Claxton, K. and Sculpher, M. and McCabe, C. and Briggs, A.H. and Akehurst, R. and Buxton, M. and Brazier, J. and O'Hagan, T. (2005) Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. *Health Economics* 14(4):pp. 339-347.

http://eprints.gla.ac.uk/4164/

Deposited on: 12 May 2008

Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra

Karl Claxton[a,*], Mark Sculpher[b], Chris McCabe[c], Andrew Briggs[d], Ron Akehurst[c], Martin Buxton[e], John Brazier[c] and Tony O'Hagan[f]

[a] *Centre for Health Economics, Department of Economics and Related Studies, University of York, UK*
[b] *Centre for Health Economics, University of York, UK*
[c] *Sheffield Health Economics Group, School of Health and Related Research, University of Sheffield, UK*
[d] *Health Economics Research Centre, Department of Public Health, University of Oxford, UK*
[e] *Health Economics Research Group, Brunel University, UK*
[f] *Department of Probability and Statistics, University of Sheffield, UK*

Summary

Recently the National Institute for Clinical Excellence (NICE) updated its methods guidance for technology assessment. One aspect of the new guidance is to require the use of probabilistic sensitivity analysis with all cost-effectiveness models submitted to the Institute. The purpose of this paper is to place the NICE guidance on dealing with uncertainty into a broader context of the requirements for decision making; to explain the general approach that was taken in its development; and to address each of the issues which have been raised in the debate about the role of probabilistic sensitivity analysis in general. The most appropriate starting point for developing guidance is to establish what is required for decision making. On the basis of these requirements, the methods and framework of analysis which can best meet these needs can then be identified. It will be argued that the guidance on dealing with uncertainty and, in particular, the requirement for probabilistic sensitivity analysis, is justified by the requirements of the type of decisions that NICE is asked to make. Given this foundation, the main issues and criticisms raised during and after the consultation process are reviewed. Finally, some of the methodological challenges posed by the need fully to characterise decision uncertainty and to inform the research agenda will be identified and discussed. Copyright

Introduction

Decisions to adopt, reimburse or issue guidance on the use of health technologies are increasingly being informed by an explicit cost-effectiveness analysis of the alternative interventions [1]. This requires an analytic framework which can repre-sent these decision problems explicitly, combine evidence from a range of sources and facilitate the extrapolation of costs and effects over time and between patient groups and clinical settings [2]. Decision analytic modelling provides such a framework and has become central to the assessment of health technologies by the National Institute of Clinical Excellence (NICE). The

*Correspondence to: Centre for Health Economics, University of York, Heslington, York YO10 5DD, UK.
E-mail: kpc1@york.ac.uk

importance of decision analytic modelling for informing decisions about the use of health technologies is reflected in the recently updated NICE guide on the methods of technology appraisal [3]. The updated guidance details what the Institute considers to be appropriate methods for estimating the cost-effectiveness of technologies, and for characterising the uncertainty surrounding these estimates. One feature of the guidance is the development of 'reference case' requirements for analysis. As with the recommendations of the Washington Panel in the 1990s [4], the reference case seeks to increase consistency in analysis and decision making by defining a set of methods which should form the basis of any cost-effectiveness analysis submitted to NICE.

A number of aspects of the NICE guidance have been controversial [5,6]. The need for probabilistic sensitivity analysis (PSA) to characterise uncertainty is one such element. The guidance requires that all estimates of input parameters in a model be specified as full probability distributions, rather than point estimates, to represent the uncertainty surrounding their values [7–9]. The distributions employed to describe the uncertainty in individual parameters are not arbitrary. Rather, the choice of distribution should be guided by the form of the data, the type of parameter and the estimation process. In most cases this would point to only one or two contending distributions. Most commonly Monte Carlo simulation is then used to propagate this *parameter uncertainty* through the model so that the imprecision of the cost-effectiveness results, and hence the *decision uncertainty*, can be represented using methods such as cost-effectiveness acceptability curves [10–12].

