VALUE IN HEALTH 16 (2013) 830–836

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Comparative Effectiveness Research/Health Technology Assessment (HTA)

Reviewing the Evidence to Inform the Population of Cost-Effectiveness Models within Health Technology Assessments

Eva Kaltenthaler, BSc, MSc, PhD, Paul Tappenden, BA, MSc, PhD, Suzy Paisley, BA, MA, PhD*

School of Health and Related Research, University of Sheffield, Sheffield, UK

ABSTRACT

Objectives: Health technology assessments (HTAs) typically require the development of a cost-effectiveness model, which necessitates the identification, selection, and use of other types of information beyond clinical effectiveness evidence to populate the model parameters. The reviewing activity associated with model development should be transparent and reproducible but can result in a tension between being both timely and systematic. Little procedural guidance exists in this area. The purpose of this article was to provide guidance, informed by focus groups, on what might constitute a systematic and transparent approach to reviewing information to populate model parameters. **Methods:** A focus group series was held with HTA experts in the United Kingdom including systematic reviewers, information specialists, and health economic modelers to explore these issues. Framework analysis was used to analyze the qualitative data elicited during focus groups. **Results:** Suggestions included the use of rapid reviewing methods and the need to consider the trade-off between relevance and quality. The need for transparency in the reporting of review methods was emphasized. It was suggested that additional attention should be given to the reporting of parameters deemed to be more important to the model or where the preferred decision regarding the choice of evidence is equivocal. **Discussion:** These recommendations form part of a Technical Support Document produced for the National Institute for Health and Clinical Excellence Decision Support Unit in the United Kingdom. It is intended that these recommendations will help to ensure a more systematic, transparent, and reproducible process for the review of model parameters within HTA.

Keywords: cost-effectiveness modeling, evidence-based decision making, health technology assessment, model parameters, systematic review methods.

Introduction

Health technology assessment (HTA) reports used to inform evidence-based decisions concerning the use of health care interventions typically involve the development of a systematic review of clinical effectiveness and the development of a cost-effectiveness model. By its very nature, the development of the model requires information beyond clinical efficacy such as health utilities, resource use, and costs. In addition, the model structure requires the use of evidence to inform judgments concerning the plausibility of relationships between intermediate and final end points, as well as other information to determine what is relevant for inclusion in the model. The way in which this evidence is used can have a fundamental impact on results of the model and ultimately the decision outcome [1]. The main groups of information needs are illustrated in Figure 1. It should be noted that the five categories presented are not mutually exclusive and there will be overlap between them. The information needs represented here include both soft contextual information and harder experimental or nonexperimental evidence.

A number of issues need to be considered when reviewing evidence to inform the specification and population of cost-effectiveness models. These include the timelines for HTA, which may be restrictive, because decisions on technologies are often needed as near to the time of licensing as possible and often before clinical effectiveness has been established. There is also a need for methods that are systematic, transparent, and reproducible to minimize the risk of bias and therefore produce more robust results. If model results are to be considered credible, researchers need to be transparent about how the model was developed and why certain inputs should be considered reliable. Sources of evidence may vary widely between models. These sources include randomized controlled trials, observational evidence and other clinical studies, registry databases, elicitation of expert clinical judgment, existing cost-effectiveness models, routine data sources, and health valuation studies. Previous work by Coyle et al. [2] looked at the most common data elements within models (clinical effect sizes, baseline clinical data, resource use, unit costs, and utilities) and developed a hierarchy of data sources for these. For clinical effect size, the authors recommend the highest level of evidence to be meta-analysis of randomized controlled trials with direct comparison between comparator therapies and the lowest ranking evidence to be expert opinion. For another evidence

* *Address correspondence to:* Eva Kaltenthaler, School of Health and Related Research, University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA, UK.

E-mail: e.kaltenthaler@sheffield.ac.uk.

