Probabilistic Analysis and Computationally Expensive Models: Necessary and Required?

Susan Griffin, MSc, Karl Claxton, PhD, Neil Hawkins, PhD, Mark Sculpher, PhD

Center for Health Economics, University of York, York, UK

ABSTRACT

Objective: To assess the importance of considering decision uncertainty, the appropriateness of probabilistic sensitivity analysis (PSA), and the use of patient-level simulation (PLS) in appraisals for the National Institute for Health and Clinical Excellence (NICE).

Methods: Decision-makers require estimates of decision uncertainty alongside expected net benefits (NB) of interventions. This requirement may be difficult in computationally expensive models, for example, those employing PLS. NICE appraisals published up until January 2005 were reviewed to identify those where the assessment group utilized a PLS model structure to estimate NB. After identifying PLS models, all appraisals published in the same year were reviewed.

Results: Among models using PLS, one out of six conducted PSA, compared with 16 out of 24 cohort models. Justification for omitting PSA was absent in most cases. Reasons for choosing PLS included treatment switching, sampling patient characteristics and dependence on patient history. Alternative modeling approaches exist to handle these, including semi-Markov models and emulators that eliminate the need for two-level simulation. Stochastic treatment switching and sampling baseline characteristics do not inform adoption decisions. Modeling patient history does not necessitate PLS, and can depend on the software used. PLS addresses nonlinear relationships between patient variability and model outputs, but other options exist. Increased computing power, emulators or closed-form approximations can facilitate PSA in computationally expensive models.

Conclusions: In developing models analysts should consider the dual requirement of estimating expected NB and characterizing decision uncertainty. It is possible to develop models that meet these requirements within the constraints set by decision-makers.

Keywords: cost-effectiveness analysis, decision uncertainty, decision-analytic modeling, patient-level simulation, probabilistic analysis.

Introduction

Economic evaluation of health-care technologies is increasingly recommended internationally for informing the allocation of health-care resources [1]. For almost every set of technologies considered, it will be necessary to combine information on costs and effects from several sources, and modeling techniques will be employed. When building any decision model, it is important to consider how to handle uncertainty and variability, because these affect the value and interpretation of the model output. Decision-makers require unbiased estimates of the costs and effects of alternative interventions for identifiable patient groups, and the ability of a model structure to provide these can depend on how uncertainty, variability, and heterogeneity are handled in the structure. There is also a need to provide an assessment of whether current evidence is sufficient to support the decision to adopt a technology. By formally estimating decision uncertainty, the value of obtaining additional information on model parameters can be assessed [2–4].

In this article we explore methods available to address uncertainty, variability, and heterogeneity within decision models. We explore how the current methods for conducting probabilistic sensitivity analysis (PSA) to characterize decision uncertainty can conflict with the use of computationally expensive model structures. The premise is that the purpose of a decision model is to provide unbiased estimates of expected cost and effects, and of decision uncertainty, in a timely fashion and within resource constraints as determined by the decision-maker that commissions the model. In this article, the focus is on one particular decision-maker, the National Institute for Health and Clinical Excellence (NICE) in the UK. This focus has a number of advantages: 1) NICE is an agency with a history of using decision-analytic cost-effectiveness models as a basis for deciding whether to support the use of particular health-care technologies in the UK National Health Service (NHS) [5]; 2) by focusing on a single decision-maker, the case studies identified in the review will have been developed with consistent...
time and resource constraints; 3) NICE has specified a reference case that defines those methods considered most appropriate for informing decisions about the adoption of new technologies [6], which makes the case studies identified in the review comparable in this respect; and 4) the potential conflict between PSA and computationally expensive model structures is pertinent, because the reference case now calls for the use of PSA as the appropriate way by which the combined implications of uncertainty in all model parameters be reflected [7].

Although the review is specific to NICE, issues relating to the provision of unbiased estimates of costs and effects and quantifying decision uncertainty, while reflecting the complexity of the disease process and treatment effect under time and resource constraints, are more general and relevant to decision models developed in different contexts.

