

**NICE DSU TECHNICAL SUPPORT DOCUMENT 13:  
IDENTIFYING AND REVIEWING EVIDENCE  
TO INFORM THE CONCEPTUALISATION  
AND POPULATION OF COST-EFFECTIVENESS MODELS**

REPORT BY THE DECISION SUPPORT UNIT

May 2011

Eva Kaltenthaler<sup>1</sup>, Paul Tappenden<sup>1</sup>, Suzy Paisley<sup>1</sup>, Hazel Squires<sup>1</sup>

<sup>1</sup>University of Sheffield, UK

Decision Support Unit, ScHARR, University of Sheffield, Regent Court, 30 Regent Street  
Sheffield, S1 4DA

Tel (+44) (0)114 222 0734

E-mail [dsuadmin@sheffield.ac.uk](mailto:dsuadmin@sheffield.ac.uk)

## **ABOUT THE DECISION SUPPORT UNIT**

The Decision Support Unit (DSU) is a collaboration between the Universities of Sheffield, York and Leicester. We also have members at the University of Bristol, London School of Hygiene and Tropical Medicine and Brunel University. The DSU is commissioned by The National Institute for Health and Clinical Excellence (NICE) to provide a research and training resource to support the Institute's Technology Appraisal Programme.

Please see our website for further information [www.nicedsu.org.uk](http://www.nicedsu.org.uk)

## **ABOUT THE TECHNICAL SUPPORT DOCUMENT SERIES**

The NICE Guide to the Methods of Technology Appraisal<sup>1</sup> is a regularly updated document that provides an overview of the key principles and methods of health technology assessment and appraisal for use in NICE appraisals. The Methods Guide does not provide detailed advice on how to implement and apply the methods it describes. This DSU series of Technical Support Documents (TSDs) is intended to complement the Methods Guide by providing detailed information on how to implement specific methods.

The TSDs provide a review of the current state of the art in each topic area, and make clear recommendations on the implementation of methods and reporting standards where it is appropriate to do so. They aim to provide assistance to all those involved in submitting or critiquing evidence as part of NICE Technology Appraisals, whether manufacturers, assessment groups or any other stakeholder type.

We recognise that there are areas of uncertainty, controversy and rapid development. It is our intention that such areas are indicated in the TSDs. All TSDs are extensively peer reviewed prior to publication (the names of peer reviewers appear in the acknowledgements for each document). Nevertheless, the responsibility for each TSD lies with the authors and we welcome any constructive feedback on the content or suggestions for further guides.

Please be aware that whilst the DSU is funded by NICE, these documents do not constitute formal NICE guidance or policy.

Dr Allan Wailoo

Director of DSU and TSD series editor.

---

<sup>1</sup> National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal, 2008 (updated June 2008), London.

## **Acknowledgements**

The authors would like to thank the ScHARR participants who made contributions to a focus group held on 25 January 2010 and a seminar held on 16 June 2010. The authors would also like to thank the parameters working group who attended a workshop on 7 February 2011 which included a series of related focus group sessions covering the components addressed in this technical brief:

Martin Hoyle, Peninsula College of Medicine and Dentistry

Rob Anderson, Peninsula College of Medicine and Dentistry

Adrian Bagust, University of Liverpool

John Brazier, University of Sheffield

Julie Glanville, University of York

Jaime Peters, Peninsula College of Medicine and Dentistry

Nerys Woolacott, University of York

Nicola Cooper, University of Leicester

Rodolfo Hernandez, University of Aberdeen

Ian Shemilt, University of East Anglia

We would also like to thank Alec Miners, Rob Anderson, Luke Vale, Chris Hyde, and Sue Golder for peer reviewing the draft manuscript.

The DSU thanks the team at NICE led by Jo Richardson for reviewing this document. The editor for the TSD series is Allan Wailoo.

### **This report should be referenced as follows:**

Kaltenthaler, E., Tappenden, P., Paisley, S., Squires, H. (2011) NICE DSU Technical Support Document 13: Identifying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models. Available from <http://www.nicedsu.org.uk>

## **EXECUTIVE SUMMARY**

### **Background**

This Technical Support Document (TSD) is concerned with methods for identifying and reviewing evidence to inform models developed to estimate the cost-effectiveness of health technologies, in particular model parameter estimates, in the NICE Technology Appraisal Process. The NICE Guide to the Methods of Appraisal states that for parameters relating to treatment effects a full systematic review is required, and that the process of assembling evidence for all parameters needs to be systematic, transparent and justifiable. Evidence must be identified, quality assessed and, when appropriate, pooled using explicit criteria and justifiable and reproducible methods. The Methods Guide goes on to state that evidence will typically be drawn from a number of different sources. These sources might include cohort studies for parameters relating to the natural history of the condition, randomised controlled trials for relative treatment effects, and cross-sectional surveys for resource use and costs. Methodological standards for reviewing the broad range of relevant evidence in the context of model development do not currently exist. The processes through which evidence is identified and selected remain largely unreported with the result that the process of using evidence within models lacks transparency. Guidance is needed regarding approaches for reviewing of evidence to inform the model development process and for informing parameter values in cost-effectiveness models.

### **Objectives**

The purpose of this document is to provide guidance and advice on what might constitute a systematic and transparent approach where there is not a requirement to use conventional systematic review methods but where little procedural guidance exists. The document covers four key objectives:

- To recommend model structuring methods including the identification and specification of relevant parameters within a model.
- To recommend appropriate methods for the systematic identification of evidence to inform model parameter estimates including relevant sources and types of searching.
- To recommend appropriate methods for the reviewing of evidence to inform model parameters in a systematic fashion including recommendations for critical appraisal and rapid review methods.

- To make recommendations on the reporting of the identification and reviewing of evidence including methods, decisions and judgements made during the process.

The guidance presented within this TSD is not intended to be prescriptive, rather the intention is to provide practical advice on how specific methods might be implemented. This support guide is set out across four sections: (1) model structuring; (2) identifying the evidence; (3) reviewing the evidence; and (4) recommendations for practice.

### **(1) Model structuring**

The role of evidence is not restricted to informing model parameters; rather it is closely linked with questions about what model parameters are considered relevant and how they should be characterised within the model. This section focuses on the conceptualisation of the decision problem and the process of model structuring. It highlights the nature of the choices that exist in developing the model structure, suggests a practical approach through which these judgements can be made explicit and highlights key issues associated with the role of evidence in the structuring elements of model development. Particular emphasis is devoted to the distinction between the development and use of problem-oriented and design-oriented conceptual models and the role of evidence in informing these.

### **(2) Identifying the evidence**

The model development process brings together diverse sources of evidence within a single analytical framework. It generates multiple complex information needs requiring different kinds of evidence drawn from various types of information resource. Evidence is used to inform all aspects of model development and information needs are not wholly pre-specified but emerge and are clarified in the process of model development. It can be difficult to apply search methods developed for the purpose of informing systematic reviews because systematic review search methods are designed to address single, focussed, predefined questions. This section considers the feasibility and applicability of systematic review search methods in the context of modelling. It highlights the main types of evidence used in modelling and suggests systematic search techniques aimed at maximising the rate of return of potentially relevant information.

### **(3) Reviewing the evidence**

This section 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 systematise and make explicit the choices involved in selecting evidence to inform models.

#### **(4) Recommendations for practice**

Recommendations regarding best practise are presented. It is hoped that these suggestions will help to improve the transparency and hence the quality of HTA work. There is a need for these recommendations to be piloted in order to determine which are most applicable within the limited time available for NICE technology appraisals.

# CONTENTS

<b>1. INTRODUCTION</b>	<b>11</b>
1.1 PURPOSE OF HIS TECHNICAL SUPPORT DOCUMENT	11
1.2 IDENTIFICATION AND REVIEWING OF EVIDENCE	11
<b>2. MODEL STRUCTURING</b>	<b>14</b>
2.1 PURPOSE OF THIS SECTION	14
2.2 INTRODUCTION - THE INTERPRETATION OF MATHEMATICAL MODELS	14
2.3 MODEL STRUCTURING ACTIVITY WITHIN THE BROADER MODEL DEVELOPMENT PROCESS	17
2.3.1 <i>Empirical research surrounding model development in HTA</i>	17
2.3.2 <i>Definition and purpose of conceptual modelling</i>	19
2.4 PRACTICAL APPROACHES TO CONCEPTUAL MODELLING IN HTA	22
2.4.1 <i>Problem-oriented conceptual modelling - disease process models</i>	22
2.4.2 <i>Problem-oriented conceptual modelling – service pathways models</i>	25
2.5 EVIDENCE SOURCES TO INFORM CONCEPTUAL MODELS	33
<b>3. IDENTIFYING AND SOURCING EVIDENCE</b>	<b>35</b>
3.1 PURPOSE OF THIS SECTION	35
3.2 BACKGROUND	36
3.3 SYSTEMATIC REVIEW SEARCH METHODS IN THE CONTEXT OF MODELLING	37
3.4 INFORMATION SOURCES TO INFORM MODELS	40
3.5 MAXIMISING THE RATE OF RETURN OF RELEVANT INFORMATION	43
3.6 DEFINING A SUFFICIENT SEARCH PROCESS	47
3.7 REPORTING OF THE SEARCH PROCESS	49
3.8 CONCLUSION	51
<b>4. REVIEWING EVIDENCE TO INFORM MODELS</b>	<b>53</b>
4.1 UNDERLYING PRINCIPLES	53
4.2 SELECTION AND PRIORITISATION OF PARAMETERS FOR REVIEWING EFFORT	53
4.3 REVIEWING METHODS	54
4.4 MINIMISING BIAS	55
4.5 HIERARCHY OF EVIDENCE SOURCES	56
4.6 ASSESSMENT OF EVIDENCE AND DATA EXTRACTION	61
4.7 SYNTHESIS AND ANALYSIS	62
4.8 REPORTING OF REVIEWING METHODS	63
4.9 PLANNING AND PROJECT MANAGEMENT	63

<b>5. RECOMMENDATIONS</b>	<b>65</b>
5.1 RECOMMENDATIONS FOR PRACTICE	65
5.2 RECOMMENDATIONS FOR FURTHER RESEARCH	67
<b>6. REFERENCES</b>	<b>69</b>

### **Figures, boxes and tables**

Figure 1: Types of evidence used to inform models	12
Figure 2: Stylised model development process <sup>12</sup>	18
Figure 3: A hierarchy of models	22
Figure 4: Illustrative disease process model	23
Figure 5: Illustrative service pathways model	26
Figure 6: Illustrative design-oriented conceptual model	30
Box 1: Main stages in the model development process (adapted from Chilcott et al <sup>12</sup> )	18
Box 2: The roles of conceptual modelling in HTA	20
Box 3: Recommendations for practice - problem-oriented models	28
Box 4: Recommendations for practice - design-oriented conceptual models	33
Box 5: Search techniques for maximising the rate of return of relevant information	44
Box 6: Overview of sources consulted in the development and population of the model	50
Box 7: Factors that impact on the approach to identifying evidence to inform the development and population of models	51
Box 8: Rapid review methods for model parameters	55
Box 9: Study selection	60
Box 10: Assessment of evidence	62
Box 11: Project management	65
Table 1: Roles and concerns regarding the use of evidence to inform alternative model structures	34
Table 2: Classification of evidence used in models of cost-effectiveness by type, source and use of evidence (Paisley, 2010 <sup>21</sup> )	41
Table 3: Example of how to report and audit a table of all sources of evidence cited in support of a model	51
Table 4: Hierarchies of data sources for health economics analyses <sup>23</sup>	57

## ABBREVIATIONS

BNF	British National Formulary
CRD	Centre for Reviews and Dissemination
HTA	Health Technology Assessment
NICE	National Institute for Health and Clinical Excellence
ONS	Office for National Statistics
QALY	Quality Adjusted Life Year
RCT	Randomised Controlled Trial
TSD	Technical Support Document

## GLOSSARY

<b>Conceptual modelling</b>	Activity related to translating the understanding of the decision problem towards a mathematical model-based solution. For the sake of this TSD, this is defined as <i>“the abstraction and representation of complex phenomena of interest in some readily expressible form, such that individual stakeholders’ understanding of the parts of the actual system, and the mathematical representation of that system, may be shared, questioned, tested and ultimately agreed.”</i>
<b>Design-oriented conceptual models</b>	Conceptual models which are focussed on the consideration of alternative potentially acceptable and feasible quantitative model designs, to specify the model’s evidence requirements and to provide a basis for comparison and justification against the final implemented model.
<b>High yield patch</b>	Any source containing a high proportion of relevant information (e.g. existing model, clinician, systematic review)
<b>Implementation/implemented model</b>	A quantitative mathematical or simulation model implemented within some software program.
<b>Indirect retrieval</b>	Retrieval of relevant information on one topic whilst searching for another topic.
<b>Information gathering</b>	Piecemeal, non-linear or non-sequential retrieval of relevant information.

<b>Information scent</b>	Perception that one source of information will lead to further potentially relevant sources of information (e.g. following up further sources by a known, relevant author)
<b>Information seeking process</b>	Any process by which information is identified or incorporated in the model development process. This may include bibliographic database searching or obtaining advice from clinicians etc.
<b>Lumping and splitting</b>	In this context, defining a broad search query in order to retrieve information relating to more focussed sub-questions.
<b>Problem-oriented conceptual models</b>	A form of conceptual model which is developed to understand the decision problem and the system in which that problem exists.
<b>Proximal cue</b>	A source of information which prompts a search for other similar or related potentially relevant information
<b>Reference source</b>	Source of information accepted or used on the grounds of its authority in the context of the decision-making (e.g. drug formulary, classification of disease, clinical guideline)
<b>Routine data source</b>	Source of information compiled primarily for administrative rather than research purposes (e.g. prescribing rates)
<b>Search</b>	Any systematic process by which information is identified and incorporated in the model development process
<b>Search filter</b>	A predefined search strategy aimed at restricting search results to studies with specific methodological or clinical characteristics
<b>Secondary retrieval</b>	See <i>indirect retrieval</i>
<b>Systematic</b>	Any form of organised approach for identifying and reviewing evidence, but not necessarily adhering to established systematic review methods and processes
<b>Systematic review</b>	A review undertaken using established, conventional systematic review methods such as those associated with the Cochrane Collaboration and the Centre for Reviews and Dissemination.
<b>Triangulation</b>	In this context, searching for information across different types of sources with the aim of capturing the level of consistency or inconsistency across the breadth of different types of source (e.g. consulting both research and non-research based sources for the same information).

## **1. INTRODUCTION**

### **1.1 PURPOSE OF HIS TECHNICAL SUPPORT DOCUMENT**

The NICE Guide to the Methods of Technology Appraisal<sup>1</sup> describes key aspects of analyses submitted to the NICE Technology Appraisal Programme. This Technical Support Document (TSD) is part of a wider initiative to produce a series of TSDs that accompany the NICE Methods Guide. Each TSD describes how to use analytical techniques recommended in the NICE Methods Guide, offers suggestions for analyses for areas not currently covered in the NICE Methods Guide and identifies areas that would benefit from further methodological research. This TSD is concerned with methods for identifying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models developed to inform the NICE Technology Appraisal Programme. This document is not intended to impose restrictions on current practice but instead offers suggestions for optimal development of cost-effectiveness models.