1098-3015/$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2013.04.009

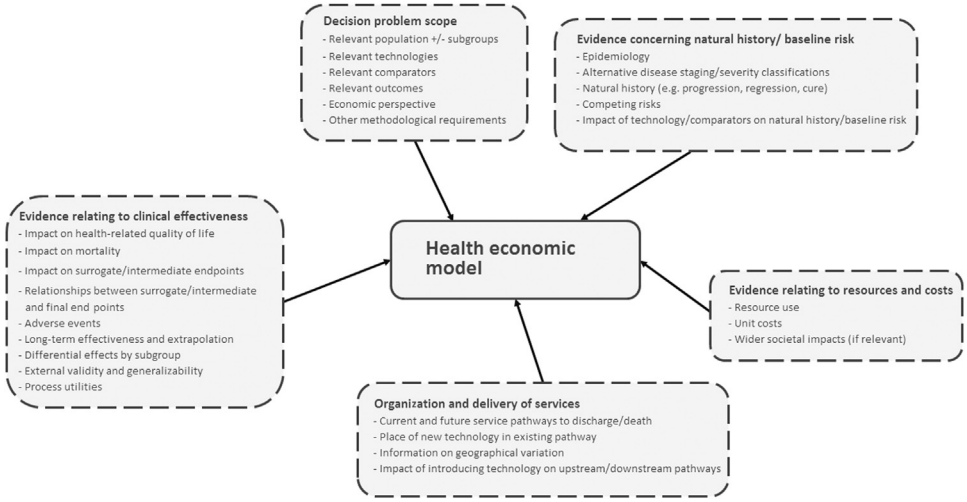

Fig. 1 – Types of evidence used to inform models.

requirement, resource use, the authors considered prospective data collection or analysis of reliable administrative data to be the highest level of evidence, with expert opinion again the lowest.

While reviewing processes are often used to identify evidence for economic models, it is less usual for model reports to describe and justify how they have identified and synthesized the evidence beyond the efficacy data or for reports to set out criteria against which the relevance and quality of the evidence are assessed [3]. While some of the issues surrounding reviewing evidence for models have been discussed previously [1,3–8], there remains very little formal guidance with respect to best practice in this area. Briggs et al. [9] in their ISPOR-SMDM Modeling Good Research Practice report recommend that analysts should conform to the broad principles of evidence-based medicine and avoid "cherry picking" the best single source of evidence. Thus, the selection of sources of evidence for model parameters should follow a systematic and transparent approach. Coyle and Lee [1] demonstrated that using different sources of evidence can have a substantial impact on the results and highlighted that there is a lack of agreement as to what constitutes good evidence for specific data inputs in economic models. It has further been argued that one potential source of errors in HTA models is the separation of information gathering, reviewing, and modeling functions [10].

Current methodological guidance regarding the reviewing of evidence to inform model parameters, apart from clinical effectiveness from the National Institute for Health and Clinical Excellence (NICE), states: "For all parameters (including effectiveness, valuation of HRQL and costs) *a* systematic consideration of possible data sources is required" [11]. This absence of clarity presents a considerable challenge to organizations submitting evidence to NICE because a full systematic review is clearly not required for each parameter, yet it is not clear what a "systematic consideration" is. A recent Technical Support Document (TSD) from the NICE Decision Support Unit [12] considers the requirements and provides methodological guidance for identifying and reviewing evidence to inform models of cost-effectiveness, in particular model parameter estimates, in the NICE Technology Appraisal Process. While this was developed to inform assessments of pharmaceutical interventions, it also has a wider relevance to the appraisal of medical devices and diagnostic techniques. Issues surrounding the identification, review, and selection of evidence to inform model parameter values are relevant to economic analyses that involve secondary data alone as well as those in which a combination of primary and secondary data is required. When economic analyses are undertaken alongside a clinical trial, it is rare that full evidence requirements to assess costs and effect of the technology would be sourced from the trial alone [13]. Part of the TSD provides guidance on methods for reviewing model parameter data in a systematic fashion. It draws distinctions between systematic reviews and reviewing in the context of informing model parameters and demonstrates how the key components of systematic review methods can be used to systematize and make explicit the choices involved in selecting evidence to inform models. Individual model parameters will have different characteristics and therefore varying evidence requirements, information availability, and reviewing needs. The purpose of this article was to provide guidance, informed by a series of focus groups, on what might constitute a systematic and transparent approach to reviewing information to populate model parameters where there is no requirement to use conventional systematic review methods and where little procedural guidance exists. While precise methods that should be used to review individual evidence types should be judged on a case-by-case basis, issues that need to be considered should not. This article highlights what these considerations are and emphasizes the importance of being transparent in how such judgments are reached. The article describes the methods used and identifies seven key themes related to the reviewing of evidence for model parameters and provides further analysis of key themes in the discussion.

Methods

A series of focus groups was used to gather information on issues around reviewing for model parameters and provide the basis of recommendations covered in the TSD. An initial pilot focus group was held with 18 researchers who had extensive experience in HTA including 6 systematic reviewers, 2 information specialists, and 10 health economic modelers in January 2010. The researchers were all from the School of Health and Related Research at the University of Sheffield, which is a major provider of HTA reports in the United Kingdom. A range of people with different areas and of expertise in HTA was invited to attend the focus group; these individuals were identified purposively to reflect the breadth of input into the model development process. A topic guide was developed to structure the discussion within the focus group and was informed through discussion with experts in the field of HTA and included questions identified through informal

interviews with staff. The topic guide included questions such as the following:

- How do HTA practitioners currently approach reviewing for model parameters?
- How do we know when we have adequate information and does this vary for different types of parameters?
- Does the timing at which reviewing activity takes place vary and should this commence at the beginning of the project?
- What is ideal practice and what changes could be made?
- What are the areas requiring further research?