The requirement, as part of the reference case, for the use of PSA to characterise uncertainty has led to a range of comments and criticisms from, broadly, two constituencies: those who are more generally critical of the focus on decision analytic modelling and decisions based on cost-effectiveness; and those involved in decision modelling and economic evaluation. The latter includes those directly involved in technology assessment for NICE, and others (inside and outside the UK) who see that the NICE guidance will have an impact on what is required for evaluation of heath technologies internationally. This is an important and timely and international debate which, although focused here on the recent NICE guidance, raises issues which are of much more general interest, and which have the potential to affect the future direction and pace of methodological development in the evaluation of health technologies.

The purpose of this paper is twofold. Firstly, to place the NICE methods guidance in general, and the reference case in particular, in a broader context and to explain the general approach which was taken in its development; and, secondly, to address the main issues which have been raised in the debate about the role of PSA in particular. Each of the authors has played a role in developing the new NICE guidance. The paper is intended to be an exploration of the issues around PSA, and a summary of the views of those closely involved in the production of the guidance.

Background to the guidance

The methods guidance was developed through four task groups, each dealing with different aspects of assessment and appraisal. The Economics Task Group developed the guidance which, among other things, included the specification of the reference case requirements for economic analysis, the role of decision analytic modelling and the requirement for PSA. The Economics Task Group's draft report was then incorporated into the full guide by the Methods Working Party, which included the chairs of all the task groups. It was then edited and amended in the light of comments received during a broad consultation process. The final *Guide to the Methods of Technology Appraisal* represents the outcome of this process [3]. Within the Economics Task Group and the Methods Working Party, there was a general consensus around the development of the reference case in general, the importance of decision modelling and the inclusion of PSA.

General approach

It is important to explain the approach taken by the Economics Task Group in developing its input to the document. The most appropriate starting point for developing methods guidance is to establish what is required of any analysis seeking to inform decisions which are consistent with the objectives of a health care system subject to its budget constraint. On the basis of these requirements, the next step is to identify an appropriate

framework and methods of analysis which can best meet these needs.

This means that the guidance is not prescriptive about particular methods; rather it sets out what is required of an analysis. However, the requirement for the analysis to be predicated on the needs of the decision making process was prescriptive [3]. The guidance did not start from the point of current capacities, current practice or consensus. In fact, it was part of the Economics Task Group's brief to encourage the use of appropriate methods rather than simply to reflect current practice in the field. Issues associated with the current capacity and resourcing of the academic groups undertaking technology assessment for NICE, and the implications for submissions from stakeholders, was not regarded as a constraint on the discussions and recommendations of the task group. Rather, these issues would be dealt with elsewhere in negotiation between the Department of Health, NICE, the technology assessment groups and sponsors. It was also understood that developing methods guidance would be a continual process with periodic updates leading to further revisions at some point in the future, reflecting the on-going methodological development in the field. It was also suggested by the task group that more detailed technical guides on key aspects of the guidance, such as evidence synthesis, decision analytic modelling and PSA, should be developed to provide further support and methodological guidance.

The reference case

This general approach to the development of the guidance underpinned our adoption of the reference case analysis. The justification for a reference case is the same as that of the Washington Panel which was set up by the US Public Health Service to advise on 'good practice' in the field [4]. The Washington Panel's reasoning was that if economic evaluation is to inform resource allocation decisions in health care, then a common and agreed analytic framework and set of methods is required which facilitates comparability. However, over prescription and the imposition of stifling orthodoxy in a fast developing field needs to be avoided. The primary purpose of the NICE reference case is, therefore, to ensure that the requirements for decision making within the NICE

remit are met as far as possible by specifying those methods which meet these needs.

A second purpose of the reference case is to provide some degree of comparability *within* the appraisal of a particular technology, between the submissions from stakeholders and the independent technology assessment report. The lack of comparability between submissions within an appraisal and the wide range of estimates of cost-effectiveness presented, without clear explanation of these differences, has been a constant problem for the Appraisal Committee and the assessment teams [13]. As well as comparability within an appraisal, the reference case was also intended to provide comparability *between* appraisals of different technologies and over time, thus contributing to consistency in decision making.

