

# Guidance for the Conduct and Reporting of Decision and Simulation Models in the Context of Health Technology Assessment

**Prepared for:**

Agency for Healthcare Research and Quality  
U.S. Department of Health and Human Services  
540 Gaither Road  
Rockville, MD 20850  
[www.ahrq.gov](http://www.ahrq.gov)

This information does not represent and should not be construed to represent an Agency for Healthcare Research and Quality or Department of Health and Human Services determination or policy.

**Contract No.**

**Prepared by:**

**Investigators:**

**AHRQ Publication No.**  
<MonthYear>

**This document is in the public domain and may be used and reprinted without permission except those copyrighted materials noted, for which further reproduction is prohibited without the specific permission of copyright holders.**

Suggested Citation: pending

**None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this report.**

William Lawrence, M.D., M.S.  
Center for Outcomes and Evidence  
Agency for Healthcare Research and Quality  
540 Gaither Road  
Rockville, MD 20850

This report is based on research conducted by an Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. xxx-xxxx-xxxxx-x). The findings and conclusions in this document are those of the author(s), who are responsible for its content, and do not necessarily represent the views of AHRQ. No statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help clinicians, employers, policymakers, and others make informed decisions about the provision of health care services. This report is intended as a reference and not as a substitute for clinical judgment.

This report may be used, in whole or in part, as the basis for the development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

## Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies and strategies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

To improve the scientific rigor of these evidence reports, AHRQ supports empiric research by the EPCs to help understand or improve complex methodologic issues in systematic reviews. These methods research projects are intended to contribute to the research base in and be used to improve the science of systematic reviews. They are not intended to be guidance to the EPC program, although may be considered by EPCs along with other scientific research when determining EPC program methods guidance.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality. The reports undergo peer review prior to their release as a final report.

We welcome comments on this Methods Research Project. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by e-mail to [epc@ahrq.hhs.gov](mailto:epc@ahrq.hhs.gov).

Richard Kronick, Ph.D.  
Director  
Agency for Healthcare Research and  
Quality

Yen-Ping Chiang, Ph.D.  
Acting Deputy Director, Center for  
Evidence and Practice Improvement  
Agency for Healthcare Research and  
Quality

Stephanie Chang, M.D., M.P.H.  
Director, EPC Program  
Center for Evidence and Practice  
Improvement  
Agency for Healthcare Research and  
Quality

William Lawrence, M.D., M.S.  
Task Order Officer  
Center for Evidence and Practice  
Improvement  
Agency for Healthcare Research and  
Quality

# Contents

Introduction .....	5
Scope of the guidance .....	5
Goals of modeling in EPC reports.....	6
Decision and Simulation Modeling Steps .....	7
When is Decision and Simulation Modeling Worth the Effort? .....	7
Guidance Development Process .....	8
Terminology and definitions .....	8
Principles for best practice.....	8
Recommendations for decision and simulation modeling in systematic reviews .....	11
Conceptualization and Structure.....	11
Data.....	13
Consistency: Anticipating and correcting errors, model verification .....	17
Consistency: Face validity, conceptual model validation .....	17
Consistency: Operational model validation, confronting models with data .....	17
Consistency: Predictive model validation .....	18
Reporting and interpreting results.....	18
References .....	20

# Tables

Table 1. Definition of terms .....	9
Table 2. Principles for best practice in decision and simulation modeling .....	10
Table 3. Conceptual definitions for the structure, data, consistency, reporting framework.....	10

# Introduction

Despite rigorous systematic reviews of efficacy and effectiveness of health care interventions, patients, providers and policymakers may remain in doubt about what they should do because of uncertainty, tradeoffs, and values. First, residual uncertainty may remain for meaningful patient-relevant outcomes from surrogate outcome measures or limited time followup or for subgroups from inadequate sample or inclusion/exclusion criteria. Second, tradeoffs occur, e.g., the U.S. Preventive Services Task Force (USPSTF) analysis of mammography for women in their 40s suggests a statistically significant reduction in breast cancer death but also potential harms, namely radiation exposure, overdiagnosis and overtreatment.<sup>1</sup> Thus, optimal decisionmaking for individuals and populations may depend on their values (or preferences) for the outcomes. Lastly, just as the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group considers resource utilization in guideline development, the USPSTF has examined modeling to estimate resource consumption for its recommendations on cancer screening. This document was designed to extend current Evidence-based Practice Center (EPC) guidance on developing methodological guidance for decision and simulation modeling.

We discuss principles and good practice recommendations for decision and simulation modeling. We believe this work to apply generally, but for concreteness, we place emphasis on models that could accompany systematic reviews produced by the EPC program. Such modeling exercises may be used to structure investigators' thinking, and facilitate the communication of assumptions and results; synthesize data from disparate sources; inform decisionmaking; make predictions; or infer the impact of manipulations.

The goals of the guidance are to encourage the use of good modeling and reporting practices, while not being too prescriptive about how to develop specific models. We deemphasize issues specific to economic modeling, because economic assessments are not a priority of the EPC program.

## Scope of the guidance

A model is a construct that represents salient aspects of reality in a simplified way. Models are physical (e.g., a scaled-down airplane wing tested in a wind tunnel) or theoretical constructs (e.g., a mathematical description of the flow of air around an airplane wing). Models that could be prepared in conjunction with systematic literature reviews are exclusively theoretical in nature; for this reason we do not discuss physical models further. The starting point for most theoretical models is a conceptual model, a simplified natural language or pictorial representation of reality. The analytic frameworks that are used to guide the conduct of systematic reviews prepared by the EPC program are conceptual models that often function as schematics of an underlying decision problem.<sup>2-5</sup> For example, the analytic frameworks used in reviews of diagnostic tests, which often resemble decision trees.<sup>6</sup> Although conceptual modeling is a prerequisite for the development of mathematical (quantitative) models, our current guidance focuses exclusively on the latter. Readers interested in the use of conceptual models in systematic reviews can consult relevant chapters of the methods guide.

Mathematical models are a large and diverse group of models that use variables, together with mathematical symbols that represent relationships between variables. The most common quantitative models encountered in systematic reviews are multivariable regression models (for

primary data analysis and meta-analysis). These models and other related techniques (e.g., neural networks) that aim to describe how a response (dependent variable) changes conditional on covariates (independent variables) are types of behavioral models. They describe how the response varies over values of the covariates, without necessarily referring to assumptions about the underlying mechanisms.<sup>a</sup> The literature addressing these models is vast (e.g., in statistics or computer science) and not covered in this guidance. Instead, we address structural models, which attempt to capture mechanistic relationships among their components. Structural models include declarative (e.g., Markov models), functional (e.g., compartmental models), and spatial models (e.g., geographic information systems data models). In applied work, elements of these structural model subtypes are commonly combined (multi-models).

## Goals of modeling in EPC reports

This document does not provide detailed guidance to help investigators decide whether decision or simulation modeling should be undertaken. Issues related to the appropriateness of modeling in EPC reports are addressed by existing guidance and are not covered in this document.<sup>7,8</sup> However, we briefly consider the potential goals of modeling when performed in conjunction with systematic reviews.<sup>9-18</sup>

- **To inform decisionmaking under uncertainty:** The decisions that can be informed by modeling, even in the relatively narrow context of systematic reviews, are extremely diverse.<sup>9,10</sup> They include decisions about patient-level care, the licensing of drugs or devices, healthcare policy decisions for populations, and decisions about the need to conduct additional research.
- **To structure investigator’s thinking, and facilitate the communication of data, assumptions and results:** modeling can help investigators organize their knowledge about the topic area, formalize the research question, and communicate assumptions and results to peers (e.g., topic or methodological experts) and stakeholders (e.g., patients, decisionmakers).<sup>10</sup>
- **To synthesize data from disparate sources:** evidence on a specific research question may be available from multiple sources and a single study may contribute to the estimation of more than one model parameter (or functional combinations of parameters). Modeling provides the mathematical tools for synthesizing all evidence and facilitating assessments of consistency between sources. For example, models can be used to combine information from clinical trials of the impact of a treatment on intermediate outcomes can be combined with information from long-term cohort studies that assess of the association of the intermediate outcome with a clinical outcome of interest.
- **To make predictions:** predictions can refer to conditions similar to those already observed (sometimes referred to as “interpolations”; e.g., prediction of outcomes in a new study, similar to an existing one); or the future (forecasts), other populations, or other outcomes (sometimes referred to as “extrapolations”; e.g., predicting outcomes at longer-term followup times based on results of short-term clinical trials). They can also pertain to the prioritization and planning of future research.<sup>19</sup> These predictions may be useful in themselves, even without reference to the anticipated effects of interventions.