In June 2010, a seminar was held at the School of Health and Related Research and used as a member-checking device. All the 18 researchers among others were invited to the seminar where further discussions were held on each of the key themes identified in the focus group.

The key issues, identified through the initial pilot focus group and seminar, were presented at a workshop held on February 7, 2011. This workshop included 13 participants from UK universities considered to be experts in the field of HTA including seven modelers, one health economist, one statistician, two information specialists, and two reviewers. Purposive sampling was used to identify participants. The workshop consisted of three focus group sessions. Topic guides for the workshop focus groups covered the following topics and questions that were generated from findings of the pilot focus group and the subsequent seminar:

- Model development (How does your "final" model structure arise? How do you know which parameters are relevant? How iterative is the process?)
- Time constraints (What kind of compromises are acceptable to make?)
- Sufficient evidence (What is your interpretation of sufficient evidence?)
- Communication and teamwork (What defines good communication and teamwork in this context?)
- Problem structuring (How do you decide what should be included in a model and what should be excluded?)
- Identification of evidence (Do we need to provide guidance on how to access nonstandard information? How do we handle issues related to comprehensiveness and sensitivity? What advice can we offer with respect to rapid searching?)
- Reviewing methods (How do we appraise for both quality and relevance? How does the process of inclusion/exclusion differ in the context of modeling? With rapid review methods what compromises are we willing to make?)
- Recommendations for reporting (How do we report decision making and judgments? What needs to be reported to allow judgment of the credibility of the model?)

The discussions were open, and participants were invited to discuss other points they felt were relevant but had not been included in the guide.

Ethical approval for all focus groups was obtained from the University of Sheffield. The focus groups were facilitator-led (E.K.) and were all recorded by using digital media, with the recordings transcribed verbatim. Standard qualitative research methods were used to conduct the focus groups [14]. Qualitative Framework Analysis [15], which involves using a thematic framework to classify and organize data according to key themes, subdivided into related subtopics, was used to analyze the transcribed data. An initial data familiarization process of the transcribed data was undertaken. The data were coded, and a conceptual framework was developed with a list of themes and related subthemes. Coding and identification of themes and subthemes was checked by a second researcher. Some of the themes were preidentified, and some were emergent.

The data obtained through the information gathering activities described above informed the development of guidance on the reviewing of evidence to populate model parameters. A report of the draft recommendations was shared with all workshop participants for comment and revised accordingly.

Results

Key themes related to the reviewing of model parameters identified from the focus groups were 1) selection and prioritization of data to inform parameter estimates, 2) reviewing methods, 3) minimizing bias, 4) hierarchies of evidence, 5) study selection, 6) assessment of evidence, and 7) evidence synthesis and analysis. These themes informed the statements presented below. More detail of the focus group findings is available from Kaltenthaler et al. [16,17], while more detail on the resulting recommendations can be found in the TSD [12].

Theme 1: Selection and Prioritization of Data to Inform Parameter Estimates

Every model parameter will need to be estimated; therefore, the choices made regarding the values selected need to be explained and justified. The choice of estimate will often be made according to some trade-off or weighing up of the available options, rather than according to rigid, predefined criteria. This may be because an estimate is required and there will usually be at best a range of options, all of which may fall short of what would be considered ideal to differing degrees. The nature of the trade-off between selecting alternative parameter estimates will often include considerations relating to quality versus relevance for each option. Procedures associated with undertaking systematic reviews can be used to make the process of choosing evidence more systematic and transparent. Given the differences between models and systematic reviews, however, the purposes for which these procedures are undertaken and the sequence in which they are undertaken may differ. In addition, time and resource constraints will have an impact on how they are undertaken. These processes need to be justifiable and replicable. It is important to prioritize parameters and focus reviewing resources on those most likely to have an impact on model outputs, bearing in mind that the importance of parameters is subject to change during the course of the modeling process. Although some parameters will be identified as important to the model early on in the process, that is, they have an impact on the model outputs, the importance of some other parameters will be identified only later in the process. A parameter may be considered important if the model results are heavily influenced by it or it is identified as such by a clinician or decision maker.

