Use of evidence in economic decision models: Practical issues and methodological challenges

Cooper NJ[1], Ades AE[2], Paisley S[3], Sutton AJ[1],
Longson C[4], Mugford M[5], Sculpher M[6]
on behalf of the working group on the 'Use of Evidence in Economic Decision Models'

[1]Department of Health Sciences, University of Leicester, UK.
[2]MRC Health Services Research Collaboration, University of Bristol, UK.
[3]School of Health and Related Research, University of Sheffield, UK.
[4]National Institute for Health and Clinical Excellence, UK.
[5]School of Medicine, Health Policy and Practice, University of East Anglia, UK.
[6]Centre for Health Economics, University of York, UK.

Correspondence address:
Dr Nicola Cooper
22-28 Princess Road West
Leicester
LE1 6TP
E-mail: njc21@le.ac.uk
Tel: 0116 252 5437
Fax: 0116 252 3272

Financial support was provided by a Medical Research Council (MRC) Health Service Research Collaboration Workshop grant.

WORK IN PROGRESS

Please do not cite without permission

Members of the working group on the 'use of evidence in economic decision models'

Keith Abrams, *Department of Health Sciences, University of Leicester*

Tony Ades, *MRC Health Services Research Collaboration, University of Bristol*

Laura Bojke, *Centre for Health Economics, University of York*

Meindert Boysen, *National Institute for Health and Clinical Excellence*

Andrew Briggs, Section *of Public Health & Health Policy, University of Glasgow*

Deborah Caldwell, *MRC Health Services Research Collaboration, University of Bristol*

Zaid Chalabi, *Public and Environmental Health Research Unit, London School of Hygiene and Tropical Medicine*

Jim Chilcott, *School of Health and Related Research, University of Sheffield*

Nicola Cooper, *Department of Health Sciences, University of Leicester*

Guy Freeman, *Department of Public Health and Epidemiology, University of Birmingham*

Sarah Garner, *National Institute for Health and Clinical Excellence*

Alan Girling, *Department of Public Health and Epidemiology, University of Birmingham*

Jason Gordon, *Department of Public Health and Epidemiology, University of Birmingham*

Colin Green, *Department of Public Health and Epidemiology, University of Birmingham*

Neil Hawkins, *Centre for Health Economics, Univerisity of York*

John Hutton, *MEDTAP*

Cynthia Iglesias, *Centre for Health Economics, University of York*

David Jones, *Department of Health Sciences, University of Leicester*

Richard Lilford, *Department of Public Health and Epidemiology, University of Birmingham*

Carole Longson, *National Institute for Health and Clinical Excellence*

Louise Longworth, *National Institute for Health and Clinical Excellence*

Jo Lord, *National Institute for Health and Clinical Excellence*

Borislava Mihaylova, *Health Economics Research Centre, University of Oxford*

Paul Miller, *AstraZeneca*

Miranda Mugford, *School of Medicine, Health Policy and Practice, University of East Anglia*

Suzy Paisley, *School of Health and Related Research, University of Sheffield*

Zoe Philips, *School of Economics, University of Nottingham*

Mark Sculpher, *Centre for Health Economics, University of York*

Ian Shemilt, *School of Medicine, Health Policy and Practice, University of East Anglia*

Alex Sutton, *Department of Health Sciences, University of Leicester*

Paul Tappenden, *School of Health and Related Research, University of Sheffield*

Rod Taylor, *Department of Public Health and Epidemiology, University of Birmingham*

Simon Thompson, *MRC Biostatistics Unit, Cambridge*

David Turner, *Trent Research and Development Support Unit, University of Leicester*

Rebecca Turner, *MRC Biostatistics Unit, Cambridge*

Dave Whynes, *School of Economics, University of Nottingham*

Jane Wolstenholme, *Health Economics Research Centre, University of Oxford*

ABSTRACT

Aims: To highlight the practical issues and methodological challenges related to identifying, combining and reporting the evidence to inform decision model parameters and structure.