### **1.2 IDENTIFICATION AND REVIEWING OF EVIDENCE**

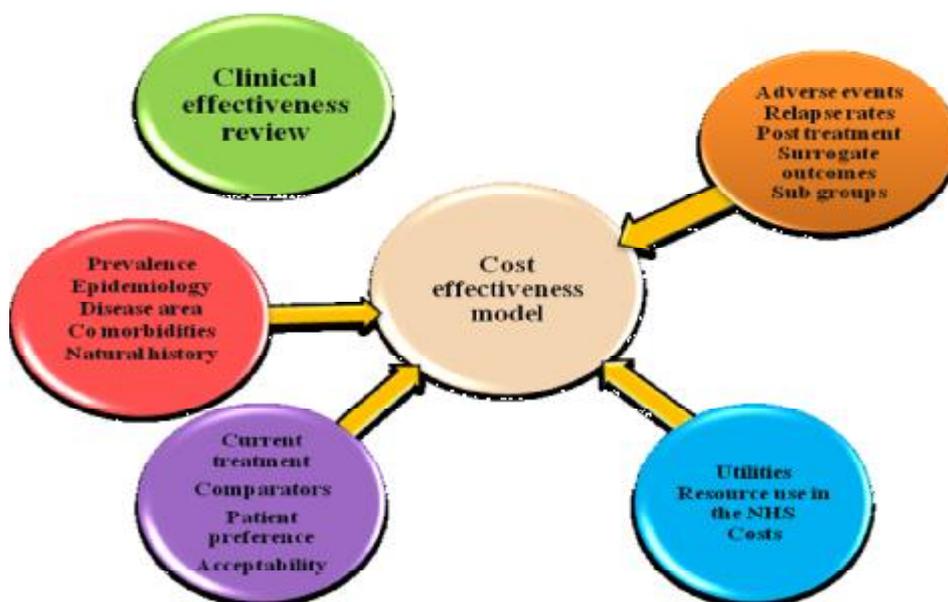
The NICE Methods Guide (Section 5.1.2) states that the process of assembling evidence needs to be systematic. Evidence must be identified; quality assessed and, when appropriate, pooled using explicit criteria and justifiable and reproducible methods. The Methods Guide focuses on the identification and assessment of evidence to inform model parameters but also points to the need to justify, with evidence, all aspects of the modelling process including structural assumptions and the extrapolation of treatment effects.

The Methods Guide (Section 3.3.4) states that for all parameters (including effectiveness, valuation of health-related quality of life [HRQoL] and costs) a systematic consideration of possible data sources is required and that the selection of sources to justify a particular outcome must be avoided. For parameters relating to treatment effects a full systematic review is required. For other types of parameter this requirement is not specified although the need for a systematic, transparent and justifiable approach is stated. The Methods Guide (Section 5.1.2) goes on to state that these principles apply to all categories of evidence that are used to estimate clinical effectiveness and cost-effectiveness, and that this evidence will typically be drawn from a number of different sources. These sources might include cohort studies for parameters relating to the natural history of the condition, randomised controlled trials (RCTs) for relative treatment effects, and cross-sectional surveys for resource use and

costs. It has long been recognised that when assembling the evidence it is essential to consider how bias can be minimised, especially when non-randomised studies are included.<sup>2</sup>

Whilst a number of issues surrounding reviewing evidence for models have been highlighted in detail<sup>3-8</sup> there is very little formal guidance with regard to best practice in this area. Cooper *et al.* highlight a number of concerns in decision models developed as part of the HTA process in the UK; these include (1) inadequate reporting on the identification of evidence, (2) a lack of justification as to how evidence was selected for use in the model, (3) a lack of quality assessment of the evidence used, and (4) a lack of transparency concerning the assumptions underpinning the model structure.<sup>3</sup> They state that this is particularly unclear for the sources of evidence used to inform adverse events, complications and resource use. The consequence of these concerns is that the process of using evidence in cost-effectiveness models lacks transparency. The quality of data inputs can directly impact on the reliability of model results. Philips *et al.*<sup>4</sup> highlight the lack of good practice guidelines on how to identify evidence for models and Paisley<sup>8</sup> argues that the role of evidence in informing not just the population of the model but the whole of the model development process. In very broad terms, Figure 1 sets out some of the more common types of information required to inform the development and population of model.

**Figure 1: Types of evidence used to inform models**



Decisions concerning the choice of evidence to inform model parameter estimates are inextricably linked with the process of model development. How a model is developed impacts on what parameter values are needed. Once these are decided, it is necessary to identify possible values, review the available evidence, make choices and report the whole process as transparently as possible. This document addresses the following questions:

1. Section 1 - What methods for model structuring can be recommended? This includes the identification, specification and prioritisation of relevant parameters. How can/should one determine which parameters are relevant for inclusion in a model?
2. How should evidence be systematically identified to inform the process of model development and, in particular, model parameter estimates? This includes consideration of the applicability of systematic review search methods to the modelling process and includes recommendations on relevant sources and on focussed searching.
3. What guidance can be provided about the methods to use for reviewing model parameter data in a systematic fashion, and how does this differ from conventional reviews of clinical effectiveness? This includes recommendations for critical appraisal and rapid review methods.
4. What recommendations concerning reporting can be made? This includes the reporting of the model structure and evidence used to inform it, the review methods and the decisions and judgments made regarding the identification and selection of evidence.

The document is set out as follows. Section 2 explores issues surrounding model conceptualisation and structuring and how this process relates to identifying evidence needs. Section 3 covers the sourcing of evidence for models. Section 4 discusses issues surrounding reviewing evidence. Finally, Section 5 presents recommendations for practice.

## **2. MODEL STRUCTURING**

### **2.1 PURPOSE OF THIS SECTION**

The purpose of this section is not to rigidly prescribe how model development decisions should be made, nor is it intended to represent a comprehensive guide of “how to model.” The former would undoubtedly fail to reflect the unique characteristics of each individual technology appraisal and could discourage the development of new and innovative modelling methods, whilst the latter would inevitably fail to reflect the sheer breadth of decisions required during model development. Rather, the purposes of this section are threefold:

- (1) To highlight that structural model development choices invariably exist;
- (2) To suggest a generalisable and practical approach through which these alternative choices can be exposed, considered and assessed, and;
- (3) To highlight key issues associated with the role of evidence in the structuring elements of model development.

### **2.2 INTRODUCTION - THE INTERPRETATION OF MATHEMATICAL MODELS**

A mathematical model is a “*representation of the real world... characterised by the use of mathematics to represent the parts of the real world that are of interest and the relationships between those parts.*”<sup>9</sup> The roles of mathematical modelling are numerous, including extending results from a single trial, combining multiple sources of evidence, translating from surrogate/intermediate endpoints to final outcomes, generalising results from one context to another, informing research planning and design, and characterising and representing decision uncertainty given existing information.<sup>10</sup> At a broad level, mathematical or simulation models in Health Technology Assessment (HTA) are generally used to simulate the natural history of a disease and the impact of particular health technologies upon that natural history in order to estimate incremental costs, health outcomes and cost-effectiveness. All mathematical models require evidence to inform their parameters. Such evidence may include information concerning disease natural history or baseline risk of certain clinical events, epidemiology, resource use and service utilisation, compliance/participation patterns, costs, health-related quality of life (HRQoL), survival and other time-to-event outcomes, relative treatment effects and relationships between intermediate and final endpoints (see Figure 1). However, the role of evidence is *not* restricted to informing model parameters. Rather, it is closely intertwined with questions about which model parameters should be considered relevant in the first place and how these parameters should be characterised. The

consideration of how best to identify and use evidence to inform a particular model parameter thus firstly requires an explicit decision that the parameter in question is “relevant”, the specification or definition of that parameter, and some judgement concerning its relationship to other “relevant” parameters included in the model. This often complex and iterative activity is central to the process of model development and can be characterised as a series of decisions concerning (a) what should be included in the model, (b) what should be excluded, and (c) how those phenomena that are included should be conceptually and mathematically represented.

The need for these types of decisions during model development is unavoidable, rather it is a fundamental characteristic of the process itself. Whilst this activity already takes place in model development, it is often unclear how this process has been undertaken and how this may have influenced the final implemented model. In practice, the reporting of model structures tends to be very limited<sup>11</sup> and, if present, usually focuses only on the final model that has been implemented. In such instances, the reader may be left with little idea about whether or why the selected model structure should be considered credible, which evidence has been used to inform its structure, why certain abstractions, simplifications and omissions have been made, why certain parameters were selected for inclusion (and why others have been excluded), and why the included parameters have been defined in a particular way. This lack of systematicity and transparency ultimately means that judgements concerning the credibility of the model in question may be difficult to make. In order to produce practically useful guidance concerning the use of evidence in models, it is firstly important to be clear about the interpretation of abstraction, bias and credibility in the model development process.<sup>†</sup>

#### *i) Credibility of models*

A model cannot include every possible relevant phenomena; if it could it would no longer be a model but would instead be the real world. The value of simplification and abstraction within models is the ability to examine phenomena which are complex, unmanageable or otherwise unobservable in the real world. As a direct consequence of this need for

---

<sup>†</sup> *This philosophical view of the nature and role of mathematical models mirrors the general position of Subtle Realism (see Hammersley M. Ethnography and realism. In: Huberman AM, Miles MB, editors. The qualitative researchers companion. California: Sage; 2002)*

simplification, all models will be, to some degree, wrong. The key question is not whether the model is “correct” but rather whether it can be considered to be useful for informing the decision problem at hand. This usefulness is directly dependent upon the credibility of the model’s results, which is, in turn, hinged upon the credibility of the model from which those results are drawn. Owing to the inevitability of simplification and abstraction within models, there is no single “perfect” or “optimal” model. There may however exist one or more “acceptable” models; even what is perceived to be the “best” model could *always* be subjected to some degree of incremental improvement (and indeed the nature of what constitutes an improvement requires some subjective judgement). The credibility of potentially acceptable models can be assessed and differing levels of confidence can be attributed to their results on the basis of such judgements. The level of confidence given to the credibility of a particular model may be determined retrospectively – through considerations of structural and methodological uncertainty *ex post facto*, or prospectively – through the *a priori* consideration of the process through which decisions are made concerning the conceptualisation, structuring and implementation of the model.

#### *ii) Defining relevance in models*

The purpose of models is to represent reality, not to reproduce it. The process of model development involves efforts to reflect those parts of reality that are considered relevant to the decision problem. Judgements concerning relevance may differ between different modellers attempting to represent the same part of reality. The question of “*what is relevant?*” to a particular decision problem should not be judged solely by the individual developing the model; rather making such decisions should be considered as a joint task between modellers, decision-makers, health professionals and other stakeholders who impact upon or are impacted upon by the decision problem under consideration. Failure to reflect conflicting views between alternative stakeholders may lead to the development of models which represent a contextually naïve and uninformed basis for decision-making.

#### *iii) The role of clinical input*

Clinical opinion is essential in understanding the relevant facets of the system in which the decision problem exists. This clinical opinion is not only relevant, but essential, because it is sourced from individuals who interact with this system in a way that a modeller cannot. This information forms the cornerstone of a model’s contextual relevance. However, it is important to recognise that health professionals cannot fully detach themselves from the

system in which they practise; their views of a particular decision problem may be to some degree influenced by evidence they have consulted, their geographical location, local enthusiasms, their experience and expertise, together with a wealth of other factors. Understanding why the views of stakeholders differ from one another is important, especially with respect to highlighting geographical variations. As such, the use of clinical input in informing models and model structures brings with it the potential for bias. Bias may also be sourced from the modeller themselves as a result of their expertise, their previous knowledge of the system in which the current decision problem, and the time and resource available for model development. Where possible, potential biases should be brought to light to inform judgements about a model's credibility.

## **2.3 MODEL STRUCTURING ACTIVITY WITHIN THE BROADER MODEL**

### **DEVELOPMENT PROCESS**

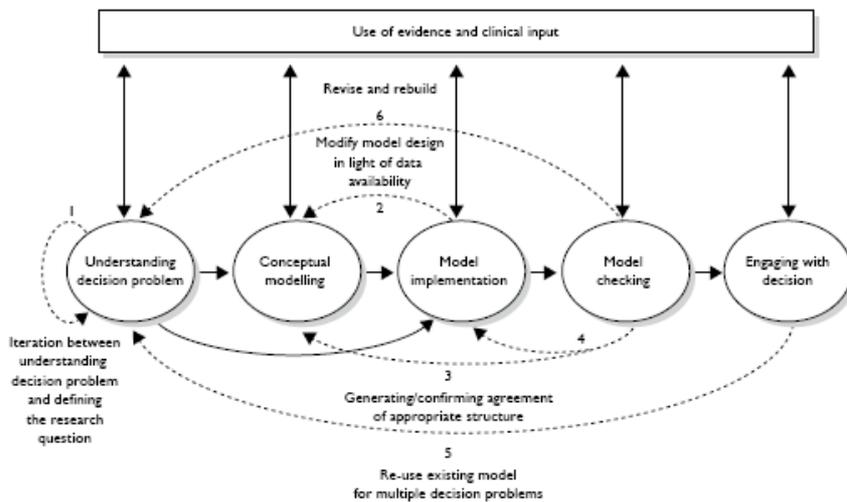
#### *2.3.1 Empirical research surrounding model development in HTA*

To date, there has been little empirical study concerning the development of HTA models in practice. A recent qualitative research study was undertaken to examine techniques and procedures for the avoidance and identification of errors in HTA models.<sup>12</sup> Interviewees included modellers working within Assessment Groups involved in supporting NICE's Technology Appraisal Programme as well as those working for outcomes research groups involved in preparing submissions to NICE on behalf of pharmaceutical companies. A central aspect of these interviews involved the elicitation of a personal interpretation of how each interviewee develops models. These descriptions were synthesised to produce a stylised model development process, comprising five broad bundles of activities (see Box 1 and Figure 2).

**Box 1: Main stages in the model development process (adapted from Chilcott et al<sup>12</sup>)**

- 1. Understanding the decision problem:** Activities including immersion in research evidence, defining the research question, engaging with clinicians, decision-makers and methodologists, and understanding what is feasible.
- 2. Conceptual modelling:** Activity related to translating the understanding of the decision problem towards a mathematical model-based solution.<sup>8</sup>
- 3. Model implementation:** Implementation of the model within a software platform.
- 4. Model checking:** Activity to avoid and identify model errors. This includes engaging with experts, checking face validity, testing values, structure and logic, checking data sources etc.
- 5. Engaging with decision:** Model reporting and use by the decision-maker(s).

**Figure 2: Stylised model development process<sup>12</sup>**



One particular area of variability between interviewees concerned their approaches to conceptual model development. During the interviews, respondents discussed the use of several approaches to conceptual modelling including documenting proposed model structures, developing mock-up models in Microsoft Excel, developing sketches of potential structures, and producing written interpretations of evidence. For several respondents, the model development process did not involve any explicit conceptual modelling activity; in these instances, the conceptual model and implementation model were both developed in parallel. This is an important distinction to make with respect to model credibility and

validation (see Section 2.3.2 below) and the processes through which evidence is identified and used to inform the final implemented model.

