Pharmacoeconomics 2000 May; 17 (5): 479-500
1170-7690/00/0005-0479/$20.00/0

Handling Uncertainty in Cost-Effectiveness Models

Andrew H. Briggs

Health Economics Research Centre, University of Oxford, Oxford, England

Contents

Abstract

The use of modelling in economic evaluation is widespread, and it most often involves synthesising data from a number of sources. However, even when economic evaluations are conducted alongside clinical trials, some form of modelling is usually essential. The aim of this article is to review the handling of uncertainty in the cost-effectiveness results that are generated by the use of decision-analytic-type modelling. The modelling process is split into a number of stages: (i) a set of methods to be employed in a study are defined, which should include a 'reference case' of agreed methods to enhance the comparability of results; (ii) the clinical and demographic characteristics of the patients the model relates to should be specified as carefully as in any experimental study; and (iii) the data requirements of the model should be estimated using the principles of Bayesian statistics, such that prior distributions are specified for unknown model parameters. Monte Carlo simulation can then be employed to sample from these prior distributions to obtain a distribution of the cost effectiveness of the intervention. Such probabilistic analyses are related to parameter uncertainty. In addition, modelling uncertainty is likely to add a further layer of uncertainty to the results of particular analyses.

The purpose of economically evaluating healthcare interventions is to provide information suitable to allow scarce healthcare resources to be allocated efficiently. Most evaluative studies involve the use of modelling techniques to synthesise data from various sources to produce the cost-effectiveness

results of interest. Although there is an increasing trend towards conducting economic analysis alongside clinical trials (in so-called stochastic evaluations), such studies still account for only a small proportion of all economic evaluation studies. For example, in a recent review,[1] only 6 out of 104 (6%) cost-effectiveness studies published in 1996 were conducted alongside a clinical trial, while 88 (85%) were judged to be principally modelling-based studies. Furthermore, given the limitations in the length and scope of most clinical trials, some form of modelling is usually required to progress from the within-trial results to the economic result of interest.[2]

A number of recently released guidelines for analysts have emphasised the need for studies to include an assessment of the implications of uncertainty for their results, either through statistical analysis (when patient-level data are available) or through sensitivity analysis.[3-6] However, very little guidance is given to analysts on exactly how this should be done and how the results of any analysis of uncertainty should be presented. The second edition of the Canadian Coordinating Office for Health Technology Assessment (CCOHTA) guidelines[6] probably goes furthest in giving analysts specific guidance for handling uncertainty and recommends the use of probabilistic sensitivity analysis (but does not give details of the method).

The aim of this article is to present a case for how models of cost effectiveness should be analysed and presented to best provide decision-makers with the information they require to make resource allocation decisions.

In the first section of the paper, the different types of uncertainty that can arise in cost-effectiveness models are discussed. Following the recent categorisation by the US Panel on Cost-Effectiveness Analysis, 2 types of uncertainty are distinguished: modelling uncertainty and parameter uncertainty.[7] Three different levels of 'parameters' in modelling studies are distinguished, and it is argued that they need to be dealt with differently in a cost-effectiveness analysis.

Firstly, a case is made for adopting a 'reference case' of analytical methods to be employed in cost-effectiveness studies to improve the comparability of results[1] (and thereby reduce uncertainty in the relative cost effectiveness of interventions estimated from different studies).

Secondly, parameters relating to the clinical and demographic characteristics of patients are required to estimate the healthcare 'production function' for a particular intervention. Knowledge of this production function is required to estimate the appropriate scope of implementation for a particular intervention. These parameters are not 'random' in any meaningful sense; rather, these characteristics relating to the relevant patient population should be specified as carefully in a cost-effectiveness model as they are in an experimental study.

The remaining parameters of a model, once a reference set of methods and the characteristics of the target population have been set, can be considered to be parameters that could, in principle, be estimated by sampling from the relevant patient population. Therefore, a distribution should be assigned to each of these parameters reflecting prior beliefs concerning their uncertainty. A probabilistic sensitivity analysis can then be undertaken by sampling from each of these distributions simultaneously using Monte Carlo simulation. Finally, the additional level of modelling uncertainty is considered together with potential remedies.

In the second section of the paper, the use of probabilistic sensitivity analysis methods employing an illustrative model is demonstrated. In particular, the shortcomings associated with performing first-order versus second-order simulations are highlighted. The use of cost-effectiveness acceptability

1 Note that in this article, cost effectiveness is used generically to describe results of economic evaluations that are commonly compared in league tables (i.e. those reporting results in terms of cost per life-year and cost per quality-adjusted life-year. Studies reporting results in so-called 'natural' units, such as cases detected or 'symptom-free days avoided', are also known as cost-effectiveness studies. However, since these natural units cannot be compared between different disease areas, the usefulness of these studies for allocating resources in the health sector is limited.

curves is advocated for presenting cost-effectiveness results from probabilistic sensitivity analyses.

The third section offers a discussion of the issues raised in the paper and a final section offers some conclusions.

1. Uncertainty in Cost-Effectiveness Models

Cost-effectiveness models require information to populate them, and these informational requirements are often referred to as the parameters of the model. An analogy with statistical techniques is employed to distinguish between different 'levels' of parameters: (i) parameters relating to analytical methods (e.g. the discount rate) employed in an evaluation; (ii) parameters that describe the characteristics of a patient sample (e.g. age/gender composition or clinical characteristics such as blood pressure); and (iii) parameters that could, in principle, be sampled if an appropriate study were designed to collect the relevant data. In the following sections, each of these levels is taken in turn to describe how a cost-effectiveness model should best be structured. Finally, modelling uncertainty is considered in a hierarchical fashion as adding further layers of uncertainty to the process of assessing cost effectiveness.

1.1 Methodological Uncertainty: The Case for a 'Reference Case'

The analytical methods used in an economic evaluation consist of a range of techniques including methods of measurement and valuation, and the choice of costs and benefits to include in an evaluation. There exists, in a number of these areas, disagreement amongst practitioners about the most appropriate analytical method.[8] An example of one such area of disagreement is the debate about the preferred way to incorporate time preference into economic evaluation and, in particular, the role of differential discounting of costs and benefits.[9-15][2] Uncertainty also exists concerning the methods selected to value the resource and health outcome consequences in an evaluation. The problems involved in estimating unit costs that accurately represent the opportunity cost of resources are well known. There has also been extensive debate over the choice of instruments to value health outcomes.[17-20] It should be noted, however, that uncertainty relating to the validity and reliability of measurement instruments exists in the clinical, as well as the economic domain.[21]

Perhaps a less obvious lack of consensus exists in issues such as whether or not to include in economic assessments the cost of healthcare resources consumed (due to unrelated illness) during extra years of life generated by the intervention under evaluation.[22,23] Similarly, there is some debate concerning whether (and how) to include the cost of production losses from time away from work[3] and/or time losses from general activities that may not receive a wage but which may be valued by society or the individual nonetheless.[1,25,26]

Given that choices have to be made about the methods that are to be used in a particular study, the existence of disagreement about the correct methods to employ inevitably introduces an important source of uncertainty into all economic evaluation studies. Some of this uncertainty can be handled by authors stating their study methods explicitly at the outset. However, this is only externalising uncertainty, as different decision-makers may not share the analysts' views and may desire information about whether the conclusions of the study would have been altered by using alternative methods. If it is accepted that this sort of uncertainty should be internalised in the study itself, sensitivity analysis may provide a useful role. Of course, this requires

2 For UK analysts, this issue has become all the more important recently because of the updated HM Treasury guidelines,[16] which the Department of Health has interpreted as recommending differential discounting for health outcomes in cost-effectiveness analysis (by removing the component of discounting assumed by the Treasury to relate to the combination of annual growth of income and the marginal utility of income).

3 These costs are often referred to as indirect costs; however, this term is avoided in this article since Drummond and colleagues[24] have argued that this terminology can cause confusion through the use of the same term in accountancy to mean overhead costs.

that the data are available to the analysts to generate results using more than one method.

As noted by Drummond et al.,[8] Russell has argued for a set of core methods to be employed to facilitate comparisons between evaluations.[27] This idea has been adopted by the recent US panel on cost effectiveness,[22] which recommended the use of a 'reference case' of core methods to be used by analysts when conducting economic evaluations: this could then be supplemented by additional analyses employing other methods thought appropriate by the authors. In a recent review of all UK cost-effectiveness results from all studies published before 1997 presenting cost per life-year and cost per quality-adjusted life-year (QALY) results, this idea of a reference case was applied retrospectively to improve the comparability of the results presented in different studies.[28]

The use of a reference case of methods has great appeal in cost-effectiveness analysis, where results of a study only have meaning in comparison to the results of other studies. It is likely that individual countries will want to develop their own reference case of methods rather than simply adopting the recommendations of the US panel. However, the process by which the reference case was agreed (by convening a panel of experts to achieve a consensus statement) has much to commend it.

While sensitivity analysis may be useful for considering whether conclusions would change under the application of alternative methods, sensitivity analysis used at this level is different from sensitivity analysis used to examine the effects of 'unknown' parameters. In fact, it would be useful to have interval estimates accompanying point estimates of cost effectiveness for all of the methodological scenarios presented as part of the study results.