However, it is not the purpose of the reference case to limit the methods that can be used as part of a technology assessment or stakeholder submission. It is explicitly recognised that additional non-reference case analyses may be appropriate. Indeed, such analyses are positively encouraged as long as they are clearly justified, and any differences in cost-effectiveness results fully explained. The methods guidance makes clear that, in its considerations, the Appraisal Committee will not necessarily regard the reference case as providing the 'best' analysis, for a particular technology, but will consider the justifications for any non-reference case analysis in coming to a view about the most reliable estimates of cost-effectiveness. There is also a recognition in the guidance that in some circumstances a complete reference case analysis may not be possible for a variety of reasons. For example, in some instances, the data required to present reference case results may not be available or there may be other barriers (including computational challenges) to applying reference case methods. In these cases, the reasons for a failure to provide a full reference case analysis should be clearly explained, and the likely implications should, as far as possible, be quantified so that the Appraisal Committee can decide what weight it will attach to the results.

So the general approach to developing the guidance and specifying the reference case was to balance the need for comparability within and between appraisals with the desire to avoid being overly prescriptive about methods. It was anticipated that, in some appraisals, a well-justified

non-reference case analysis will be regarded as providing the best estimates of cost-effectiveness. Therefore, it is important to see the guidance in general, and on PSA in particular, in the context of this overall approach, the purposes of the reference case and the consideration of a justified non-reference case analysis.

Dealing with uncertainty

The guidance states that 'it is important for the Appraisal Committee to know about the uncertainty associated with clinical and cost effectiveness information' (p. 20) [3]. Nevertheless, there are strong arguments for basing decisions about resource allocation on *expected* cost-effectiveness rather than the traditional and arbitrary rules of inference [14]. This had led a number of commentators to suggest that this makes the consideration of uncertainty and the use of PSA redundant. However, decision making based on expected utility (in this context expected net benefit) in no way implies that decision uncertainty is unimportant. Indeed, an assessment of the implications of decision uncertainty is an essential part of a decision making process that is consistent with objectives and constraints of any health care system [14].

An honest and transparent characterisation of the uncertainty is needed for a number of reasons. Firstly, NICE does make recommendations about further research, and can issue guidance which is conditional on additional evidence being provided [15], on the conduct of pilot studies before wider adoption [16] or on a technology only being used within a clinical trial [17]. Secondly the date when the guidance will be reconsidered must be specified which may well be informed by the uncertainty surrounding the decision and the anticipation of further evidence being available in the future. Finally, the Appraisal Committee requires some means of assessing the costs associated with a possible change in the decision about a technology in the future (e.g. about the consequences of irreversibility) and about the impact on ongoing research of issuing guidance.

In principle, many of these issues could be formally addressed using value of information analysis [14–18] and real options pricing [19]. However, these formal approaches were not specified as part of the reference case although

value of information analysis was recommended. Nevertheless, some assessment of whether existing evidence is sufficient to support the use of a technology, the appropriate length of time until the reconsideration of the guidance and the implications of irreversibility when costs and effects of a decision are uncertain must be made as part of any decision making process. The issue is one of who should be responsible for such an assessment. The opinion of the task group was that, at the present time, these assessments should be made by the Appraisal Committee supported by NICE, and not necessarily by the analysts making submissions to the Institute. Since these assessments can only be made on the basis of a clear and transparent characterisation of decision uncertainty, the focus of the guidance was on the use of PSA to provide this characterisation.

Finally, it should also be recognised that, for decision models in which there is a non-multi-linear relationship between inputs and outputs (e.g. Markov models), the correct calculation of *expected* costs and effects will need the full uncertainty around parameters to be expressed. Therefore, probabilistic analysis of the model also ensures adequate estimates of *expected* cost and effects.

Probabilistic sensitivity analysis

The following discussion focuses on those comments and arguments regarding PSA that have been made during the consultation process, and which have been raised elsewhere in discussion about the implementation of the NICE methods guidance.