Methods: Two workshops consisting of 37 experts from a range of disciplines (i.e. decision-makers, health economists, information specialists, operations researchers, and statisticians) and funded by the Medical Research Council (MRC) Health Services Research Collaboration were held to establish the i) current 'state of play' regarding the use of evidence in decision modelling and ii) recent methodological developments that may aid decision-modellers in the future.

Results: Practical issues and methodological challenges identified included:

i) Definition and identification of 'relevant' evidence to inform decision model structure and inputs (e.g. Are systematic review techniques useful for identifying evidence to inform model parameters other than clinical effectiveness? Should evidence be limited to the 'most relevant' or include 'all relevant' and how should 'relevance' be defined?);
ii) Assessment of quality/reliability of different sources of evidence (e.g. How should we assess study quality? Are hierarchies of evidence useful?); and
iii) Synthesis of evidence (e.g. How can potential biases in, say, observational studies be identified and adjusted for? What is the added value of combining indirect evidence with direct evidence to go beyond pairwise treatment comparisons?).

There was also the additional issue of transparent reporting of evidence identification and selection methods to aid reproducibility and updating of decision models.

Conclusions: This paper highlights both practical issues and methodological challenges associated with the appropriate use of evidence in decision modelling. Where applicable, potential solutions are discussed and gaps in the established methodology identified for future research.

INTRODUCTION

To conduct an economic evaluation, evidence on costs and effects of healthcare interventions/screening are required. These may be derived from clinical trials, observational studies, administrative databases, case series, expert opinion and/or secondary analyses (such as meta-analysis) of any of the above. All the evidence required for an economic evaluation is rarely extractable from one data source(1); for this reason, mathematical (decision-analytical) models are often developed to synthesise the data to allow the cost-effectiveness of alternative clinical strategies/interventions to be assessed.

Policymakers, such as National Institute of Health and Clinical Excellence (NICE) in the UK, are increasingly being required to utilise economic decision models when making decisions regarding which health care interventions and programmes to fund from the limited resources available to health care systems. As with any model, the results obtained from decision models depend on the appropriateness of the model structure and the quality of the data inputs.

This paper outlines the current 'state of play' regarding the use of evidence in decision modelling and highlights both practical issues and methodological challenges in terms of identification, synthesis and transparent reporting. Where applicable, potential solutions are discussed and gaps in the established methodology identified for future research.

WORKSHOPS

Two workshops, funded by the MRC Health Services Research Collaboration, were organised consisting of 37 experts from a range of disciplines (i.e. decision-makers, health economists, information specialists, operations researchers, and statisticians). These workshops aimed to derive a multi-disciplinary standpoint on the appropriate use of evidence in economic decision models. The first workshop focused on establishing the current 'state of play' regarding the use of evidence in decision modelling and the second workshop focused on recent methodological developments in evidence synthesis that may aid decision-modellers in the future. Gaps in the established methodology knowledge base in this area were identified and form the basis of future research agendas. A discussion of the practical issues and methodological challenges identified in the two workshops are provided below.

CURRENT SITUATION

The recent methods guidelines for health technology assessments (HTA) from NICE(2) acknowledge the need *"…to construct an analytical framework within which to synthesise the available evidence in order to estimate clinical and cost effectiveness relevant to the clinical decision making context. The framework …may be a decision-analytic model using aggregated data or a statistical model using patient level data*". They recommend "*all relevant evidence must be identified*" and "*evidence must be identified, quality assessed and, where appropriate, pooled using explicit criteria and justifiable and reproducible methods*". The guidelines also state that "*explicit criteria by which studies are included and excluded*" should be presented. However, the guidelines lack specific procedural guidance and provide no clear definition of 'relevant evidence'.

A recent study(3) of 42 decision models developed as part of the Department of Health Research and Development HTA programme in the UK between 1997 and 2003 (22 (52%) of which also informed a NICE appraisal) highlighted that the evidence used in these decision models was extremely varied. The study showed parameter estimates (e.g. clinical effectiveness, resource use, utilities, natural history, etc.) for decision models to be obtained from diverse sources of evidence ranging from randomised controlled trials (RCTs) to expert opinion often with a lack of justification for the choice of evidence used. The use of data from diverse sources raises concerns about varying degrees of bias due to confounding, patient selection, and/or methods of analysis(4). In addition, there was inadequate reporting on how evidence was identified (e.g. literature databases searched) and a lack of quality assessment of the evidence used.