1.2 Patient Characteristics and the Importance of the Margin

The importance of the margin is paramount in economic thinking, particularly when interpreting incremental cost-effectiveness ratios (ICERs). Consider figure 1, which shows the results of an economic evaluation of the implantable cardioverter defibrillator (ICD) for patients who survive cardiac arrest.[29] Figure 1 shows that the incremental cost effectiveness of the ICD device across the whole patient group is pounds sterling (£)57 000 per life-year saved (1991 values). However, this amount masks important differences between patients with different clinical characteristics. For patients with a low ejection fraction and inducible arrhythmia that is not controlled by drugs, the cost effectiveness of the ICD device is £22 000 per life-year saved. By contrast, in patients with high ejection fractions and inducible arrhythmia that is controlled by drugs, the incremental cost effectiveness is nearly £700 000 per life-year saved. This is not simply due to the effectiveness of the ICD device itself in the different categories of patients, it is also due to the effectiveness of the alternative (medical) management of these patients

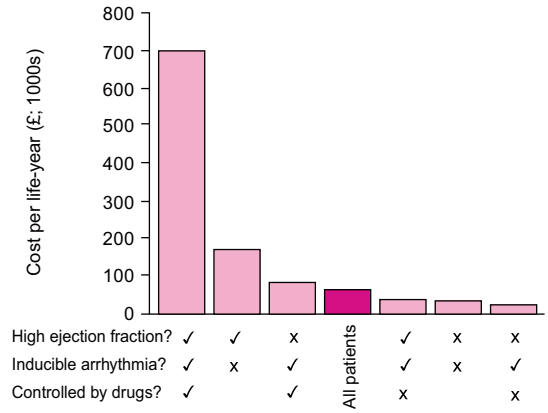

Fig. 1. Changing incremental cost effectiveness at the clinical margin for an implantable cardioverter defibrillator fitted in patients who survive cardiac arrest (1991 values).[29]

Figure 1 is an example of changes in the ICER taking place at the 'clinical margin', that is, as the same intervention is expanded to individuals with the same clinical condition, but with less capacity to benefit owing to their clinical characteristics, the estimated cost-effectiveness ratio increases. Age, gender or risk factors such as prior medical history could also be seen as clinical margins when ex-

panding programmes. For example, in a recent study of statin therapy for the reduction of cholesterol levels, the average incremental cost effectiveness for patients with pre-existing heart disease and a cholesterol level of >5.4 mmol/L was £32 000 per life-year gained (1995 values).[30] As the authors pointed out, however, this average value hides enormous differences in subgroups of patients at risk, ranging from £6000 per life-year gained for men aged 55 to 64 years who have had a myocardial infarction and have cholesterol levels of >7.2 mmol/L to £361 000 per life-year gained for women aged 45 to 54 years with angina and a cholesterol level between 5.5 and 6.0 mmol/L.

The clinical margin is not the only margin of interest when comparing incremental cost-effectiveness ratios. Interventions may be offered at different levels of intensity to the same patient groups, for example annual versus 3-yearly breast screening or low-dose versus high-dose antiviral therapy. In these cases, the 'intensity margin' is of interest, while the patient group remains the same.[4] Furthermore, the incremental cost-effectiveness ratio must be calculated along this intensity margin. For example, in a cancer screening evaluation the analyst should be interested in comparing screening every 3 years with no screening, screening every 2 years with every 3 years, and every year with every 2 years. To compare an annual screening programme with no programme will be misleading, as many of the benefits of annual screening could potentially be achieved by a 2-year programme, that is, at a lower intensity point at the margin.[31]

The 2 examples above highlight the distinction between independent and mutually exclusive programmes.[32] Independent programmes can be implemented either singly or jointly, so in the case of the ICD, the device could be implanted into patients with a high ejection fraction or those with a low ejection fraction, or into both types of patient. Such a decision should be based on the incremental cost effectiveness of the ICD for these 2 types of patients compared with alternative (medical) management. Mutually exclusive programmes involve the same group of patients and, therefore, one or the other must be chosen. Patients can be screened yearly or biennially, but not both. Similarly, patients can receive either a 40 or 80mg regimen for a particular drug. Hence, each of these margins is subtly different.

Figure 2 illustrates an example of the clinical margin on the cost-effectiveness plane. Each point on the figure relates to the costs and effectiveness of statin therapy for men aged 55 to 65 years with a previous myocardial infarction by cholesterol level.[30] Clearly, while cholesterol level is a continuous variable, for practical purposes it is necessary to define discrete categories for this. Hence, the 4 points on the figure can be considered as estimated points on a continuous 'cholesterol level production function' that shows how the costs and effects of statin therapy vary with cholesterol level. Importantly, it is not appropriate to compare the incremental costs and effects of independent programmes. In each case, the relevant incremental comparison is between the costs and effects with statin therapy and the costs and effects without statin therapy. Hence, the marginal incremental cost effectiveness increases as we move along the cholesterol production function in the direction of decreasing levels.

In comparison, the intensity margin is different since it is the same group of patients receiving the intervention, which means that the alternative intensities of treatment are mutually exclusive. Figure 3 shows the effect of decreasing the screening interval on the incremental costs and of a cervical cancer screening programme and the life-years gained[31] relative to the no-screening programme. Since only one screening interval is possible, it is appropriate to consider the additional benefits received from decreasing the screening interval relative to the

4 Another margin that might be identified relates to 'returns to scale'. Standard economic textbooks emphasise falling costs in the presence of returns to scale. However, in terms of the exposition given here, the issue of returns to scale is considered to be a technical efficiency problem. Similarly, where capital equipment results in a discontinuous cost function, it is assumed that the location of such equipment is organised to maximise throughput and therefore minimise overall cost.

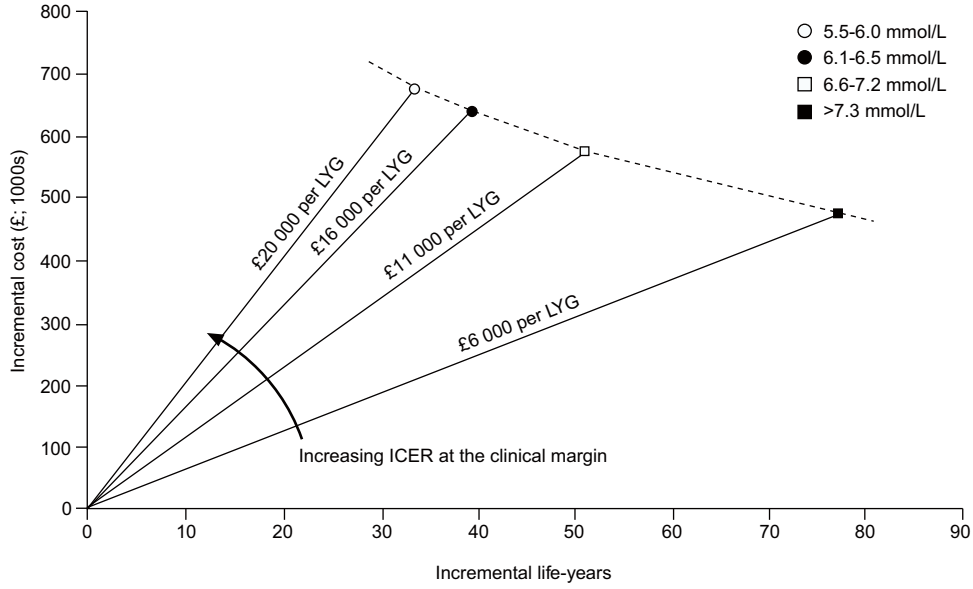

Fig. 2. Changing incremental cost effectiveness at the clinical margin on the cost-effectiveness plane (the dotted line represents the estimated cholesterol level production function) [1995 values].[30] **ICER** = incremental cost-effectiveness ratio; **LYG** = life-year gained.

additional costs. In practice, this is achieved by calculating the incremental cost effectiveness between the alternative approaches (joining the dots) to effectively map out a 'cost-effectiveness frontier' given by the screening interval production function (see fig. 3). However, since the screening interval is effectively a continuous variable, the cost-effectiveness frontier will also be continuous. In principle, screening intervals could be measured in months or even days, although in practice, intervals <1 year are unlikely to be practical or cost effective in cervical cancer screening. Incremental analysis, in the case of mutually exclusive programmes, is a practical method of obtaining an estimate of the marginal cost effectiveness, which is equal to the slope of the cost-effectiveness frontier (the estimated production function) at any point.

The point of this discussion of marginal issues in economic analysis is to emphasise that the characteristics of a patient sample (and indeed of the intervention itself in terms of intensity of use) are important parameters in a cost-effectiveness model, and that consideration of these parameters will lead to important conclusions concerning the appropriate scale of activity for a particular intervention. Estimation of the 'production function' for a given intervention must be an important part of cost-effectiveness modelling, and it is clear that in this regard cost-effectiveness models have the potential to estimate much more of the relevant production function than may be possible in a single clinical trial.

Naturally, the extent to which different patient characteristics are associated with different capacities to benefit (and therefore different cost-effectiveness values) could be seen as a statistical problem to be verified in trial-based evaluations. However, this should not prevent analysts from constructing models that seek to address these marginal issues when a clear case for differing degrees of benefit among patients with different characteristics can be made. Indeed, commentators have been stressing the importance of the clear specification of patient characteristics in cost-effectiveness models for many years precisely because of the potentially important implications for the cost-effectiveness results.[33-35]

What is clear, is that these marginal issues should be addressed outside of a formal sensitivity analy-

sis. Although there may be uncertainty regarding the associated cost effectiveness for a particular patient group and issues associated with whether particular patient groups can be distinguished in terms of cost effectiveness, patient characteristics are not sources of uncertainty in themselves. Although parameters such as age or cholesterol level affect cost effectiveness, they do so as a result of a fundamental process (the one that has been modelled) not as a result of 'uncertainty'. All too often, parameters such as age or cholesterol level appear in the sensitivity analysis section of cost-effectiveness models and are discussed alongside the 'data-driven' parameters discussed below (section 1.3). Consequently, the characteristics of a patient sample should be made equally clear in economic evaluation models as they are in clinical trial–based evaluations.