When model parameters are deemed to be important later in the development process, time constraints necessitate the use of rapid searching and review methods. Some caution is advised with respect to using simple sensitivity analysis to prioritize certain model parameters for detailed review, especially before the model is finalized. If the model structure changes temporally throughout its development process, or if other evidence is identified to inform other model parameters, the original conclusions regarding the importance of a given model parameter may no longer hold.

Theme 2: Reviewing Methods

Owing to time and resource constraints, it may be necessary to use rapid review methods to identify and select evidence to inform certain model parameters. Rapid review methods differ

from systematic review methods in that they involve a shorter time frame and may have limited searching, appraisal, and synthesis [18–20]. Timescales for reviewing may be shortened by carefully focusing on the research question, conducting a review of reviews, and extracting only key variables, among other strategies [21]. The term *rapid review*, however, does not have one single definition and neither the Cochrane Handbook [22] nor the UK Centre for Reviews guidance [23] provides guidelines on conducting them [24]. Other potentially relevant rapid review methods in this context include reduced formal quality assessment and reduced levels of synthesis.

Rapid review methods are not ideal because they risk missing relevant information. It is therefore essential that methods be reported in a transparent manner and that the limitations and potential biases of the chosen approaches be addressed. Some rapid methods used for reviewing clinical effectiveness evidence may also be applicable for reviewing evidence to inform model parameters, including the use of restricted review questions and restrictions on included study designs. There may however be issues with identifying suitable review methods for some model parameters. For example, for reviews of health state utility values, relevance needs to be assessed alongside the quality of studies [25].

Theme 3: Minimizing Bias

A variety of potential biases may be introduced through the process of rapidly reviewing evidence to inform model parameter estimates. These may include biases introduced through the use of less thorough searching, potentially resulting in publication bias or limiting appraisal or quality assessment, which may result in more emphasis placed on poorer quality research or lack of attention to synthesis, resulting in overlooking inconsistencies or contradictions in the evidence [21]. Bias may also be introduced through the purposive selection of evidence to create more or less favorable results. One option to reduce such bias is to ensure and to demonstrate that more than one member of the team is involved with making decisions where choices about values need to be made. This is partly because there may be more than one plausible option, and a joint decision may provide a more robust and systematic approach to considering the advantages and disadvantages of each. Those involved in this decision-making process may include clinical advisors, information specialists, systematic reviewers, and other modelers on the team. In such circumstances, the process of documenting the review methods used and highlighting their limitations should be considered important [21].

Theme 4: Hierarchy of Evidence

Types of evidence used to populate models will vary considerably. Hierarchies of evidence sources as suggested by Coyle et al. [2] may be of use as a means of judging the quality of individual parameter estimates and aid the study selection process. While hierarchies of evidence may be useful, there are other issues to consider, including the quality of the individual studies. The Grading of Recommendations Assessment, Development and Evaluation (GRADE) [26] system potentially provides a framework for rating the quality of evidence from all potential sources of all data components that may be used to populate model parameters. It allows flexibility in the quality assessment process to include additional considerations alongside internal validity, including (crucially for most data components used to populate model parameters) applicability to the specific decision problem at hand, which is part of the "indirectness" criterion in GRADE. Thus, the extent to which the available evidence reflects levels and combinations of resource use appropriate to the decision under consideration is a key component of the quality assessment.

Theme 5: Study Selection

The definition of what is required may be based on an initial understanding of what constitutes "relevant" evidence. The objective is to identify a set of possible options from which choices will be made. One option is to initially apply strict selection. If no relevant studies are identified, the selection criteria may then be broadened. It is important to explain the process used and why it was chosen to justify the choices and to maintain transparency. For many parameters there may be very few sources and potential studies to use or alternatively many good quality studies to choose from. If several potentially relevant studies are identified, slightly stricter selection criteria may be applied. In instances whereby a large number of sources are identified, study selection using standard systematic review processes of screening for titles, abstracts, and full texts may be most appropriate. It is important to be as transparent as possible about the judgments being made when selecting studies, for example, stating which studies were deemed to be most relevant to the setting under consideration.

Evidence used to inform model parameters will need to be assessed on the basis of relevance to the context of the decision problem, as well as quality. By assessing relevance first, a large number of studies may be eliminated. Ideally, criteria for relevance should be established *a priori*. However, it is important to recognize that it is not possible to have prespecified criteria for every parameter as information needs may change (e.g., if the model structure evolves) and information that was not expected may be identified iteratively. Relevance criteria may therefore change throughout the project; hence, flexibility is essential. What remains important is that the criteria or factors that inform the choice of evidence remain clear. Anticipated evidence requirements, as perceived during the earlier stages of model development, may be identified by adopting an explicit stage of conceptual model development before embarking on the mathematical model [27]. When the final model is developed, it is important to be clear how this deviated from the initial plan and why to justify abstractions and simplifications driven by the available evidence.