Guidelines for good practice in decision-analytic modelling were recently reviewed by Philips et al(5). The authors identified 15 publications that provided general guidance on the elements of a good model or explicit criteria against which to assess the quality of a model. The review concluded that although the different guidelines did provide a consistent message regarding some aspects of modelling (e.g. the need to be explicit in presenting an analysis), in other areas conflicting guidance was presented (e.g. the extent to which the model structure, and data inputs, should be determined by data availability). The report also highlights some areas of relevance to health technology assessment that are not currently addressed in any of the published guidelines; for example, established methodology for the identification of parameter estimates from the literature, and

adjustments for potential biases related to observational studies or registries. Unlike other areas of health technology assessment (e.g. meta-analysis – QUOROM statement(6)) there is currently no recommended checklist to assess the quality of published decision models, although the authors of the aforementioned review did compile synthesised guidelines and a checklist in an attempt to provide a framework for critical appraisal of decision models by all parties involved in the HTA process.

PRACTICAL ISSUES AND METHODOLOGICAL CHALLENGES

The practical issues and methodological challenges related to the use of evidence in economic decision models discussed within the workshops can be classified into the following three categorises and are discussed in more detail below:

i) Definition and identification of 'relevant' evidence to inform decision model structure and inputs
ii) Assessment of quality/reliability of different sources of evidence
iii) Synthesis of evidence

i) Definition and identification of 'relevant' evidence to inform decision model structure and inputs

Relevant evidence

In the NICE guide to the methods for technology assessment(2) it is stated that "*all relevant evidence must be identified*". Under this requirement, all evidence from trials to expert opinion and beyond should be identified and then selected or rejected based on some pre-specified criteria (e.g. study population, outcomes and interventions) to inform the input parameters of a decision model. However, how is the relevance of different sources of evidence to be decided and should it be classified as a dichotomy (i.e. relevant or not relevant) or as a continuum (i.e. most relevant to least relevant)? A perhaps more interesting question is when do we have 'sufficient' evidence to obtain reliable results from the decision model whereby 'sufficient' may be defined in terms of parameter uncertainty (i.e. the incorporation of additional evidence has little effect on reducing the parameter uncertainty) or precision-bias trade-off. The issue of 'sufficient' evidence is important because generating additional information draws on a limited pool of resources and it may be inefficient to require they be used for further research that is not expected to change the decision based on existing data. For example, imprecision of, say, a resource use estimate

may or may not contribute to uncertainty in the decision. Quantification of how further information will reduce the decision uncertainty may be achieved through a decision theoretic framework such as value of information(7). To date, little is known about current practice regarding evidence identification and selection strategies except that it is neither uniform nor transparent(3). However, it is suspected that researchers use what they perceive to be the most relevant or convenient evidence for each model input parameter rather than all relevant evidence available.

Model structure

The process by which a model structure is derived is important as an inappropriate model structure may lead to unreliable model results. In their review of Department of Health HTA reports, Cooper et al.(3) found that the development of model structure and assumptions were only discussed in 5 out of the 42 (12%) reports and in these reports only discussed very briefly (e.g. 'the model structure was developed in conjunction with clinical experts'). Therefore, although dimensions of model structure (e.g. justification of modelling approach, structural assumptions, model type, etc.) are included in many of the published good decision modelling practice guidelines, little is known about how model structure is decided upon in practice. It is suspected that the first step of most modellers is to identify previous natural history or economic models developed in the clinical area of interest and, where appropriate, use these to inform the structure of the decision model designed to answer their pre-specified objectives in conjunction with clinical 'experts'. However, is there a need in future to formalise the process, and if so how and what evidence should be used? One possibility is the application of problem structuring methods (PSM)(8) from the operations research field. That is applying 'soft' qualitative structuring processes such as soft systems methodology (SSM)(9), strategic options development and analysis (SODA)(10) or journey making(11) to inform the quantitative decision-modelling process. Thus making the process of developing the model structure more formal and explicit. This type of approach, if reported appropriately, might aid decision makers when assessing the appropriateness of the model structure and therefore the reliability and accuracy of the model results. Adopting such an approach would mean decision modellers learning new skills such as construction of cognitive maps and interviewing. For example, PSM relies on skilled facilitators to help construct cognitive maps as a way of capturing individuals' views and then merge these maps to provide a single model based on explicit choices and assumptions(8). There would also be a requirement for new methodology to integrate PSM within a decision analytic modelling framework; that is, integrate qualitative