The necessity of PSA

If the characterisation of decision uncertainty is required for decision making, then the question is whether PSA is the most appropriate framework to achieve this. For example, it has been suggested that in many cases a simple series of one-way sensitivity analyses may be sufficient [20]. However, simple sensitivity analysis cannot provide enough insight into the scale of decision uncertainty. There is a strong incentive for manufacturers, when submitting cost-effectiveness models to NICE, to claim that their sensitivity analyses show their conclusions are robust to parameter

uncertainty. In many models, the uncertainty in the individual parameters may be very unlikely to change a decision (i.e. the value of additional information for individual parameters can be zero). In combination with the other parameters, however, a complete picture of the parameter uncertainty in the model may generate considerable decision uncertainty and the value of additional information may be very high.

The way to understand the implications for decision uncertainty of imprecisely estimated parameters is to include all of those parameters subject to uncertainty in the sensitivity analysis, and to use the full distributions of those parameters based on all available evidence. This is not possible with simple one-way sensitivity analysis. Although multi-way sensitivity analysis is sometimes used to explore combined uncertainty, with a large number of parameters this can be markedly more time- and computer-intensive than PSA. More importantly, it is generally very difficult to interpret correctly and becomes impossible if some parameters are correlated. For example, the statistical estimation of parameters, often requiring methods of evidence synthesis [21], generates complex correlation structures between parameters which makes it impossible to locate a fixed (set of) parameter value(s) which can be regarded as 'extreme'. The Economics Task Group was clear that one-way sensitivity analysis will underestimate decision uncertainty, and that a well conducted PSA will engender a more realistic representation of uncertainty in a model's results.

It was recognised that, in making an assessment of the implications of decision uncertainty when issuing guidance, the Appraisal Committee will also need an understanding of the contribution of specific parameters (or combinations of parameters). This contribution to decision uncertainty can be assessed by using value of information methods for individual and groups of parameters [22]. Indeed, these methods are recommended in the guidance but were not made part of the reference case for the reasons given above.

Additional assumptions of PSA

Some commentators have criticised PSA on the grounds that it introduces further assumptions into the decision model. In particular, the choice of distribution to represent uncertainty and the common assumption of independence between parameters have been identified as limitations of the probabilistic approach [23].

A large number of potential distributions are available within commonly used modelling packages, and this may be seen to reinforce the apparently arbitrary choice of distribution. However, if the principle of using the standard statistical approach to the estimation of a parameter is followed, such that the distribution employed reflects the statistical uncertainty in that parameter's estimation, then typically only one or two distributional forms are candidates. For example, probability parameters are bounded on the interval zero-one, so it would be inappropriate to specify a distribution that gave a non-negligible probability to obtaining a parameter value outside of that range. With this in mind, the appropriate choice of distribution will be closely related to the nature of the parameter, the form of the data and the method of estimation. Where a probability parameter is estimated from a proportion, the beta distribution (which is bounded zero-one) is the natural choice of distribution. However, if the probability parameter is estimated from a logistic regression, then the parameters of interest are the coefficients on the log-odds scale and multivariate normality on this scale would be the appropriate assumption. For probabilities estimated from time-to-event data, the parameters would be the coefficients from a survival analysis estimated on the log hazard scale and, again, the appropriate assumption would be multivariate normality on this scale. Application of this general approach to using the distributional form that relates to the estimation of the parameter of interest is likely to make the depiction of uncertainty in PSA less arbitrary than one-way sensitivity analysis rather than more arbitrary.

While it is true that examples of PSA often assume independence between parameters, it is not a requirement of the approach. The use of multivariate normal distributions is common in regression analysis and illustrates a situation where the covariance between parameters can easily be estimated from the covariance matrix. Where regression analyses inform the estimation of parameters in a decision model, therefore, it is clearly appropriate that correlations between regression coefficients are included in that model and methods exist to allow the correlation of parameters in the multivariate normal case.

Computational burden of PSA

An important area of discussion regarding the role of PSA has been in the context of micro simulation. This type of model can be particularly useful for modelling diseases where patients' future prognoses are strongly time-dependent or are influenced by earlier events. In such models the progress of individual patients through the model is simulated in order to allow for a patient's case history in the model to impact the probability of, or the time until, future events. In order to characterise decision uncertainty, a second level of simulation is required and this would, for some models, require considerable computation time. This problem is accentuated by the tight time constraints which exist for NICE appraisals which, in particular, limit the technology assessment groups in developing their analyses.