### 2.3.2 Definition and purpose of conceptual modelling

Whilst others have recognised the importance of conceptual modelling as a central element of the model development process, it has been noted that this aspect of model development is probably the most difficult to undertake and least well understood.<sup>13,14</sup> Part of the problem stems from inconsistencies in the definition and the role(s) of conceptual modelling, and more general disagreements concerning how such activity should be used to support and inform implementation modelling. The definition and characteristics of conceptual modelling are dependent on the perceived purposes of the activity; existing literature around this subject has been reviewed in detail elsewhere.<sup>13</sup> For the purpose of this document, conceptual modelling is taken as: “*the abstraction and representation of complex phenomena of interest in some readily expressible form, such that individual stakeholders’ understanding of the parts of the actual system, and the mathematical representation of that system, may be shared, questioned, tested and ultimately agreed.*”

Whilst there is inevitable overlap associated with processes for understanding the decision problem to be addressed, conceptual modelling is distinguishable from these activities in that it is targeted at producing *tangible* outputs in the form of one or more conceptual models. In the context of HTA, conceptual model development may be used to achieve a number of ends, as shown in see Box 2. Broadly speaking, these roles fall into two groups: (1) those associated with developing, sharing and testing one’s understanding of the decision problem and the system in which this exists, and (2) those associated with designing, specifying and justifying the model and its structure. Therefore it seems sensible to distinguish between *problem-oriented* conceptual models and *design-oriented* conceptual models; this distinction has been made elsewhere outside of the field of HTA.<sup>15</sup> The characteristics of these alternative types of conceptual model are briefly detailed below. Both of these types of model may be useful approaches for informing the relevant characteristics of the final implemented model.

## Box 2: The roles of conceptual modelling in HTA

### *Problem-oriented conceptual models*

- To ensure that health professionals understand how the model will capture the impact of the interventions under consideration on costs and health outcomes
- To ensure that the proposed model will be clinically relevant - that all relevant events, resources, costs and health outcomes have been included and that these reflect current knowledge of disease and treatment systems
- To ensure that the proposed model will meet the needs of the decision-maker
- To provide a reference point during model implementation
- To highlight uncertainty and variation between healthcare practitioners

### *Design-oriented conceptual models*

- To provide a common understanding amongst those involved in model development regarding model evidence requirements prior to model implementation
- To provide an explicit platform for considering and debating alternative model structures and other model development decisions prior to implementation (including the *a priori* consideration of structural uncertainties)
- To provide a reference point during model implementation
- To provide the conceptual basis for reporting the methods and assumptions employed within the final implemented model
- To provide a basis for comparison and justification of simplifications and abstractions during model development

**Problem-oriented conceptual models:** This form of conceptual model is developed to understand the decision problem and the system in which that problem exists. The focus of this model form concerns fostering communication and understanding between those parties involved in informing, developing, and using the model. In health economic evaluation, this type of conceptual model is primarily concerned with developing and agreeing a description of the disease and treatment systems: (a) to describe the current clinical understanding of the relevant characteristics of the disease process(es) under consideration and important events therein, and; (b) to describe the clinical pathways through which patients with the disease(s)

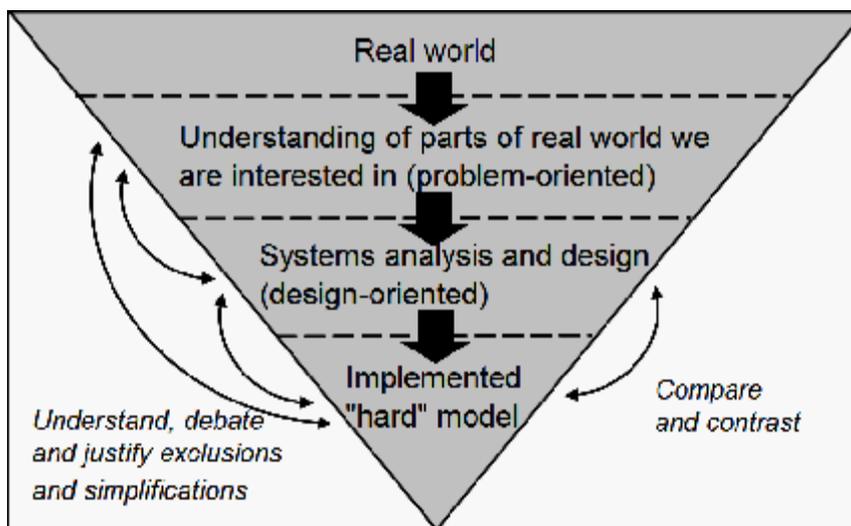
are detected, diagnosed, treated and followed-up. This type of conceptual model is therefore solely concerned with unearthing the complexity of the decision problem and the system in which it exists; its role is *not* to make assertions about how those relevant aspects of the system should be mathematically represented. The definition of “*what is relevant?*” for this type of conceptual model is thus primarily dependent on expert input rather than the availability of empirical research evidence. In this sense, this type of conceptual model is a problem-led method of enquiry.

**Design-oriented conceptual models:** This form of conceptual model is focussed on the consideration of alternative potentially acceptable and feasible quantitative model designs, to specify the model’s anticipated evidence requirements, and to provide a basis for comparison and justification against the final implemented model. In order to achieve these ends, it draws together the problem-oriented conceptual views of relevant disease and treatment processes and interactions between the two. The design-oriented conceptual model sets out a clear boundary around the model system, defines its breadth (how far down the model will simulate certain pathways for particular patients and subgroups) and sets out the level of depth or detail within each part of the model. It therefore represents a platform for identifying and thinking through potentially feasible and credible model development choices prior to actual implementation. Within this context, the definition of “*what is relevant?*” is guided by the problem-oriented models and therefore remains problem-led, but is mediated by the question of “*what is feasible?*” given the availability of existing evidence and model development resources (available time, money, expertise etc.).

Conceptual modelling activity, however defined, is directly related to model credibility and validation.<sup>16</sup> The absence of an explicit conceptual model means that a specific point of model validation is lost. As a model cannot include everything, an implemented model is inevitably a subset of the system described by the conceptual model. This hierarchical separation allows simplifications and abstractions represented in the implemented model to be compared against its conceptual counterpart, thereby allowing for debate and justification.<sup>12</sup> However, in order to make such comparisons, conceptual model development must be overt: the absence or incomplete specification of a conceptual model leads to the breakdown of concepts of model validation and verification. Without first identifying and considering the alternative choices available, it is impossible to justify the appropriateness of any particular model. The benefit of separating out conceptual modelling activity into distinct

problem-oriented and design-oriented components is that this allows the modeller (and other stakeholders) to firstly understand the complexities of the system the model intends to represent, and then to examine the extent to which the simplifications and abstractions resulting from alternative “hard” model structures will deviate from this initial view of the system. Figure 3 shows the relationship between the real world, the conceptual models and the implemented model.

**Figure 3: A hierarchy of models**



## 2.4 PRACTICAL APPROACHES TO CONCEPTUAL MODELLING IN HTA

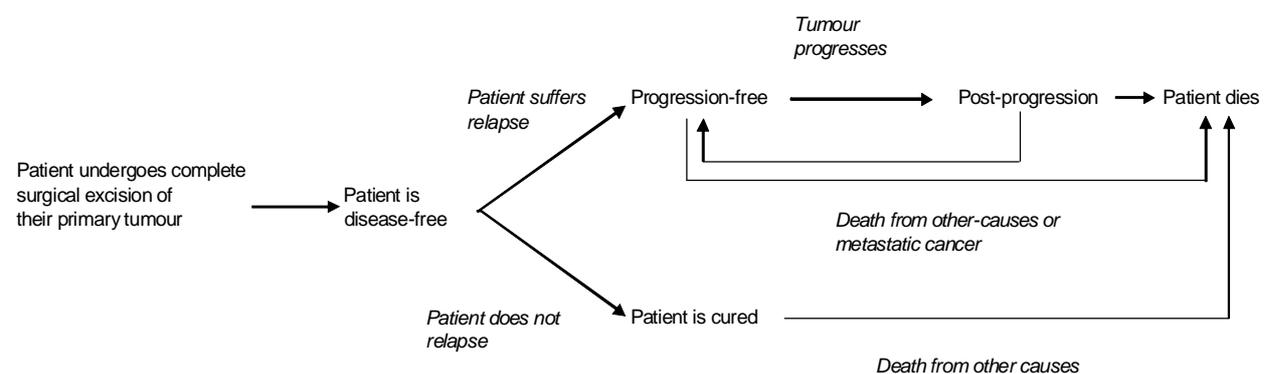
This section suggests how conceptual modelling could be undertaken and which elements of model development activity should be reported. Practical considerations surrounding conceptual model development are detailed below with reference to a purposefully simple model to assess the cost-effectiveness of adjuvant treatments for a hypothetical cancer area. It should be noted that the illustrative model is only intended to suggest how the alternative conceptual models forms may be presented and used. The problem-oriented model is divided into two separate conceptual model views; a disease process model and a service pathways model.

### 2.4.1 Problem-oriented conceptual modelling - disease process models

Figure 4 presents a simple example of a conceptual disease process model for the hypothetical decision problem. The focus of this type of model is principally on relevant disease events and processes rather than on the treatments received. At each point in the pathway, the focus should therefore relate to an individual patient’s true underlying state

rather than what is known by healthcare professionals at a particular point in time. It should be reiterated that this type of conceptual model does not impose or imply any particular decision concerning modelling methodology or outcome measure; it is solely a means of describing the relevant clinical events and processes within the system of interest. Furthermore, much of the information required to develop this type of model is already required for manufacturer’s submissions and assessment reports produced to inform NICE’s Technology Appraisals. It should also be noted that such conceptual models should be accompanied by textual descriptions to support their interpretation.

**Figure 4: Illustrative disease process model**



The following non-exhaustive set of issues and considerations may be useful when developing and reporting this type of problem-oriented conceptual model:

***Inclusion/exclusion of disease-related events***

- What are the main relevant events from a clinical/patient perspective? Does the conceptual model include explicit reference to all clinically meaningful events? For example, could a patient experience local relapse? Or could the intervention affect other diseases (e.g. late secondary malignancy resulting from radiation therapy used to treat the primary tumour)?
- Can these relevant events be discretised into a series of mutually exclusive biologically plausible health states? Does this make the process easier to explain?
  - If so, which metric would be clinically meaningful or most clinically appropriate? Which discrete states would be clinically meaningful? How do clinicians think about the disease process? How do patients progress between these states or sequences of events?
  - If not, how could the patient’s preclinical trajectory be defined?

- Do alternative staging classifications exist, and if so can/should they be presented simultaneously?
- Are all relevant competing risks (e.g. relapse or death) considered?
- For models of screening or diagnostic interventions, should the same metric used to describe preclinical and post-diagnostic disease states?
- Is the breadth of the conceptual model complete? Does the model represent all relevant states or possible sequences of events over the relevant patient subgroups lifetime?
- What are the causes of death? When can a patient die from these particular causes? Can patients be cured? If so, when might this happen and for which states does this apply? What is the prognosis for individuals who are cured?

***Impact of the disease on HRQoL and other outcomes***

- Is there a relationship between states, events and HRQoL? Which events are expected to impact upon a patient's HRQoL?
- Does the description of the disease process capture separate states in which a patient's HRQoL is likely to be different?
- Does the description of the disease process capture different states for prognosis?

***Representation of different-risk subgroups***

- Is it clear which competing events are relevant for particular subgroups?
- Does the description of the disease process represent a single patient group or should it discriminate between different subgroups of patients?
- Are these states/events likely to differ by patient subgroup?

***Impact of the technology on the conceptualised disease process***

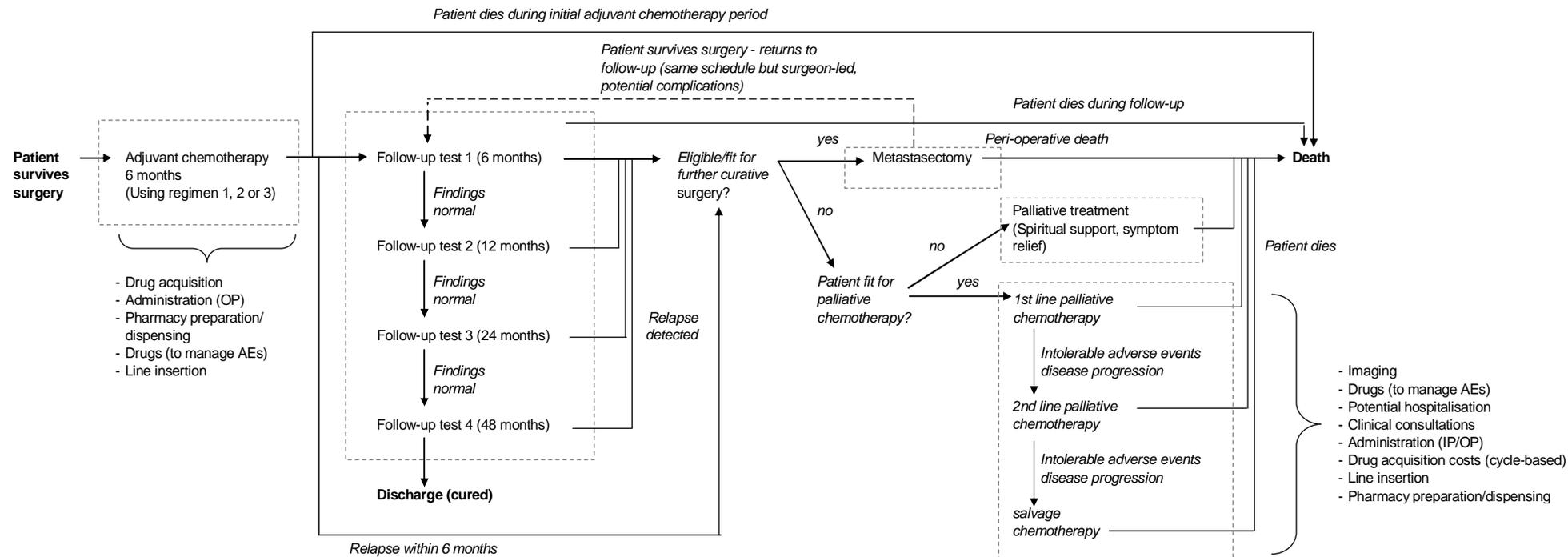
- Have all competing technologies relevant to the decision problem been identified?
- Can the conceptual model be used to explain the impact(s) of the technology or technologies under assessment? Do all technologies under consideration impact upon the same set of outcomes in the same way?
- Are there competing theories concerning the impact(s) of the technology upon the disease process? Can these be explained using the conceptual model?

- Does the use of the health technology result in any other impacts upon health outcomes that cannot be explained using the conceptual disease process model?

#### *2.4.2 Problem-oriented conceptual modelling – service pathways models*

Figure 5 presents an example of a conceptual model of service pathways for the example decision problem. In contrast to the disease process model, the focus of the service pathways model is principally concerned with the treatments received based upon what is known or believed by healthcare practitioners at any given point in time. Again, such conceptual models should be accompanied by textual description to ensure clarity in their interpretation.

**Figure 5: Illustrative service pathways model**



*IP – inpatient; OP - outpatient*

The following issues and considerations may be useful when developing and reporting this type of conceptual model:

***Relationship between risk factors, prognosis and service pathways***

- Is it clear where and how patients enter the service? Is it clear where patients leave the service (either through discharge or death)?
- Does the model make clear which patients follow particular routes through the service?
- Are any service changes occurring upstream in the disease service which may influence the casemix of patients at the point of model entry? E.g. if surgical techniques were subject to quality improvement might this change patient prognosis further downstream in the pathway?
- Does the model highlight the potential adverse events resulting from the use of particular interventions throughout the pathway? What are these? Do they apply to all competing technologies under consideration?
- Are there any potential feedback loops within the system (e.g. resection→follow-up→relapse→re-resection→follow-up)?
- Which patients receive active treatment and which receive supportive care alone? What information is used to determine this clinical decision (e.g. fitness, patient choice)?

***Distinction between what is true and what is known***

- How does the pathway change upon detection of the relevant clinical events, as defined in the conceptual disease process model? For example, at what point may relapse be detected?
- Is the occurrence of certain events likely to be subject to interval censoring?