There is a tension between identifying all relevant evidence and producing a model that is useful for informing decision making. The definition of "relevance" is a subjective judgment, and this may change over the course of developing a model. This may be best avoided by determining evidence requirements early on in the process by following a formal conceptual modeling approach enabling dialogue between clinicians, decision makers, and other stakeholders as to whether there is an agreed definition of relevance. The use of *a priori* criteria for evidence review and appraisal is also helpful.

Theme 6: Assessment of Evidence

After appraising studies for relevance, they can then be assessed for quality, preferably by using standardized quality assessment tools. In this context, quality assessment may be difficult due to the absence of standardized methods for all types of information used to populate the model. Also, some studies may be poorly reported. It may be possible to establish quality assessment criteria *a priori*. An example of this is data collected for utility studies, which may include study recruitment procedures, inclusion and exclusion criteria, descriptions of the background characteristics of the sample population from whom values are obtained, response rates, and follow-up data [25]. Other issues to consider include the type of reporting (self or proxy), follow-up rates, number of patients, location, and methods of elicitation among other issues. Establishing very broad *a priori* criteria may be necessary, initially making quality assessment closely lined with selection of evidence. Criteria may change according to the

availability and relevance of existing evidence. For example, "there were five options and we chose option 1 because of the reasons a, b and c." This also captures the necessary trade-off between relevance and quality. It is important to be clear about the factors or criteria that drive the choice and to examine the implications of that choice. This level of transparency will allow judgments to be made as to whether or not a reasonable choice has been made. Because it can be very time consuming to judge the quality of all potentially relevant studies, adjusting them according to relevance and rigor may not be practical. Some types of data are of potentially very poor quality, and it can be very difficult to identify appropriate sources of information, for example, for cost data. These are not limitations of the cost-effectiveness model but rather of the evidence base and as such these evidence gaps should be exposed and clearly reported.

Data to be extracted from studies may include study date, information on disease area and patients (age, sex, comorbidities), study methods, outcomes, and other important descriptive details. This can be set out *a priori* and presented in a way to make it easy for the reader to compare and contrast the characteristics of the available studies from which a selection has been made, for example, using tables and/or graphs. This level of detail is not appropriate for all parameter estimates but should be reserved for those decisions whereby none of the available studies is clearly superior or whereby evidence available to inform a particular parameter or set of parameters is notably weak. When extracting data from studies, it is important to provide information for all the potentially relevant studies. By providing a summary of all potentially relevant studies, the reader is able to assess the study differences and heterogeneity more accurately and to examine the spread of evidence. Information from the studies that are not selected may be used to inform the sensitivity analysis. Inconsistencies between different estimates should be represented. Although the results presented may be wide when using the available studies, it is important to show how the range of values between disease stages or different baseline event rates, for example, are driving the model results. It is recognized that these suggestions may be quite time consuming and there may be time and reporting constraints within an HTA report. However, the overriding objective should be to present the information and its implications on model results as clearly as possible.

Theme 7: Evidence Synthesis

For many types of model parameters, the issue of synthesis may not be considered relevant because of study heterogeneity. Often only one or two values are appropriate for use in populating a model parameter. Where there are more than one or two potentially relevant studies, the issue of synthesis becomes important. Guidance is available regarding evidence synthesis techniques for model parameters [28–35]. A decision needs to be made as to whether complex synthesis methods will provide a meaningful value for a parameter. In some instances, however, it may be simpler and more defensible to select the value from the most appropriate and relevant study as opposed to using a weighting system for pooling estimates. In some cases, it may be reasonable to inflate the uncertainty associated with that parameter estimate to reflect the range of other estimates that have not been used. This should be judged on a case-by-case basis. In instances whereby a quantitative synthesis is not undertaken, however, this should be justified explicitly. The choice of available evidence should be made clear and the implication of choosing one source from a number of available options explored through sensitivity analyses. While transparency and clear reporting are essential in exposing these types of subjective model development decisions, it is essential that this includes consideration of the evidence that has been used to inform the model parameters as well as the evidence that has not been used.

The issues identified in this research have been incorporated into a checklist, shown in Table 1. The purpose of the checklist is to guide researchers through considerations regarding evidence requirements, availability and selection of evidence, rapid review options, and reporting throughout the model development process.