---

<sup>a</sup> In some cases behavioral models (e.g., regression) are used to estimate the parameters of structural models.

- **To support causal explanations and infer the impact of interventions:** Related to the above is prediction of the effects of possible or even hypothetical manipulations.<sup>20-22</sup> When used this way, at least implicitly, models are claimed to encode structural causal mechanisms or emulate such mechanisms sufficiently well.

Conveying assumptions, synthesizing evidence, and informing decisions are probably the primary goals of decision and simulation models that would be developed in conjunction with systematic reviews.

## Decision and Simulation Modeling Steps

Because decision and simulation models are used to achieve multiple goals and address diverse research questions, the model development and evaluation process is bound to differ across specific applications. Nonetheless, the key steps to develop quantitative models that are within the scope of this guidance can be identified<sup>23-29</sup>:

1. **Definition of the research question:** specify an answerable research question for the relevant stakeholder.
2. **Model conceptualization:** determine which components of a disease or process need to be represented in the model to address the research question, and describe their relationships.
3. **Data collection and processing:** identify data sources and process data to inform the model.
4. **Model implementation and mathematical manipulation:** ‘solve’ the model using mathematical or numerical analysis methods, or simulation techniques.
5. **Model evaluation:** detect model shortcomings by examining the model, and by comparing its output with data and other similar models.
6. **Reporting and interpretation of results:** present the model findings in a way that addresses the research question.

Model development is an iterative and dynamic process.<sup>27,30</sup> Multiple iterations are typically needed between the phases outlined above because, at each step, the need for changes at earlier phases may become apparent. For example, the availability of some data (possibly preliminary or incomplete) often provides an incentive for modeling and simulation; as the model is conceptualized additional data needs may arise that require further data collection. Similarly, deficiencies that are detected at the evaluation phase may require restructuring of the model, supplemental data collection, or other modifications of the modeling strategy.

## When is Decision and Simulation Modeling Worth the Effort?

Developing decision and simulation models, especially models that can be used to inform complex decisions or understand complex disease processes, is a demanding task. Similarly, choosing between alternative modeling approaches can be difficult because the correct choice is not always obvious early in the modeling process. Also, the same research question may be amenable to multiple modeling approaches, each with distinct strengths and disadvantages.

Although this document, and the cited references, can provide guidance on decision and simulation modeling methods, it is harder to define with certainty the circumstances under which

modeling is worth the investment of time and resources beyond those required for a systematic evidence review. In general, modeling is most useful when the research question is complex, data sources are multiple (and possibly conflicting), outcomes involve trade-offs, and choices are value-laden. The details of the research question, the availability of resources (e.g., analyst time and experience with the related methods), and the potential impact of modeling on future research, clinical practice and policy decisions, should also be considered when deciding about whether modeling efforts are likely to be beneficial.

## **Guidance Development Process**

Development of this guidance document is the culmination of a multistep process of summarizing existing recommendations and soliciting stakeholder input. We first updated and expanded two systematic reviews of recommendations for the conduct and reporting of decision and simulation models, as described in detail in our companion paper.<sup>7,31</sup> This was done with input from a multidisciplinary team of clinical, policymaking, and decision analysis experts. The results of our systematic review were discussed in-person with a panel of 28 stakeholders including patient representatives, providers of care, purchasers of care, payers, policy makers (including research funders and professional societies), and principal investigators. Stakeholders commented on available recommendations and identified gaps, limitations and areas for expansion. The stakeholders reviewed, added to, and prioritized the list of future research recommendations. We subsequently reviewed the websites of 126 international agencies and institutes conducting health technology assessments for their guidance or standards for how they conduct and report decision and simulation modeling, with an emphasis on how systematic reviews incorporated modeling. We [are] solicit[ing] input from senior researchers at EPCs and AHRQ with experience in the methods of conducting decision and simulation models.]

Based on the gathered systematic review evidence on modeling recommendations and guidance, we developed a list of recommendations to serve as guidance for developing models in conjunction with systematic reviews. There are two major types of recommendations: (1) those that follow from principles and are not amenable to empirical testing, and (2) those that can be tested empirically or through simulation.<sup>32</sup> We provide the rationale for each guidance recommendation, evidence that the recommendation should be preferred, or best judgment where adequate evidence is lacking. We have also categorized the recommendations as proposed by Sculpher (2000),<sup>33</sup> Philips (2004),<sup>31,34</sup> and Kuntz (2013),<sup>35,36</sup> by whether they pertain to the model structure, model data, or consistency, and reporting.

## **Terminology and definitions**

**Table 1** defines terms used in the recommendations.

## **Principles for best practice**

We begin by outlining general principles for the conduct and reporting of decision and simulation modeling studies found consistently in our systematic review (**Table 2**). We believe that these principles represent generally accepted rules for sound practice and have used them to guide our more specific recommendations, which are presented in the next section.

Modelers should consider (1) the goals of the modeling exercise; (2) the nature of the phenomena-to-be-modeled; (3) their own abilities in math and computation; and (4) objective constraints in terms of available time, data, or resources. Further, when developing, implementing and ‘solving’ models, one makes many methodological decisions. One should report them explicitly, justify them, and subject them to appropriate stability and sensitivity analyses.<sup>37-39</sup>

**Table 1. Definition of terms**

<b>Term</b>	<b>Explanation and elaboration</b>	<b>Comments</b>
Model	A simplified representation of reality.	We focus on models which represent reality by means of mathematical relationships.
Simulation	A typical process of “solving” the model, especially when analytical solutions are cumbersome or intractable.	Often even analytically tractable models are ‘solved’ with simulation methods.
Model component	An element of a model. Model components may include variables (parameters), health states, agents, processes, and so on.	The descriptor is on purpose generic, to encompass all model types.
Stochastic (aleatory) uncertainty	Statistical uncertainty around the estimates of model variables that are informed by empirical data.	See below.
Structural uncertainty	Uncertainty secondary to our incomplete understanding of the modeled phenomenon. Typically it pertains to functional forms of relationships between model variables. At a more fundamental level, structural uncertainty will always exist, because the ‘true’ relationship between variables in the real world cannot be uncovered from data.	In some cases, the distinction between stochastic and structural uncertainty is a matter of definition. Some structural uncertainty would become stochastic uncertainty, if appropriate data were available to the modeler.
Propagation of (stochastic) uncertainty	The process of obtaining the stochastic uncertainty in model outputs. This derives from the stochastic uncertainty in model inputs. Propagation of uncertainty can be done: <ul style="list-style-type: none"> <li>Analytically, exactly, or up to an approximation (e.g., first order delta method)</li> <li>Numerically, e.g., with forward Monte-Carlo simulations or equivalently with Markov Chain Monte Carlo (MCMC) methods</li> </ul>	It is customary to use the term “probabilistic sensitivity analysis” (PSA) to refer to numerical propagation of uncertainty by means of forward Monte-Carlo methods.  We do not use this term (PSA) in this work to avoid confusion.
Sensitivity analysis	The process of varying model variables over all elements of a set of interesting values and examine impact on results.	Sensitivity analysis has a ‘continuous’ aspect.
Stability analysis	Performing discrete actions and evaluating their impact on results. Examples include <ul style="list-style-type: none"> <li>Changing the structure of the model, e.g., use alternative statistical distributions or different functional forms for relationships between variables</li> <li>Systematically excluding input data (e.g., leave one-study-out in a meta-analysis)</li> </ul>	Stability analysis involves discrete actions.