***Geographical variations***

- How do the service pathways represented in the model likely to vary by geographical location or local enthusiasms? What are these differences and which parts of the pathway are likely to be affected most?

### *Nature of resource use*

- What are the relevant resource components across the pathway and what is the nature of resource use at each point of intervention? E.g. routine follow-up dependent on relapse status, once-only surgery (except for certain relapsing patients), cycle-based chemotherapy, doses dependent on certain characteristics, dose-limited radiation treatment etc.
- Does the conceptual service pathways model include all relevant resource components?
- Which resources are expected to be the key drivers of costs?

### *Impact of the technology on the service pathway*

- Which elements of the conceptual model will the intervention under assessment impact upon? E.g. different costs of adjuvant treatment, different mean time in follow-up, different numbers of patients experiencing metastatic relapse? What are expected to be the key drivers of costs?

Box 3 presents recommendations for developing and reporting problem-oriented conceptual models.

#### **Box 3: Recommendations for practice - problem-oriented models**

- (1) Develop the structure of the problem-oriented conceptual model using clinical guidelines and input from health professionals
- (2) Use other health professionals not involved in model development to provide peer review and to check understanding of the conceptual models
- (3) The precise graphical approach for presenting the conceptual models is important only in that they should be easily understood by health professionals and other decision stakeholders
- (4) For the sake of clarity, it may be beneficial to present the model in both diagrammatic and textual forms using non-technical, non-mathematical language
- (5) Develop the problem-oriented models before developing the design-oriented model. The feasibility and acceptability of the design-oriented conceptual model should have no bearing on the adequacy of the problem-oriented conceptual models.

### *Practical considerations – design-oriented conceptual models*

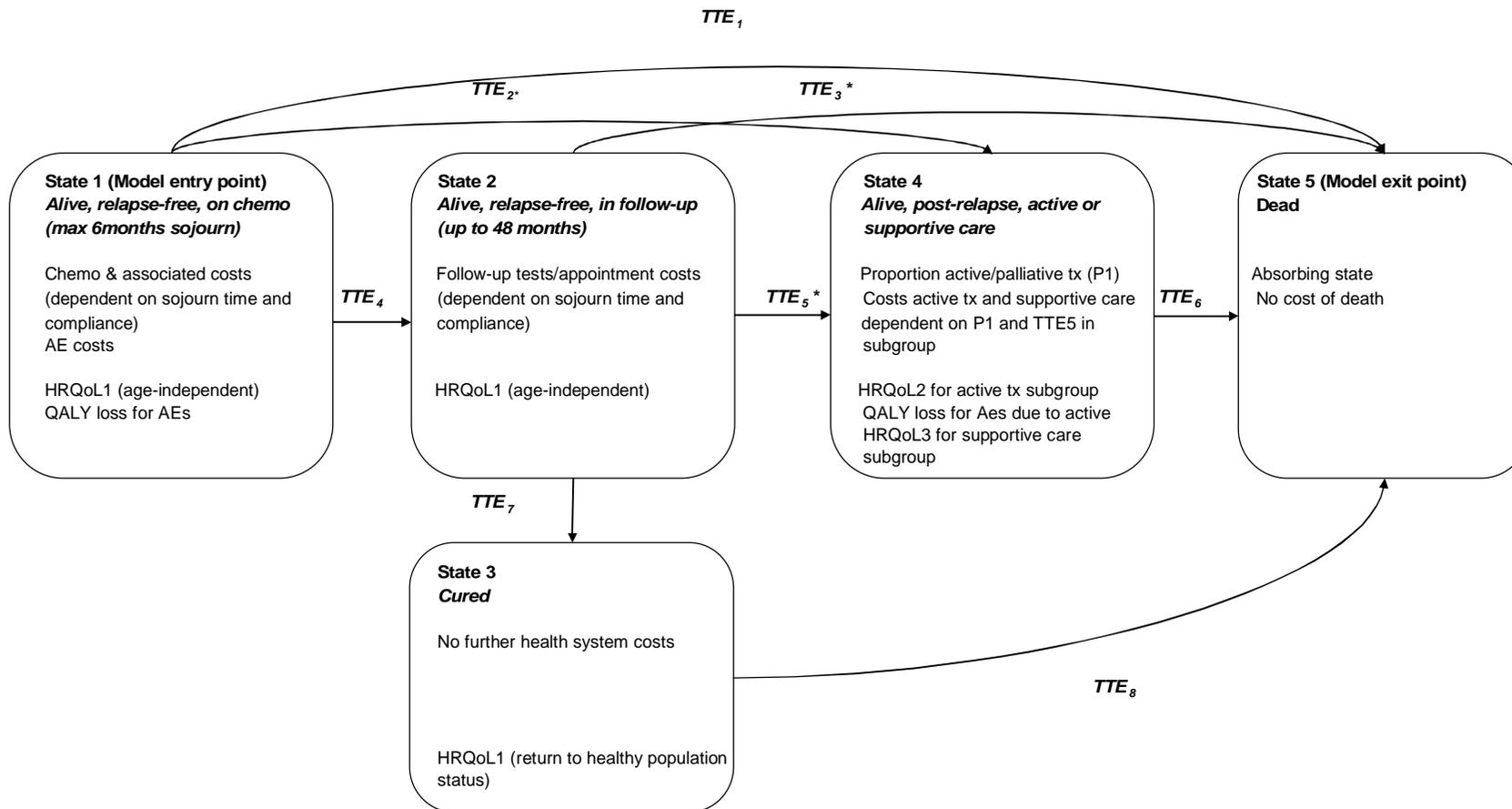
Figure 6 presents an example of a design-oriented conceptual model for the hypothetical decision problem (again, note that this is not intended to represent the “ideal” model but merely illustrates the general approach). This type of model draws together the problem-oriented model views with the intention of providing a platform for considering and agreeing structural model development decisions. By following this general conceptual approach it should be possible to identify the anticipated evidence requirements for the model at an early stage in model development.

Anticipated evidence requirements to populate the proposed illustrative model are likely to include the following types of information:

- Time-to-event data to describe sojourn time/event rates and competing risks in States 1-4 for the current standard treatment
- Relative effect estimates for the intervention(s) versus comparator (e.g. hazard ratios or independent hazards time-to-event data)
- Information relating to survival following cure
- HRQoL utilities for cancer and cured states
- Estimates of QALY losses or utility decrements and duration data for adverse events
- Information concerning the probability that a relapsed patient undergoes active/palliative treatment
- Survival and other time-to-event outcomes for relapsed patients
- Resource use and costs associated with:
  - Chemotherapy (drug acquisition, administration, pharmacy/dispensing, drugs to manage adverse events, line insertion)
  - Resource use and unit costs for follow-up
  - Supportive care following relapse
  - Active treatments following relapse

It may be helpful to consider the following issues when developing design-oriented conceptual models.

**Figure 6: Illustrative design-oriented conceptual model**



*TTE = time to event; AE = adverse event*

### ***Anticipated evidence requirements***

- What clinical evidence is likely to be available through which to simulate the impact of the new intervention(s)? How should these parameters be defined and what alternatives are available? Should independent or proportional hazards be assumed?
- Are all relevant interventions and comparators compared within the same trial? If not, is it possible for outcomes from multiple trials to be synthesised? How will this be done?
- What evidence is required to characterise adverse events within the model? What choices are available?
- Beyond the baseline and comparative effectiveness data relating to the technology itself, what other outcomes data will be required to populate the downstream portions of the model (e.g. progression-free survival and overall survival by treatment type for relapsed patients, survival duration in cured patients)?
- Will any intermediate-final relationships be modelled? What external evidence is there to support such relationships? What are the uncertainties associated with this approach and how might these be reflected in the model?
- Which descriptions of HRQoL states are possible and how will these parameters be incorporated into the final model?
- Will all model parameters be directly informed by evidence or will calibration methods (e.g. Markov Chain Monte Carlo) be required? Which calibration methods will be used and why should these be considered optimal or appropriate?
- What pre-model analysis will be required to populate the model? Which parameters are likely to require this?

### ***Modelling clinical outcomes***

- Which outcomes are needed by the decision-maker and how will they be estimated by the model?
- How/should trial evidence be extrapolated over time?
- If final outcomes are not reported within the trials, what evidence is available concerning the relationship between intermediate and final outcomes? How might this information be used to inform the analysis of available evidence?
- How will the impact(s) of treatment be simulated? How will this directly/indirectly influence costs and health outcomes? What alternative choices are available?

### ***Modelling approach***

- Which methodological approach (e.g. state transition, patient-level simulation) is likely to be most appropriate? Why?
- Is the proposed modelling approach feasible given available resources for model development?
- How does the approach influence the way in which certain parameters are defined? What alternatives are available (e.g. time-to-event rates or probabilities)?
- Does the proposed modelling approach influence the level of depth possible within certain parts of the model?

### ***Adherence to a health economic reference case***

- Will the proposed model meet the criteria of the reference case specific to the decision-making jurisdiction in which the model will be used? If not, why should the anticipated deviations be considered appropriate?

### ***Simplifications and abstractions***

- Have any relevant events, costs or outcomes been purposefully omitted from the proposed model structure? Why? For what reason(s) may these omissions be considered appropriate?
- Are there any parts of the disease or treatment pathways that have been excluded altogether? Why?
- What is the expected impact of such exclusion/simplification decisions? Why?
- What are the key structural simplifications? How does the design-oriented model structure differ from the problem-oriented conceptual models? Why should these deviations be considered appropriate or necessary? What is the expected direction and impact of these exclusions on the model results?

#### **Box 4: Recommendations for practice - design-oriented conceptual models**

- The design-oriented conceptual model should be developed initially prior to the development of the final implementation model. It may, however, be revisited and modified within an iterative process during the development of the quantitative model.
- Model development involves making a large number of decisions and judgements. Not every decision or judgement made during model development will be important. The key decisions are likely to be those whereby the implemented model clearly deviates from the problem-oriented models (e.g. a part of the system is excluded) or whereby several alternative choices exist but none of which are clearly superior (i.e. structural uncertainties). These decisions should be clearly documented and reported.
- The sources of evidence used to inform model structure and the methods through which this information is elicited should be clearly reported.
- Where possible, alternative model development choices drawn out at this stage should be later tested using the quantitative model to assess their impact upon the model results. This will not however always be possible or feasible.

## **2.5 EVIDENCE SOURCES TO INFORM CONCEPTUAL MODELS**

A number of potential evidence sources may be useful for informing these types of conceptual model. Whilst the evidence requirements for any model will inevitably be broader than that for traditional systematic reviews of clinical effectiveness, the task of obtaining such evidence should remain a systematic, reproducible process of enquiry. Possible sources of evidence to inform conceptual models include: (1) clinical input; (2) existing systematic reviews; (3) clinical guidelines; (4) existing efficacy studies; (5) existing economic evaluations or models, and; (6) routine monitoring sources. Table 1 puts forward some pragmatic concerns which should be borne in mind when using these evidence sources to inform conceptual model development. These concerns have been drawn from discussions held during the focus groups used to inform this document (see Acknowledgements).

The next section moves on to discuss methods for the identification of evidence to inform cost-effectiveness models.

**Table 1: Roles and concerns regarding the use of evidence to inform alternative model structures**

Existing economic evaluations / models	Expert input (including clinicians and potentially patients/service users)	Clinical guidelines / previous TA guidance / local treatment protocols	Empirical clinical studies and reviews (e.g. RCTs, cohort studies)
<b><i>Principal role(s) in conceptual model development</i></b>			
<ul style="list-style-type: none"> <li>To apply previously developed model structure to the current decision problem under consideration</li> <li>To use existing economic analyses to highlight key evidence limitations</li> <li>To identify possible options for model development decisions</li> <li>To identify relevant treatment pathways</li> </ul>	<ul style="list-style-type: none"> <li>To inform problem-oriented conceptual model development</li> <li>To scrutinise the credibility of alternative model structures</li> <li>To elucidate uncertainty regarding geographical variation</li> </ul>	<ul style="list-style-type: none"> <li>To identify existing treatment/management pathways</li> <li>To highlight gaps in the existing evidence base</li> </ul>	<ul style="list-style-type: none"> <li>To identify available evidence to inform relationships between intermediate and final endpoints</li> <li>To investigate what evidence is available</li> </ul>
<b><i>Issues and caveats associated with use</i></b>			
<ul style="list-style-type: none"> <li>Existing models should not be relied upon without considerable scrutiny.</li> <li>Publication or other forms of dissemination of an existing model does not guarantee that the previous model was either appropriate or credible.</li> <li>Advances in knowledge may render an existing model redundant</li> <li>There may exist a gap between the decision problem that the model was developed to address and the current decision-problem under consideration</li> </ul>	<ul style="list-style-type: none"> <li>Seek input from more than one health professional to capture the spectrum of clinical opinion</li> <li>Use multiple experts located in different geographical locations</li> <li>There exists a trade-off between seeking support from individuals with considerable expertise and standing (may not have much time but more experience/knowledge) and less experienced clinicians (may have more time to engage but lesser knowledge of evidence base).</li> <li>Health professionals cannot be completely objectively detached from the system the model intends to represent</li> <li>May be difficult to distinguish between conflict and geographical variations</li> <li>Potential conflicts of interest</li> <li>Potential ethical restrictions</li> </ul>	<ul style="list-style-type: none"> <li>Current practice may have evolved since publication of guidance</li> <li>Such evidence sources may not provide sufficient detail to inform the current decision problem</li> <li>Local protocols may not reflect geographical variations between centres</li> <li>Local protocols and guidelines may not be evidence-based</li> <li>There may exist a gap between what should happen and what does in happen in clinical practice</li> </ul>	<ul style="list-style-type: none"> <li>Potential reliance on the availability of evidence rather than the structure of the problem</li> <li>Differences between studies may suggest competing theories regarding (a) the nature of the disease process and (b) the relevance of particular events. This is not a problem as such but should be drawn out during conceptual model development.</li> <li>Treatments and comparators may reflect historical rather than current or best practice</li> </ul>

### **3. IDENTIFYING AND SOURCING EVIDENCE**

#### **3.1 PURPOSE OF THIS SECTION**

The NICE Reference Case<sup>1</sup> requires that evidence to inform parameters of clinical effectiveness should be identified by systematic review. In its specification or definition of ‘systematic review’ for these parameters, the NICE Methods Guide refers to the systematic review methods of the Centre for Review and Dissemination (CRD).<sup>17</sup> For all other types of evidence, including utilities, costs, and baseline risks of events, the need to be systematic and transparent is specified several times but the requirement for a systematic review of these types of evidence is not specified. There is an implication that a systematic and transparent process should be used but that this should not or cannot necessarily adhere to conventional systematic review methods.

In an area so dominated by the systematic review methodology it is perhaps useful to establish some definitions for the terms used in this section of the TSD. ‘Systematic review’ is used in the context of established, conventional systematic review methods such as those associated with the Cochrane Collaboration and the Centre for Reviews and Dissemination. ‘Systematic review search methods’ refer to the search methods associated with systematic reviews, characterised by an extensive search approach, directed by a clearly focussed search question and aimed at achieving 100% sensitivity. ‘Systematic’ is used in its generic sense to mean any kind of organised approach to identifying evidence. ‘Evidence’ is used to describe any kind of information, including research, routine and expert information, used to support the development or credibility of a model. ‘Searching’ is used to describe any information seeking process or any process by which information is incorporated in the modelling process. As such, information seeking processes are not restricted to database searching and can include any form of systematic enquiry resulting in the identification or obtaining of information for use in the modelling process.