Discussion

This article has presented seven key themes exploring issues around the reviewing of evidence to inform model parameters informed by a series of information gathering activities with experts in HTA in the United Kingdom. This research has focused on issues of importance in NICE technology appraisals, which involve mainly pharmaceutical interventions although the issues raised are also relevant for HTAs of nonpharmaceutical interventions as well as other HTAs. The processes of selecting and prioritizing data to inform parameter estimates were considered to be important. There was agreement that reviewing effort should be prioritized around the model parameters considered to be important and reviewing methods chosen commensurate with the parameter's importance. Caution was advised however as the importance of certain model parameters may change as other parts of the model are developed and refined. Also mentioned was the applicability of rapid reviewing methods. As suggested by Watt et al. [19] and Ganann et al. [20], rather than developing a formalized methodology to conduct rapid reviews, which may be inappropriate and oversimplified, emphasis should be placed on the transparent reporting of methods. The inclusion of more than one team member in the process of choosing appropriate parameter estimates was considered an option to help to minimize bias.

Table 1 – Checklist of points to consider when reviewing for model parameters.

Evidence requirements
- Identification of parameters requiring evidence to inform them (this may be supported by formal conceptual model development)
- Prioritization of evidence requirements by deciding on most important parameters and revisiting this throughout the project
- Consideration of time available

Availability and selection of information
- Selection of appropriate search strategy
- Identification of appropriate sources of evidence
- Consideration of the use of hierarchies of evidence
- Consideration of both quality and relevance when selecting information

Rapid reviewing options
- Restricting review questions
- Establishing level of quality assessment needed
- Reducing level of data extraction
- Data extraction of key outcomes only

Reporting
- Transparent reporting of methods used
- Transparent reporting of choices about which evidence sources have been used *and* those sources that have not
- Description of limitations of chosen methods

Hierarchies of evidence as suggested by Coyle et al. [2] were considered to be a potentially useful tool for guiding the choice of evidence to inform parameter estimates although concerns were raised because hierarchies do not take into account the quality of the evidence. GRADE [26] may be useful in this context in that it allows flexibility in the quality assessment process to include additional considerations alongside internal validity. Thus, the extent to which the available evidence reflects levels and combinations of resource use appropriate to the decision under consideration is a key component of the quality assessment. Sources of data that could be incorporated by using GRADE include national disease registers, claims, prescriptions or hospital activity databases, or standard reference sources such as drug formularies or collected volumes of unit costs [36,37].

Both quality and relevance were considered to be important when undertaking study selection. Quality assessment of the evidence was thought to present some challenges due to the absence of quality assessment tools for many types of evidence used to populate cost-effectiveness models. With regard to choices made related to evidence synthesis and analysis, these need to be made clear and explicitly justified. Study selection processes need to be clearly reported, and there should be transparency around what judgments have been made regarding study selection.

There are some limitations to this study. First, focus groups were used as the main method of data collection in this study. Other qualitative methods such as in-depth interviews [38] or consensus techniques [39] could have been used in this research and may have yielded different results. The initial pilot focus group that informed later elements of the study included researchers only from one institution. Using a wider group of researchers for the pilot focus group may have ensured that the results were more representative of practice in this field. The later workshop did however include HTA experts from several institutions in the United Kingdom. The study was focused on the needs of NICE health technology appraisals, and more research is needed to determine the reviewing needs for model parameters within HTAs for other decision makers.

The issues raised in this research are important in the field of HTA. The principles of evidence-based medicine lie at the heart of HTA; hence, questions regarding the identification and selection of evidence inevitably influence the structure, conclusions, and consequences of any HTA model. Whether the model results are credible is always a matter of subjective judgment and is dependent on the interpretation of the reader/decision maker. It is for precisely these reasons that we would advocate the need to make clear the subjective nature of the approach used and the other alternatives available that have not been followed and to examine, where possible, the likely impact of these choices on the model results. There is a need for agreed standardized practice in this area while still maintaining flexibility and adaptability to suit the needs of individual HTAs. As Cooper et al. [3] state, "it is imperative that evidence for all model parameters is identified systematically, quality assessed and where applicable pooled using explicit criteria and reproducible methods." The findings from this research support this statement. A systematic, transparent, and reproducible process is essential for the development of cost-effectiveness models to support HTAs. Further research is needed in this area and includes the need for the development of appropriate rapid reviewing methods, quality assessment tools for nonstandard sources of evidence, investigation into the use of hierarchies of evidence and GRADE, development of the methods used for the selection of evidence, and development of reporting standards.