1.3 Parameters that Could (in Principle) be Sampled

Remaining parameters of the model relate to: (i) probabilities (for example, conditional probabilities for branching pathways in a decision tree model, and transition probabilities for movement between states over time in a Markov model); (ii) resource use consequences and health outcome consequences of the programmes under evaluation; and (iii) data necessary to value those consequences (unit cost/price information for resource use and quality-of-life weightings for cost-utility analyses). Having carefully specified parameters relating to the characteristics of the patient group to which the results of the cost-effectiveness analysis will apply, these remaining parameters could, in principle, be estimated by sampling from patients with the specified characteristics.

The increasing use of the clinical trial as a vehicle to collect economic data prospectively encourages the analyst to describe distributions of data and to represent uncertainty as a point estimate accompanied by a confidence interval through the use of standard statistical techniques. Suppose that we are interested in a particular parameter for a model, say the proportion of patients who move between defined stages of a disease in a given time period or the effectiveness of a particular drug in a

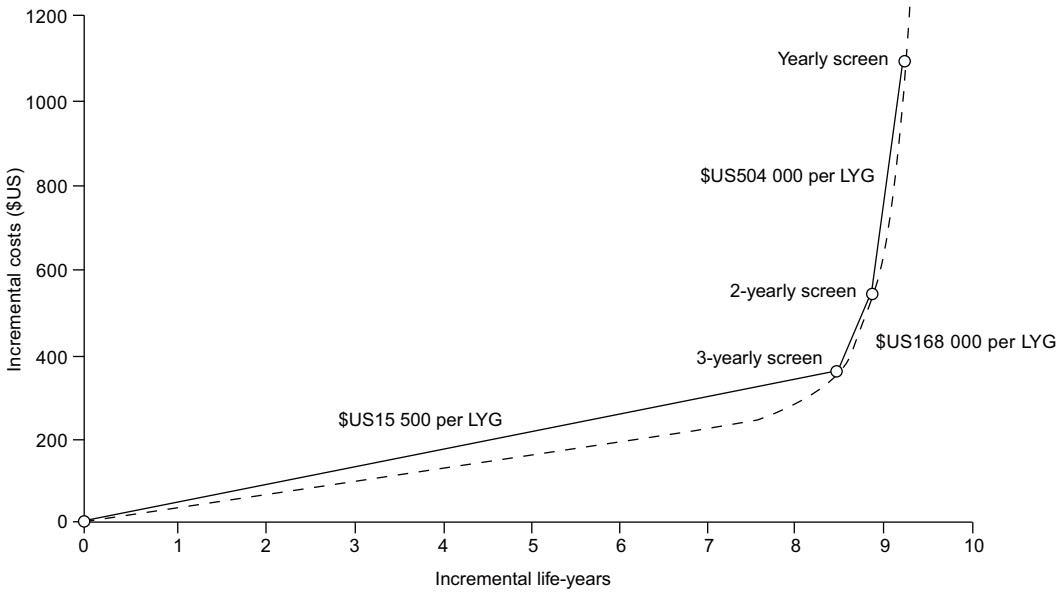

Fig. 3. Changing marginal cost effectiveness at the intensity margin on the cost-effectiveness plane (the dotted line represents the estimated screening interval production function) [year of costing not stated].[31]. **LYG** = life-year gained.

certain patient group, then statistical information may be available on that parameter if a previous study has addressed the issue. In such a case, it is possible to estimate both the point estimate and the variation in that point estimate through standard statistical methods.

What we are interested in when considering the variability in any point estimate of the parameter of interest is its standard error (the standard deviation of the estimator). This is known as second-order uncertainty, which is in contrast to variability in the underlying population from which the sample is drawn (first-order uncertainty).[36] This becomes important when we consider how to handle uncertainty when sample data do not exist for the parameter of interest.

In the absence of data from which to estimate a parameter we know to be important for a model, we have to estimate that parameter by other means. Usually, this involves some form of literature review and/or consultation with medical experts. In traditional sensitivity analyses, analysts tend to report a range of values for parameters in their model and use this range to consider the effect of altering a parameter value either individually or in combination with others. While this sort of analysis is certainly a minimum requirement of any competent economic analysis, such an approach does not encourage analysts (or the intended audience) to think about the statistical issues associated with parameter estimation.

A brief consideration of the problem from a statistical point of view suggests that decision-analytic–type modelling of the kind considered here is inherently Bayesian in perspective. In standard statistical methods (such as those practised in almost all clinical trials), parameters to be estimated from the data are considered to have true values and do not vary. Probabilities attached to confidence limits relate to long-run coverage probabilities of the intervals, were the same experiment to be repeated many times.

In modelling the cost effectiveness of interventions, this approach is not taken. Instead, parameters are considered as random variables, which can take a range of values. Explicit recognition of this fact should lead analysts to specify prior distributions for parameters of interest for economic evaluation models, in line with standard Bayesian methods. Although these prior distributions will represent 'degrees of belief' in the parameters of interest, it does not necessarily follow that the analysis will become automatically 'subjective' (the great fear of many of those who object to Bayesian methodology). Indeed, it would seem that the normative economics of the situation require an empirical Bayes approach, whereby prior distributions are constructed from available data. Where data are lacking and it becomes necessary to engage experts to provide information on prior distributions then, where feasible, a number of experts should be consulted so that the prior distribution employed in the final analysis reflects uncertainty between experts as to the distribution of the unknown parameter rather than representing the subjective beliefs of a single expert. Eddy et al.[37,38] have outlined such an approach to synthesising data based on empirical Bayes methods, termed the 'confidence' profile technique.

The relevant uncertainty we hope to capture in the formation of the prior distribution is second-order uncertainty, not the variability in the values observed in a particular population (first-order uncertainty). This is helpful as it can reduce the number of distributions we may consider using to represent prior beliefs concerning parameters, in contrast to the very large number of distributions on offer in some software packages. This suggests that these distributions are available to capture the shape of population distributions. For example, it is widely accepted that for a particular group of patients, healthcare costs follow a skewed distribution.[39] Yet, in a review of 5 data sets of patient-level data exhibiting skewness, it was found that the central limit theorem is powerful enough such that the assumption of a normal sampling distribution for the mean per patient cost in each of the data sets was not unreasonable.[40] Even where some residual skewness remained in the sample distributions (estimated by nonparametric bootstrapping), this skewness was slight compared with that in the original sample. In

fact, statistical theory tells us that the skewness in the sampling distribution will be less by a factor $1/\sqrt{n}$ where n is the sample size. Hence, it may be possible to exploit the power of the central limit theorem when specifying prior distributions for parameters that could, in principle, be estimated as averages for a particular patient population.

Although an assumption of normality (together with the binomial estimate of standard error) is often employed to calculate a confidence interval for a proportion, the normality assumption is rarely appropriate for a probability since a probability must be bounded on the interval 0-1. However, the beta distribution can be employed, which is bounded on the interval 0-1 and resembles a normal distribution for some parameterisations. Furthermore, the beta distribution is widely employed in the Bayesian paradigm since a beta prior combined with binomial data leads to a beta posterior – a property known as conjugacy.[41]

In summary, for those parameters of a cost-effectiveness model that could, in principle, be estimated from observed data, consideration should be given to the prior distribution of these parameters to reflect uncertainty. Where possible, this should be based on the available data from studies, supplemented where necessary by expert opinion. The specified prior distributions should relate to second-order uncertainty rather than the variability in parameter values, and care should be taken to ensure that the prior distributions chosen are consistent with any logical bounds on the parameter values.

In specifying prior distributions, it is important to consider the level of generalisability required for the model. Generalisability relates to the setting of the study: for a given population of patients, would the resource use and health outcome consequences observed in one hospital, region or country be replicated in other locations? This area of uncertainty is linked to known variations in clinical practice within and between countries.[42-44] Note that generalisability relating to setting is essentially data-driven, in that this issue can be handled by relating the specification of prior distributions to the scope of the evaluation. For example, an evaluation of a proposed national screening programme should include prior distributions for unit costs/resource use that reflect the variation between clinical centres in different parts of the country. By contrast, an evaluation of an intervention designed to inform a particular institution might have very little variation in the prior distributions related to unit costs if that institution is able to provide accurate information on those costs. These examples emphasise how cost-effectiveness modelling might be made much more dynamic than a single publication summarising the results of a study would suggest. Through the use of personal computers, or even the internet, there is no reason why cost-effectiveness models could not be tailored to local circumstances if appropriate local information exists.

1.4 Modelling Uncertainty

The US Panel on cost-effectiveness analysis[22] has recently reported, and the chapter[7] on reflecting uncertainty in cost-effectiveness analysis categorises uncertainty as parameter-related or modelling-related.[7] Modelling-related uncertainty is further categorised as related to the structure of the chosen model, for example through the use of a particular functional form for a transition probability over time in a Markov model, or related to the overall process of modelling, for example the choice of the appropriate states themselves in a Markov model. Having reviewed uncertainty as it applies to parameters of a model, it is worth setting this in the context of overall model uncertainty. In particular, there is a clear hierarchy of uncertainty, with uncertainty related to parameters that could in principle be sampled being conditional on the model structure, which in turn is conditional on the overall modelling process.