The purpose of this section is to provide advice on what might constitute a systematic and transparent approach to identifying sources of evidence where the NICE Methods Guide does not specify a requirement to follow conventional systematic review methods, but where no procedural guidance is given. It should be noted that the implications of *not* applying established systematic review methods and of *not* searching exhaustively in the context of

decision-analytic models has not been tested empirically. In the absence of exhaustive searching and the absence of the assessment of the quality of a search by its perceived sensitivity, the guiding principles of the advice given here are:

- § The processes by which potentially relevant sources are identified should be systematic, transparent and justified such that it is clear that sources of evidence have not been identified serendipitously, opportunistically or preferentially.
- § All sources of information used to support the conceptualisation, specification or parameterisation of a model should be identified and supported by appropriate referencing.
- § The factors influencing the choice of evidence source should be identified and justified.

### **3.2 BACKGROUND**

Decision-analytic models bring together, within a relevant analytical framework, the full range of available evidence required to address a decision problem. As such they have multiple information needs, requiring different kinds of evidence, drawn from various types of information resource. Whilst search methods exist for the identification of RCT evidence on treatment effects, the development of equivalent search methods for other types of evidence is not as advanced. In addition, the search methods developed for the identification of RCT evidence form part of the set of methods developed to support the task of undertaking systematic reviews. As such they have not been designed to take into account activities specific to the task of decision-analytic modelling.

The objective of systematic review search methods is to identify as complete a population as possible of relevant studies. The purpose is to maximise the sample population on which the estimate of clinical effectiveness is based and to seek to minimise the risk of bias in terms of publication, language, time-to-publication and study selection bias. The search procedures involve two key requirements. The first of these is the definition of a clearly focussed search question, typically structured according to the PICO question (populations, interventions, comparators, outcomes) that the review is seeking to address. The second is the searching of multiple sources of information using a number of different search techniques. The focus of this approach is on maximising the sensitivity of the search in an attempt to identify every study that matches the PICO question.<sup>18</sup>

In practice, systematic review search methods are not applied consistently across the modelling process. Beyond the population of clinical effectiveness parameters, extensive searching to support the full range of information requirements of the model is rarely, if ever reported.

The debate on whether exhaustive search methods, such as those used in systematic reviews, should be used to support decision-analytic models of cost-effectiveness has thus far been inconclusive. Anecdotally, a consensus appears to have emerged that, ideally exhaustive searches should be undertaken but that in practice this is not feasible. The theoretical methods literature questions the applicability of systematic review search methods in the context of decision-analytic models on several levels, one of these being that the value of searching should be measured not in terms of sensitivity but in terms of the impact of information on the understanding of uncertainty in the model and ultimately on the decision-making process.<sup>19</sup> Conversely, there is concern about the extent to which decision-analytic models are open to the risk of bias.<sup>2</sup> In either case it is important that the processes through which evidence is identified are transparent and that the criteria or judgments that contribute to the choice of evidence should be justified. Currently, there are no methodological reporting standards for searches for decision-analytic models and the information seeking processes that support model development remain frequently unreported.<sup>11</sup>

The remainder of this section will consider the reasons why it might be difficult to apply the systematic review search approach in the context of modelling, what the main types of information used in modelling are, what search techniques might be used to maximise the rate of return of potentially relevant evidence and what might constitute sufficient evidence and a sufficient search process.

### **3.3 SYSTEMATIC REVIEW SEARCH METHODS IN THE CONTEXT OF MODELLING**

The feasibility and applicability of systematic review search methods in the context of decision-analytic modelling is open to discussion on a number of pragmatic and theoretical grounds. In practical terms, it is often argued that there is not sufficient time or resource to undertake a systematic review on every information need generated by the model.<sup>4</sup> On theoretical grounds it has been argued that this should not be a requirement and that search activities should focus on aspects of the model that impact on model outputs and ultimately

the decision that the model is seeking to support.<sup>19</sup> The emphasis here is not on comprehensive searching to minimise the risk of bias, rather it concerns identifying enough or sufficient information to maximise the understanding of the implications of uncertainty in the model.

There are a number of other considerations relating to the activities involved in developing a decision-analytic model which point to fundamental differences between systematic reviews and models. At one level, these can lead to difficulties in fitting the systematic review search approach to the task of developing and populating a model. At another level, they raise questions concerning the applicability of systematic review search methods in the context of modelling.

In addressing ‘real world’ decision problems and in seeking to provide a credible representation of the real world, models draw extensively on a range of information that is specific to the context of the decision and that includes non-research based information. Such sources include registries, administrative or routine data sources and expert advice. In practical terms, the location and format of these types of information makes retrieval difficult using the techniques associated with the systematic review search approach and the completeness of a search in terms of sensitivity is difficult to assess. In theoretical terms, concepts such as publication bias, language bias and time-to-publication bias might not be applicable or might be interpreted differently. For example, the retrieval of information on all-cause mortality to support a decision in the UK would not be at risk of language bias. Likewise the retrieval of information from an authoritative, routine data source to support estimates of health service costs would not require extensive searching to minimise the risk of time-to-publication bias but it would be necessary to demonstrate that the most up-to-date version of the source had been used.

Systematic review search methods rely on capturing the research question or decision problem in a clearly focussed search query. This however raises several problems when applied to the context of model development. The multiple information needs of a model cannot be captured in a single search query. A study to assess the feasibility of using the systematic review search approach for all information needs associated with populating a typical model of cost-effectiveness identified 42 search questions.<sup>20</sup> The process of structuring a search according to the PICO question is not adequate in the context of

modelling as the underlying PICO question cannot be translated directly into search queries relevant to the scope of the model, and it is generally accepted that multiple searches of some form are required to identify the full scope of relevant information. The manageability of this task is further complicated by the complexity and lack of clarity of some information needs. For example the task of defining relevant treatment scenarios requires information on current practice. Current practice is a complex information need and is difficult to translate into a focussed, single search query. An analysis of the evidence used to inform the specification of treatment scenarios in a model of treatments for hypercholesterolaemia identified six sources of evidence including empirical effectiveness studies, expert advice and routine data.<sup>21,22</sup>

The activities associated with conceptualising the decision problem and translating this into a relevant model specification inevitably involve an iterative process of clarification and analysis. In addition, information needs cannot be dealt with individually and sequentially. The generation of information needs is inextricably linked to the process of model development and multiple information needs have to be managed simultaneously. As such, the model development process does not begin with a complete, known set of neatly defined, discrete search questions. Whilst some information needs may be clear at the outset (and following the conceptual modelling approach in Section 2 may help to expose these), others will emerge and become clearer in the course of building the model and in the course of consulting evidence identified through an iterative search process.

Much of the theoretical discussion on searching for evidence for models relates to the population of model parameters. However an analysis of evidence cited in typical cost-effectiveness models concluded that evidence is cited in support of a range of activities that form part of the process of developing a model, other than the population of model parameters.<sup>8</sup> These are outlined in more detail in the previous section and have been observed to include the use of evidence to generate a conceptual understanding of the decision problem, to inform the judgments involved in specifying a relevant model structure and to justify the analytical approach of the model, for example in extrapolating short-term evidence over a relevant time horizon. The information seeking processes that inform this type of modelling activity are often implicit but should nonetheless involve a systematic and transparent consideration of the available evidence such that the acceptability of the sources used can be assessed. The applicability of systematic review search methods to this type of activity is however open to debate. For example, the specification of a given method for

extrapolating short-term evidence should be based on some process of identifying and exploring the available options. It is not necessary however that a search be undertaken to identify every occurrence of the chosen approach in the published and unpublished literature.

If there is a dearth of evidence with which to inform a model, extensive searching in an attempt to address multiple, individual information needs becomes impracticable. In the absence of ‘ideal’ information, the selection of evidence frequently involves the weighing up and trading off of the attributes of different sources, including the use of informed assumptions, in an attempt to find the most acceptable close match to the required information. Rather than focus on the sensitivity of searching, as is the case in systematic review searching, a more efficient approach might be to focus on precision in order to maximise the retrieval of possible ‘close matches.’ Using this approach, the scope or extensiveness of a search can be decided by the volume, characteristics and quality of the close matches being retrieved.

### **3.4 INFORMATION SOURCES TO INFORM MODELS**

The NICE Methods Guide refers to a number of types of information need in its guidance on model-based cost-effectiveness analyses. The guide also points to the need to identify different study designs to address different types of information need and to use both research and non-research based information.

A content analysis of the evidence cited in typical models of cost-effectiveness identified fourteen types of information drawn from seven types of sources.<sup>8</sup> (see Table 2). This included the citing of sources of evidence to support the full scope of modelling activities. In terms of populating model parameters, major types of input have been defined as treatment effects, costs, utilities, baseline risk of clinical events<sup>11,23</sup> and adverse events.<sup>19</sup> The main types of sources from which evidence to populate model parameters is drawn has been classified as evidence syntheses (including secondary economic analyses), RCTs, observational studies (including both longitudinal and cross-sectional studies), routine data sources (including registries and routine and administrative data sources), references sources (including drug formularies) and expert judgment.<sup>8</sup> The range of types and sources of evidence used in models has implications for the retrieval of information.

**Table 2: Classification of evidence used in models of cost-effectiveness by type, source and use of evidence (Paisley, 2010<sup>21</sup>)**

<b>Types of information</b>	<b>Types of source of evidence</b>	<b>Uses of evidence within model</b>
Adverse effects	Evidence synthesis	Design and specification of model framework
Compliance	Expert judgment	Model validation
Current practice	Methodological theory and empirical evidence	Modelling and analytical approach
Epidemiology	Observational research	Population of model parameters
Modelling methods	RCT (clinical and economic)	Sensitivity and uncertainty analysis
Natural history	Reference sources	
Patient preferences	Routine data sources	
Prescribing rates		
Prognosis		
Resource use		
Results and methods from other models		
Clinical outcomes		
Unit costs		
Utilities		

General biomedical bibliographic databases, including Medline and Embase remain important sources. The practice of using validated filters to restrict search results to RCTs is well established. Filters also exist for other types of study design. However, many of these have been designed pragmatically and few have been widely validated. This should be borne in mind when performing searches aimed at achieving high sensitivity. Despite this caveat, filters for study types other than RCTs are widely used and are an important resource in identifying the range of information required for models. The Information Specialists' Sub-Group of InterTASC, the network of academic centres undertaking health technology assessments for NICE, have developed an extensive resource of critically appraised search filters.<sup>17,18:24</sup> In terms of non-RCT evidence the Campbell and Cochrane Economic Methods

Working Group provides advice on searching for economic evaluations and on utilities evidence.<sup>19,25</sup> Another in the series of NICE DSU Technical Support Documents focuses on the identification of utilities evidence.<sup>26</sup> Research on the retrieval of evidence on adverse events is currently being undertaken at the University of York.<sup>27</sup> These sources provide advice both on searching general biomedical databases and on specialist resources specific to their topic areas.

Compilations of resources bringing together routine data sources such as registries and administrative data sources and reference sources such as drug formularies and unit costs have not been developed. Part of the value of this type of information is its relevance to the context of the decision-making process, such as geographical context. For example the British National Formulary (BNF), the Office for National Statistics (ONS) and the NHS Reference Costs provide relevant information in the context of NICE decision-making but do not carry the same authority in all decision-making jurisdictions. As such it would be difficult to develop a generic, comprehensive resource useful to all HTA decision-making systems. However this type of information source is difficult to locate systematically. The compilation of resources relevant to specific, local decision-making processes is an important area for further development. Many of these resources are already mentioned in the NICE Methods Guide. Further methods of identifying this type of information include the use of expert advisors in identifying important sources in their clinical area and brief exploratory searches of the internet and of bibliographic databases to identify leads or 'proximal cues' that can be followed to potentially relevant sources.

The role of expert judgment as a source of evidence in decision-analytic models is more complex than in systematic reviews. As a source of estimates with which to populate model parameters, the expert judgment of clinicians is placed at the bottom of the evidence hierarchy, as in systematic reviews.<sup>23</sup> However, in the absence of relevant evidence, the elicitation of expert information is preferable to an unqualified assumption. Beyond the population of model parameters, the role of expert judgment in interpreting the available evidence and in assessing the face validity or credibility of the model as an acceptable representation of the decision problem is recognised as an important form of validation.<sup>9</sup> In terms of retrieving or obtaining information from experts, the potential for bias is high and the guiding principles pertaining to the more familiar methods of information retrieval remain in place. Formal methods of elicitation, including Bayesian elicitation methods, qualitative

operational research techniques<sup>22,28</sup> and consensus methods<sup>29</sup> have been used in the context of decision-analytic models. Other forms of good practice include the use of a number of experts in order to capture a range of opinion and variation and taking precautions to ensure that the full range of clinical expertise relevant to the decision problem is represented (see Table 1).

In the context of different types of information (e.g. epidemiology, clinical effectiveness, resource use) it is perhaps worth a further consideration of the scope of evidence required to inform a model. Systematic review search methods focus on the identification of direct evidence for inclusion in a review. That is, the search aims to identify studies that seek to address the same question as the review. In order to do so, the scope of the search strategy is structured according to the specification of the PICO question. Whilst a model might address the same PICO question as a systematic review, it does so within a broader analytic framework that aims to reflect the complexity of the decision problem. In bringing together evidence within this broader framework, models consider a wide range of evidence both directly and indirectly related to the specified PICO question. In particular, models require information in order to assess the impact of a technology over the course of a disease. As such, the scope of relevant evidence is dictated by the definition of health states and events within the model rather than by the underlying PICO question. For example, in terms of evidence on effectiveness, costs and resource use, utilities, adverse events and baseline risk of events, the scope of information required is not reflected explicitly in the PICO question but is dictated by the treatment options, management of the disease and population characteristics at all other stages of the disease pathway included in the model.

### **3.5 MAXIMISING THE RATE OF RETURN OF RELEVANT INFORMATION**

The constraint of limited time and resources is probably the most frequently discussed barrier to implementing a systematic process of searching to inform decision-analytic models. The emphasis of systematic review search methods is on searching widely in order to maximise sensitivity. As a result very little has been written in the HTA literature on how to minimise systematically the retrieval of irrelevant information. As mentioned above, the implications of not searching exhaustively have not been tested in the context of decision-analytic model development. However, given that it is not common practice to apply comprehensive, systematic review search methods consistently across the modelling process and given that

the model development process requires generally the assimilation of a broad range of information within a short time, the following techniques, summarised in Box 5 are suggested as a means of maximising the retrieval of potentially relevant information and of minimising the opportunity costs of managing irrelevant information. The techniques are suggested with the caveat that, compared with the systematic review search approach, they all increase the risk of missing potentially relevant sources of information.

**Box 5: Search techniques for maximising the rate of return of relevant information**

- One-line filter searching
- Restricting the number of sources searched
- Restricting search terms to within specific fields
- Lumping and splitting
- Triangulation of different types of sources
- High yield patches
- Proximal cue or information scent searching
- Information gathering using secondary or indirect retrieval

The inclusion of methodological filters to restrict search results to certain study types is an established approach aimed at increasing the relevance of a search yield. Search filters can be designed to maximise either the sensitivity or precision of this restriction. The choice of high precision filters, sometimes referred to as ‘one-line filters’ can be used to maximise the relevance of searches. The Hedges project at McMaster University has developed and tested a set of searches filters, including one-line filters.<sup>30</sup> The ISSG filters website is also an important resource.<sup>24</sup>

Minimising the number of databases searched will reduce the number of references retrieved. In addition, opting to search specialist databases should focus search results to studies with specific characteristics. Such sources include CENTRAL, the Cochrane trials database, NHS EED for economic evaluations. The Cochrane Handbook and the CRD guide to systematic review methods provide information on the coverage of different databases.<sup>18</sup>

Restricting searches to specific fields within bibliographic databases is a commonly recognised information retrieval technique aimed at maximising precision. For example,

searching for relevant terms within the title of journal articles should minimise the retrieval of irrelevant information. Depending on the nature and amount of information retrieved, a decision or judgment can be made as to whether to extend the search across other fields, such as the abstract, with a view to increasing sensitivity.