**Table 2. Principles for best practice in decision and simulation modeling**

The research question and model scope should be clearly defined
The model structure and assumptions should be explicated and justified
Model components and the relationships between them should be defined; the chosen relationships between model components should be justified
The model should be informed by data; data selection, analysis and interpretation should be aligned with the research question and model scope
The model should incorporate stochastic uncertainty in inputs
Sensitivity and stability analyses to evaluate the impact of modeling decisions should be undertaken and reported
Models should be evaluated with respect to their ability to address the research question and within their stated scope
Modeling methods should be transparent; adequate details about the structure, data, and evaluation methods should be reported so that the modeling process is replicable

The following subsections provide detailed guidance on the conduct of decision and simulation modeling in the context of systematic reviews, organized by ‘conceptualization and structure’, ‘data’, ‘consistency’, and ‘reporting’.<sup>7;31;33;34;36</sup> Briefly, structure and data are what constitute the model proper; consistency refers to an assessment of the model against its stated goals; and reporting considers issues related to results reporting and presentation. **Table 3** provides operational definitions and examples for these areas of modeling.

**Table 3. Conceptual definitions for the structure, data, consistency, reporting framework**

<b>Recommendation Areas</b>	<b>Description of what is encompassed</b>	<b>Examples</b>
<b><i>Conceptualization and structure</i></b>	<i>Conceptualization</i> pertains to the decision to use modeling, and the delineation of the perspective and scope. <i>Structure</i> pertains to variables, health states and other components of the model, and how they relate to each other (i.e., the mathematical scaffold of the model).	Consider a discrete-time Markov model. The disease states, variables informing transition probabilities, the mathematical relationships among the variables, the time horizon of the model and so on characterize the model’s structure.
<b><i>Data</i></b>	Model inputs. May be obtained through empirical investigation, systematic elicitation of opinion, or best judgment/introspection.	Estimates for variables in the model, e.g., treatment effects, transition probabilities, costs, and utility weights.
<b><i>Consistency</i></b>	Whether the model achieves its stated goals, and the processes of assessing this.	Determination of whether the model has logical errors and whether the model output is consistent with expert opinion, observed data, or other models.
<b><i>Reporting</i></b>	Summarizing model output to achieve modeling (e.g., further one’s understanding the topic, inform decisionmaking).	Incremental cost effectiveness curves (to present the results of cost-effectiveness analyses), risk diagrams (to represent model-based risk assessments), and tornado diagrams (to summarize sensitivity analyses).

This guidance is provided to facilitate the use of decision and simulation modeling in conjunction with systematic reviews, particularly as they are prepared within the AHRQ EPC program. The recommendations provide general guidance about conceptualizing, specifying, implementing, and evaluating decision and simulation models. It is not possible to provide detailed recommendations about which structures to use in which cases, or instructions about the (technical) implementation. Interested readers should consult some of the numerous books, technical reports available on this topic (several of which are cited below), including the vast literature on decision and simulation modeling in areas other than healthcare. Other general guidance documents, which were used as sources for the current set of recommendations, should also be consulted.<sup>7;27;31;33;34;36;40-48</sup>

## **Recommendations for decision and simulation modeling in systematic reviews**

### **Conceptualization and Structure**

**The research question should be explicit. The decision to use modeling for addressing the research question should be described and justified.**

As described earlier, modeling will be useful for many research questions, especially when the research questions are not directly answerable by existing empirical data. Defining the question at hand and the objective of the analysis may require using literature-based information, expert knowledge, and input from Key Informants and Technical Experts.<sup>30</sup>

**The choice of perspective depends on the research question to be addressed and the relevant decisionmaker. There is no *a priori* preferred modeling perspective.**

The modeling perspective determines the methods for choosing and handling consequences, values, and costs in the model; thus it should depend on the research question and the decisionmaker.<sup>49</sup> For example, when modeling a specific clinical interaction where the decisionmaker is an identifiable patient, the appropriate perspective is that of the individual patient. In contrast, when the goal is to inform the decisionmaking of a public payer or a federal agency, one should prefer a payer or societal perspective.<sup>50</sup>

That said, the societal perspective (which considers impact on sectors beyond healthcare and includes time costs, opportunity costs, and community preferences) may allow for a more complete accounting of benefits and costs.<sup>51</sup> For this reason, it has been recommended as the appropriate perspective in “base case” analyses.<sup>52</sup> Obtaining appropriate data for modeling from a societal perspective is challenging and may be difficult to do well.<sup>51;53;54</sup> Simplifications have been advocated, such as ignoring impact on outside non-healthcare sectors, to represent a ‘partial societal perspective’.

**The model’s scope should be described and justified. The model’s scope should be consistent with the research question and the model’s perspective.**

A model represents only some aspects of the phenomena under study. The research question defines how elaborate the modeling should be, and what aspects of reality it chooses to represent or omit (for parsimony). For example, many research questions in healthcare pertain to length of life, and mortality outcomes should be within the scope of models answering these questions.

More broadly, the scope of the model includes defining the condition or disease of interest, populations, risk factors, diagnostic or therapeutic interventions. For decision models one should also define the decision-relevant quantities, the decision (optimality) criteria, and the decisionmaking perspective. This is akin to defining a systematic reviews' eligibility criteria (Population, Intervention, Comparator, Outcomes).

**The mathematical structure of the model and its implementation (computation) should correspond to the research question, the model's scope, and the decisionmaking perspective. The rationale for the choice of a mathematical structure should be provided, and structural assumptions should be explicated and justified.**

The preferred model structure depends on the research question and the model scope. The model structure should reflect the current understanding of the topic being modeled (e.g., disease prognosis and treatment effects, diagnostic test application, public health interventions) Disease states and transitions, or functional relationships should reflect understanding of the course of the disease. Detailed guidance on choosing among alternative mathematical structures (and computational implementations) is beyond the scope of this work. Readers are referred to the extensive technical literature available in healthcare and other fields.<sup>16;25;26;40;48;55-73</sup> Of note, relatively simple models (e.g., decision trees, time-homogeneous Markov chain-based models) may be appropriate for use in the setting of most EPC evidence reports, particularly when the goal of modeling is to contextualize the evidence and extend review findings.

**The model should allow for comparisons among all interventions that are relevant to the research question, model scope, and decisionmaking context.**

In many cases the goal of modeling is to inform decisionmaking about the implementation of an intervention (e.g., a specific treatment or policy) or to assess the impact of modifying the levels of risk factor or an exposure (e.g., reducing cholesterol, or eradicating a disease agent from the environment). In such cases, the model should allow the inclusion of all relevant and feasible interventions (or exposures). In general, feasible options should not be excluded from the model; in the rare case that such exclusions are deemed necessary, they should be carefully justified.

**The time horizon should be long enough to allow all meaningful outcomes to be evaluated fully.**

When comparing alternative interventions, the time horizon should be long enough to allow the manifestation of differences in important outcomes. In some cases, a short time horizon may be adequate to compare interventions (e.g., when modeling the effectiveness of interventions for alleviating symptoms of the common cold); in many cases a life-time horizon is needed, particularly when modeling the effects of long-term treatment of chronic disease. Choice of long-term time-horizons has implications for the data used to populate models (e.g., life-time horizons almost always require the extrapolation of treatment effects well beyond the followup duration of available clinical trials).