**Review Methods**

The review allows an exploration of the potential conflict between the use of computationally expensive model structures and the implementation of PSA in models submitted to NICE. For the purposes of this review, computational expense refers to the limited resources available with which to produce model results, given the constraints of the decision-making process. Models may be computationally expensive for a number of reasons. For our review, the use of patient-level simulation (PLS) was selected as an indicator of computational expense because its use should be readily apparent in models submitted to NICE and, for a given model structure, analysis using PLS is more computationally expensive than a cohort analysis.

All technology appraisals detailed on the NICE Web site that were published up until January 2005 were reviewed to identify those where the independent technology assessment group had utilized a model structure involving PLS to provide an estimate of cost-effectiveness. After identifying models that employed PLS, appraisals published in the same year were identified and reviewed. The benefit of comparing models that employed PLS with models published in the same period is the consistency of constraints facing model developers. Given the computing expense of PSA, it might be expected that it would be more common among models employing a cohort framework. We identified common stated reasons for choosing a PLS rather than a cohort framework, and assessed their implications for model structure with respect to uncertainty, variability, and heterogeneity. For each case study, we ascertained whether a PSA had been undertaken. In cases where a PSA had not been undertaken, we explored the availability of alternative modeling techniques, such as less computationally expensive modeling structures or emulators. In those cases where PSA was performed, we discuss the techniques used.

**Uncertainty, Variability, and Heterogeneity**

In this section, we distinguish uncertainty, variability, and heterogeneity, explore the implications of each for model structure and interpretation of results, and briefly review the methods available for characterizing decision uncertainty.

**Decision Uncertainty**

Decision uncertainty can be regarded as epistemic uncertainty [8] which relates to model parameters that have a definite value, but which cannot be known with certainty. For example, the risk of future individual developing cervical cancer has a definite value, but one which we can only estimate with uncertainty. This uncertainty can be characterized with a distribution and can be reduced with further investigation. Epistemic uncertainty is not confined to model parameters, and may exist in the determination of model structure. This discussion concentrates on parameter uncertainty, although it can be generalized to other sources of uncertainty [9].

The decision to adopt a particular technology should be based on expected net benefit (NB) so that, when comparing mutually exclusive treatment strategies for a particular disease area, the optimal strategy is simply the one with the highest expected NB [10]. Nevertheless, uncertainty is important for two reasons: 1) in nonlinear models, or multilinear models with correlated parameters, unbiased estimates of expected NB require a characterization of uncertainty; and 2) decisions based on expected NB are only appropriate if there is also some consideration of whether current evidence is sufficient for allocating health-care resources, based on an assessment of the consequences of decision uncertainty [11]. If the decision uncertainty, and/or the consequences of adopting a suboptimal treatment strategy are large, the decision-maker may require further evidence on which to base the adoption decision [4].

**Variability**

Expected costs and effects are not only uncertain but also vary across individuals with identical observed characteristics. This variability can be regarded as aleatory uncertainty [8] which arises as a result of stochastic variation. It cannot be reduced through measurement, but can be characterized with empirical distributions. For example, the rate at which an individual’s cervical cancer develops will vary between patients. We can describe the distribution of the rate of cancer progression by counting the number of patients who progress at different rates. Nevertheless, further investigation would not reduce variation in the rate of
progression. Another example is where, given a probability of an event occurring, such as death, the realization of that event can be imagined as being governed by a lottery. So we may know with certainty that the probability of death is, for example, 5%, but we do not know which 5% of people will die.

Variability in itself is not relevant to an adoption decision based on expected NB. Nevertheless, it may be necessary to explicitly represent variability in model structures to obtain an unbiased estimate of expected NB if, within a patient population which is homogeneous in observed baseline characteristics, there is a nonlinear relationship between a characteristic that varies between patients and NB.