Conversely, searches with a broad scope, aimed at identifying evidence to satisfy a number of more focussed information needs might be considered. Here, a decision about the granularity of the search is made. A typical example in the context of decision-analytic models is to search for health utility values across a whole disease in order to identify specific values relevant to different health states within the pathway. Another example is to search for ‘costs’ as a broad concept instead of or preceding more focussed searches on specific cost components within a detailed cost analysis. This technique, comparable to the practice of ‘lumping and splitting’ in the definition of review questions, is useful for conditions for which there is not a large volume of evidence and where it would not be efficient to pursue a separate search for each individual information need.

The triangulation of different types of information source might be used in preference to in-depth searching within one type of source or might be used to inform a decision as to whether to extend the search process. For example, cost information from a published source might be compared with or judged alongside similar information from routine sources and information derived from clinical experts. The purpose of triangulation would be to capture the level of consistency or inconsistency across the breadth of a number of different types of information, particularly in terms of research-based and non-research based information.

Existing cost-effectiveness models in the same disease area are frequently reported, anecdotally, as important sources of information in gaining an understanding of the disease area, for identifying possible modelling approaches and for identifying possible sources with which to populate a model. This can be described as a ‘rich patch’ or ‘high yield patch’ whereby one source of information is used to satisfy a relatively high number of information needs within the model. The previous experience of a project team in the same disease area or similar type of decision problem is a variation of this type of high yield patch. Other high yield patches might include systematic reviews, clinical guidelines and clinical experts. The use of high yield patches can provide useful short cuts or can help cover a lot of ground quickly in terms of gaining an understanding of a decision problem and a view of how it

might be addressed within an analysis. However, it is important that consideration is given to the limitations of a potentially rich source. For example, a previously published model might provide useful information in terms of model structure or analytical approach but might be poor in terms of sources with which to populate model parameters. It might limit the potential to explore alternative analytical approaches to a similar decision problem. Lastly, it would be necessary to justify the quality of any sources of information being used as a high yield patch.

Following trails of potentially relevant sources can provide an alternative approach to searches that cast the net wide using broad keyword strategies. One form of this technique, sometimes referred to as 'snowballing', is used in systematic review searching, whereby trails of cited references are followed prospectively and retrospectively from a single or a series of index sources. However, it is also possible to use any information from the source documents as the starting point of an information trail. As such, a starting point might take the form of an idea or concept, an author or a set of keywords. The starting point or points act as 'proximal cues' which can be followed to further, similar, potentially relevant information. In the field of information seeking behaviour this is referred to as following 'information scents.'<sup>31</sup>

It has already been stated that information needs do not arise sequentially but that multiple information needs might be identified at the outset of the modelling process or might arise simultaneously during the course of developing the model. A useful way of handling multiple information needs is to consider information retrieval as a process of information gathering alongside a more directed process of searching. The pursuit of one information need might retrieve information relevant to a second or third information need. The yield of this 'secondary' or indirect retrieval can be saved and added to the yield of a later more directed retrieval process. Good communication and an understanding of the requirements of the model across the whole project team will maximise the usefulness of the information gathering approach.

The above techniques have not been formally tested in terms of their impact on the sensitivity and precision of searches to inform models. They are however commonly accepted strategies for managing the process of information retrieval. Whilst they are suggested as a means of maximising the precision of a search, none of the techniques prohibits extensive searching aimed at achieving a high level of sensitivity as all the techniques can be used iteratively to extend the scope of the search until no further relevant items are being retrieved. They are

suggested here as a means of encouraging a more systematic exploration of the broad scope of evidence of potential relevance to a model and of providing an auditable search process through which the judgements that direct the scope of information seeking processes can be made explicit and more transparent.

### **3.6 DEFINING A SUFFICIENT SEARCH PROCESS**

In considering what might constitute a systematic search process, this section has questioned whether searches for models can or indeed should aim to achieve maximum sensitivity. If this is not a practical or appropriate measure of the quality of a search then it is important to consider what might constitute sufficient evidence or what might be justified as being a sufficient search process in the context of modelling. Previous discussions have suggested that a suitable definition might be not that searches should aim to identify every occurrence of a predetermined focussed question, but rather that they should aim to identify the breadth of information needs relevant to a model and sufficient information such that further efforts to identify more information would add nothing to the analysis.<sup>19</sup> This is more akin to the idea of saturation than sensitivity and questions the idea that using a predefined, focussed question to determine the scope of a search process is sufficient to support the inductive and iterative process of model development.

The concept of sensitivity as a measure of the quality of a search or effectiveness of the search process has a substantial influence on the field of search methods in HTA and dominates the scope of empirical methods research in this area. As a consequence, very little has been discussed or tested in relation to additional possible interpretations of what constitutes a 'good' or sufficient search. In reality, very few searches undertaken to support systematic reviews are or can be tested in terms of their true sensitivity. Instead, a transparent and detailed description of the extensiveness of the search process acts as a proxy by which the sufficiency of the search can be judged.

As mentioned previously in this section, the implications of not undertaking extensive searching in terms of increasing the risk of bias in the model have not been tested empirically in the context of modelling. However, the purpose of this document is to give guidance on process in the absence of undertaking full systematic review searches. As such it is important

to offer some discussion on the factors that might influence decisions to stop searching and that might inform judgments as to whether a sufficient search process has been undertaken.

As previously suggested, the definition of a sufficient search process might be judged according to the use of information within the model. The search process should be fit for the purpose for which the information is being used. For example, the population of clinical effectiveness parameters demands an extensive search process whilst the understanding of the underlying condition or disease requires sufficient evidence but not every occurrence of evidence to inform a relevant and acceptable representation of the disease pathway in the model structure.

A sufficient search process might also be defined by the availability or lack of available relevant information. If a relatively systematic search process exploring a number of different search options has retrieved no relevant information it might be acceptable to assume that further extensive searching will not be of value.

A further definition, which is suggested several times in the NICE Methods Guide, is to focus on understanding the implications of the uncertainty generated by the evidence used, including an analysis of any alternative sources identified but not used in the baseline analysis of the model. The process of bringing together, within one framework, multiple and diverse sources of evidence will bring with it unavoidable uncertainty that cannot fully be understood or removed by comprehensive searching on every information need within the model. An extension of this idea would be to use sensitivity analysis or some form of value of information analysis to understand which information needs might have the greatest impact on the outcome of the model. Sensitivity analysis could be used to determine where search resources should be focussed. It could act as a device to prioritise areas for further rounds of searching during the course of a modelling project or to make research recommendations for more in depth searching and reviewing on specific topics to inform future models. It is however important to note that the usefulness of this approach is dependent on the timing at which this approach is undertaken within the model development process. A particular model parameter which appears unimportant during the early stages of model development may become more important as the model, and the other parameters specified therein, are further developed and refined over time.

### **3.7 REPORTING OF THE SEARCH PROCESS**

It is recognised that the development of a model is not necessarily a reproducible process and that different groups addressing the same decision problem might build different models, neither of which is wrong and both of which can be an acceptable representation and analysis of the decision problem (see Section 2.2). The aim of model reporting is to make the process of model development transparent such that users of the model can judge whether it is a plausible and acceptable explanation and analysis of the decision problem. The conceptual model development approaches in Section 2 are intended to support this aim.

This issue is also relevant to the search process. Searches are not directed by a focussed, structured question. The scope of relevant evidence is not fully predefined but emerges in the course of model development. As such there is not a “right” or “wrong” set of evidence, but an interpretation of what evidence is relevant to the scope of the decision problem. The purpose of reporting searches is to make explicit the information seeking processes that have underpinned the development of the model such that users of the model can understand how sources of evidence came to be incorporated into the process and can judge whether the model is based on a plausible or acceptable set of evidence.

The search process that supports the development of a model is not a series of sequential, discrete information retrieval activities. The NICE Methods Guide describes a process of ‘assembling’ evidence and this reflects an iterative, emergent process of information gathering.<sup>1</sup> As such, it is difficult to report searches in a way that makes transparent all the procedures or activities associated with assembling the evidence base. It is useful therefore to provide an account of the main search activities together with an audit trail of how the sources of evidence cited came to be identified. Two assessments undertaken to support NICE appraisal decisions, and cited here as examples, have attempted to report both the main search activities and provide an audit of all sources identified.<sup>22,32</sup> Similar guidelines on providing an audit of sources have also been developed.<sup>33</sup>

A comprehensive account of how every source of evidence came to be part of the modelling process can be time consuming to generate and to read. In the absence of formal reporting standards, judgements have to be made concerning how much information to include. The

following are suggestions of reporting devices aimed at improving the transparency of the process.

- § A brief table providing an overview of the main sources and searches from which the evidence has been drawn. This can include searches to inform the systematic reviews of clinical and cost-effectiveness, searches undertaken specifically for the model, important 'high yield' patches such as existing models and expert opinion. A hypothetical example is provided in Box 6.
- § Inclusion, if possible of judgements illustrating why specific search approaches have been used, and why and when decisions have been made to stop searching.
- § An appendix of any directed searching, including keyword search strategies.
- § An appendix containing an audit table of all sources cited in the model, including a brief reference, purpose (or purposes) for which the source has been used and brief description of how the source was identified (e.g. from searches for systematic review of clinical effectiveness etc.) A hypothetical example is provided in Table 3.

**Box 6: Overview of sources consulted in the development and population of the model**

- Searches undertaken for review of clinical effectiveness (refer to relevant parts of report)
- Existing models and economic analyses (provide references)
- Studies identified through the review of cost-effectiveness (refer to relevant parts of report)
- Studies identified through searches undertaken to inform the model (provide appendix)
- Reference sources (e.g. BNF, NHS Reference Costs) (provide references)
- Expert opinion

**Table 3: Example of how to report and audit a table of all sources of evidence cited in support of a model**

Source	How the source is used in the model	Identification process
Reference to source	Estimate of cost X	Existing published model (provide reference)
Reference to source	Baseline risk of event X	Searches undertaken to inform model
Personal communication	Specification of treatment scenarios	Discussion with clinical advisor to project

### 3.8 CONCLUSION

Models of cost-effectiveness draw on a complex set of information comprising different types of evidence. The nature of the sources to be searched, the need to manage simultaneously multiple information needs and the need to work within the iterative modelling process make it difficult to apply a discrete, sequential search approach based on clearly defined search questions. The guidance in this section aims to give advice on factors that impact on the way evidence is identified. These are summarised in Box 7. The section also provides suggestions concerning how these factors might be managed within the search process in order that the process of model development, and in particular the process of reviewing evidence for model parameters, has access to a broad a range as possible of the available options within the available time and resources.

**Box 7 Factors that impact on the approach to identifying evidence to inform the development and population of models**

- The process of model development generates multiple information needs that cannot be satisfied by a single search strategy. Multiple information seeking strategies are required.
- Information needs are not wholly predefined but emerge and become clear during the course of model development. The retrieval of information requires focussed, directed searching but also relies on indirect retrieval and information gathering techniques.

- The process of identifying and assembling evidence to support the network of information needs should be managed in order to audit how evidence used to inform the model is identified and in order to make decisions about when an information need has been satisfied.
- A separate search is not required for each information need. A single search might identify evidence to satisfy a number of information needs. Likewise, a single piece of evidence might satisfy a number of information needs.
- It is not possible to identify all evidence on every information need. However, decisions about when to stop searching or what constitutes sufficient evidence should be transparent and justified. Such decisions might be influenced by the impact of information on the outputs of the model and by the availability of evidence. Implications regarding the choice of evidence could be explored using sensitivity analysis.
- Information retrieval techniques aimed at high precision can be used with a view to making searches more efficient, particularly given the need to manage multiple, interrelated information needs. This does not preclude the possibility of searching in more depth on information needs judged as being important to the outputs or credibility of the model.
- The process of model development will draw on different types of information and evidence will take many different forms. This requires consultation and interrogation of a wide range of information sources using a range of information retrieval techniques and methods.
- Information is used across the whole process of model conceptualisation and population. The purpose for which evidence is used might determine the type of information and the level of searching required.
- The reporting of the search process should aim to provide an overview of the scope of the evidence consulted and of the main information retrieval activities. In addition it is useful to provide an audit of how evidence, cited in the model report, came to be identified.

## **4. REVIEWING EVIDENCE TO INFORM MODELS**

### **4.1 UNDERLYING PRINCIPLES**

The NICE Methods Guide<sup>1</sup> states that the process of assembling evidence for health technology assessment needs to be systematic. That is evidence must be identified, quality assessed and when appropriate pooled using explicit criteria and justifiable and reproducible methods. The Methods Guide goes on to state that evidence will be drawn from a variety of sources, depending on the evidence need: “When assembling the evidence it is essential to consider how bias can be minimised, especially when non-randomised studies are included.”<sup>1</sup> The methods of systematic reviewing have been developed to minimise bias when assessing the relative effectiveness of interventions. Some, but not all of these are relevant in this context. Reviewing is essentially comprised of four key components: (1) searching, (2) appraisal, (3) synthesis and (4) analysis, in this context that is the interpretation of the findings and how successfully they have addressed the initial question to be answered by the review.<sup>34</sup> However, an important distinction exists between the contexts of conventional systematic reviews of clinical effectiveness and reviewing activity to inform cost-effectiveness models. Whilst a clinical effectiveness review can remain inconclusive, especially when faced with bias, this situation is not the same within the context of reviewing information for model development. Instead, choices *must* be made regarding the identification, selection and use of evidence to inform models, and inevitably may lead to questions regarding bias.

### **4.2 SELECTION AND PRIORITISATION OF PARAMETERS FOR REVIEWING EFFORT**

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, pre-defined criteria. This may be because an estimate is required and there will usually at best be a range of options, all of which may fall short of what would be considered ideal to differing degrees. In addition, it is not possible or appropriate to synthesise all the available options due to heterogeneity and other factors. The nature of the trade-off between selecting alternative parameter values will often include elements relating to quality versus relevance for each option. The process of modelling draws out uncertainty and examines the impact of that uncertainty on the decision problem under consideration. The bringing together of so many components brings uncertainty. It is usual practice to explore, understand and explain

the implications of uncertainty; this can include explaining the implications of choosing one source of evidence over another. Given the above, procedures associated with undertaking systematic reviews can be used to make the process of choosing evidence systematic and transparent. However, given the differences between models and systematic reviews, the purposes for which these procedures are undertaken and the sequence in which they are carried out may differ. In addition, time and resource constraints will also impact on how they are undertaken. A full systematic review is certainly not possible or even preferable for all evidence needs within cost-effectiveness models due to time and resource constraints within NICE technology appraisal process. In addition systematic reviewing approaches may not be applicable for the selection of certain parameter values, such as costs due to very limited appropriate sources of information. However, systematic approaches for the identification and reviewing of evidence are needed. It is important that these processes are transparent, justifiable and replicable. This is particularly important as the use of different data sources to populate parameter values has been shown to have a marked impact on cost-effectiveness results.<sup>23</sup>

Different types of evidence will be needed to populate model parameters including RCTs, expert judgement, observational research, reference sources, such as drug formularies and routine data among others.<sup>19</sup> The complexities associated with identifying these data sources have been described in detail in Section 3. Systematic reviewing techniques will not be entirely appropriate for the reviewing of some of these types of evidence although some aspects of the process may still be applicable, for example the use of a pre-defined quality assessment tool. In order to provide information within the time and resource constraints of NICE technology appraisals, it is important to be pragmatic about which parameters to prioritise with regard to the allocation of reviewing time. Although some parameters will be identified as important to the model early on in the process, the importance of some other parameters will only be identified later.