Modelling uncertainty has recently become a topic of some interest in the statistical literature, where it has been argued that focusing only on parameter uncertainty conditional on a given model will underestimate the true level of uncertainty and lead to poor inference.[45] The suggested solution is to run repeated analyses utilising different mod-

els and specify prior probabilities of different models across this model space. This is the solution proposed by the US Panel in relation to modelling structure uncertainty by appropriately weighting analyses employing different assumptions concerning the functional form of particular elements of the model structure.[7] At this level, there is a clear analogy with the statistical literature, where different forms of statistical model in terms of functional form and error structure are available. However, it is less clear that this sort of approach could be taken for the holistic process of generating cost-effectiveness models. The plethora of assumptions and judgement calls that go in to creating a cost-effectiveness model could not reasonably be addressed by the same analyst(s). In practice, only by commissioning the same cost-effectiveness model from different teams of analysts will any understanding of the uncertainty in the overall process of modelling be generated. If this were done then, in principle, modelling process uncertainty could be handled in a similar way to structural uncertainty by weighting the different models available and combining the results.

2. An Illustrative Model

The purpose of this section is to illustrate the principles outlined in section 1 using a hypothetical Markov cost-effectiveness model of a drug therapy. Although modelling uncertainty is important, it is not addressed within the context of this example. Instead, the focus is on parametric uncertainty, in particular on the appropriate way of implementing probabilistic sensitivity analysis and the limitations of available software in relation to the principle of second-order versus first-order parametric uncertainty.

2.1 Introduction to the Model

The model employed here has previously been described in relation to introducing Markov models for economic evaluation,[46] therefore only a brief description is given here. Figure 4 shows a state transition diagram for a chronic disease. Patients are assumed to begin in an asymptomatic phase and progress through a symptomatic phase to death. It is assumed that administering a drug therapy during the asymptomatic phase can delay progression of the disease and therefore improve survival and quality of life among treated patients. The costs are estimated from a profile of resources used at each stage of the disease and health outcomes are estimated by QALYs obtained by weighting time spent in each state by a utility value representing the quality of life in each state.

Table I shows the parameter values required to populate the model and table II shows the results of the baseline cost-effectiveness analysis indicating that the drug therapy has a cost per QALY ratio of approximately £8000. A typical 1-way sensitivity analysis is reported in table III (suggesting a range for the cost effectiveness of drug therapy of approximately £6000 to £12 000 per QALY) to serve as a comparison for the probabilistic method developed in section 2.2.

Not all of the parameters from table I are included in table III; the age parameter is excluded as a parameter relating to the patient characteristics

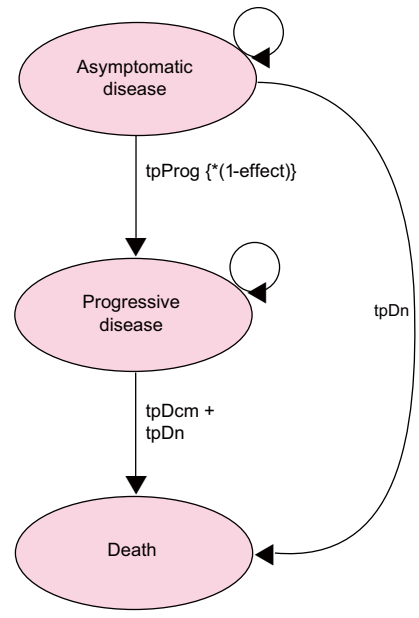

Fig. 4. Illustrative Markov model of disease progression (see table I for definitions of the variable names).

Table I. Parameter values for the illustrative Markov model of disease progression

Name	Value	Description
Transition probabilities		
tpProg	0.01	Coefficient of increase for probability of entering the progressive disease state[a]
tpDcm	0.15	Probability of dying from the disease in a single cycle
tpDn	Taken from standard life tables	Background all-cause mortality
Costs		
cAsymp	£500	Cost of 1 cycle in the asymptomatic disease state
cProg	£3000	Cost of 1 cycle in the progressive disease state
cDrug	£1000	Cost of drug for 1 cycle
cDeath	£1000	Cost associated with transition to the dead state
Quality-of-life adjustments		
uAsymp	0.95	Quality-of-life weight for 1 cycle in the asymptomatic disease state
uProg	0.75	Quality-of-life weight for 1 cycle in the progressive disease state
Other parameters		
cycle	1 year	Length of 1 cycle
effect	50%	Effectiveness of drug in terms of reducing disease progression
ini_age	55 years	Initial age at which patients are deemed to start the model
oDR	6%	Discount rate for outcomes
cDR	6%	Discount rate for costs

a Coefficient of increase for probability of entering the progressive disease state (note that this parameter is multiplied by the cycle of the model to provide an increasing probability of disease progression over time).

and the discount rate parameter is excluded since this relates to the underlying methods of analysis. The age-specific natural death rates are excluded to keep the analysis simple[5] and the cycle length chosen is for convenience and will not affect the results of the analysis.

2.2 Probabilistic Sensitivity Analysis

A distinction is made in this section between 2 approaches to conducting a simulation. The first approach, based on individual patients passing through the model, is known as first-order simulation and is analogous to first-order uncertainty. This method is usually employed for evaluating complex models where an analytical solution may be intractable. The second approach of second-order simulation is argued to be the appropriate method for probabilistic sensitivity analysis.

5 In any case, these rates are taken from large national studies and their variation can be assumed to be trivial compared with the other model parameters.

2.2.1 First-Order Simulation

A common method for the evaluation of Markov models, which appears to generate some measure of uncertainty is individual (Monte Carlo–based) simulation, and this method has been described in an influential report on the use of Markov models for medical decision-making.[47] The ability to conduct large numbers of individual-based Monte Carlo simulations is a standard feature of dedicated decision analysis software packages.[48,49]

In first-order simulation, a large number of patients are followed through the model individually,

Table II. Results of the baseline cost-effectiveness for the illustrative model

Strategy	Cost (£)	QALYs
No drug	9265	7.76
Drug	16 155	8.62
difference	6891	0.87
ICER	**7931 per QALY**	

ICER = incremental cost-effectiveness ratio; **QALY** = quality-adjusted life-year.

Table III. One-way sensitivity analysis of individual parameters in the model (see table I for parameter definitions)

Name	Value	Variable range (low-high)	ICER range (low-high[a])
tpProg	0.01	0.005-0.015	£12 211-6486
tpDcm	0.15	0.10-0.20	£8532-7605
cAsymp	£500	£250-750	£7534-8328
cProg	£3000	£2000-4000	£8611-7252
cDrug	£1000	£800-1200	£6076-9787
cDeath	£1000	£500-1500	£7982-7880
uAsymp	0.95	0.90-1.00	£8616-7347
uProg	0.75	0.60-0.90	£7198-8831
effect	50%	40-60%	£10 430-6251

a The low-high range relates to the value of the input parameter such that the ICER range can be reversed indicating that a low value of the input parameter increases the ratio and vice-versa.

ICER = incremental cost-effectiveness ratio.

and since an individual patient can only be in one state at a given time they may or may not transit between states in any given cycle.[6] Hence, the path followed by different patients will differ due to chance. Following the patient through the model allows an overall profile of costs and outcomes to be generated for that patient's path. For example, the first patient through the model may live for 15 years in the asymptomatic phase of the disease and a further 5 years in the progressive disease state before dying. This individual is therefore predicted to cost £10 423 and to accrue 10.5 QALYs. By contrast, the second patient through the model may live for only 5 years in the asymptomatic phase of disease and for only 1 year in the progressive disease state before dying. This patient will have a predicted cost of £4886 and will have accrued 4.5 QALYs. Averaging these costs and effects over a large number of patients gives the overall estimate of the average costs and effects in each arm.

The results of 10 000 simulations of individual patients through each arm of the model are given in table IV. It is clear that the (per patient) estimates of the average costs and effects for each arm of the model are very similar between table II (based on the analytical solution to the Markov model) and table IV. The individual simulation method will not give the same results on any 2 occasions because of the random nature of the simulation; however, providing the number of simulations over which the results are averaged is very large, the differences between the simulation method and the analytical solution will be small.

One advantage of the individual simulation method is that it gives an estimate of the likely variance associated with the costs and effects in each arm of the model given by the standard deviations in parentheses in table IV. However, this representation of uncertainty in the estimated cost and effects relates simply to the inherent uncertainty of the probabilistic structure of the model, and is often termed a 'first-order' Monte Carlo simulation. Importantly, nowhere in this analysis have any of the parameter values of the model been allowed to vary. The variation observed is purely a product of the alternative pathways through the model and is akin to the population variability of outcome in a clinical trial. Second-order Monte Carlo simulations can also be undertaken. In addition to allowing for uncertainty due to the way individuals travel through the model, the underlying model variables are allowed to vary over a given range with a given distribution.

Although individual-level Monte Carlo simulations of this sort can estimate the variability in the costs and effects of the different arms of the model,

Table IV. Results of 10 000 (first-order) simulations of individual patients for the illustrative model

Strategy	Cost (£) [SD]	QALYs [SD]
No drug	£9344 [5779]	7.76 [2.89]
Drug	£16 112 [6039]	8.61 [3.11]
difference	£6768 [?]	0.85 [?]
ICER	**£7944 per QALY [n/a]**	

ICER = incremental cost-effectiveness ratio; **n/a** = not applicable; **QALYs** = quality-adjusted life-years; **SD** = standard deviation; **?** indicates that it is not clear how the standard deviation of the estimated differences can be calculated.

6 Note that the same individual simulation approach could be used for decision tree–type models, since a given individual can only pass down 1 branch of the tree at a given chance node.

it is not clear how these results can be interpreted.[7] As the cost and effect pairs are simulated at the individual level, the simulated data can be seen as equivalent to patient-level data from clinical trials. Hence, one way to interpret the simulated data would be to apply standard parametric methods of confidence interval estimation to estimate the cost effectiveness of drug therapy.[50] However, the problem with such an approach is that the effective sample size is directly under the analyst's control through the number of simulations undertaken.