problem structuring with quantitative problem solving. Evidence from existing models (e.g. natural history, epidemiological, economic), literature searches and expert opinion would be used as sources to inform the PSM process.

Having developed a model, another issue of contention is whether a model's structure should be adapted according to data availability. It may be argued that reducing the parameter space to match the available data is usually a result of modellers not knowing how to appropriately integrate the available data with the model developed(12); thus leading to hidden uncertainty. A good example of where available evidence is used efficiently is the model for HIV prevalence in the UK where synthesis methods are used in which the number of parameters exceeds the number of data points required for the model(13).

Guidance on how to report the process of developing the model structure is required to ensure transparency.

Model inputs

The appropriate methodology for the identification of evidence to inform model parameters, other than RCTs of clinical effectiveness where established literature database searches and published guidance of quality assessment exist, was a major issue of discussion. Some guidance is provided in the NICE guide to methods of technology appraisal(2); for example, "*all evidence on outcomes should be obtained in a systematic review…*" and "*where cost data are taken from the literature, the methods used to identify the sources should be defined. Where several sources are available, a justification for the costs chosen should be provided…*". However, in the recent review of HTA reports(3), only 22 out of 42 (52%) reports clearly stated the sources of evidence used to inform the different model parameters. Particularly unclear were the sources of evidence used to inform adverse events and complications, and resource use. As search strategies for identifying the relevant evidence, other than clinical effectiveness, were rarely made explicit in the HTA reports reviewed it is difficult to obtain a clear view of what is currently done in practice. For example, where it was stated that expert opinion was used it was often unclear whether this was for convenience or because no other sources of evidence were available for that particular parameter at the time of the analysis.

The extensiveness of the evidence identification process to inform model parameters will undoubtedly be very much limited by the financial and time constraints of the investigators. Recent research by Golder et al.(14) investigated the feasibility and efficiency of literature database (e.g. MEDLINE, EMBASE) searching for individual model parameters including baseline event rates, resource use, unit costs, health-related quality of life and outcomes as well as relative clinical effectiveness. Through searching the databases, the authors identified 1,237 records of which only 48 contained data relevant to populate the model. This type of systematic searching to populate decision-analytic models requires continual dialogue between the information specialist and decision modeller(s) to establish the appropriate search strategies defined in terms of inclusion and exclusion criteria which may not be apparent from the outset (i.e. difficult to pre-specify). This has substantial time and budget implications, and for this reason is rarely carried out in practice. Difficulties in appropriately specifying inclusion and exclusion criteria, especially where the evidence required is not the primary focus of a paper, may lead to a failure to identify the evidence of interest. A more efficient approach may be to conduct focused but systematic structured searches around those parameters expected to have the largest influence on the model results.

The appropriateness of exhaustive searching compared with undertaking searches which reflect the complexity of the evidence-base used in modelling was discussed. Another alternative, and perhaps more efficient and pragmatic strategy, supported by many of the workshop participants, was to take a more iterative approach to searching. This would reflect more accurately the way evidence is used in modelling where evidence informs model development which in turn defines more search queries and so on. Such approaches to searching include citation searches, 'snowballing' and 'berrypicking'. This would still be classified as a systematic structured identification strategy although not exhaustive.

Interestingly, Egger et al(15) investigated the importance of comprehensive literature searches and the assessment of trial quality in systematic reviews. From this empirical study the authors found that trials which were difficult to locate were often of lower quality and therefore raised a concern that rather than preventing bias through extensive literature searches there was the potential that biases may be introduced by including trials of low methodological quality. They concluded that '*in situations where resources are limited, thorough quality assessments should take precedence over extensive literature searches and translations of articles'*.