### **4.3 REVIEWING METHODS**

Due to time and resource constraints it may be necessary to use rapid review methods to identify and select evidence to inform certain model parameters. Although rapid review methods are not ideal due to the potential for missing relevant information, it is essential that methods are reported in a transparent manner and that the limitations and potential biases,

such as study selection bias introduced by these reviewing approaches are addressed.<sup>35</sup> Various rapid methods have been described for the review of clinical effectiveness and some of these may be applicable for the reviewing of model parameters. These include restricted research questions, truncated search strategies and reduced use of peer review.<sup>36</sup> Rather than developing a formalised methodology to conduct rapid reviews, which may be inappropriate and oversimplified, the authors suggest that transparent reporting of methods is essential. Other potentially relevant rapid review methods in this context include reduced formal quality assessment, data extraction of key outcomes only and reduced levels of synthesis.

**Box 8: Rapid review methods for model parameters**

- Restricted research questions
- Truncated search strategies
- Reduced use of peer review
- Data extraction of key outcomes only
- Reduced levels of synthesis
- Transparent reporting of methods including limitations and potential biases

#### **4.4 MINIMISING BIAS**

A variety of potential biases may be introduced through the process of reviewing evidence to inform model parameters values. This may include biases introduced through the use of less thorough searching and reviewing methods as well as biases through the selection of evidence to create more or less favourable results. An important factor to reduce such bias is to ensure 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. This may include clinical advisors, information specialists, systematic reviewers and modellers on the team. Through the use of the Reference Case, the NICE Methods Guide provides guidance on the acceptable types of information and sources to populate cost-effectiveness models within the technology appraisal process; the identification of these sources is covered in some detail in Section 3 of this document.

## 4.5 HIERARCHY OF EVIDENCE SOURCES

Types of model parameters vary considerably, hence the sources of evidence appropriate to each type will also vary considerably (see Table 2). Some have proposed the use of a hierarchy of evidence sources for use in cost-effectiveness models as a means of judging the quality of individual evidence inputs and informing processes of study selection<sup>23</sup> (see Table 4). The hierarchy covers five common data elements for model parameters (1) clinical effect sizes, (2) baseline clinical data, (3) resource use, (4) unit costs and (5) health utilities. Those data sources with a rank of 1 are considered to be of the highest quality with 6 as the lowest quality, although there is a lack of consensus on preferred sources of health utility values.<sup>23</sup>

However, whilst such hierarchies of evidence may be useful, they fail to incorporate the quality of the individual studies identified,<sup>3</sup> and contribute only to one part of the study selection process, namely the assessment of quality. This represents only a narrow view of what should be considered when selecting model parameter values. Various instruments and checklists have been developed to inform assessments of risk of bias in RCTs<sup>37-39</sup> and non-randomised studies<sup>40,41</sup> of effects that may generate components of data used to populate parameters in economic models. Perhaps the most prominent amongst these instruments is the Cochrane Risk of Bias tool, which can be applied both to RCTs and non-randomised studies.<sup>42,43</sup> However, these instruments are generally not applicable to the diverse range of potential data sources that may be used to inform the model development process at different stages. The Grading of Recommendations Assessment, Development and Evaluation (GRADE) system offers some promise in this respect, as it provides a consistent framework and criteria for rating the quality of evidence collected (or derived) from all potential sources of all data components that may be used to populate model parameters, including research-based and non-research based sources (e.g. national disease registers, claims, prescriptions or hospital activity databases, or standard reference sources such as drug formularies or collected volumes of unit costs).<sup>44-46</sup> Consistent with the hierarchy proposed by Coyle and colleagues, GRADE 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.<sup>47</sup>

It is also important to note that the NICE Methods Guide<sup>1</sup> (Sections 5.3 to 5.5) sets out the Institute’s preferences for certain parameters such as for effectiveness data, utilities and resource costs.

**Table 4: Hierarchies of data sources for health economics analyses<sup>23</sup>**

<b>Rank</b>	<b>Data Components</b>
<b>A</b>	<b>Clinical effect sizes</b>
1+	Meta-analysis of RCTs with direct comparison between comparator therapies, measuring final outcomes
1	Single RCT with direct comparison between comparator therapies, measuring final outcomes
2+	Meta-analysis of RCTs with direct comparison between comparator therapies, measuring surrogate outcomes*  Meta-analysis of placebo-controlled RCTs with similar trial populations, measuring final outcomes for each individual therapy
2	Single RCT with direct comparison between comparator therapies, measuring surrogate outcomes*  Single placebo-controlled RCTs with similar trial populations, measuring final outcomes for each individual therapy
3+	Meta-analysis of placebo-controlled RCTs with similar trial populations, measuring surrogate outcomes*
3	Single placebo-controlled RCTs with similar trial populations, measuring surrogate outcomes* for each individual therapy
4	Case-control or cohort studies
5	Non-analytic studies, for example, case reports, case series
6	Expert opinion
<b>B</b>	<b>Baseline clinical data</b>
1	Case series or analysis of reliable administrative databases specifically conducted for the study covering patients solely from the jurisdiction of interest
2	Recent case series or analysis of reliable administrative databases covering patients solely from the jurisdiction of interest
3	Recent case series or analysis of reliable administrative databases covering patients solely from another jurisdiction
4	Old case series or analysis of reliable administrative databases. Estimates from RCTs

5	Estimates from previously published economic analyses: unsourced
6	Expert opinion
<b>C</b>	<b>Resource Use</b>
1	Prospective data collection or analysis of reliable administrative data from same jurisdiction for specific study
2	Recently published results of prospective data collection or recent analysis of reliable administrative data – same jurisdiction
3	Unsourced data from previous economic evaluations – same jurisdiction
4	Recently published results of prospective data collection or recent analysis of reliable administrative data – different jurisdiction
5	Unsourced data from previous economic evaluation – different jurisdiction
6	Expert opinion
<b>D</b>	<b>Unit costs</b>
1	Cost calculations based on reliable databases or data sources conducted for specific study – same jurisdiction
2	Recently published cost calculations based on reliable databases or data sources – same jurisdiction
3	Unsourced data from previous economic evaluation – same jurisdiction
4	Recently published cost calculations based on reliable databases or data sources – different jurisdiction
5	Unsourced data from previous economic evaluation – different jurisdiction
6	Expert opinion
<b>E</b>	<b>Utilities</b>
1	<p>Direct utility assessment for the specific study from a sample:</p> <ul style="list-style-type: none"> <li>a) of the general population</li> <li>b) with knowledge of the disease(s) of interest</li> <li>c) of patients with the disease(s) of interest</li> </ul> <p>Indirect utility assessment from specific study from patient sample with disease(s) of interest: using tool validated for the patient population</p>
2	Indirect utility assessment from a patient sample with disease(s) of interest: using a tool not validated for the patient population
3	<p>Direct utility assessment from a previous study from a sample:</p> <ul style="list-style-type: none"> <li>a) of the general population</li> <li>b) with knowledge of the disease(s) of interest</li> </ul>

	c) of patients with the disease(s) of interest Indirect utility assessment from previous study from patient sample with disease(s) of interest: using tool validated for the patient population
4	Un sourced utility data from previous study – method of elicitation unknown
5	Patient preference values obtained from a visual analogue scale
6	Delphi panels, expert opinion
Adapted from Coyle & Lee and Cooper and colleagues <sup>5,11</sup> * Surrogate outcomes = an endpoint measured in lieu of some other so-called true endpoint (including survival at end of clinical trial as predictor of lifetime survival).	

The definition of what is required may be based on an initial understanding of what will constitute relevant evidence. The objective is to identify a set of possible options from which choices will be made. Alternatively, strict selection criteria may initially be applied. If no relevant studies are identified these criteria can be broadened. The approach can be flexible depending on the types of information needed. Both approaches may be useful, however it is important to explain the process used and why it was chosen in order to justify the choices and 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 can be applied. For large numbers of sources, study selection using standard systematic review processes of screening for titles, abstracts and full texts can be used as described in a review of health state utility values for osteoporosis.<sup>25</sup> It is important to be as transparent as possible about the judgements being made when selecting studies (for example, stating which studies were deemed to be most relevant to UK clinical practice or the patients were most similar to those described in the scope). That is, the decision will include issues of relevance as well as other factors such as methodological quality. Study selection should ideally be made by more than one person in order to reduce the potential for study selection bias. As with systematic reviews of clinical effectiveness, the reliability of the decision process is increased if all papers are independently assessed by more than one researcher.<sup>17</sup> This potentially encourages a systematic thinking through of the different factors underlying selection.

Usually, as part of the clinical effectiveness systematic review, a comprehensive database is developed which will often contain references which may be considered relevant for informing a model's parameters. It may be helpful for the reviewer to identify or tag

references of potential relevance to the model. These may either form a set of useful sources or be a useful starting point for identifying further sources as described in Section 3.5. The danger associated with this approach is that the modeller makes an assumption that all relevant studies have been identified from the searches when in fact some may have been missed. Good communication within the team can help to prevent this. A system of checking a sample of references to ensure a common understanding between the reviewer and modeller may also be helpful.

**Box 9: Study selection**

- The NICE Methods Guide (sections 5.3 to 5.5) sets out its preferences for certain parameters such as for effectiveness data, utilities and resource costs
- Hierarchies of evidence can be considered when selecting studies
- Judgements around study selection should be as transparent as possible
- Selection criteria may need to be flexible
- Study selection should ideally be made by more than one person

#### **4.6 ASSESSMENT OF EVIDENCE AND DATA EXTRACTION**

Selection of evidence will often be a series of steps, first an initial selection of potential evidence followed by further selection choices. Evidence for model parameters will need to be assessed on the basis of relevance as well as quality. By assessing relevance first, a large number of studies may be eliminated. Criteria for relevance are ideally established *a priori* for example, studies which report disease-free survival for patient group A in population B might be considered relevant for a specific model parameter. However, it is important to recognise that it is not possible to have pre-specified criteria for every parameter as information needs will change and information that was not expected may be identified iteratively. Relevance criteria may also change throughout the project and flexibility is essential. Anticipated evidence requirements, as perceived during the earlier stages of model development, can be set out during the conceptual modelling stage. When the final model is developed it is important to be clear how this deviated from the initial plan and why. For example, a modeller may want to identify EQ-5D health valuation studies undertaken within a particular population but only identify one such study in a slightly different population. The

modeller is then faced with a choice of using this study or perhaps mapping available SF-36 data even though these studies may have initially been deemed to be less appropriate.

After appraising studies for relevance, they can then be assessed for quality, preferably using standardised quality assessment tools. In this context, quality assessment may be difficult due to the absence of standardised 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* and may include for example study recruitment procedures, inclusion and exclusion criteria, description of the background characteristics of the sample population from whom values are obtained, response rates and follow-up data for utility studies.<sup>25</sup> Other issues to consider may include the type of reporting (self or proxy), follow-up rates, number of patients, location, method of elicitation among other issues. Establishing very broad *a priori* criteria may be necessary initially; these criteria may change according to the availability and relevance of existing evidence. It is important to be clear about the factors or criteria that drove the choice and to examine the implications of that choice. For example, “*there were five options and we chose one because of the reasons a, b and c.*” This level of transparency will allow judgements to be made as to whether or not a reasonable choice has been made. As it can be very time consuming to judge the quality of all potentially relevant studies, adjusting them according to relevance and rigour 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 reported.

Data extraction is undertaken in this context to potentially inform the choice of information source. Data to be extracted from studies may include study date, information on disease area and patients (age, sex, co-morbidities), 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 included studies, for example using tables and/or graphs. This level of detail is not appropriate for all parameter values but should be reserved for those decisions whereby none of the available studies are clearly superior or whereby evidence available is notably weak. When extracting data from studies it is important to provide information for all of the potentially relevant studies. By providing a summary of potentially relevant studies, the reader is able to assess the study differences and heterogeneity more

accurately and see the spread of evidence. Information from the studies not chosen may be used to inform the sensitivity analysis. In evidence synthesis there will be some heterogeneity associated with the chosen studies; sensitivity analyses around these estimates may be necessary. The difference in inclusion and exclusion criteria should be sufficiently clear and concise to make the process reproducible. Inconsistencies between different estimates should be represented. It may be useful to provide a table of potential studies and graphically display values describing the sources from which they were derived. Although the results presented may be wide, this can show where the differences in the utility values between disease stages or different baseline event rates for example, are driving the model results. It is recognised that these suggestions may be quite time consuming and there may be time and reporting constraints within a technology assessment report. However, the overriding objective should be to present the information as transparently as possible.

**Box 10: Assessment of evidence**

- Assessing relevance before quality of studies can reduce the number of studies that need to be assessed
- Where available standard quality assessment tools can be used to assess study quality
- Selected studies can be presented in a tabular format
- The more important model parameters or those where the choice is not clear should be given additional attention during model reporting

## **4.7 SYNTHESIS AND ANALYSIS**

Practical guidance surrounding evidence synthesis is covered in detail elsewhere in the series (TSDs 1-7) and is therefore not reproduced here. For many types of model parameters, the issue of synthesis may not be considered relevant due to study heterogeneity. Often only one or two values are appropriate for use in populating a model parameter. The issue of synthesis obviously becomes important when there are more than one or two potentially relevant studies. A decision needs to be as to whether a complex synthesis method may 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 instances whereby a quantitative synthesis is not undertaken, this should be justified explicitly. Exploration of the use of alternative values is best explored via sensitivity analyses as described above.

With respect to analysis, which in this context refers to the interpretation of the findings, it is important to give some consideration to all of the potentially relevant values identified to inform the model parameters. This would include issues such as whether or not the patient population in the selected studies was representative of the population in the model, and if not, what were the differences as well as other concerns or issues with the included data. This information provides the foundation for sensitivity analyses which will be undertaken to examine the importance of uncertainty within the model.

#### **4.8 REPORTING OF REVIEWING METHODS**

Owing to time constraints, it may not be possible to provide a full description of the review process and criteria used for every model parameter, it is good practice to provide some description of the process used. This may include information such as the number of researchers involved in the study selection process and the justification for the selection of particular studies. It is particularly important to provide descriptions of the process where decisions have been particularly difficult, or where other credible or plausible values might have been chosen. A balance is required between providing sufficient information for the process to be replicable and transparent, and producing a document that is readable, useful and of a manageable size. One means of ensuring that the Assessment Report or submission is of a manageable size is through the use of appendices. The reader can be directed to appendices to find the process for the identification and choice of specific parameters. A form could also be used for each parameter value listing key information for each. Suggested reporting requirements include a list of the parameters in the model together with the method of evidence identification, details of the selection process of evidence and information on the quality and relevance of selected evidence.<sup>33</sup> More thorough reporting is required for those parameters deemed to be most important in the model. Some information, such as changes to the model, the introduction or removal or parameter values may not necessarily be presented in the final report however it is useful to retain this information as a record of the process for future reference. This is particularly useful when the project is revisited.