For example, the 10 000 simulation results reported in table IV are equivalent to having 10 000 individual patient cost and effect pairs in each arm of a study. The standard deviations reported in table IV indicate the overall variability in this (pseudo) population. However, in estimating the incremental cost-effectiveness ratio together with confidence limits, it is the mean cost and effect differences that are important, and their associated standard errors. The standard error of a mean value is equal to the standard deviation in the sample population divided by the square root of the sample size. This causes a problem since, unlike in a clinical trial situation, the sample size is under the direct control of the analyst when deciding how many simulations to run.[8] Hence, at very high numbers of simulations, such as the 10 000 reported in table IV, the standard error of the mean value collapses, and virtually no variation in the mean effects is apparent. For example, the standard error of the mean cost in the 'no drug' arm of table IV is 5779/100 (= 57.79). The coefficient of variation of the estimator is therefore 57.79/9344 (= 0.006) – representing a much greater level of precision than would ever occur in practice.

This effect can be clearly seen from figure 5, which presents the same approach applied to the first-order Monte Carlo simulations summarised in table IV. Individual Monte Carlo simulations based on 10, 100 and 1000 simulations are repeated in a series of 1000 experiments. The resulting cost and effect differences for each experiment are plotted on the cost-effectiveness plane. It is clear that for the experiments based on just 10 simulated individuals, the resulting uncertainty is huge. This reduces for 100 individual simulations and for the experiments based on 1000 individual simulations, there is very little variation left in the cost and effect differences (and therefore the ICER).

At the design stage of an economic evaluation to be conducted prospectively alongside a clinical trial, this individual simulation approach could be used to generate the variance data required for a sample-size calculation. By contrast, at the analysis stage, it is not clear how to determine the appropriate number of simulations (and therefore sample size) for estimating quasi-confidence limits. Some analysts have opted for choosing a given sample size to reflect, for example, the number of patients that might be expected to be treated in a given centre each year.[51,52] Although, this sort of solution may be of interest to decision-makers at the individual centre level, it does not give a general prescription for policy-making across different centres. Indeed, such decision-making at the individual centre level should be discouraged since overall uncertainty may appear high in some centres, owing to small numbers of patients, even though the overall cost effectiveness at a population level might be known with some precision.

2.2.2 Second-Order Simulation (in Isolation)

To generate appropriate results in a probabilistic sensitivity analysis it is necessary to dispense with the notion of uncertainty at the individual level. The ICER statistic is concerned with the difference in the mean treatment costs and the mean treatment effects across patient populations.

7 One might be tempted to calculate the individual-level cost-effectiveness ratios and average across the patients. Stinnett and Paltiel[36] examined the implications of calculating 'mean ratios' versus the 'ratio of means' and demonstrated that the analysis of 'mean ratios' is inappropriate.

8 Of course, at the design stage of a clinical trial the sample size for an evaluation is under the control of the design team, and is subject to cost and other logistical constraints. However, it is clear that in a Monte Carlo simulation situation there are no such constraints and the effective sample size (the number of simulations) can be set at the analysis stage.

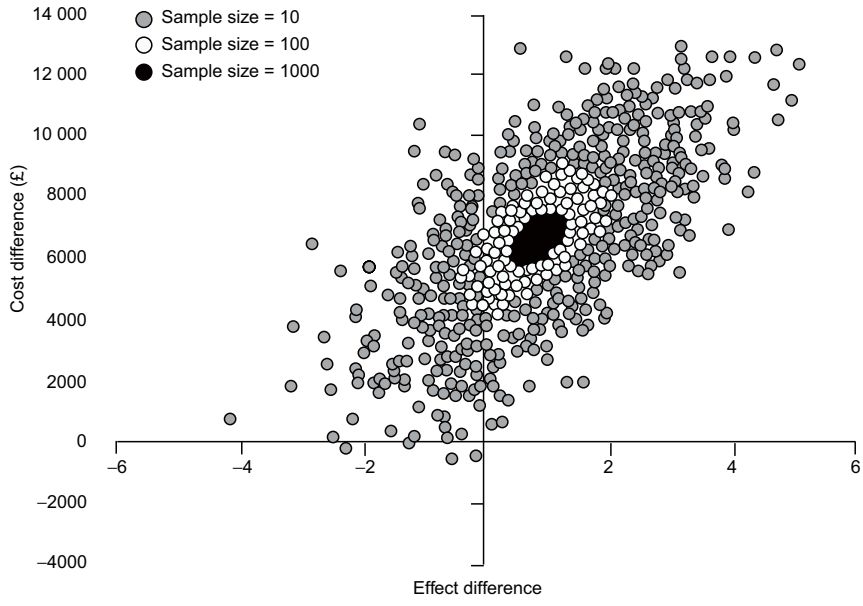

Fig. 5. 1000 experiments based on first-order Monte Carlo simulations of 10, 100 and 1000 individuals, plotted on the cost-effectiveness plane.

Although for an individual there may be considerable uncertainty associated with the effect (and cost) of a particular treatment, the expected effect (and expected cost) will be the average outcome across the group of patients receiving the treatments. Therefore, it seems appropriate to consider 'second-order' type uncertainty in isolation from 'first-order' uncertainty.[9]

Consideration of second-order uncertainty in isolation in this way is equivalent to repeating the analytical solution of the model employing different values for the underlying parameters sampled from specified ranges and distributions. This was undertaken for the Markov model, employing the assumption that the range of values for each of the parameters previously defined in table III represent 95% confidence intervals. Therefore equation 1 can be employed to get an estimate of the standard error (se):

$$se \approx \frac{u - l}{2 \times 1.96} \quad \text{(Eq. 1)}$$

where u and l are the upper and lower limits of the range, respectively.

The distributional assumptions made for each variable relate to the nature of that variable. Although cost data are technically bound to be positive, the point estimate of costs compared with the range were sufficiently far from zero for the normal distribution to be employed. In each case for cost parameters, the normal distribution with mean equal to the baseline parameter value and standard deviation estimated from equation 1 was employed. If there are concerns over the skewed nature of cost, the Log-Normal or gamma distributions could be employed for cost parameters. Since transition probabilities and utilities are bounded on the 0-1 interval these parameters were assumed to have a beta distribution, which can be bounded on the in-

9 Note that this causes problems for the dedicated decision analysis software used to generate the individual simulation results.[49] Although it is possible to run only a first-order analysis, the second-order uncertainty can only be included as an adjunct to the first-order uncertainty. This problem should be rectified in future releases of the software.

terval 0-1, has parameters α and β, with mean and standard deviation (SD) given in equation 2.

$$\text{mean} = \frac{\alpha}{\alpha+\beta} \quad \text{sd} = \sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}} \qquad \text{(Eq. 2)}$$

If the probabilities for the model are based on proportions and the original data are available, then Bayesian methods can be applied directly and a beta posterior distribution for the probability can be obtained by updating a beta prior with the binomial data.[41] If only the point estimate and confidence limit for a proportion are available, then a beta distribution can be fitted by equating the point estimate and the estimated standard deviation from equation 1 with the expressions in equation 2 and solving to give the parameter values[53] (this is known as method of moments estimation). This approach was employed to fit beta distributions to the parameters in the illustrative model.

The process of resampling from each of the distributions and recalculating the cost effectiveness from the model was repeated 1000 times to generate a distribution for the estimated ICER. The results of these 1000 simulated ICERs are presented in figure 6 in the form of a histogram. Overlaid on the histogram is the normal distribution with the same mean and variance as the simulated ICER values. It is clear that the simulated ICER distribution is heavily skewed; a common feature due to the non-negligible probability of the denominator of the ratio in the neighbourhood of 0.

Uncertainty intervals[10] can be estimated from this simulated data by taking the 2.5 and 97.5 percentile values to represent the endpoints (for a 95% interval). In other words, 2.5% of the simulated ICERs are excluded from the extreme ends of the distribution to get the interval. This approach is illustrated in figure 7 and the estimated interval for the ICER of £4560 to £15 043 therefore represents the probabilistic sensitivity analysis interval for the model estimate of the cost effectiveness of the drug intervention. In figure 8, this interval is compared to the intervals implied by the standard 1-way sensitivity analysis (from table III) and an extreme scenario analysis – obtained by combining

10 Uncertainty intervals are employed here as a generic term rather than employing the frequentist 'confidence interval' or the Bayesian equivalent 'credible interval'.

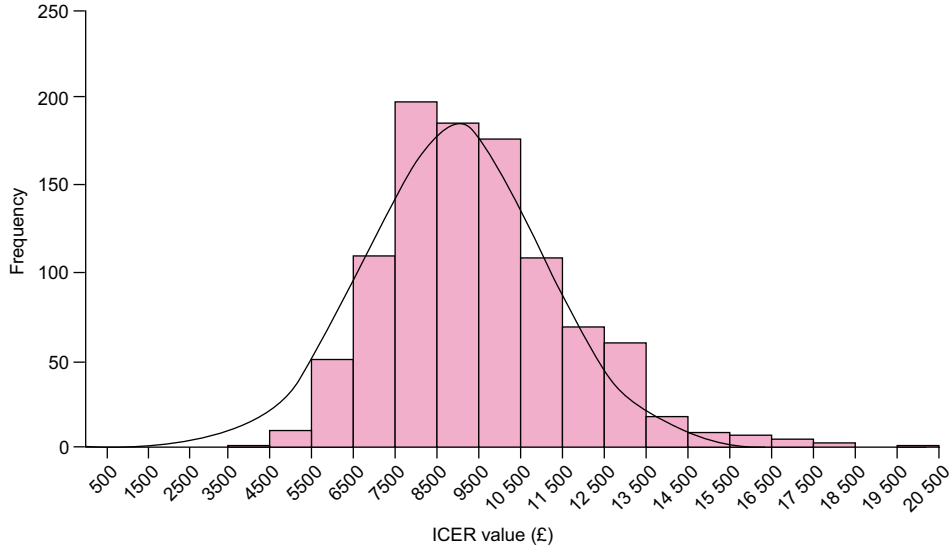

Fig. 6. Distribution of the simulated incremental cost-effectiveness ratio (ICER) values in a 1000 cohort simulation of the Markov model.

