behavior.* New York, Princeton University Press.

Weinstein, M. C. and Fineberg, H. V. (1980) *Clinical decision analysis.* Philadelphia, PA, WB Saunders Company.

This page intentionally left blank