For example, suppose the outcome of interest, cost \((C)\), is a nonlinear function of some patient characteristic \((x)\) with mean, \(\mu\), that varies between patients \((i)\) according to a normal distribution with variance \(\sigma\): \[C_i = kx_i^2\] \[x_i \sim N(\mu, \sigma)\] (1) (2)
The expected value of \(x\) across all patients, \(E[x]\), cannot be used to derive the expected value of \(C\); an estimate of \(E[x^2]\) is required because \(E[x]^2 \neq E[x^2]\). In this instance, an analysis which failed to account for the variability in \(x\) across patients would provide biased estimates of expected costs. The use of PLS accounts for variability in all included parameters, regardless of whether there exist any nonlinear relationships between these parameters and model output. Nevertheless, there are a number of other methods by which we could also address this issue.

For example, by repeatedly sampling from \(\mu\) and \(\sigma\) we can estimate \(E[x^2]\) as an input to the model. Alternatively it may be possible to derive a linear approximation to the model. That is, we find a mathematical function of \(\mu\) and \(\sigma\) that gives us \(E[x^2]\), that is, determine function \(G\) such that: \[E[x^2] = G(\mu, \sigma^2)\] (3)
The requirement to account for variability in model structure under these circumstances does not negate the need to estimate decision uncertainty. Failure to account for uncertainty will also lead to biased estimates of cost and effect in a nonlinear model. Consequently, under these circumstances, both variability and uncertainty must be characterized in order properly to inform an adoption decision. In model structures which are linear with respect to variability but nonlinear with respect to uncertainty, unbiased estimates of NB require the characterization of uncertainty but not variability. Where models are linear or multilinear (with independent parameters), it is not necessary to represent variability or uncertainty to obtain unbiased estimates of expected NB. Nevertheless, it will still be necessary to represent uncertainty in a model structure to address the question of whether current evidence provides sufficient basis for the adoption decision.

**Heterogeneity**

Heterogeneity can be regarded as variation as a result of observed characteristics on which it is possible to condition model parameters and therefore expected NB. Such heterogeneity must be observable at the time at which the treatment decision is taken. For example, the risk of developing cervical cancer may depend on family history. In principle, this can be observed and subsequent decisions, such as the decision of whether to screen, based on this observation. This contrasts with variation in the rate of disease progression which is unobservable at the time at which the decision to screen is made. Thus, one could not decide to screen only those patients whose cancer would develop at a fast rate. When estimating the cost-effectiveness of an intervention for a heterogeneous population, one can condition on the observed characteristics and separate the overall group into homogenous subpopulations within which patients have identical observed characteristics. Adoption decisions can then made for each of these mutually exclusive and identifiable patient groups [12].

**Probabilistic Sensitivity Analysis**

In this section, we consider the role of PSA in dealing with uncertainty in decision models. One- and multiway sensitivity analyses cannot reflect the combined uncertainty in all model parameters, and so are inappropriate for informing decision uncertainty. PSA, when conducted properly, provides a more rigorous approach by requiring that all input parameters in a model be specified as full probability distributions, rather than as point estimates, to indicate the uncertainty of the estimates [13]. PSA can be used accurately to estimate expected NB in a nonlinear model, and also to reveal the effect of the combined uncertainty in all model parameters.

Decisions based on expectation are only appropriate if the consequences of the uncertainty surrounding the decision are also considered. This informs the necessary question about whether current evidence is sufficient or whether further research is needed. Formal methods are available to estimate this value of information [2] and these are now recommended, although not required, by NICE. Nevertheless, for the purposes of this review it is sufficient to note that an appropriate characterization of decision uncertainty is a prerequisite for any assessment of the consequences of decision uncertainty, whether or not this achieved using formal
methods. Clearly, decisions based on point estimates without any consideration of uncertainty will lead to the adoption of technologies with inadequate and poor quality evidence [3,10]. Therefore, adoption decisions cannot be separated from an assessment of whether the evidence is sufficient to support such decisions. For example, as well as making the adoption decision, NICE also makes recommendations for future research, specifies a review date for guidance, in part based on when new evidence is expected to be available, and has issued guidance conditional on further evidence being collected to inform future decisions.