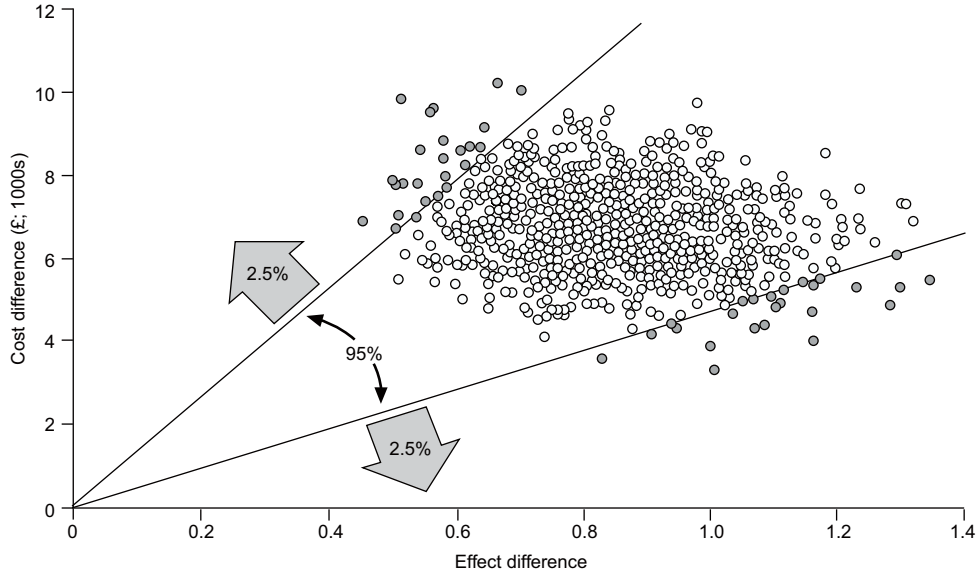

Fig. 7. Results of the 1000 cohort simulations on the cost-effectiveness plane with the percentile uncertainty interval endpoints.

the most optimistic/pessimistic values from table III to form a best/worst case scenario (from the point of view of the manufacturer of the drug therapy).

In the probabilistic sensitivity analysis reported here, no allowance was made for any covariance that may exist between the variables, instead it was assumed that all parameters were varying independently. There is no reason why a probabilistic sensitivity analysis could not be undertaken with 2 or more variables having some covariance. Rather, the problem is likely to be a lack of information on whether or not such covariant relationships exist. However, in some situations it should be very clear that such a relationship does exist (for example, the well documented relationship between the sensitivity and the specificity of a test), in which case these parameters should follow an appropriate joint distribution in a probabilistic analysis. Whether covariance between variables will result in greater or less variance in cost effectiveness is impossible to say since it will depend on the variables and on the direction of covariance. It is clear, however, that covariance could potentially increase the overall uncertainty in the ICER value.

2.3 Presenting Results: Cost-Effectiveness Acceptability Curves

The description of the method for calculating confidence intervals for ICERs above comes with a 'health warning'. When uncertainty covers more than 1 quadrant of the cost-effectiveness plane, negative ICERs will result. This causes enormous problems of interpretation for confidence intervals[54] since negative ICERs due to negative cost differences (favouring the intervention under evaluation) and negative ICERs due to negative effect differences (evidence against the intervention) become conflated.[11] The net-benefit framework[55] is a new approach to handling uncertainty in cost-effectiveness analysis that overcomes these problems by using the decision rule to rescale the cost-effectiveness decision onto either the cost or the health scale.

11 The intervals presented in the illustrative example are valid as all the simulated results were in the positive quadrant of the cost-effectiveness plane.

This problem of the interpretation of confidence intervals when ICERs become negative can also be overcome by employing cost-effectiveness acceptability curves to summarise uncertainty.[56] Such curves also directly summarise the evidence in favour of the intervention being cost effective for all possible values of the maximum acceptable cost-effectiveness ratio appropriate for decision-making. For this reason, cost-effectiveness acceptability curves have been argued to be a more useful way of presenting probabilistic information than confidence intervals,[50] and exactly the same arguments apply to the results obtained from probabilistic sensitivity analysis. Hence, rather than employing the data presented in figures 6 and 7 to estimate an interval, the same data could be used to estimate a cost-effectiveness acceptability curve instead. Such a curve is generated by considering how likely it is that the true costs and effects of an intervention fall to the right of a line on the cost-effectiveness plane, with slope equal to the maximum acceptable (or ceiling) ICER appropriate for decision-making purposes. Rotating the line anticlockwise from the horizontal through to the vertical gives all values of this ceiling ratio from 0 to ∞, and the associated probability that the intervention is cost-effective[12] at each point gives the cost-effectiveness acceptability curve. Such a curve is therefore presented in figure 9 for the results of the second-order simulation. For each value of cost-effectiveness decision rule, the proportion of observed simulation results from figure 7 lying below and to the right of the line on the cost-effectiveness plane representing this rule gives the estimate of the probability that the intervention is cost effective. The point of inflexion on the curve occurs at the 50% point and corresponds to the baseline estimate of cost effectiveness.

12 It is worth noting that this interpretation of cost-effectiveness acceptability curves as showing the probability that an intervention is cost effective (probability of the hypothesis given the data) necessarily requires a Bayesian interpretation.[57]

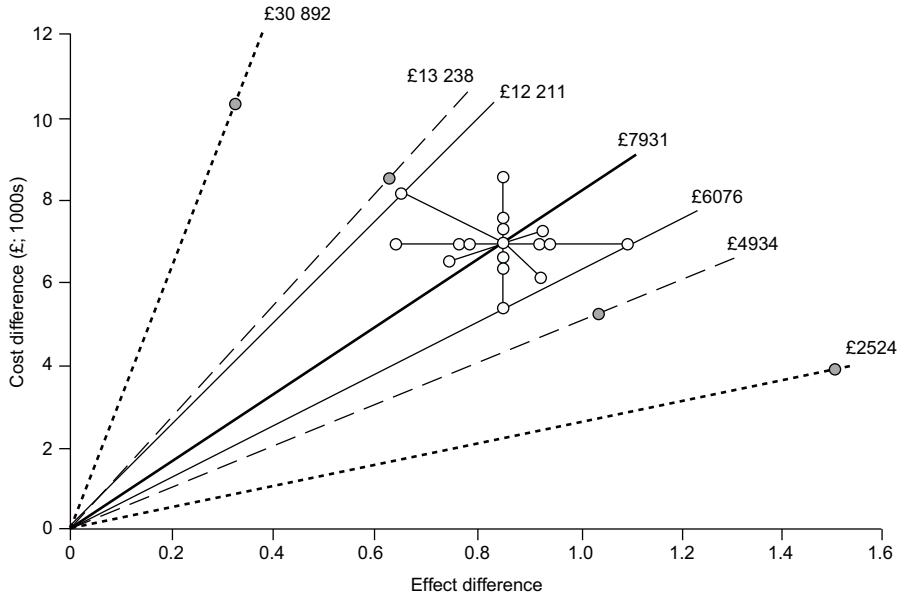

Fig. 8. Probabilistic sensitivity analysis on the cost-effectiveness plane compared with the standard sensitivity analysis methods. The thick solid line represents the baseline cost effectiveness of drug therapy and the thin solid lines represent the interval implied by a 1-way sensitivity analysis. The black dotted lines represent the interval implied by an extreme scenario analysis and the black dashed lines represent the probabilistic sensitivity analysis interval falling between the other 2 intervals.The slopes of the lines, which give the estimated ICER values, are also shown in the figure.

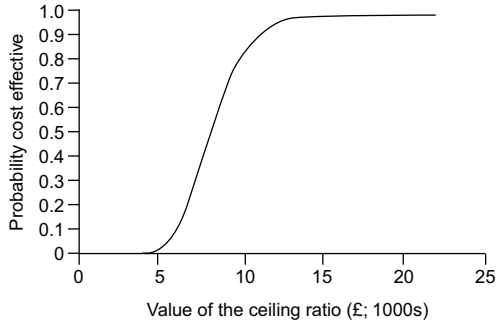

Fig. 9. Cost-effectiveness acceptability curve for the illustrative model.

It was also argued that probabilistic analysis is only appropriate for those parameters that could, in principle, be estimated from sample data. Other parameters, relating to patient characteristics and to the analytical methods employed, should not be included in the probabilistic analysis since uncertainty in these parameters is not related to sampling variation. In the sensitivity analysis of the model reported in this article, the discount rate was not incorporated into the sensitivity analyses used to generate interval estimates. Hence, although the implications of uncertainty in the rate of discount for the results of the analysis need to be examined, ideally for each value of the discount rate employed there should also be some representation of uncertainty as well as the point estimate of cost effectiveness.