The choice about model structure and complexity is always a trade-off between descriptive realism and tractability in terms of computational burden and data requirements. The key issue is whether the model is sufficiently realistic to inform the decision. Some research has been undertaken to compare micro-simulation and cohort models [24]. However, identifying the 'best model' for a given appraisal is difficult. More research is required to develop criteria for selecting an appropriate structure for a decision model given available evidence on disease prognosis and the impact of alternative interventions.

Until such research is undertaken, the selection of an appropriate model structure will have to be based on a careful balance between a reasonable simplification of reality and computational and data burden. It is likely that micro-simulation will play a continued role in some technology assessments where there is a need to incorporate evidence of the impact of patients' histories in a way that simply cannot be captured in more simple cohort models. Nevertheless, given the importance of decision uncertainty for the Appraisal Committee's deliberations, it seems reasonable for NICE to expect PSA to be undertaken. Some micro-simulation models will be limited in complexity and number of parameters, and PSA will not be computationally expensive. For more complex models, methods have recently been described which can avoid two-level simulation, cutting down on computation time [25]. Indeed, these methods have been used in a recent technology assessment for NICE [26]. In those situations where a mixture of time constraints and model complexity preclude the implementation of PSA, a non-reference case analysis might be justified on the basis that a micro-simulation was the best means of structuring the decision problem appropriately.

Presenting and interpreting probabilistic analysis

Several methods have been discussed and used to present the results of PSA, many of which have been developed in the context of stochastic analysis of patient-level data, usually from randomised trials [27]. In the updated NICE methods guidance, the use of cost-effectiveness acceptability curves (CEACs) is recommended. These present the probability that (proportion of simulations in which) a given intervention is more cost-effective than the alternatives for a range of maximum thresholds regarding NICE's willingness to pay for an additional QALY [10–12]. Some have argued that non-specialists on the Appraisal Committee find the CEAC difficult to interpret and that the presentation of the joint uncertainty in differential costs and effects on the cost-effectiveness plane [28] is more intuitive. However, NICE appraisals increasingly require the comparison of several interventions. As a means of comparing the cost-effectiveness of more than two competing options, the cost-effectiveness plane becomes impossible to interpret correctly. NICE is embarking on a series of training sessions for the Appraisal Committees, including a session on the methods of presenting the results of PSA.

It is true, however, that a standard CEAC does not contain a decision rule; that is, it will not in itself suggest the 'best intervention'. This is because the intervention with the highest probability of being cost-effective is not always the one with the highest expected (i.e. mean) cost-effectiveness. The reason for this is that the distributions of costs and effects are often asymmetrical and is analogous to the difference between the median and a mean of a standard frequency distribution based on patient-level data. The concept of the cost-effectiveness frontier has been presented as a way of overcoming this problem [12], where the intervention on the frontier is always the one with the highest expected cost-effectiveness.

Hence the Economics Task Group took the view that the CEAC is the best way of communicating decision uncertainty, particularly when more than two options are being compared. Other ways of presenting the uncertainty in an analysis are not precluded, however. As described above, a key advantage of PSA is that, by quantifying decision uncertainty, it can help prioritise future research, and the CEAC together with an appropriate frontier has a direct association with the expected value of perfect information [12]. This is particularly important given that NICE has recently appointed a Research and Development Advisory Committee and that it is thinking more carefully about the appropriate dates on which to consider the revision of guidance about particular technologies, and about research recommendations.

Handling other sources of uncertainty and sub-group variability

It is explicitly recognised in NICE's updated methods guidance [3] that the uncertainty surrounding parameter estimates is not the only source of uncertainty in a cost-effectiveness model. As noted above, an important source of uncertainty relates to the assumptions regarding the structure of the model. In addition, the appropriate interpretation of evidence, given its variable quality and likely heterogeneity, is another source of uncertainty. In principle, these types of uncertainty can be handled as part of the PSA by ascribing probabilities to alternative assumptions about model structure to reflect the analysts' or the Appraisal Committee's views about their appropriateness. An alternative approach, and the one favoured in the methods guidance, is to conduct scenario analysis where the probabilistic analysis is run several times, each scenario conditional on different assumptions about model structure or interpretations of the available evidence. The Appraisal Committee is then responsible for making an assessment of which of the scenarios is the most credible.