Implementing PSA

PSA is commonly conducted using Monte Carlo simulation. The model is run repeatedly, and each run uses new random draws from distributions describing the uncertainty surrounding the value of each of the model parameters. This propagates parameter uncertainty through the model, which is then reflected in the results, and can be used to describe the likelihood that a treatment decision is optimal. This is distinct from the use of Monte Carlo simulation in PLS where the model parameters are fixed and a random number is used to determine the path of each individual patient. This propagates variability, and sometimes heterogeneity, into the model results. Enough simulations must be run to ensure that this variation does not affect identification of the optimal treatment decision. Thus, executing PSA within PLS requires a two-level simulation where each set of probabilistic inputs is held constant and the required number of patients is simulated through the model [14]. This can make PSA an order of 1,000 or 10,000 more computationally expensive in a PLS structure as compared with a cohort structure, and it is sometimes for this reason that PSA is omitted from models employing PLS.

Alternative methods to conduct PSA exist in the form of analytical model solutions and emulators. A closed form solution of expected NB, and possibly the associated uncertainty, may be tractable. That is, the analyst could define a closed-form approximation or simplification that gives the expectation of nonlinear functions using the model parameters. One such example would be the use of a Taylor series expansion [15]. Emulators take the form of nonparametric statistical models of the outputs of a model, such that those outputs can be recalculated with minimal time and computational expense when varying the model inputs according to the associated uncertainty [14]. Nevertheless, there are some limitations associated with the use of emulators; for example, in the number of uncertain parameters that may be included. It is not yet known whether they provide a directly exchangeable alternative to Monte Carlo simulation. Therefore, the implementation of PSA through Monte Carlo simulation can be viewed as current practice.

Results of the Review

The review identified six assessment reports submitted to NICE where the estimates of expected costs and outcomes were based on a model structure using PLS. Details of the included appraisals are given in Table 1.

We identified four reasons for choosing a PLS structure over a cohort framework. These were treatment switching, sampling from patient characteristics, dependence on patient history, including previous events and time-in-state, and uncertainty and variability. The following review section provides a more detailed description of these reasons in the context of their use in the case studies, and assesses their implications for model structure.

Treatment Switching

In some disease areas, patients will be treated with a sequence of interventions. These may involve different drugs or different dosages of the same drugs. The decision regarding whether to move a patient to the next treatment in a sequence may be based on patient characteristics or patient history and therefore subject to variability. Nevertheless, for a given set of eligibility criteria for treatment switching, there will not be uncertainty around whether a patient proceeds to the next treatment. In other words, patients may vary in their characteristics and history, and as a result there will be variation in the number of patients switching treatment, but for a given set of characteristics and history there will not be random variation in the number of patients switching.

Case study 1 assessed the cost-effectiveness of imatinib for gastrointestinal stromal tumors (GIST) [16]. Current guidelines at the time of the assessment recommended an initial dose of 400 mg daily, with the option of proceeding to a higher dose in the event of a poor response or disease progression, and withdrawal of treatment in the absence of benefit after 8 weeks. Nevertheless, because of a paucity of data, the best starting dose of imatinib and best treatment pattern were highly uncertain.

The model had four health states: progressive disease, treatment with 400 mg imatinib, treatment with 600 mg imatinib and death. Patients in the imatinib
treatment group began with 400 mg daily: patients whose disease progressed could move to 600 mg daily, or move to the progressive disease state. Patients who failed treatment with 600 mg imatinib daily moved to the progressive disease state, from which the only transitions were to remain in state or die. Patients could die at any stage in the model. Patients in the control group (i.e., no imatinib) began the model in the progressive disease state, and could remain in state or die. The cycle length was 4 weeks, and the time horizon was 10 years.