Cost-effectiveness acceptability curves (in contrast to confidence intervals) are ideally suited to this purpose, as is shown in figure 10. If the ceiling cost-effectiveness ratio appropriate for decision-making purposes is £10 000 per QALY, there is a 76% chance that the drug intervention is cost effective at baseline discount rates of 6%. Varying the discount rate from 0 to 10% for both costs and health outcome effects has an important effect on the results of the analysis leading to a 100% (99.7%) and a 38% chance, respectively, that the intervention is cost effective. Discounting costs at a rate of 6% without discounting health outcomes leads to highly improved cost effectiveness as shown by the shifting of the cost-effectiveness acceptability curve to the left in figure 10. Clearly, a similar set of curves could be generated with respect to age (or other parameters related to patient characteristics).

3. Discussion

Modelling-based economic evaluations make up the majority of economic analyses of healthcare interventions undertaken to date. Such analyses are still subject to uncertainty, although the lack of sample data on patient-level costs and effects means that standard statistical methods for handling uncertainty cannot be used. Instead, sensitivity analysis is the recommended method for handling uncertainty when sample data are unavailable. However, there are problems with standard sensitivity analyses. One-way sensitivity analysis, by considering the effect of parameters individually, is likely to underestimate overall uncertainty, while the opposite is true of extreme scenario analysis. The solution described in this article is to employ a probabilistic sensitivity analysis by specifying prior distributions for unknown data-related parameters, and the importance of excluding parameters related to analytical method and patient characteristics is emphasised.

Very few probabilistic analyses have been presented in the health economic evaluation literature. Although the methods have been described in relation to medical decision-making,[58,59] in a review of cost-effectiveness studies published up to 1996,[28] only 7 studies (out of 492 reviewed) reported any kind of probabilistic analysis.[30,60-64] It is worth noting that 2 of these recent studies[30,63] included the discount rate as a variable in the probabilistic analysis, whereas it is argued here that it is more appropriate to separate out methodological variables in order to enhance the comparability of results. Furthermore, one of the studies was based on a first-order Monte Carlo simulation employing a sample size assumption similar to that reported in clinical trials in the area.[64] This suggests that the methods of probabilistic sensitivity analysis have not been widely employed outside of methodology-type journals, perhaps because the principles underlying probabilistic analysis are not well understood.

O'Brien et al.[65] describe 3 major limitations associated with sensitivity analysis: (i) that the variables and ranges of those variables included in an analysis are under the discretion of the analyst, creating the potential for bias; (ii) that interpretation of sensitivity analysis is essentially arbitrary due to the absence of guidelines or standards concerning what is robust; and (iii) that variation of uncertain parameters individually in a 1-way sensitivity analysis carries the risk that interactions between parameters are ignored. They further argue that these limitations of traditional sensitivity analysis methods in the clinical literature have led to the development of probabilistic sensitivity analysis methods based on Monte Carlo simulation methods.[65]

In discussing and developing the concept of probabilistic sensitivity analysis applied to economic evaluation, it is clear that the method only addresses the third limitation, and only to a limited extent. While interactions between parameters can be modelled in a probabilistic analysis, it is most likely that such interactions are simply unknown. What probabilistic analysis does show, is that even where the individual variables in an analysis are assumed to be independent, 1-way sensitivity analysis is likely to underestimate the range of uncertainty.

There is no way of avoiding the fact that most economic evaluations are, and are likely to continue to be, deterministic in nature. This means that sensitivity analysis methods must continue to be employed and that they will necessarily be under the discretion of the analyst. However, by adopting a formal (Bayesian) statistical perspective, probabilistic sensitivity analysis encourages the analyst (and the audience) to think more carefully about the relevant uncertainty in the parameter that is being described. The specification of a prior distribution for a particular parameter (including justification for the formulation of that prior) within the framework described here could be seen as meeting some of the challenges for cost-effectiveness modelling in terms of increasing the clarity of modelling.[66] Also note that the specification of prior distributions for parameters within a Bayesian model gives a straightforward mechanism for updating parameter distributions as more data becomes avail-

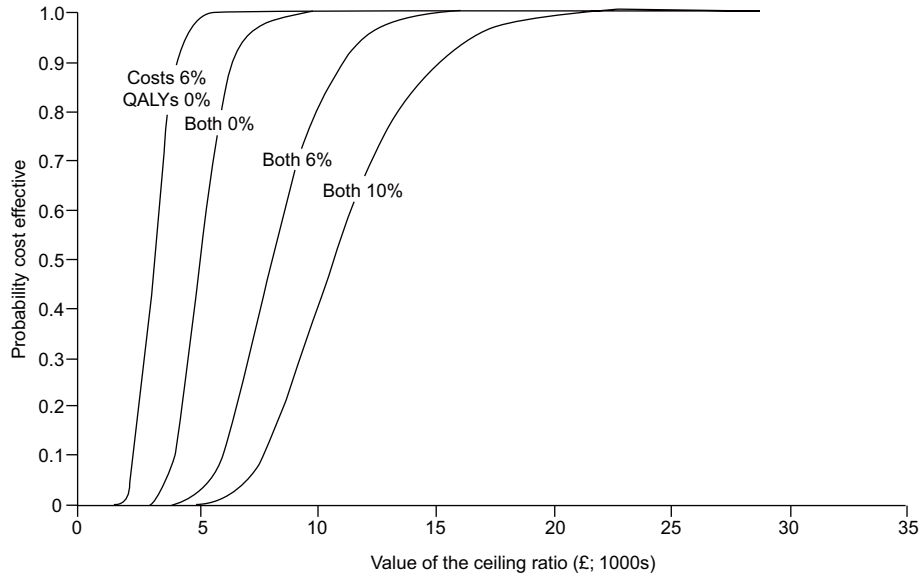

Fig. 10. Discount rate sensitivity analysis and cost-effectiveness acceptability curves. **QALYs** = quality-adjusted life-years.

able, which links well with the interative approach to cost-effectiveness analysis suggested by some commentators.[67]

The formal use of Bayesian methods may also go some way to helping with the interpretation of the resulting distribution of cost-effectiveness results [68]. The result is that we should have more confidence in the validity of such models – even if (or perhaps especially if), as a consequence of following the approach outlined here, the resulting uncertainty is so wide as to give no clear policy guidance.[66]

In some ways, the arguments proposed here are only half the story. A fully Bayesian decision-making approach would also look at the loss of function associated with decision-making and examine the expected value of perfect information (EVPI) with regard to collecting more information on uncertain parameters. Important advances in economic evaluation modelling are being made by a number of researchers in this area.[69-71] It should be emphasised, however, that the probabilistic nature of the EVPI approach relies heavily on a valid underlying probabilistic sensitivity analysis, and it is important therefore that the mechanisms of creating such an analysis are well understood.

4. Conclusion

Many commentators are agreed that probabilistic sensitivity analysis of cost-effectiveness models offers advantages over traditional deterministic methods for handling uncertainty in cost-effectiveness models. This paper has reviewed those types of uncertainty that can arise in cost-effectiveness models and has outlined the probabilistic sensitivity analysis approach within a Bayesian statistical framework. It is argued that uncertainty in model parameters relating to methodological issues be handled outside of a probabilistic analysis using a 'reference case' of methods. Parameters relating to characteristics of the patient or intervention are not unknown, but should be stated as clearly as in any experimental evaluation. Uncertainty in the remaining parameters of a model should be ascribed distributions reflecting 'degrees of belief' concerning the possible values of those parameters. The model can then be repeatedly evaluated by drawing values at random from the parameter distributions in a probabilistic sensitivity analysis. This generates a distribution over the cost-effectiveness result of interest that can be used to summarise uncertainty.

Appendix

The illustrative model used in this paper will be made available for download from www.ihs.ox.ac.uk/herc/.

Acknowledgements

I am grateful to Dr Alastair Gray for comments and suggestions on earlier versions of this work, the participants of the Expert Workshop on 'Validating Cost-Effectiveness Models', 22 to 23 April 1999, held in Sheffield, England, and particularly to Dr Dennis Fryback for his insightful comments on a previous draft of this article, and finally to 2 anonymous referees for their comments. Of course, the responsibility for errors and inaccuracies in this article is all my own.

References

1. Posnett J, Jan S. Indirect cost in economic evaluation: the opportunity cost of unpaid inputs. Health Econ 1996; 5 (1): 13-23
2. Buxton MJ, Drummond MF, van Hout BA, et al. Modelling in economic evaluation: an unavoidable fact of life. Health Econ 1997; 6: 217-27
3. Association of British Pharmaceutical Industries. Pharmaceutical industry and Department of Health agree guidelines for the economic analysis of medicines [press release]. London: Association of British Pharmaceutical Industries, 1994
4. Australian Commonwealth Department of Human Services and Health. Guidelines for the pharmaceutical industry on preparation of submissions to the Pharmaceutical Benefits Advisory Committee: including major submissions involving economic analysis. Canberra: Australian Government Publishing Service, 1995
5. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. BMJ 1996; 313: 275-83
6. Canadian Coordinating Office for Health Technology Assessment. Guidelines for the economic evaluation of pharmaceuticals: Canada. 2nd ed. Ottawa: Canadian Coordinating Office for Health Technology Assessment (CCOHTA), 1997
7. Manning WG, Fryback DG, Weinstein MC. Reflecting uncertainty in cost-effectiveness analysis. In: Gold MR, Siegel JE, Russell LB, et al., editors. Cost-effectiveness in health and medicine. New York (NY): Oxford University Press, 1996: 247-75
8. Drummond M, Brandt A, Luce B, et al. Standardizing methodologies for economic evaluation in health care: practice, problems, and potential. Int J Technol Assess Health Care 1993; 9 (1): 26-36