**When deciding how to handle time, space, interagent interactions, and health states, one should consider the “nature” of the modeled phenomena, and the**

### **convenience of, and calculation errors associated with, alternative modeling choices.**

For example, when deciding how to deal with time, we have three options: (1) do not model it explicitly (integrate it out, as for example in decision trees); (2) model it as a continuous quantity (as in differential-equation-based dynamic systems); (3) model it as a discrete quantity (as in discrete-time Markov processes). Whether time is modeled as continuous or discrete should be guided by the specifics of the system being modeled and the process for making decisions (e.g., whether decisions are made in a continuous fashion or only at specific timepoints).<sup>74</sup> In some cases where discrete modeling may be appropriate (e.g., modeling the occurrence of an outcome when measurement is possible only at specific intervals), continuous time models may offer convenient mathematical approximations. The converse may be the case in problems of a continuous nature that can be approximated by more tractable discrete-time models (e.g., models describing the development of epidemics). For discrete-time models, the cycle length should match the speed of changes in the system being modeled (e.g., the natural history of the disease, or the anticipated temporal evolution of a system).

Analogous considerations pertain to modeling vs not modeling spatial location; inter-agent interactions (interactions between people); and modeling health states in various degrees of granularity (e.g., disease severity).

### **The targeted level of complexity (or parsimony) should be determined by the research question and the model's scope.**

Models should be complex enough to capture all pertinent aspects of the system being modeled but not more ('rule of reason').<sup>75</sup> At the same time, models should be as simple as possible to facilitate timely development, error checking, and validation. Simple models are also generally more accessible to nontechnical stakeholders of the modeling process, and their results can be communicated more easily. The tradeoff between simplicity and complexity should be driven by considerations related to the research question and the context in which model results will be used.

## **Data**

### **Methods for identifying, and analyzing data should be described. Data choices should be driven by the research question and the model's scope and structure. All data sources should be reported clearly and appropriate references should be provided.**

To enhance transparency and face validity, the source of each data element should be identified fully. Particularly for data that is not derived from systematic review and meta-analysis, the rationale for why the given value was chosen. This applies both to the base case data and, for each data element, the range of values to be tested in sensitivity analyses.

### **Estimates for influential variables should be obtained from systematic processes (systematic review).**

For decision and simulation models prepared jointly with a systematic review of studies of interventions, the summary estimates of treatment effects and related parameters should provide the data to inform the relevant model parameters. In particular, model parameters likely to have a large influence on model results should be identified through a systematic and replicable process

that aims to minimize bias.<sup>76-80</sup> In most cases this will mean conducting systematic literature reviews to inform influential parameters. However, in many cases only part of the evidence retrieved by the systematic review will be appropriate for use in the model. The research question, decisional context, and goals of modeling should inform the choice of which studies to be included and the choice of synthesis methods.<sup>80-87</sup>

Decision and simulation models typically require data on multiple parameters that are not collected in systematic reviews, such as prevalence, incidence, costs, and utilities. Appropriate sources of such data can include registries and other large observational studies, studies found through a nonsystematic approach, stakeholder panel opinions, and domain expert judgment. When retrieving and processing data, modelers make a large number of methodological decisions that can appreciably impact results. Thus, all such decisions should be reported and justified. Supplementary material describing detailed methods and data sources can be made available electronically.

### **Obtaining estimates for model variables should follow epidemiological and statistical principles.**

All major assumptions and methodological choices should be reported and justified. Several excellent sources provide detailed guidance on data management and manipulation, exploratory data analysis, inference, estimation, and related computational techniques. When multiple studies contribute information on a parameter of interest (e.g., treatment effectiveness, prevalence of disease, accuracy of a diagnostic test) evidence should be synthesized across studies using appropriate meta-analytic methods.<sup>77;78</sup> Detailed guidance on the conduct of quantitative synthesis for different types of data structures is beyond the scope of this document; interested readers should refer to the relevant EPC guidance, and the many sources on meta-analysis and evidence synthesis.<sup>82;88-113</sup>

### **A “best evidence approach” should be used when selecting data sources for model parameters.**

Data from randomized trials cannot be used to inform all model parameters because (1) some parameters are best estimated from alternative study designs (e.g., the prevalence of a risk factor is best estimated from a sampling survey of a representative population; the test performance of a diagnostic test is best estimated from a cohort study); (2) available randomized trials may not be sufficiently applicable to the population to be modeled (e.g., trials may enroll highly selected populations, provide inadequate information for subgroups of interest, or have short followup duration); and (3) trials may not be available. In all these cases, evidence from other study designs will have to be included in the model. Researchers contemplating decision and simulation modeling in the context of systematic reviews will have to select and appraise appropriate study types for each model parameter.<sup>114</sup> General guidance on “best evidence” strategies in systematic reviews is provided by a recent EPC Methods Research report.<sup>115</sup>

### **The risk of bias of the available data should be assessed and accounted for when obtaining estimates for model parameters.**

One should avoid using unadjusted, incompletely adjusted, or inappropriately adjusted results that are potentially biased simply because no other information is available.<sup>116;117</sup> Instead, modelers should consider using methods that allow the adjustment of study results to account for all sources of bias and related uncertainty (i.e., multiple bias adjustment, possibly in the setting of sensitivity analysis).<sup>118;119</sup>

The factors that contribute to a study's risk of bias depend on the specific modeling context and the study designs considered (e.g., sources of bias for surveys are distinct for those of randomized and non-randomized studies). These factors should be assessed for each study. For modeling, it is generally not adequate to assess the risk of bias of individual studies (or entire bodies of evidence). Because models typically are specified with respect to "true" parameters, it is desirable that model inputs be 'corrected' (adjusted) for biases.

The direction and magnitude of the bias associated with each item, the uncertainty around their effect, and the relationship between bias items (e.g., whether bias effects are additive or nonadditive effects), should be incorporated in the analyses. In most cases, the effect of bias items (also known as "bias parameters") cannot be identified from study data; thus, one should use methods that incorporate external information (empirical or judgmental). Extensive literature exists on the assessment of specific risk of bias items for individual studies, as well as methods for multiple bias modeling (i.e., bias adjustment and risk analysis).<sup>116;118;120-132</sup>

### **Formal elicitation methods should be used to quantify expert opinion and the associated uncertainty.**

When no empirical evidence is available for parameters of interest, modelers will have to rely on expert opinion. Current EPC processes for Key Informant and Technical Expert engagement (during the development, refinement and conduct of systematic reviews), or similar processes, can be leveraged to incorporate formal opinion elicitation methods. Modelers should be aware that elicitation methods (e.g., the framing of questions) can influence the information that is obtained, particularly when the subjects of the elicitation process have labile values for the parameters of interest.<sup>133</sup> Interested readers should consult the extensive literature on elicitation methods for different types of parameters.<sup>134-143</sup>

### **Assumptions required for extrapolating from existing data should be reported and justified. They should also be subjected to stability and sensitivity analyses.**

A particular challenge arises when there is need to extrapolate beyond the observed data (e.g., to longer followup periods, or to other populations). Such extrapolations are based on assumptions about unobserved data. These assumptions should be reported and justified; they should also be subjected to sensitivity analyses (e.g., assessing a range of values for the parameters of the chosen survival distributions) and stability analyses (e.g., using alternative survival distributions when extrapolating survival times).<sup>144</sup>

### **The assumptions required for transporting information across studies to a common (new) setting should be described, justified and subjected to stability and sensitivity analyses.**

Decision and simulation models often use data obtained from diverse sources.<sup>15</sup> In fact, modeling is often used with the explicit goal of synthesizing information from diverse domains (e.g., treatment effect estimates from trials of selected populations may be combined with natural history information from large observational cohorts). In such cases, the validity of modeling results depends on the validity of assumptions about the transportability of effects across domains. These assumptions should be identified explicitly and justified based on theoretical considerations and the understanding of the underlying mechanisms.<sup>145;146</sup> Consideration should be given to formal (causal) methods for assessing the transportability of results across domains.<sup>147-152</sup>

## **Analyses should take into account heterogeneity (nonrandom variation) in all parameters.**

As a general principle, decision and simulation models should account for clinical heterogeneity, defined as nonrandom (systematic) variation in parameters of interest.<sup>153;154</sup> Attempts should be made to explain heterogeneity via appropriate statistical methods (e.g. subgroup or regression analyses) by incorporating information on determinants of variability. Because our current understanding of any topic is likely to be incomplete (e.g., important modifiers of effect may be unknown) and because data unavailability may limit our ability to explore heterogeneity (e.g., well-known modifiers may not be measured or reported in published studies), models should also allow for residual (unexplained) variation.