The issue of how to deal with possible variation in cost-effectiveness between patient sub-groups was also considered in the updated guidance. This is an important issue given NICE's objective to maximise health gain from available resources. If available evidence suggests that there is variability between sub-groups in one or more parameters, which could lead to important variation in cost-effectiveness, and where sub-groups can be adequately identified in routine clinical practice before management decisions are taken, then this should be dealt with using sub-group analysis. That is, separate estimates of cost-effectiveness should be presented for each sub-group, in a manner analogous to scenario analysis for structural uncertainty and interpretation of evidence. However, the use of sub-group analysis emphasises why PSA is so important for decision making. This is because one implication of 'dividing up' the available evidence into specific sub-groups is that the samples become smaller and, all things being equal, the sub-group-specific estimates of parameters are less precise than the group-level estimates. Hence, although expected cost-effectiveness may show a degree of variation between sub-groups that is considered relevant for decision making, the implications for decision uncertainty should not be ignored, particularly in terms of recommendations for future research.

Conclusions

The purpose of this paper is to explain the rationale for making PSA part of the reference case for NICE appraisal. It is hoped that there will be greater acceptance of PSA given the explanation of how it was seen to fit with the requirements for NICE decision making. There are some good examples of PSA in the literature to provide a guide to the technical implementation of the methods [9,13,29] [Palmer S, Sculpher M, Philips Z *et al*. Management of non-ST-elevation acute coronary syndromes: how cost-effective are glycoprotein IIb/IIIa antagonists in the UK National Health Service? *Int J Cardiol*, in press, Briggs A, Sculpher M, Dawson J, FitzPatrick R, Murray D, Malchau H. The use of probabilistic decision models in technology assessment: the case of hip replacement. *J Appl Health Econ Policy*, in press]. Increasingly these methods are being seen in manufacturers' submissions to NICE and in the models developed by the assessment teams. More detailed technical guides are likely to emerge from NICE to provide greater clarity about the use, and critical appraisal, of PSA. In due course, the methods guidance itself will be updated, and this is likely to reflect the continued development of

methods of cost-effectiveness. However, the need to quantify the implications of the imprecision in parameter estimates for decision uncertainty is likely to remain.

There are, of course, many challenges when attempting to estimate costs and effects across a range of possible interventions, over a relevant time horizon and for specific patient groups, while attempting fully to represent the uncertainty surrounding the decision. The issues of interpretation of evidence, synthesis, potential bias, exchangeability and appropriate model structure have always been present in any informal and partial review of evidence. In fact, until quite recently, these challenging issues could be conveniently ignored by both policy makers, clinicians and analysts while decision making was opaque and based on implicit criteria and unspecified 'weighing' of the evidence. These challenges must be faced as more explicit and transparent approaches to decision making are being taken. Indeed, one of the many advantages of taking a more transparent and explicit approach to decision making and the characterisation of uncertainty is that it exposes many important methodological issues which have previously been avoided by presenting partial analysis which do not directly address the decisions which must be made in any health care system.