More research is required to compare the results obtained from the different evidence identification strategies as well as the implications on the final model results. There is a need to acknowledge that methodology for identifying evidence for decision model parameters may be different than that established for systematic reviews not just for reasons of resource. Currently, the identification strategy for model inputs other than clinical effectiveness are poorly reported(3), if at all, and therefore it is important that whichever method of identification is used, it is clearly documented enabling reproducibility and updating. In addition, formal methods of appraisal regarding the quality and rigour of the evidence used to inform model parameters, which is usually informal at present, would be advantageous as it would aid decision-makers when assessing the reliability of the model results and also updating the analyses.

Another interesting possibility is the development of methods that would be able to determine in advance whether, given an existing body of ,say, high quality RCT-based 'direct' evidence identified by standard literature review methods, there would be any material reduction in remaining uncertainty by seeking to identify additional 'indirect', or lower quality observational evidence. Such methods would be especially valuable as they could be used *in advance* of carrying out searches, thus avoiding the requirement to identify and then sift through potentially thousands of articles without substantially affecting decision uncertainty.

ii) Assessment of quality/reliability of different sources of evidence

In the NICE guidance(2) it is stated that evidence must be quality assessed. Evidence may lack quality for a number of reasons; for example, RCTs may suffer from inadequacy of randomisation or allocation concealment, non-compliance and/or drop-out whereas observational studies may suffer from selection bias, non-response bias, and/or confounding. Therefore assessing the quality of the evidence to be combined within a decision model is important as the results obtained will be dependent on the quality of the data inputs.

Instruments such as the Jadad score(16) have been developed for assessing the quality of RCTs and Newcastle-Ottawa scale(17) for assessing the quality of non-randomised studies but it becomes more difficult to assess quality across study designs(18). Hierarchies of

evidence may be useful as they rank the different study designs but they fail to incorporate quality of the different studies; for example, is a good quality case-control study better than a poor quality RCT? It is also important to note that developed quality scales are usually applied based on the study publications where the methods reported may be limited due to journal space restrictions thus assessing the quality of study reporting.

If all relevant evidence is to be identified then being able to assess quality across study designs will be essential. Currently, for clinical effectiveness it appears acceptable to limit the evidence to RCTs thus assigning a weight of zero to all other available evidence (i.e. a value judgement) whereas for other model parameters a wide range of evidence may be considered (e.g. from RCT to observational to expert opinion). More guidance is required on how to quality assess evidence and then incorporate this information into the analysis.

Another issue beyond quality, but equally important, relates to the *relevance* of the evidence to the question of interest. For example, are the study population, outcomes and interventions different to those being considered by the appraisal? And if so, how do we appropriately extrapolate between the data required and the actual data provided in the literature?

iii) Synthesis of evidence

Adjusting for quality/bias

Leading on from the previous section, how can quality/potential biases be adjusted for? For example, it is acknowledged that observational studies are often of poorer quality than RCTs as they are prone to internal biases such as selection bias, non-response bias, and/or confounding but can this be adjusted for in the analysis? Possible solutions from the statistics literature include: i) exclusion of studies below a pre-specified quality threshold, ii) adjustment of the weight given to a study in the analysis by quality(19), iii) random effects modelling of bias(20) by choosing a parameter to describe the plausible size (and direction) of the bias for each study or study design, and iv) full bias modelling(21) that attempts to identify all sources of potential bias in the available evidence and to obtain external information on the likely form of each bias and then construct a model to correct the data analysis accordingly. The first method is very simplistic with an arbitrary threshold whereas the other methods are subjective regarding the judgement about how study quality/design affects the reliability and precision. These methods ideally require

information being available to inform the extent of the biases, and are complex methods which may be difficult for decision makers to interpret the results from. Empirical work to quantify the uncertainty associated with potential study-bias has recently been commissioned by the NHS Research and Development Methodology Programme (http://pcpoh.bham.ac.uk/publichealth/nccrm/index.htm).