This article goes some way in ensuring that those involved in modeling within HTAs are able to engage with the issues associated with the selection and review of evidence to inform model parameters. Potential solutions have been offered to ensure that this is done in a systematic fashion and to improve the transparency of reporting of the modeling process.

Source of financial support: Funding for this work was received from the NICE Decision Support Unit.

REFERENCES

[1] Coyle D, Lee K. Evidence-based economic evaluation: how the use of different data sources can impact results. In: Donaldson C, Mugford M, Vale L, eds., Evidence-Based Health Economics: From effectiveness to efficiency in systematic review. London: BMJ Books, 2002.
[2] Coyle D, Lee K, Cooper N. Use of evidence in decision models. In: Shemilt I, Mugford M, Vale L, et al., eds., Evidence-Based Decisions and Economics: Health Care, Social Welfare, Education and Criminal Justice. Oxford: Wiley-Blackwell, 2010.
[3] Cooper N, Sutton AJ, Ades AE, et al. Use of evidence in economic decision models: practical issues and methodological challenges. Health Econ 2007;16:1277–86.
[4] Cooper N, Coyle D, Abrams K, et al. Use of evidence in decision models: an appraisal of health technology assessments in the UK to date. J Health Serv Res Policy 2005;10:245–50.
[5] Philips Z, Ginnelly L, Schulpher M, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technol Assess 2004;8:1–179.
[6] Shemilt I, Mugford M, Byford S, et al. Incorporating economics evidence. In: Higgins JPT, Green S, eds., Cochrane Handbook for Systematic Reviews of Interventions. Chichester, UK: John Wiley and Sons, 2008: chapter 15.
[7] Marsh K. The role of review and synthesis methods in decision models. In: Shemilt I, Mugford M, Vale L, et al, eds., Evidence-Based Decisions and Economics: Health Care, Social Welfare, Education and Criminal Justice. Oxford: Wiley-Blackwell, 2010.
[8] Paisley S. Classification of evidence used in decision-analytic models of cost effectiveness: a content analysis of published reports. Int J Technol Assess Health Care 2010;26:458–62.
[9] Briggs AH, Weinstein MC, Fenwick E, et al. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. Value Health 2012;15:835–42.
[10] Chilcott JB, Tappenden P, Rawdin A, et al. Avoiding and identifying errors in health technology assessment models: qualitative study and methodological review. Health Technol Assess 2010;14:1–152.
[11] National Institute for Health and Clinical Excellence. NICE Guide to the methods of technology appraisal 2008. Available from: http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf. [Accessed July 10, 2012].
[12] Kaltenthaler E, Tappenden P, Paisley S, Squires H. NICE DSU Technical Support Document 13: identifying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models. Report by the NICE Decision Support Unit, May 2011. Available from: http://www.nicedsu.org.uk/TSD%2013%20model%20parameters.pdf. [Accessed July 10, 2012].
[13] Buxton M, Drummond MF, Van Hout BA, et al. Modelling in economic evaluation: an unavoidable fact of life. Health Econ 1997;6:217–27.
[14] Finch H, Lewis J. Focus groups. In: Ritchie J, Lewis J, eds., Qualitative Research Practice: A Guide for Social Science Students and Researchers. London: Sage, 2003.
[15] Ritchie J, Spencer L, O'Connor W. Carrying out qualitative analysis. In: Ritchie J, Lewis J,eds., Qualitative Research Practice: A Guide for Social Science Students and Researchers. London: Sage, 2003.
[16] Kaltenthaler E, Tappenden P, Paisley S. Reviewing the evidence used in cost-effectiveness models in health technology assessment: a qualitative investigation of current concerns and future research priorities. HEDS Discussion Paper Number 12/01, 2012. Available from: http://www.shef.ac.uk/polopoly_fs/1.165496!/file/12.01.pdf. [Accessed July 10, 2012].
[17] Kaltenthaler E, Essat M, Tappenden P, Paisley S. The identification and review of cost effectiveness model parameters: a qualitative study. HEDS Discussion Paper Number 13/08 2013. Available from: http://www.sheffield.ac.uk/scharr/sections/heds/discussion-papers/1308-1.268573 [Accessed May 21, 2013].
[18] Khangura S, Konnyu K, Cushman R, et al. Evidence summaries: the evolution of a rapid review approach. Syst Rev 2012;1:10.
[19] Watt A, Cameron A, Sturm L, et al. Rapid reviews versus full systematic reviews: an inventory of current methods and practice in health technology assessment. Int J Technol Assess Health Care 2008;24:133–9.
[20] Gannon R, Ciliska D, Thomas H. Expediting systematic reviews: methods and implications of rapid reviews. Implement Sci 2010;5:1–10.
[21] Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Info Libr J 2009;26:91–108.