Acknowledgements

References

1. Hjelmgren J, Berggren F, Andersson F. Health economic guidelines – similarities, differences and some implications. *Value Health* 2001; **4**(3): 225–250.
2. Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute for Clinical Excellence. *Lancet* 2002; **360**: 711–715.
3. National Institute for Clinical Excellence (NICE). *Guide to the Methods of Technology Appraisal*. NICE: London, 2004.
4. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-Effectiveness in Health and Medicine*. Oxford University Press: New York, 1996.
5. Langley PC. The NICE reference case requirement: implications for drug manufacturers and health systems. *Pharamacoeconomics* 2004; **22**: 267–271.
6. Gafni A, Birch S. The NICE reference case requirement: more pain for what, if any, gain? *Pharamacoeconomics* 2004; **22**: 271–273.
7. Critchfield GC, Willard KE, Connelly DP. Probabilistic analysis of decision trees using Monte Carlo simulation. *Med Decision Making* 1986; **6**: 85–92.
8. Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation. *Med Decision Making* 1985; **5**: 157–177.
9. Briggs AH, Goeree R, Blackhouse G, O'Brien BJ. Probabilistic analysis of cost-effectiveness models: choosing between treatment strategies for gastro-esophageal reflux disease. *Med Decision Making* 2002; **22**: 290–308.
10. Van Hout BA, Al MJ, Gordon GS, Rutten FFH. Costs, effects and c/e-ratios alongside a clinical trial. *Health Econ* 1994; **3**: 309–319.
11. Briggs AH. A Bayesian approach to stochastic cost-effectiveness analysis. *Health Econ* 1999; **8**: 257–262.
12. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Econ* 2001; **10**: 779–789.
13. Chilcott J, McCabe C, Tappenden P, O'Hagan A, Cooper NJ, Abrams K *et al.* Modelling the cost effectiveness of interferon beta and glatiramer acete in the management of multiple sclerosis. *Br Med J* 2003; **326**: 522.
14. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999; **18**: 342–364.
15. National Institute for Clinical Excellence. *Guidance on the Selection of Prostheses for Primary Total Hip Replacement* (www.nice.org.uk). National Institute for Clinical Excellence: London, 2000.
16. National Institute for Clinical Excellence. *Health Technology Assessment Report on Liquid Based*

Cytology (www.nice.org.uk). National Institute for Clinical Excellence: London, 2000.

17. National Institute of Clinical Excellence (NICE). *Guidance on the Use of Laparoscopic Surgery for Colorectal Cancer* (www.nice.org.uk). National Institute for Clinical Excellence: London, 2000.

18. Claxton K, Posnett J. An economic approach to clinical trial design and research priority-setting. *Health Econ* 1996; **5**: 513–524.

19. Palmer S, Smith PC. Incorporating option values into the economic evaluation of health care technologies. *J Health Econ* 2000; **19**: 755–766.

20. Meltzer D. Addressing uncertainty in medical cost-effectiveness analysis. Implications of expected utility maximization for methods to perform sensitivity analysis and the use of cost-effectiveness analysis to set priorities for medical research. *J Health Econ* 2001; **20**: 109–129.

21. Ades AE, Cliffe S. Markov Chain Monte Carlo estimation of multi-parameter decision model: consistency of evidence and the accurate assessment of uncertainty. *Med Decision Making* 2002; **22**: 359–371.

22. Ades AE, Lu G, Claxton K. Expected value of sample information in medical decision modelling. *Med Decision Making* 2003; **24**: 207–227.

23. O'Brien BJ, Drummond MF, Labelle RJ, Willan A. 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): 150–163.

24. Karnon J. Alternative decision modelling techniques for evaluation of health care technologies: Markov processes verses discrete event simulation. *Health Econ* 2003; **12**: 837–848.

25. Stevenson MD, Oakley J, Chilcott JB. Gaussian process modelling in conjunction with individual patient simulation modelling: a case study describing the calculation of cost-effectiveness ratios for the treatment of established osteoporosis. *Med Decision Making* 2004; **24**: 89–100.

26. Stevenson M, Jones LM, de Nigris E, Brewer N, Oakley J. *The Clinical Effectiveness and Cost Effectiveness of Prevention and Treatment of Osteoporosis* (www.nice.org.uk). National Institute for Clinical Excellence: London, 2003.

27. Briggs AH, O'Brien BJ, Blackhouse G. Thinking outside the box: recent advances in the analysis and presentation of uncertainty in cost-effectiveness studies. *Ann Rev Public Health* 2002; **23**: 377–401.

28. Black WC. A graphic representation of cost-effectiveness. *Med Decision Making* 1990; **10**: 212–214.

29. Claxton K, Neuman PJ, Araki SS, Weinstein MC. The value of information: an application to a policy model of Alzheimer's disease. *Int J Technol Assess Health Care* 2001; **17**: 38–55.