#### **4.9 PLANNING AND PROJECT MANAGEMENT**

This section puts forward some suggestions for planning and project management. Many of these suggestions are relevant for all aspects of the model development process. Approaches

for the identification and reviewing of evidence will differ across each individual decision problem, centre and team. However, good communication within the project team and planning are usually key components of the process. This is an iterative process, hence there needs to be a clear understanding of roles and responsibilities at the beginning of the project for all team members.

Although it is important that the process is not “data driven” changes will need to be made to the model structure depending on what data is or is not identified to populate parameters. The model will be modified and understanding will change as the project progresses. A considerable amount of time is spent on the clinical effectiveness review and the importance of some parameters only becomes apparent during the later stages of model development. Due to time and resource constraints it may be necessary to use very rapid reviewing methods, that is, data extraction and quality assessment may need to be limited.

As suggested in Sections 3.3 and 3.4, the anticipated information needs for the model can be set out as part of conceptual model development. This will need to be revisited throughout the project as information needs change as an understanding of the disease area is improved and the model is developed. The whole team can be involved in identifying potentially useful sources of information. It may help to develop a plan or protocol with a timetable for every part of the project. Although this may be time consuming, everyone is clear what is expected of them and what information needs are their responsibility. It is very useful to plan in contingency time for unexpected information needs as the project progresses.

Clinical experts are important in the identification and choice of parameters. This may include a wider group of people who are involved in caring for patients. Such input may be particularly useful in instances whereby searches fail to identify any relevant information. Clinicians may be able to provide values or identify areas to search for evidence. They and other researchers can serve as a reality check to ensure that important information has not been missed and that the values used are the most appropriate. Peer review may also be used to challenge the model parameter values employed within the model.

### Box 11: Project management

- Good team communication and planning is essential
- Protocols or project plans may be useful
- Use of clinical experts is crucial
- Peer review should be considered

## 5. RECOMMENDATIONS

### 5.1 RECOMMENDATIONS FOR PRACTICE

The following are a set of suggested recommendations for practice. These are not intended to be prescriptive but offer practical options to help improve the systematicity and transparency of the model development process.

#### *Recommendations related to problem-oriented disease process and service pathways models*

- Develop the structure of the conceptual model using clinical guidelines and clinical experts.
- Use other clinical experts not involved in model development to provide peer review and to check understanding of the conceptual model.
- The precise graphical approach for presenting the conceptual model is important only in that the model should easily understood by clinical experts and other individuals involved in the model development process.
- It may be beneficial to present the model in both diagrammatic and textual forms using non-technical, non-mathematical language.
- The feasibility and acceptability of the design-oriented conceptual model should have no bearing on this phase (once the problem-oriented model is considered adequate there should be no need to iterate between the two types of conceptual model).

#### *Recommendations related design-oriented conceptual models*

- The design-oriented conceptual model should be developed initially prior to the development of the ‘hard’ quantitative model. It should, however, be modified within an iterative process upon development of the ‘hard’ quantitative model.

- Model development involves making a large number of decisions and judgements. Not every decision or judgement made during model development will be important. The key decisions are likely to be those whereby the implemented model clearly deviates from the problem-oriented models (e.g. a part of the system is excluded) or whereby several alternative choices exist but none of which are clearly superior (i.e. structural uncertainties). These decisions should be clearly documented and reported.
- The sources of evidence used to inform model structure and the methods through which this information is elicited should be clearly reported.
- Where possible, alternative model development choices should be tested to assess their impact upon the model results. This will not however always be possible or feasible.

*Recommendations related to the identification of evidence*

- The information seeking processes underpinning the development of a model should be systematic and explicit. This includes the identification of evidence to inform the population of model parameters and the identification of key information used for the conceptual model and the specification of the model structure or analytical framework.
- The processes of identifying and selecting evidence should be transparent. Users of the model should be able to judge whether sufficient effort has been made to identify an acceptable set of evidence on which to base the model and that sources of evidence have not been identified serendipitously, opportunistically or preferentially.
- Information to inform this judgment could include the reporting of search strategies, an audit of how individual sources of evidence were identified, information on alternative sources and an assessment of the impact of the limitations of the evidence base or of alternative sources on the outputs of the model.
- Given restrictions on time and resources, consideration should be given to the use of search techniques that are aimed at maximising the rate of return of potentially relevant information. Justification of the use of these techniques and consideration of the implications in terms of missing potentially relevant information should be given.
- Decisions relating to the prioritisation of key information needs and judgments relating to when to stop searching should be transparent.

### *Recommendations related to the reviewing of evidence*

- Reviewing effort should be prioritised around the important model parameters and reviewing methods chosen commensurate with the parameter's importance. Caution is however advised as the importance of certain model parameters may change as other parts of the model are developed and refined.
- Study selection processes should be clearly reported. There should be transparency about what judgements have been made regarding study selection.
- In selecting evidence for the value of a parameter, relevance may be more important than study quality. However, whatever basis by which the parameter is selected, sources of bias should be carefully considered and where possible taken into account in sensitivity analyses during the modelling process.
- Evidence needs may be set out at the beginning of the reviewing process and changes to these needs documented.
- More detailed reviewing methods may be reported where the evidence is notably weak or where no clearly superior study was identified.

## **5.2 RECOMMENDATIONS FOR FURTHER RESEARCH**

The following are areas where further research is needed:

- Compilation of information resources, including routine data sources and other non-research based information, aimed at improving access and exploitation of this difficult to find information.
- Development and evaluation of search procedures specific to the task of developing models of cost-effectiveness.
- Development and evaluation of procedures aimed at maximizing the rate of return of the search process. Evaluation to include the implications or impact of not undertaking searches aimed at high sensitivity.
- Exploration of the concept of sufficient evidence or a sufficient search process in the context of modelling, with a view to informing the development of search stopping rules.
- Development of reviewing methods for non-standard sources of evidence.
- Development of reporting standards for the whole process from the conceptualisation of the model to reporting the outcomes using real world examples.

- Sensitivity to a particular model parameter might be the focus for second round parameter searches and the evidence systematically reviewed as a subsequent research project.

## 6. REFERENCES

1. NICE Guide to the methods of technology appraisal. 2008.  
<http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf>.
2. Sheldon, T. Problems of using modelling in the economic evaluation of health care. *Health Economics* 1996; 5(1):1-11.
3. Cooper, N., et al. Use of evidence in economic decision models: practical issues and methodological challenges. *Health Economics* 2007; 16:1277-1286.
4. Philips, Z., et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technology Assessment* 2004; 8(36).
5. Coyle, D., Lee, K. Evidence-based economic evaluation: how the use of different data sources can impact results. Donaldson, C, Mugford, M, Vale, L, editors. 2002. London, BMJ Books. Evidence-based Health Economics.
6. Shemilt, I., Mugford, M., Byford, S., Drummond, M., Eisenstein, E., Knapp, M. et al. Incorporating economics evidence. In: Higgins J., Green S., eds. *Cochrane Handbook for systematic reviews of interventions*. John Wiley and Sons; Chichester (UK): 2008.
7. Marsh, K. The role of review and synthesis methods in decision models. In: Shemilt I., Mugford M., Vale L., Marsh K., Donaldson C., eds. *Evidence-based decisions and economics: health care, social welfare, education and criminal justice*. Wiley-Blackwell; Oxford: 2010.
8. Paisley, S. Classification of evidence used in decision-analytic models of cost-effectiveness : a content analysis of published reports. *IJTAHC* 2010; 26(4):458-462.
9. Eddy, D. Technology assessment: the role of mathematical modeling. Assessing medical technology. *National Academy Press Washington, DC* 1985.
10. Brennan, A., Akehurst, R. Modelling in Health Economic Evaluation: What is its Place? What is its Value? *Pharmacoeconomics* 2000; 17(5):445-459.
11. Cooper, N., Coyle, D., Abrams, K., Mugford, M., Sutton, A. 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):245-250.
12. Chilcott, J., Tappenden, P., Rawdin, A., Johnson, M., Kaltenthaler, E., Paisley, S. et al. Avoiding and identifying errors in health technology assessment models: qualitative study and methodological review. *Health Technology Assessment* 2010; 14[25].
13. Robinson, S. Conceptual modelling for simulation Part I: definition and requirements. *Journal of the Operational Research Society* 2008; 59:278-290.
14. Law, A., Kelton, W. Simulation Modeling and Analysis. 3rd ed. McGraw-Hill: New York; 2000.

15. Lacy, LW., Randolph ,W., Harris ,B., Young Blood ,S., Sheehan, J., Might, R., Meta, M. Developing a consensus perspective on conceptual models for simulation systems. Proceedings of the 2001Spring Simulation Interoperability Workshop. 2001. [www.sisostds.org](http://www.sisostds.org) (accessed May 2007).
16. Sargent RG. Validation and verification of simulation models. Proceedings of the 2004 Winter Simulation Conference. 2004.
17. Centre for Reviews and Dissemination. Systematic reviews : CRD's guidance for undertaking review in health care. 3 ed. University of York, Centrefor Reviews and Dissemination, York; 2009.
18. Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2 [updated September 2009]. In: Higgins,JPT, Green,S editors. The Cochrane Collaboration; Chicester Wiley-Blackwell; 2009.
19. Glanville, J., Paisley, S. Searching for evidence for cost-effectiveness decisions. In: Shemilt I., Mugford M., Vale L., Marsh K., Donaldson C., eds. *Evidence-based decisions and economics: health care, social welfare, education and criminal justice*. Wiley-Blackwell; Oxford: 2010.
20. Golder, S., et al. Populating decision-analytic models: the feasibility and efficiency of database searching for model parameters. *International Journal of Health Technology Assessment in Health Care* 2005; 21(3):305-311.
21. Paisley S. Identification of evidence for decision-analytic models of cost-effectiveness. [ University of Sheffield; 2011.
22. Ara, R., Tumor, I., Pandor, A., Duenas, A., Williams, R., Wilkinson, A. et al. Ezetimibe for the treatment of hypercholesterolaemia: a systematic review and economic evaluation. *Health Technol Assess* 2008; 12(21).
23. Coyle, D., Lee, K., Cooper, N. Use of evidence in decision models. In: Shemilt I., Mugford M., Vale L., Marsh K., Donaldson C., eds. *Evidence-based decisions and economics: health care, social welfare, education and criminal justice*. Wiley-Blackwell; Oxford: 2010.
24. InterTASC Information Specialists' Sub-Group Search Filter Resource. 2011; available from <http://www.york.ac.uk/inst/crd/intertasc/index.htm>
25. Brazier, J., Papaioannou, D., Cantrell, A., Paisley, S., Herrmann, K. Identifying and reviewing health state utility values for populating decision models. In: Shemilt I., Mugford M., Vale L., Marsh K., Donaldson C., eds. *Evidence-based decisions and economics: health care, social welfare, education and criminal justice*. Wiley-Blackwell; Oxford: 2010.
26. Papaioannou, D., Brazier, J., Paisley, S. Measurement and valuation of health: identification, review and synthesis of HSUV values- a technical support guide : report by the Decision Support Unit. 2011.

27. Golder, S., Loke, Y. Sources of information on adverse effects: a systematic review. *Health Information & Libraries Journal* 2010; 27:176-190.
28. O'Hagan, A., Buck, C., Daneshkhah, A., Eiser, J., Garthwaite, D., Oakley, J. et al. *Uncertain Judgements: Eliciting Experts' Probabilities*. Wiley, Chichester; 2006.
29. Stein, K., Dyer, M., Crabb, T., Milne, R., Round, A., Ratcliffe, J. et al. An Internet "Value of Health" Panel: recruitment, participation and compliance. *Health and Quality of Life Outcomes [Accepted]* 2011.
30. Health Information Research Unit. Hedges Project. *HIRU* 2011; Available from [http://hiru.mcmaster.ca/hiru/HIRU\\_Hedges\\_home.aspx](http://hiru.mcmaster.ca/hiru/HIRU_Hedges_home.aspx)
31. Pirolli, P., Card, S. Information foraging. *Psychological Review* 1999; 106(4):643-675.
32. Pandor, A., Eggington, S., Paisley, S., Tappenden, P., Sutcliffe, P. The clinical and costeffectiveness of oxaliplatin and capecitabine for the adjuvant treatment of colon cancer: systematic review and economic evaluation. *Health Technol Assess* 2006; 10(41).
33. Longworth L. Guidelines for the reporting of evidence identification and selection in decision models: the NICE perspective. *MRC HSRC Funded Working Group on the Use of Evidence in economic Decision Models, Leicester* 2005 Available at <http://www2.le.ac.uk/members/njc21/mrc-hsrc-workshop/>
34. Grant, M., Booth, A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal* 2009; 26:91-108.
35. Gannon, R., Ciliska, D., Thomas, H. Expediting systematic reviews: methods and implications of rapid reviews. *Implementation Science* 2010; 5(56).
36. Watt, A., et al. Rapid reviews versus full systematic reviews: an inventory of current methods and practice in health technology assessment. *International Journal of Technology Assessment in Health Care* 2008; 24(2):133-139.
37. Moher, D., Jadad, A., Nichol, G., Penman, M., Tugwell, P., Walsh, S. Assessing the quality of randomized controlled trials: An annotated bibliography of scales and checklists. *Controlled Clinical Trials* 1995; 16:62-73.
38. Jadad, A., Moore, R., Carroll, D., Jenkinson, C., Reynolds, D., Gavaghan, D. et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Controlled Clinical Trials* 1996; 17(1):1-12.
39. Moher, D., Jadad, A., Tugwell, P. Assessing the quality of randomized controlled trials: Current issues and future directions. *International Journal of Technology Assessment in Health Care* 1996; 12:195-208.
40. Deeks, J., Dinnes, J., D'Amico, R., Sowden, A., Sakarovich, C., Song, F. et al. Evaluating non-randomised intervention studies. *Health Technology Assessment* 2003; 7:27.

41. Wells, G., Shea, B., O'Connell, D., Petersen, J., Welch, V., Losos, M. et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses. *Ottawa: Department of Epidemiology and Community Medicine, University of Ottawa, Canada* 2005; Available from [http://www.ohri.ca/programs/clinical\\_epidemiology/oxford.htm](http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm)
42. Higgins J, Altman D, Sterne J. Assessing risk of bias in included studies. In: Higgins J, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration: Chicester, 2011 Available from [www.cochrane-handbook.org](http://www.cochrane-handbook.org).
43. Reeves B, Deeks J, Higgins J, Wells G. Including non-randomized studies. In: Higgins J, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration; Chicester, 2011 Available from [www.cochrane-handbook.org](http://www.cochrane-handbook.org).
44. Guyatt, G., Oxman, A., Schunemann, H., Tugwell, P., Knottnerus, A. GRADE guidelines: A new series of articles in the *Journal of Clinical Epidemiology*. 2011; 64 (4): 380-382.
45. Brunetti, M., Shemilt, I., Pregno, S., Vale, L., Oxman, A., Lord, J. et al. GRADE guidelines: 10. Special challenges - quality of evidence for resource use. *Journal of Clinical Epidemiology*. 2011; (in press)
46. Shemilt, I., Vale, L., Ruiz, F., Drummond, M., Brunetti, M., Lord, J. GRADE guidelines: 17. Going from evidence to recommendations - evidence profiles and decision models for economic evaluation. *Journal of Clinical Epidemiology*. 2011; (in press)
47. Shemilt, I., Mugford, M., Vale, L., Marsh, K., Donaldson, C., Drummond, M. Evidence synthesis, economics and public policy. *Research Synthesis Methods* 2010; 1(2):126-135.