For those patients who failed to respond to 400 mg imatinib, a random number was generated to determine whether they would be moved to 600 mg, or straight to the progressive disease state. The probability of receiving 600 mg was based on the number of patients who had responded after crossing over from 400 mg to 600 mg imatinib in a clinical trial [17].

In reality, the decision to move from a dose of 400 mg to 600 mg would not be based on random chance. Although there may be uncertainty about the number of patients who would respond to 600 mg after progressing on 400 mg imatinib, this could not be identified before the treatment decision being taken. The treatment strategy in practice would involve moving all eligible patients onto 600 mg. When response is

### Table 1  Examples of the use of patient-level simulation in NICE assessment reports

<table>
<thead>
<tr>
<th>Case study</th>
<th>Title</th>
<th>Gave justification for choice of PLS?</th>
<th>Estimate decision uncertainty?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Imatinib for the treatment of patients with unresectable and/or metastatic gastrointestinal stromal tumors—a systematic review and economic evaluation [18]</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>The clinical effectiveness and cost effectiveness of prevention and treatment of osteoporosis [20]</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>The effectiveness of infliximab and etanercept for the treatment of rheumatoid arthritis: a systematic review and economic evaluation [22]</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>The clinical and cost-effectiveness of anakinra for the treatment of rheumatoid arthritis in adults [23]</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>The clinical effectiveness and cost-effectiveness of newer drugs for children with epilepsy [24]</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>6</td>
<td>Coronary artery stents: rapid systematic review &amp; economic evaluation [29]</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

NICE, National Institute for Clinical Excellence; PLS, patient level simulation.

**Figure 1** Search results. *Two cost-minimization analyses, two considered the cost results alongside trial results; †One meta-model of industry submission, one model type unclear. PSA, probabilistic sensitivity analysis.
assessed after 4 weeks, those patients not responding, as observed in the clinical trial, would then move to the progressive state. This could be compared with a strategy of not moving patients to a higher dose of the drug after failure on 400 mg. This alternative approach was taken by the developers of the GIST model in a subsequent evaluation. This approach enables decisions to be made about the best treatment pattern, rather than including current variability in treatment patterns in the model results. The situation where variability in patient characteristics may introduce uncertainty into the number of patients switching treatment is discussed in the later section on patient histories.

**Sampling Baseline Patient Characteristics**

Many characteristics of interest will vary between patients with the same disease, and some will be observable at the time at which the treatment decision is made. In other words, they are observable, able to be measured with precision, at baseline. If these characteristics determine the likelihood of future events, parameters in the model will depend on that observation. It is important to separate the heterogeneous population into homogenous subgroups, and to model these groups separately; otherwise the results of the model relate to the average patient, that is, the mean of the distributions describing the variation of the characteristics of interest. In practice, the treatment decision can and should be made conditional on the observed characteristic of each patient, not the expected characteristic.

Case study 2 considers treatments for osteoporosis for primary and secondary prevention of fractures [18]. The model is an update of a previously constructed model by the same assessment group [19]. This, in effect, relaxes the time constraint normally imposed on the development of such models for NICE. The structure of the model is described as being similar to a Markov model with a cycle length of 1 year, the difference being that patients are entered into the model individually, and their history is tracked. The states in the secondary prevention model are hip fracture, wrist fracture, vertebral fracture, proximal humerus fracture, breast cancer, coronary heart disease, and death. The model also tracks the residential status of each patient to assign costs. The authors state the belief that reflecting the increased risk of recurrent fractures after an initial fracture, and tracking the residential status of patients in the model, would be difficult in a cohort model.