[22] Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from: www.cochrane-handbook.org. [Accessed January 23, 2013].
[23] Centre for Reviews and Dissemination. Systematic Reviews: CRD's Guidance for Undertaking Review in Health Care (3rd ed.). York: University of York, Centre for Reviews and Dissemination, 2009. Available from: http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf. [Accessed January 23, 2013].
[24] Harker J, Kleijnen J. What is a rapid review? A methodological exploration of rapid reviews in Health Technology Assessments. Int J Evid Based Healthc 2012;10:397–410.
[25] Brazier J, Papaioannou D, Cantrell A, et al. Identifying and reviewing health state utility values for populating decision models. In: Shemilt I, Mugford M, Vale L, et al., eds., Evidence-Based Decisions and Economics: Health Care, Social Welfare, Education and Criminal Justice. Wiley-Oxford: Blackwell, 2010.
[26] Shemilt I, Mugford M, Vale L, et al. Evidence synthesis, economics and public policy. Res Synth Methods 2010;1:126–35.
[27] Tappenden P. Conceptual modelling for health economic model development. HEDS Discussion Paper 12/05. Available from: http://www.shef.ac.uk/polopoly_fs/1.202248!/file/HEDSDP1205.pdf. [Accessed January 26, 2013].
[28] Saramago P, Manca A, Sutton AJ. Deriving input parameters for cost-effectiveness modelling: taxonomy of data types and approaches to their statistical synthesis. Value Health 2012;15:639–49.
[29] Dias S, Welton NJ, Sutton AJ, Ades AE. NICEDSU Technical Support Document 1: introduction to evidence synthesis for decision making. Report by the Decision Support Unit April 2011 (last updated April 2012). Available from: http://www.nicedsu.org.uk/TSD1%20Introduction.final.08.05.12.pdf. [Accessed January 23, 2013].
[30] Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. Report by the Decision Support Unit August 2011 (last updated April 2012). Available from: http://www.nicedsu.org.uk/TSD2%20General%20meta%20analysis.final.08.05.12.pdf. [Accessed January 23, 2013].
[31] Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 3: heterogeneity: subgroups, meta-regression, bias and bias-adjustment. Report by the Decision Support Unit September 2011 (last updated April 2012). Available from: http://www.nicedsu.org.uk/TSD3%20Heterogeneity.final%20report.08.05.12.pdf. [Accessed January 23, 2013].
[32] Dias S, Sutton AJ, Welton NJ, et al. NICE DSU Technical Support Document 4: inconsistency in networks of evidence based on randomised controlled trials. Report by the Decision Support Unit May 2011 (last updated April 2012). Available from: http://www.nicedsu.org.uk/TSD4%20Inconsistency.final.08.05.12.pdf. [Accessed January 23, 2013].
[33] Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 5: evidence synthesis in the baseline natural history model. Report by the Decision Support Unit August 2011 (last updated April 2012). Available from: http://www.nicedsu.org.uk/TSD5%20Baseline.final%20report.08.05.12.pdf. [Accessed January 23, 2013].
[34] Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 6: embedding evidence synthesis in probabilistic cost-effectiveness analysis: software choices report by the Decision Support Unit May 2011 (last updated April 2012). Available from: http://www.nicedsu.org.uk/TSD6%20Software.final.08.05.12.pdf. [Accessed January 23, 2013].
[35] Ades AE, Caldwell DM, Reken S, et al. NICE DSU Technical Support Document 7: evidence synthesis of treatment efficacy in decision making: a reviewer's checklist. Report by the Decision Support Unit January 2012. Available from: http://www.nicedsu.org.uk/TSD7%20reviewer%20checklist.final.08.05.12.pdf. [Accessed January 23, 2013].
[36] Guyatt G, Oxman A, Schunemann H, et al. GRADE guidelines: a new series of articles in the *Journal of Clinical Epidemiology*. J Clin Epidemiol 2011;64:380–2.
[37] Brunetti M, Shemilt I, Pregno S, et al. GRADE guidelines: 10, considering resource use and rating the quality of economic evidence. J Clin Epidemiol 2013;66:140–50.
[38] Legard R, Keegan J, Ward K. In-depth interviews. In: Ritchie J, Lewis J, eds., Qualitative Research Practice: A Guide for Social Science Students and Researchers. London: Sage, 2003.
[39] Jones J, Hunter D. Using the Delphi and nominal group technique in health services research. In: Pope C, Mays N, eds., Qualitative Research in Health Care (2nd ed.). London: BMJ Books, 2000.