Unexplained heterogeneity arises very often in meta-analyses of treatment effects using published (group level) level data. In such cases efforts to explain heterogeneity rely primarily on meta-regression methods and residual heterogeneity is accounted for using random effects models.<sup>87;155-158</sup> Modelers should be aware that random effects models can “average over” and obscure, important data patterns and – contrary to popular belief – are not always more conservative than fixed effect models.<sup>159;160</sup> Person-level data can allow decision and simulation models to meaningfully incorporate heterogeneity;<sup>161-168</sup> however, their use is very uncommon in systematic reviews prepared by EPCs or meta-analyses published in peer-reviewed journals.<sup>169</sup>

## **Models should propagate the uncertainty in inputs to outputs.**

Data analysis should allow unbiased parameter estimation and appropriately account for parameter uncertainty from model inputs to model outputs.<sup>27;153;154;170-179</sup> Sometimes this can be done analytically, either exactly, or approximating up to an order of error, such as with the delta method. In most cases it is computationally convenient to propagate uncertainty with numerical methods, typically with a forward-Monte-Carlo approach. It is customary to use the term “probabilistic sensitivity analysis” (PSA) to refer to numerical propagation of uncertainty by means of forward Monte-Carlo methods. We do not use this term in this work.

Detailed guidance on the conduct of probabilistic analyses is available elsewhere.<sup>153;154;173;180-186</sup> Of note, probabilistic methods for incorporating and propagating uncertainty in models do not eliminate the need for stability and sensitivity analyses. For example, the use of a specific probability distribution to represent uncertainty around a model input does not eliminate the need to assess the impact of using alternative probability distributions (stability analysis) or the need to assess the impact of evaluating permutations of the distributional parameters (sensitivity analysis).

Depending on the goal of the model, in rare cases it may not be desirable, or necessary, to perform analyses that propagate uncertainty. For example, for decisional problems where optimality is judged with minimax or maximin criteria, an analysis of bounds (extreme values) may suffice. Furthermore, if substantial uncertainty exists about the appropriate distributional form for estimates of model inputs, it may be futile to insist on probabilistic analyses, and may be appropriate to set more modest and attainable goals for the modeling exercise (e.g., use models to gain insights or to communicate implications). When such cases arise, analysts should provide the rationale for not using probabilistic analyses.

## **Consistency: Anticipating and correcting errors, model verification**

**After implementing the model, attempts should be made to detect errors in its logic and implementation.**

Errors are unavoidable in any nontrivial model.<sup>187</sup> Mistakes in the research question formulation, the model structure, incorporation of data, or software implementation can become apparent during any phase of model development, and may require revising the structure, or collecting additional data.<sup>188;189</sup> Errors in implementation can be challenging to detect, and can also have important consequences. The risk of mistakes in question formulation and model structure can be reduced by adhering to some of the principles outlined earlier in this document (e.g., consulting with topic experts, using a conceptual model to guide the mathematical model implementation), together with transparent reporting of methods and results and the use of teams with sufficient expertise. Several checking techniques have been advocated for healthcare-related models (e.g., sensitivity analysis, extreme value analysis).<sup>187</sup> In addition, software production techniques such as unit testing, code review (review of one programmer's work by another team member), paired programming (i.e., one programmer's coding is monitored by another in real time) can be considered. Duplicate implementation of the same model by an independent team can also be used to identify errors in coding. Because these strategies can substantially increase the costs of model development, their use should be balanced against the modeling goals, model complexity, and anticipated frequency and impact of errors.

## **Consistency: Face validity, conceptual model validation**

**Topic experts should be invited to review the model structure and outputs and to judge whether they appear consistent with their expectations. Counterintuitive model results should be described and explained.**

An evaluation of the model and its results by a group of topic experts can alert modelers to the presence of deficiencies in model structure or data.<sup>188</sup> Counterintuitive model results ("paradoxical findings") may indicate errors. If an error has been ruled out, they should be described and explained, with reference to model structure, available data, and current understanding of the modeled phenomena.

## **Consistency: Operational model validation, confronting models with data**

**The consistency between model outputs and the data on which the model was based should be evaluated.**

A combination of graphical and statistical methods should be used to compare model outputs with expected results.<sup>190-195</sup> For parameters that are identifiable using available data, model validation is essentially an assessment of model fit. As such, comparisons of observed versus expected values (graphical or statistical) can be used to identify potential areas of improvement in model structure and assumptions.

### **Data should not be withheld from model development with the sole purpose of assessing model validity.**

Generally, data should not be withheld during model development for the purpose of using the data for model validation. Using all available data during model development (for parameter estimation) is more efficient (because all information is incorporated), allows appropriate handling of correlated inputs, and permits the assessment of consistency across available sources of evidence.<sup>196</sup> Resampling-style methods to compare the “fit” of the model to available data and the detection of outlying or influential observations can be used. Additional model validation methods are available under a Bayesian framework. Even when parameters are not identifiable by available data (e.g., parameters related to the rate of tumor growth in cancer microsimulation models) holding out data on identifiable parameters (or on functional combinations of identifiable and nonidentifiable parameters) is, in general, less efficient than joint modeling.

### **If multiple models addressing the same research question are available their results should be compared and any discrepancies explained.**

Results from independently developed models addressing the same research question can be available by design (comparative modeling) or happenstance (e.g., multiple teams working on the same research question simultaneously).<sup>197</sup> If such independent models are known about or identified through literature review, then outputs from different models should be compared as part of cross-model validation and any discrepancies need to be explained, with reference to the model structure and data inputs of each model.

## **Consistency: Predictive model validation**

### **The appropriateness of using future observations to evaluate a model depends on the research question and the intended use of the model.**

A comparison of model output with future empirical results (unavailable at the time model development) is not an appropriate method of evaluation for some models.<sup>33</sup> In general, models used to guide decisionmaking or to contextualize and synthesize evidence at a specific point in time should generally not be evaluated with respect to their ability to predict the results of future empirical research.<sup>18;33;198</sup> The majority of models developed in conjunction with systematic evidence reviews are likely to belong to this category. However, for models intended as predictive or forecasting tools, predictive validation is an important component of model assessment.

### **Models should be updated as understanding of disease improves, and as new interventions and empirical data become available.**

Models should be updated as our understanding of disease mechanisms (causal agents, natural history), potential interventions (e.g., new treatments or variations of existing treatments) and their associated benefits and costs evolve. The model structure and its software implementation need to be flexible enough to accommodate this updating process.