The secondary prevention model considers the cost-effectiveness of treatment strategies for women presenting at baseline with hip, vertebral, wrist or proximal humerus fractures. Thus, the PLS is employed, in part, to track the presenting fracture site for each woman, because the model assumes a different baseline risk of subsequent fractures for each initial fracture site. Because of the difference in baseline risk of future events for each initial fracture site, and the fact that this characteristic is known when the treatment decision is taken, these could have been assessed as separate subgroups; it is important to be able to make separate treatment recommendations for each group. This would considerably reduce the number of states required to represent that portion of the model in a cohort framework. This alternative represents a different characterization of the decision problem which justifies a simplification of the model structure. This conditioning on baseline characteristics is distinct from conditioning on events that occur throughout the model process, for example, the location of new fracture sites after treatment has been initiated. In this example, the PLS model structure was also used to record patient history in the model, and this issue is the focus of the next section on patient histories.

**Dependence on Patient Histories**

Observable variation within groups homogenous in baseline characteristics may arise as a result of subsequent events that occur within the model structure. To condition model parameters on this observed variation, it is necessary to record these events in some way. The method with which to do this will depend on the choice of model structure.

Case studies 3 and 4 examined treatments for rheumatoid arthritis [20,21]. There is a low likelihood of long-term use for any one disease-modifying antirheumatic drug (DMARD), because they are not always effective, lose effectiveness over time, or cause adverse effects. Case study 5 assessed the cost-effectiveness of “newer” antiepileptic drugs (AEDs) in children [22]. Lack of effect on seizure rate and intolerable side effects means that many patients with epilepsy are treated with a sequence of drugs. The models compared fixed treatment sequences. The discontinuation rate of each treatment was modeled as a Weibull distribution with a shape parameter not equal to one (i.e., the hazard rate was not constant). In other words, the probability of discontinuing each treatment was dependent on the time spent on that treatment. Also, the availability of future treatment options was affected by toxic reactions to previous drugs. The PLS structure allows the analyst to record the realized time spent receiving each drug for each individual patient.

A separate assessment examined the use of newer AEDs in adults [23]. The decision problem and the events to be reflected in the model were very similar to the model of AEDs in children. As with case studies 3, 4, and 5, time to treatment discontinuation was a function of time spent on the drug, but this was facilitated in a cohort model by employing a semi-Markov framework. This semi-Markov model was built in the statistical programming language R [24], which can manipulate n-dimensional arrays and track the time.
spent in each state. This alternative model structure enabled PSA to be undertaken to provide an estimate of decision uncertainty, without sacrificing the time-dependent structure of the model.

**Probabilistic Sensitivity Analysis with PLS**

As can be seen in Table 1, case study 2 formally assessed decision uncertainty [18]. This was made feasible through the use of an emulator employing Gaussian processes [14]. The full model was run 80 times using different values for the inputs to estimate a non-parametric relationship between the input parameters and the outputs of the model. This “model of a model” could then be analyzed relatively quickly to produce an estimate of decision uncertainty. The use of 80 runs of the model, as compared with perhaps 10,000 runs for PSA, considerably reduces the computing power and time required to estimate decision uncertainty within a PLS model structure.

**Discussion**

If it is accepted that adoption decisions should be made with consideration of the associated decision uncertainty, then we may say that models submitted to decision-makers have a dual requirement to estimate expected NB and characterize decision uncertainty. The use of PLS is often justified with reference to the first of these. In other words, the claim is made that it would not be possible to estimate NB accurately using a cohort framework for that particular decision problem. In our review, we have identified four reasons for choosing PLS, and showed that none of these necessarily preclude the use of a cohort framework. We have also identified an example where a formal estimate of decision uncertainty was obtained alongside a computationally expensive PLS model structure, showing that the two are not mutually exclusive.