There is also the issue that if we are going to adopt a method of adjusting for study quality should this be presented as the main analysis or, given the above limitations of the methods developed to date, a sensitivity analysis? If presented as a sensitivity analysis then how is this to be interpreted by decision-makers especially if the results change significantly depending on the thresholds or weightings assigned? There is also an issue regarding when it would be acceptable for a health technology assessment to conclude that the quality of the available evidence is too poor and therefore high-quality studies need to be undertaken before an informed decision can be made.

Mixed (multiple) treatment comparisons

Often systematic reviews and health technology assessments are limited to pairwise comparisons of interventions where direct evidence exists. However, there maybe an array of candidate interventions relevant to the clinical question of interest in which case an analysis comparing all the candidate interventions maybe more appropriate and useful to decision-makers. To address this latter question, methodology known as mixed (multiple) treatment comparisons(22) or network meta-analysis(23) has been developed which combines both direct and indirect evidence within a single analysis thus enabling the different intervention options to be compared in terms of effectiveness and cost-effectiveness. For example, if we were interested in the comparison of treatment A vs. B then trials of A vs. B would provide direct evidence and trials of A vs C combined with trials B vs C would provide indirect evidence but it would be more useful for decision makers to present a 3-way comparison of A vs. B vs. C. Even if the comparison of interest is considered by the decision maker to be a pairwise comparison of 2 active treatments or an active treatment vs. placebo, mixed treatment comparison methods allow the incorporation of additional related data thus reducing the parameter, and ultimately the decision, uncertainty around the pooled treatment effect estimate.

The main advantages of mixed treatment comparison methodology are i) when combining RCT evidence randomisation is maintained, ii) transparent framework (i.e. no need for

'under the table' indirect comparisons to be made based on a series of pairwise comparisons), iii) reduced uncertainty through inclusion of more data, and iv) greater robustness of analysis as inclusion of more data tends to make analysis more representative and also can allow consistency of evidence to be checked. There is a need for greater awareness of mixed treatment comparisons by commissioners of research when 'scoping' for health technology assessments.

iv) Other issues

Use of patient level data

NICE guidelines(2) state that the construction of an analytical framework will be necessary and that *"This may be a decision-analytic model using aggregated data or a statistical model using patient level data"*. Access to patient level data allows more in depth analyses of the data than is possible from summary data usually extracted from published study reports. For example, patient level data can provide information on the i) correlation structure between model parameters, ii) subgroup analyses, and iii) internal validity not usually available from the summary reports. However, obtaining patient level data from trial investigators can be time consuming and further research is required to investigate the added value of obtaining patient level data. It is also unlikely that patient level data will be supplied by primary investigators for all relevant studies; therefore, methods for combining summary data with patient level data are required as limiting the analysis to those studies where patient level data is provided may result in a non-representative sample of studies and thus bias the model results(24).

Reporting of evidence sources/identification methodology

In order for health technology appraisals to be reproducible and updateable it is essential that the methods used to identify sources of evidence for inclusion in the decision models are made explicit in the final report. To date, although the identification process, including inclusion and exclusion criteria, for the RCTs to inform the clinical effectiveness parameter are usually well reported, the identification process for other decision model parameters is less clear. This is mainly because search strategies for the identification on clinical effectiveness studies are more established and therefore easier to report than for the other model parameters. Suggested reporting requirements for all model parameters include: i) List of parameters in the decision model together with method of evidence identification (e.g. systematic review, non-systematic review, primary research, expert opinion, etc.), ii)

Details of selection process of evidence (i.e. inclusion and exclusion criteria), and iii) Quality and relevance of selected evidence (use of validated checklists, comment of generalisability of evidence to specific decision problem, etc.). More guidance on the most appropriate way to report the above information is the subject of further research. Due to space restrictions of journal articles, there may be an argument for published decision models to include supplementary material on a website,

SUMMARY

It is imperative that evidence for all model parameters are identified systematically, quality assessed and, where applicable, pooled using explicit criteria, and reproducible methods. More formal and replicable approaches to identification and assessment of the quality of model inputs may reduce the 'black box' concept of decision models and lead to less scepticism regarding the model outputs. However, to achieve the above goal further research is required into appropriate methodology for i) identifying the *relevant* evidence for the different model parameter, ii) using the evidence appropriately to inform model structure and parameter estimates, iii) quality assessing and synthesising evidence from different study designs, and iv) reporting all of the aforementioned in a transparent and reproducible way. More research is also required into the opportunity cost (financial as well as time) of delaying the decision-making process to employ a more thorough evidence identification and selection process, more sophisticated evidence synthesis methodology, obtain patient level data from primary study investigators, and so on.