## **Reporting and interpreting results**

### **The model structure, data used to populate it, and results should be transparent.**

Information about the model structure and data used to populate it should meet the standards of reproducible research.<sup>43;46;47;188;199</sup> This is particularly important for models that are supported

by public funds (e.g., models that can be created in conjunction with EPC evidence reports) or models used to inform decisions that affect public policy. Transparent reporting will generally involve a detailed technical description of the model structure, an implementation of the model in computer code (or equivalent formats, such as spreadsheet files), a detailed tabular presentation of model inputs (e.g., probability distributions and their parameters) together with the data sources used to estimate these parameters.<sup>188</sup> This level of transparency allows rigorous external peer review of the model, increases public trust in the modeling enterprise, and facilitates future research in the content area (e.g., extensions of the model to incorporate new data or to make it transferable to new settings) and in modeling methodology (e.g., cross-model type comparisons or technical extensions of the model).<sup>200;201</sup>

### **Reported results and their interpretation should convey uncertainty in model outputs.**

Results should be reported in a way that conveys uncertainty in model output.<sup>171;172</sup> This may include the use of graphical and statistical summaries that convey the degree of uncertainty in model results (e.g., confidence bands, credible intervals, scatterplots of multiple model runs), together with summaries of sensitivity and stability analyses. Given the large number of methodological choices made at every step of model development, and the inherent subjectivity of drawing conclusions from complex research activities, we believe that general purpose algorithmic approaches cannot be developed or recommended for summarizing model results. Instead, we recommend complete reporting of model structure and data, coupled with transparency in presenting the modelers' rationale for their decisions.

### **Results should be reported in a way that addresses user needs.**

Because models have many different functions, reporting of results should be tailored to the goals of the modeling effort, while remaining faithful to the model structure and assumptions, and conveying all uncertainty in the results.<sup>33</sup> Every effort should be made to report the model findings and analyses in a manner that will be most useful to the stakeholders who would be expected to use the report.<sup>202</sup> It is impossible to provide specific guidance to address all possible model types and uses of modeling in this document. Interested readers are referred to the many available texts on healthcare modeling, the reporting of statistical and simulation analyses, and graphing quantitative information.<sup>110;113;203-208</sup>

### **All individuals who provided input to, developed and analyzed the model and interpreted its results should fully disclose any perceived conflicts of interest.**

As with any research, all investigators should provide full disclosures of any interests that can reasonably be perceived as a conflict. Both financial and nonfinancial conflicts of interest should be provided.<sup>209-213</sup> For models produced for the AHRQ EPC program and many other HTA groups, it is necessary that conflicts of interest be avoided. Modelers should adhere to established guidance for managing conflicts of interest for EPC products (e.g., Institute of Medicine recommendations; existing EPC guidance).<sup>214;215</sup>

## References

- (1) Mandelblatt JS, Cronin KA, Bailey S et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. 2009;151:738-747. PMID: 19920274
- (2) Harris RP, Helfand M, Woolf SH et al. Current methods of the US Preventive Services Task Force: a review of the process. 2001;20:21-35
- (3) U.S.Preventive Services Task Force. U.S. Preventive Services Task Force (USPSTF) Procedure Manual: Section 3. Topic Work Plan Development. 2011.

### Ref Type: Report

- (4) Whitlock EP, Orleans CT, Pender N, Allan J. Evaluating primary care behavioral counseling interventions: an evidence-based approach. 2002;22:267-284. PMID: 11988383
- (5) Woolf SH, DiGuseppi CG, Atkins D, Kamerow DB. Developing evidence-based clinical practice guidelines: lessons learned by the US Preventive Services Task Force. 1996;17:511-538. PMID: 8724238
- (6) Samson D, Schoelles KM. Developing the Topic and Structuring Systematic Reviews of Medical Tests: Utility of PICOTS, Analytic Frameworks, Decision Trees, and Other Frameworks. 2012. PMID: 22834028
- (7) Kuntz K, Sainfort F, Butler M, Taylor B, Kulasingam S, Gregory S et al. Decision and Simulation Modeling in Systematic Reviews. Methods Research Report. (Prepared by the University of Minnesota Evidence-based Practice Center under Contract No. 290-2007-10064-I.) AHRQ Publication No. 11(13)-EHC037-EF. 2013. Rockville, MD, Agency for Healthcare Research and Quality.

### Ref Type: Generic

- (8) Trikalinos TA, Kulasingam S, Lawrence WF. Deciding Whether To Complement a Systematic Review of Medical Tests With Decision Modeling. 2012. PMID: 22834021
- (9) Albert DA. Decision theory in medicine: a review and critique. *Milbank Memorial Fund Quarterly - Health & Society* 1978;56:362-401
- (10) Brennan A, Akehurst R. Modelling in health economic evaluation. What is its place? What is its value? 2000;17:445-459
- (11) Buxton MJ, Drummond MF, Van Hout BA et al. Modelling in economic evaluation: an unavoidable fact of life. *Health Economics* 1997;6:217-227

- (12) Eddy D. Technology assessment: the role of mathematical modeling. *Assessing medical technologies*. 1985;144-175.
- (13) Halpern EF, Weinstein MC, Hunink MG, Gazelle GS. Representing both first- and second-order uncertainties by Monte Carlo simulation for groups of patients. 2000;20:314-322
- (14) Hodges JS. Six (or so) things you can do with a bad model. *Operations Research* 1991;39:355-365
- (15) Mulrow C, Langhorne P, Grimshaw J. Integrating heterogeneous pieces of evidence in systematic reviews. *Annals of internal medicine* 1997;127:989-995
- (16) Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Medical Decision Making* 1993;13:322-338
- (17) Soto J. Health economic evaluations using decision analytic modeling. Principles and practices--utilization of a checklist to their development and appraisal. *Int J Technol Assess Health Care* 2002;18:94-111
- (18) Weinstein MC, Toy EL, Sandberg EA et al. Modeling for health care and other policy decisions: uses, roles, and validity. 2001;4:348-361
- (19) Chilcott J, Brennan A, Booth A, Karnon J, Tappenden P. The role of modelling in prioritising and planning clinical trials. *Health Technology Assessment (Winchester, England)* 2003;7:iii-125
- (20) Berk RA. Causal inference as a prediction problem. *Crime & Just* 1987;9:183
- (21) Greenland S. Causal inference as a prediction problem: assumptions, identification, and evidence synthesis. *Causal Inference: Statistical Perspectives and Applications New York, NY: Wiley* 2012;43-58
- (22) Pearl J. *Causality: models, reasoning and inference*. 29 ed. Cambridge Univ Press, 2000.
- (23) Anderson MP, Woessner WW. *Applied groundwater modeling: simulation of flow and advective transport*. Academic press, 1992.
- (24) Bender EA. *An introduction to mathematical modeling*. Courier Dover Publications, 2012.
- (25) Detsky AS, Naglie G, Krahn MD, Redelmeier DA, Naimark D. Primer on medical decision analysis: Part 2--Building a tree. *Medical Decision Making* 1997;17:126-135
- (26) Fishwick PA. *Simulation model design and execution: building digital worlds*. Prentice Hall PTR, 1995.

- (27) Habbema JD, Bossuyt PM, Dippel DW, Marshall S, Hilden J. Analysing clinical decision analyses. *Statistics in Medicine* 1990;9:1229-1242
- (28) Meyer WJ. *Concepts of mathematical modeling*. Courier Dover Publications, 2012.
- (29) Simon Fraser University.Complex Systems Modelling Group. *Modelling in Healthcare*. American Mathematical Soc., 2010.
- (30) Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--2. *Value Health* 2012;15:804-811
- (31) Philips Z, Ginnelly L, Sculpher M et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;8:iii-xi, 1
- (32) Trikalinos TA, Dahabreh IJ, Wallace BC, Schmid CH, Lau J. Towards a Framework for Communicating Confidence in Methodological Recommendations for Systematic Reviews and Meta-Analyses. 2013. Agency for Healthcare Research and Quality.