This review indicates that the most common justifications for choosing a PLS are the need to incorporate time and history dependence in transition probabilities. In a Markov model, these would be handled using tunnel states, and if the number of states required is very large, the Markov framework may become unwieldy and inefficient. An alternative was illustrated whereby time- and state-dependent transitions were represented in a cohort framework using semi-Markov processes [25,26]. To employ this method to track elements of patient history within a cohort framework, it may be necessary to build the model in appropriate software. The use of alternative software can provide flexibility to design an alternative model structure, and may also reduce analysis time. This is no panacea, because there will be limits to the gain in analysis time available with alternative software, and improved hardware. For example, in the updated rheumatoid arthritis case study [21], the use of Borland Delphi [27] instead of TreeAge DATA 3.5 [28] sped up the analysis. Nevertheless, in this particular case the gain in analysis time was not enough to facilitate PSA. Dissemination and training are also required to allow further use of alternative software, and this may represent an additional constraint.

Time- and state-dependent transitions are easily handled within PLS. If the use of Monte Carlo simulation to conduct PSA is too computationally expensive, the requirement to characterize decision uncertainty could still be met, as evidenced by the use of an emulator in case study [18]. Nevertheless, the use of emulators is still in development, and there are currently some limitations, for example, in the number of uncertain input parameters that can be included. More research is necessary to validate such models before it is known whether they are exchangeable with conducting PSA by means of Monte Carlo simulation.

We have shown that there is a nonlinear relationship between a characteristic that varies between patients and the model output, it is necessary to account for this variability. This is distinct from the issue of baseline patient characteristics that confer a different baseline risk of subsequent events (heterogeneity), but refers to variability within homogenous patient subgroups. Importantly, this does not counter the need to address decision uncertainty because any nonlinear model which requires an assessment of variability will also need to assess uncertainty to estimate expected NB. A PLS can address the issue of variability, but this can also be addressed within a cohort framework by employing a two-level simulation. A third way to address this issue would be to find a closed-form approximation to the model which simplifies the analysis greatly. The alternatives identified here should allow the same level of detail in characterizing the decision problem as is possible with a PLS framework, but with lower computational expense. Nevertheless, they do require specialized knowledge to be able to correctly model that decision problem.

So how generalizable are the results of the review detailed here? An increasing number of models in the literature have been developed for—or to influence—a specific decision-making body at a particular point in time. Such models are typically developed with similar constraints as the models for NICE described here. Some models are developed over a longer-time period, and perhaps with more generous funding. Usually such analyses are not concerned with the decision problem of a particular organization, but have a wider set of aims and objectives. Nevertheless, even models developed in this way still face binding resource and time constraints, not least from the research funder. Furthermore, the purpose of all decision models is ultimately to inform real decisions at some point in time. Therefore, all analysts will simultaneously have to tackle the tasks of quantifying decision uncertainty
while reflecting the complexity of the disease process and effect of the intervention under time and resource constraints. Hence, the general issues highlighted by the NICE case studies are relevant to decision models developed in different contexts.

In conclusion, if the dual requirement of models to estimate NB and to characterize decision uncertainty is accepted, then the failure to fulfill the latter requirement will limit its value for decision-making. The claim that a model has been structured to provide a more appropriate estimate of expected NB but is too complex for PSA given the analysts’ constraints would leave the decision-maker without a key element of information. There are often alternative modeling approaches to handle particular characteristics of a decision problem that can be used to reduce the complexity of the model, or to facilitate the conduct of PSA within a complex model structure, and both of these do not reduce the ability of the analyst to estimate NB correctly. This makes it possible to produce probabilistic models to estimate expected NB and characterize decision uncertainty within the constraints of the decision-making process.

Source of financial support: The work was funded by the UK Medical Research Council as part of a Health Services Research Collaboration program entitled “Methodological issues relating to decision analysis for resource allocation in health care.” Mark Sculpher also receives funding via a Career Award in Public Health funded by the UK NHS Research and Development Programme. An earlier version of this article was presented at the UK Health Economists’ Study Group in Oxford in January 2005. The authors are grateful for colleagues’ comments and suggestions. All views and any errors are the sole responsibility of the authors.

References

21 Clark W, Jobanputra P, Barton P, Burls A. The clinical and cost-effectiveness of anakinra for the treat-