ACKNOWLEDGEMENTS

Thank you to the Medical Research Council Health Services Research Collaboration for funding the workshops and all workshop participants, without which this work would not have been possible. A special thank you to all of the workshop presenters who shared their thoughts and current research ideas to help stimulate the discussions. Slides of talks presented at the workshops are available from: http://www.hs.le.ac.uk/personal/njc21/

REFERENCES

1. Mugford M. Using systematic reviews for economic evaluation. In: Egger M, Davey Smith G, Altman D, editors. Systematic reviews in health care: Meta-analysis in context. 1 ed. London: BMJ Publishing; 2001. p. 419-28.
2. NICE. Guide to the methods of technology appraisal. London: National Institute of Clinical Excellence; 2004.
3. Cooper NJ, Coyle D, Abrams KR, Mugford M, Sutton AJ. Use of evidence in decision models: An appraisal of health technology assessments in the UK to date. Journal of Health Services Research and Policy. 2005;10(4):245-50.
4. Kuntz KM, Weinstein MC. Modelling in economic evaluation. In: Drummond M, McGuire A, editors. Economic evaluation in health care: Merging theory with practice. Oxford: Oxford University Press; 2001. p. 141-71.
5. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technology Assessment. 2004;8(36).
6. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. The Lancet. 1994;354:1896-900.
7. Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty - When is there sufficiennt evidence? Value in Health. 2005;8(4):433-46.
8. Pidd M. Tools for Thinking : Modelling in Management Science. 2nd ed. Chichester: John Wiley & Sons; 2003.
9. Checkland PB. Systems thinking, systems practice. Chichester: John Wiley & Sons; 1981.
10. Eden CL. Using cognitive mapping for strategic options development and analysis (SODA). In: Rosenhead J, editor. Rational analysis for a problematic world. Chichester: John Wiley & Sons; 1989.
11. Eden CL, Ackermann F. Strategy making: The journey of strategic management. London: Sage; 1998.

12. Ades AE, Cliffe S. Markov Chain Monte Carlo estimation of a multiparameter decision model: Consistency of evidence and the accurate assessment of uncertainty. Medical Decision Making. 2002;22:359-71.

13. Goubar A, Ades AE, De Angelis D, McGarrigle CA, Mercer C, Tookey P, et al. Estimates of HIV prevalence and proportion diagnosed based on Bayesian multi-parameter synthesis of surveillance data. JRSS. Submitted.

14. Golder S, Glanville J, Ginnelly L. Populating decision-analytic models: The feasibility and efficiency of database searching for individual parameters. International Journal of Technology Assessment in Health Care. 2005;21(3):305-11.

15. Egger M, Jüni P, Bartlett C, Holenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. Health Technology Assessment. 2003;7(1).

16. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clinical Trials. 1996;17(1):1-12.

17. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. http://wwwohrica/programs/clinical_epidemiology/oxfordhtm Accessed 21st Nov 2005.

18. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. Journal of Epidemiology and Community Health. 1998;52:377-84.

19. Tritchler D. Modelling study quality in meta-analysis. Statistics in Medicine. 1999;18:2135-45.

20. Spiegelhalter DJ, Best NG. Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. Statistics in Medicine. 2003;22(23):3687-709.

21. Greenland S. Multiple-bias modelling for analysis of observational data. Journal of the Royal Statistical Society: Series A (Statistics in Society). 2005;168(2):267-306.

22. Caldwell DM, Ades AE, Higgins JPT. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ. 2005;331:897-900.

23. Lumley T. Network meta-analysis for indirect treatment comparisons. Statistics in Medicine. 2002;21:2313-24.

24. Sutton A. Exploring the effect of patient characteristics on effectiveness using a combination of individual subject and aggregate level data. Cochrane Colloquium; 2005; Melbourne, Australia; 2005.