Ref Type: Report

- (33) Sculpher M, Fenwick E, Claxton K. Assessing quality in decision analytic cost-effectiveness models. A suggested framework and example of application. 2000;17:461-477
- (34) Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. *Pharmacoeconomics* 2006;24:355-371
- (35) Kuntz KM, Lansdorp-Vogelaar I, Rutter CM et al. A systematic comparison of microsimulation models of colorectal cancer: the role of assumptions about adenoma progression. *Med Decis Making* 2011;31:530-539. PMID: 21673186
- (36) Sainfort F, Kuntz KM, Gregory S et al. Adding decision models to systematic reviews: informing a framework for deciding when and how to do so. *Value in Health* 2013;16:133-139
- (37) Andronis L, Barton P, Bryan S. Sensitivity analysis in economic evaluation: an audit of NICE current practice and a review of its use and value in decision-making. *Health Technol Assess* 2009;13:iii, ix-61
- (38) Karnon J, Brennan A, Akehurst R. A critique and impact analysis of decision modeling assumptions. *Med Decis Making* 2007;27:491-499
- (39) Rosenbaum PR. *Observational studies*. Springer, 2002.

- (40) Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. *Journal of Health Services & Research Policy* 2004;9:110-118
- (41) Canadian Agency for Drugs and Technologies in Health. *Guidelines for the Economic Evaluation of Health Technologies: Canada*. 3 ed. Ottawa: 2006.
- (42) Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling good research practices--overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--1. *Value Health* 2012;15:796-803
- (43) Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996;313:275-283
- (44) Hay J, Jackson J. Panel 2: methodological issues in conducting pharmacoeconomic evaluations--modeling studies. *Value Health* 1999;2:78-81
- (45) Hjelmgren J, Berggren F, Andersson F. Health economic guidelines--similarities, differences and some implications. *Value in Health* 2001;4:225-250
- (46) Husereau D, Drummond M, Petrou S et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMJ* 2013;346:f1049. PMID: 23529982
- (47) Nuijten MJ, Pronk MH, Brorens MJ et al. Reporting format for economic evaluation. Part II: Focus on modelling studies. *Pharmacoeconomics* 1998;14:259-268
- (48) Weinstein MC, O'Brien B, Hornberger J et al. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. *Value Health* 2003;6:9-17
- (49) Russell LB, Fryback DG, Sonnenberg FA. Is the societal perspective in cost-effectiveness analysis useful for decision makers? *Joint Commission Journal on Quality Improvement* 1999;25:447-454
- (50) Roy S, Madhavan SS. Making a case for employing a societal perspective in the evaluation of Medicaid prescription drug interventions. *PharmacoEconomics* 2008;26:281-296
- (51) Garrison LP, Jr., Mansley EC, Abbott TA, III, Bresnahan BW, Hay JW, Smeeding J. Good research practices for measuring drug costs in cost-effectiveness analyses: a societal perspective: the ISPOR Drug Cost Task Force report--Part II. *Value in Health* 2010;13:8-13
- (52) Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. *JAMA* 1996;276:1253-1258. PMID: [8849754](https://pubmed.ncbi.nlm.nih.gov/8849754/)

- (53) Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G. *Methods for the Economic Evaluation of Health Care Programmes*. 3 ed. Oxford: Oxford University Press, 2005.
- (54) Neumann PJ. Costing and perspective in published cost-effectiveness analysis. *2009;47:S28-S32*
- (55) Ademi Z, Kim H, Zomer E, Reid CM, Hollingsworth B, Liew D. Overview of pharmaco-economic modelling methods. *British Journal of Clinical Pharmacology* 2013;75:944-950
- (56) Banks J, Carson II J, Nelson B, Nicol D. *Discrete-Event System Simulation*. 5 ed. Prentice Hall, 2009.
- (57) Beck JR, Pauker SG. The Markov process in medical prognosis. *Medical Decision Making* 1983;3:419-458
- (58) Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ* 2006;15:1295-1310
- (59) Briggs A, Sculpher M. An introduction to Markov modelling for economic evaluation. *Pharmacoeconomics* 1998;13:397-409
- (60) Caro JJ, Moller J, Getsios D. Discrete event simulation: the preferred technique for health economic evaluations? *Value in Health* 2010;13:1056-1060
- (61) Jun JB, Jacobson SH, Swisher JR. Application of discrete-event simulation in health care clinics: A survey. *Journal of the operational research society* 1999;50:109-123
- (62) Karnon J, Brown J. Selecting a decision model for economic evaluation: a case study and review. *Health Care Management Science* 1998;1:133-140
- (63) Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. *Health Economics* 2003;12:837-848
- (64) Karnon J, Haji Ali AH. When to Use Discrete Event Simulation (DES) for the Economic Evaluation of Health Technologies? A Review and Critique of the Costs and Benefits of DES. 2014. PMID: 24627341
- (65) Law A. *Simulation Modeling and Analysis*. 4 ed. McGraw-Hill Science/Engineering/Math, 2006.
- (66) Naimark D, Krahn MD, Naglie G, Redelmeier DA, Detsky AS. Primer on medical decision analysis: Part 5--Working with Markov processes. *Medical Decision Making* 1997;17:152-159

- (67) Pitman R, Fisman D, Zaric GS, Postma M, Mirjam K, John E. Dynamic Transmission Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-5. 2012;Value Health:range
- (68) Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens D, Cohen D. State-Transition Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. 2012;15:range
- (69) Soares MO, Canto E Castro. Continuous time simulation and discretized models for cost-effectiveness analysis. *Pharmacoeconomics* 2012;30:1101-1117
- (70) Stahl JE. Modelling methods for pharmacoeconomics and health technology assessment: an overview and guide. *Pharmacoeconomics* 2008;26:131-148
- (71) Standfield L, Comans T, Scuffham P. Markov Modeling and Discrete Event Simulation in Health Care: A Systematic Comparison. *International Journal of Technology Assessment in Health Care* 2014;1-8. PMID: 24774101
- (72) van Rosmalen J, Toy M, O'Mahony JF. A mathematical approach for evaluating Markov models in continuous time without discrete-event simulation. *Medical Decision Making* 2013;33:767-779
- (73) Zeigler BP, Praehofer H, Kim TG. *Theory of modeling and simulation: integrating discrete event and continuous complex dynamic systems*. Academic press, 2000.
- (74) Koopman JS, Jacquez G, Chick SE. New data and tools for integrating discrete and continuous population modeling strategies. *Annals of the New York Academy of Sciences* 2001;954:268-294
- (75) Siegel JE, Weinstein MC, Russell LB, Gold MR. Recommendations for reporting cost-effectiveness analyses. Panel on Cost-Effectiveness in Health and Medicine. *JAMA* 1996;276:1339-1341. PMID: <http://www.ncbi.nlm.nih.gov/pubmed/8861994>
- (76) Brettell AJ, Long AF, Grant MJ, Greenhalgh J. Searching for information on outcomes: do you need to be comprehensive? *Quality in Health Care* 1998;7:163-167
- (77) 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 since 1997. *Journal of Health Services & Research Policy* 2005;10:245-250
- (78) Novielli N, Cooper NJ, Abrams KR, Sutton AJ. How is evidence on test performance synthesized for economic decision models of diagnostic tests? A systematic appraisal of Health Technology Assessments in the UK since 1997. *Value in Health* 2010;13:952-957
- (79) Royle P, Waugh N. Literature searching for clinical and cost-effectiveness studies used in health technology assessment reports carried out for the National Institute for Clinical Excellence appraisal system. 2003;7:iii