9. Lipscomb J. Time preference for health in cost-effectiveness analysis. Med Care 1989; 27 (3 Suppl.): S233-53
10. Parsonage M, Neuburger H. Discounting and health benefits. Health Econ 1992; 1 (1): 71-6
11. Cairns J. Discounting and health benefits: another perspective [comment]. Health Econ 1992; 1 (1): 76-9
12. Coyle D, Tolley K. Discounting of health benefits in the pharmacoeconomic analysis of drug therapies. Pharmacoeconomics 1992; 2 (2): 153-62
13. Katz DA, Welch HG. Discounting in cost-effectiveness analysis of healthcare programmes. Pharmacoeconomics 1993; 3 (4): 276-85
14. Lipscomb J. The proper role for discounting: search in progress. Med Care 1996; 34 (12 Suppl.): DS119-23
15. Lipscomb J, Weinstein M, Torrance, et al. Time preference. In: Gold MR, Siegel JE, Russell LB, et al., editors. Cost-effectiveness in health and medicine. New York (NY): Oxford University Press, 1996: 214-46
16. HM Treasury. Appraisal and evaluation in central government. London: H.M. Stationery Office, 1997
17. Donaldson C. Willingness to pay for publicly-provided goods: a possible measure of benefit. J Health Econ 1990; 9: 103-18
18. Llewellyn TH, Sutherland HJ, Tibshirani R, et al. The measurement of patients' values in medicine. Med Decis Making 1982; 2 (4): 449-62
19. Mehrez A, Gafni A. Quality-adjusted life years, utility theory, and healthy-years equivalents. Med Decis Making 1989; 9: 142-9
20. Torrance GW. Measurement of health state utilities for economic appraisal: a review. J Health Econ 1986; 5: 1-30
21. Conroy RM, O'Brien E, O'Malley K, et al. Measurement error in the Hawksley random zero sphygmomanometer: what damage has been done and what can we learn? BMJ 1993; 306 (6888): 1319-22
22. Gold MR, Siegel JE, Russell LB, et al. Cost-effectiveness in health and medicine. New York (NY): Oxford University Press, 1996
23. Meltzer David, Accounting for future costs in medical cost-effectiveness analysis, J. Health Econ 1997; 16 (1): 33-64
24. Drummond MF, O'Brien B, Stoddart GL, et al. Methods for the economic evaluation of health care programmes. 2nd ed. Oxford: Oxford University Press, 1997
25. Koopmanschap MA, Rutten FFH. Indirect costs: the consequence of production loss or increased costs of production. Med Care 1996; 34 (12 Suppl.): DS59-68
26. Koopmanschap MA, Rutten FFH, Van Ineveld BM, et al. The friction cost method for measuring indirect costs of disease. J Health Econ 1995; 14: 171-89
27. Russell LB. Is prevention better than cure? Washington, DC: The Brookings Institution, 1986
28. Briggs AH, Gray AM. Handling uncertainty when performing economic evaluation of health care interventions. Health Technol Assess 1999; 3 (2): 1-134
29. Anderson MH, Camm AJ. Implications for present and future applications of the implantable cardioverter-defibrillator resulting from the use of a simple model of cost efficacy. Br Heart J 1993; 69 (1): 83-92
30. Pharoah PD, Hollingworth W. Cost effectiveness of lowering cholesterol concentration with statins in patients with and without pre-existing coronary heart disease: life table method applied to health authority population. BMJ 1996; 312 (7044): 1443-8
31. Eddy DM. Screening for cervical cancer . Ann Intern Med 1990; 113 (3): 214-26
32. Johannesson M, Weinstein MC. On the decision rules of cost-effectiveness analysis. J Health Econ 1993; 12: 459-67
33. Weinstein M, Fineberg HV. Clinical decision analysis. Philadelphia (PA): WB Saunders Company, 1980
34. Williams A. The economics of coronary artery bypass grafting. BMJ 1985; 291: 326-9
35. Pauker SG, Kassirer J. Decision analysis. N Engl J Med 1987; 316 (5): 250-8
36. Stinnett AA, Paltiel AD. Estimating CE ratios under second-order uncertainty: the mean ratio versus the ratio of means. Med Decis Making 1997; 17 (4): 483-9
37. Eddy DM, Hasselblad V, Shachter R. A Bayesian method for synthesizing evidence: the Confidence Profile Method. Int J Technol Assess Health Care 1990; 6 (1): 31-55
38. Eddy DM, Hasselblad V, Shachter R. An introduction to a Bayesian method for meta-analysis: the confidence profile method. Med Decis Making 1990; 10 (1): 15-23
39. Zhou XH, Melfi CA, Hui SL. Methods for comparison of cost data. Ann Intern Med 1997; 127 (8 Pt 2): 752-6
40. Briggs A, Gray A. The distribution of health care costs and their statistical analysis for economic evaluation. J Health Serv Res Policy 1998; 3: 233-45
41. Gelman A, Carlin J, Stern H, et al. Bayesian data analysis. London: Chapman & Hall, 1995
42. Wennberg J, Gittelsohn A. Variations in medical care in small areas. Sci Am 1982; 4: 120-4
43. Andersen TF, Mooney G. The challenge of medical practice variations. Basingstoke: Macmillan, 1990
44. Cleary PD, Greenfield S, Mulley AG, et al. Variations in length of stay and outcomes for six medical and surgical conditions in Massachusetts and California. JAMA 1991; 266 (1): 73-9
45. Draper D. Assessment and propagation of model uncertainty. J R Stat Soc Br 1995; 57 (1): 45-97
46. Briggs A, Sculpher M. An introduction to Markov models for economic evaluation. Pharmacoeconomics 1998; 13 (4): 397-409
47. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Med Decis Making 1993; 13: 322-38
48. SMLTREE [computer program]. 2.9. Hollenberg J, 1989
49. Decision Analysis by TreeAge (DATA) [computer program]. TreeAge Software Inc. v3.5. Williamstown (MA): TreeAge, 1998
50. Briggs AH, Fenn P. Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. Health Econ 1998: 7 (8): 723-40
51. Sharples LD, Briggs A, Caine N, et al. A model for analyzing the cost of main clinical events after cardiac transplantation. Transplantation 1996; 62 (5): 615-21
52. Hunink MG, Bult JR, de Vries J, et al. Uncertainty in decision models analyzing cost-effectiveness: the joint distribution of incremental costs and effectiveness evaluated with a nonparametric bootstrap method. Med Decis Making 1998; 18 (3): 337-46
53. Pratt JW, Raiffa H, Schlaifer R. Introduction to statistical decision theory. Cambridge (MA): MIT Press, 1995
54. Stinnett AA, Mullahy J. The negative side of cost-effectiveness analysis [letter]. JAMA 1997; 277 (24): 1931-2
55. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Med Decis Making 1998; 18 (2 Suppl.): S65-80
56. van Hout BA, Al MJ, Gordon GS, et al. Costs, effects and C/E-ratios alongside a clinical trial. Health Econ 1994; 3 (5): 309-19
57. Briggs AH. A Bayesian approach to stochastic cost-effectiveness analysis. Health Econ 1999; 8 (3): 257-61

58. Doubilet P, Begg CB, Weinstein MC, et al. Probabilistic sensitivity analysis using Monte Carlo simulation: a practical approach. Med Decis Making 1985; 5: 157-77
59. Critchfield GC, Willard KE, Connelly DP. Probabilistic sensitivity analysis methods for general decision models. Comput Biomed Res 1986; 19: 254-65
60. Berwick DM, Cretin S, Keeler E. Cholesterol, children, and heart disease: an analysis of alternatives. Pediatrics 1981; 68 (5): 721-30
61. Hornberger JC. The hemodialysis prescription and cost effectiveness: Renal Physicians Association Working Committee on Clinical Guidelines. J Am Soc Nephrol 1993; 4 (4): 1021-7
62. Gabriel SE, Campion ME, O'Fallon WM. A cost-utility analysis of misoprostol prophylaxis for rheumatoid arthritis patients receiving nonsteroidal antiinflammatory drugs. Arthritis Rheum 1994; 37 (3): 333-41
63. Fiscella K, Franks P. Cost-effectiveness of the transdermal nicotine patch as an adjunct to physicians' smoking cessation counseling. JAMA 1996; 275 (16): 1247-51
64. Oh PI, Maerov P, Pritchard D, et al. A cost-utility analysis of second-line antibiotics in the treatment of acute otitis media in children. Clin Ther 1996; 18 (1): 160-82
65. O'Brien BJ, Drummond MF, Labelle RJ, et al. In search of power and significance: issues in the design and analysis of stochastic cost-effectiveness studies in health care. Med Care 1994; 32 (2): 150-63
66. Sheldon TA. Problems of using modelling in the economic evaluation of health care [editorial]. Health Econ 1996; 5 (1): 1-11
67. Sculpher MJ, Drummond MF, Buxton MJ. The iterative use of economic evaluation as part of the process of health technology assessment. J Health Services Res Policy 1997; 2: 26-30
68. Sonnenberg FA, Roberts MS, Tsevat J, et al. Toward a peer review process for medical decision analysis models. Med Care 1994; 32 (7 Suppl.): JS52-64
69. Claxton K, Posnett J. An economic approach to clinical trial design and research priority-setting. Health Econ 1996; 5 (6): 513-24
70. Felli JC, Hazen GB. Sensitivity analysis and the expected value of perfect information. Med Decis Making 1998; 18 (1): 95-109
71. Fenwick E, Claxton K, Sculpher M, et al. Improving the efficiency and relevance of health technology assessment: the role of decision analytic modelling. Health Economists' Study Group Conference; 1999 Jan 6-8; Birmingham

Correspondence and offprints: Dr *Andrew H. Briggs*, Health Economics Research Centre, University of Oxford, Institute of Health Sciences, Headington, Oxford OX3 7LF, England.
E-mail: andrew.briggs@ihs.ox.ac.uk
URL: www.ihs.ox.ac.uk/herc/