- (80) Sutton AJ, Cooper NJ, Goodacre S, Stevenson M. Integration of meta-analysis and economic decision modeling for evaluating diagnostic tests. *Medical Decision Making* 2008;28:650-667
- (81) Ades AE, Lu G, Higgins JP. The interpretation of random-effects meta-analysis in decision models. *Medical Decision Making* 2005;25:646-654
- (82) Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence synthesis for decision making 6: embedding evidence synthesis in probabilistic cost-effectiveness analysis. *Med Decis Making* 2013;33:671-678. PMID: 23804510
- (83) Graham PL, Moran JL. Robust meta-analytic conclusions mandate the provision of prediction intervals in meta-analysis summaries. *Journal of Clinical Epidemiology* 2012;65:503-510
- (84) Oppe M, Al M, Rutten-van MM. Comparing methods of data synthesis: re-estimating parameters of an existing probabilistic cost-effectiveness model. *Pharmacoeconomics* 2011;29:239-250
- (85) Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011;342:d549
- (86) Trikalinos TA, Siebert U, Lau J. Decision-analytic modeling to evaluate benefits and harms of medical tests: uses and limitations. *Med Decis Making* 2009;29:E22-E29
- (87) Vemer P, Al MJ, Oppe M, Rutten-van Molken MP. A choice that matters? Smulation study on the impact of direct meta-analysis methods on health economic outcomes. *Pharmacoeconomics* 2013;31:719-730
- (88) Ades AE, Cliffe S. Markov chain Monte Carlo estimation of a multiparameter decision model: consistency of evidence and the accurate assessment of uncertainty. 2002;22:359-371. PMID: 12150601
- (89) Ades AE. A chain of evidence with mixed comparisons: models for multi-parameter synthesis and consistency of evidence. *Statistics in Medicine* 2003;22:2995-3016
- (90) Ades AE, Sutton AJ. Multiparameter evidence synthesis in epidemiology and medical decision-making: current approaches. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2006;169:5-35
- (91) Ades AE, Sculpher M, Sutton A et al. Bayesian methods for evidence synthesis in cost-effectiveness analysis. *Pharmacoeconomics* 2006;24:1-19
- (92) Ades AE, Welton NJ, Caldwell D, Price M, Goubar A, Lu G. Multiparameter evidence synthesis in epidemiology and medical decision-making. *J Health Serv Res Policy* 2008;13 Suppl 3:12-22. PMID: 18806188

- (93) Borenstein M, Hedges LV, Higgins JP, Rothstein HR. *Introduction to meta-analysis*. John Wiley & Sons, 2011.
- (94) Cooper H, Hedges LV, Valentine JC. *The handbook of research synthesis and meta-analysis*. Russell Sage Foundation, 2009.
- (95) Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making* 2013;33:607-617. PMID: 23104435
- (96) Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. *Med Decis Making* 2013;33:641-656. PMID: 23804508
- (97) Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence synthesis for decision making 3: heterogeneity--subgroups, meta-regression, bias, and bias-adjustment. *Med Decis Making* 2013;33:618-640. PMID: 23804507
- (98) Eddy DM. The confidence profile method: a Bayesian method for assessing health technologies. *Operations Research* 1989;37:210-228
- (99) Eddy DM, Hasselblad V, Shachter R. An introduction to a Bayesian method for meta-analysis: The confidence profile method. *Medical Decision Making* 1990;10:15-23
- (100) Eddy DM, Hasselblad V, Shachter R. A Bayesian method for synthesizing evidence. The Confidence Profile Method. *International Journal of Technology Assessment in Health Care* 1990;6:31-55
- (101) Egger M, Smith GD, Altman D. *Systematic reviews in health care: meta-analysis in context*. John Wiley & Sons, 2008.
- (102) Fu R, Gartlehner G, Grant M et al. Conducting Quantitative Synthesis When Comparing Medical Interventions: AHRQ and the Effective Health Care Program. 2008. PMID: 21433407
- (103) Hartung J, Knapp G, Sinha BK. *Statistical meta-analysis with applications*. 738 ed. John Wiley & Sons, 2011.
- (104) Kaizar EE. Estimating treatment effect via simple cross design synthesis. *Statistics in Medicine* 2011;30:2986-3009
- (105) Lau J, Terrin N, Fu R. Expanded Guidance on Selected Quantitative Synthesis Topics. 2008. PMID: 23596640
- (106) Olkin I. Statistical methods for meta-analysis. *San Diego, CA: Academic* 1985;
- (107) Petitti DB. *Meta-analysis, decision analysis, and cost-effectiveness analysis: methods for quantitative synthesis in medicine*. Oxford University Press, 1999.

- (108) Schmid CH. Using Bayesian inference to perform meta-analysis. *Evaluation & the Health Professions* 2001;24:165-189
- (109) Spiegelhalter DJ, Best NG. Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Statistics in Medicine* 2003;22:3687-3709
- (110) Spiegelhalter DJ, Abrams KR, Myles JP. *Bayesian approaches to clinical trials and health-care evaluation*. 13 ed. John Wiley & Sons, 2004.
- (111) Stangl D, Berry DA. *Meta-analysis in medicine and health policy*. CRC Press, 2009.
- (112) Sutton AJ, Abrams KR, Jones DR, Jones DR, Sheldon TA, Song F. *Methods for meta-analysis in medical research*. J. Wiley, 2000.
- (113) Sutton AJ, Cooper NJ, Abrams KR, Ades AE. *Evidence synthesis for decision making in healthcare*. 132 ed. John Wiley & Sons, 2012.
- (114) Naglie G, Krahn MD, Naimark D, Redelmeier DA, Detsky AS. Primer on medical decision analysis: Part 3--Estimating probabilities and utilities. *Medical Decision Making* 1997;17:136-141
- (115) Chopra A. <sup>111</sup>In/<sup>125</sup>I/<sup>131</sup>I-Labeled anti-mucin-1 murine, chimeric or humanized antibody hPAM4. 2004. PMID: 21834178
- (116) Greenland S. Basic methods for sensitivity analysis of biases. *International Journal of Epidemiology* 1996;25:1107-1116
- (117) Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology* 2009;20:488-495. PMID: 19525685
- (118) Braithwaite RS, Roberts MS, Justice AC. Incorporating quality of evidence into decision analytic modeling. *Annals of internal medicine* 2007;146:133-141
- (119) Goldhaber-Fiebert JD. Accounting for biases when linking empirical studies and simulation models. *Medical Decision Making* 2012;32:397-399
- (120) Dias S, Welton NJ, Marinho VCC, Salanti G, Higgins JPT, Ades AE. Estimation and adjustment of bias in randomized evidence by using mixed treatment comparison meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2010;173:613-629
- (121) Doi SA, Barendregt JJ, Onitilo AA. Methods for the bias adjustment of meta-analyses of published observational studies. *Journal of Evaluation in Clinical Practice* 2013;19:653-657

- (122) Greenland S. Bayesian perspectives for epidemiologic research: III. Bias analysis via missing-data methods.[Erratum appears in *Int J Epidemiol.* 2010 Aug;39(4):1116]. *International Journal of Epidemiology* 2009;38:1662-1673
- (123) Greenland S. Multiple-bias modelling for analysis of observational data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2005;168:267-306
- (124) Gustafson P, McCandless LC. Probabilistic approaches to better quantifying the results of epidemiologic studies. *International Journal of Environmental Research & Public Health [Electronic Resource]* 2010;7:1520-1539
- (125) Hofler M, Lieb R, Wittchen HU. Estimating causal effects from observational data with a model for multiple bias. *International Journal of Methods in Psychiatric Research* 2007;16:77-87
- (126) Kuroki M, Pearl J. Measurement bias and effect restoration in causal inference. 2011. DTIC Document.

Ref Type: Report

- (127) Lash TL, Fox MP, Fink AK. *Applying quantitative bias analysis to epidemiologic data.* Springer, 2011.
- (128) Maldonado G. Adjusting a relative-risk estimate for study imperfections. *Journal of Epidemiology & Community Health* 2008;62:655-663
- (129) Molitor N-T, Best N, Jackson C, Richardson S. Using Bayesian graphical models to model biases in observational studies and to combine multiple sources of data: application to low birth weight and water disinfection by products. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2009;172:615-637
- (130) Thompson S, Ekelund U, Jebb S et al. A proposed method of bias adjustment for meta-analyses of published observational studies. *International Journal of Epidemiology* 2011;40:765-777
- (131) Turner RM, Spiegelhalter DJ, Smith G, Thompson SG. Bias modelling in evidence synthesis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2009;172:21-47
- (132) Welton NJ, Ades AE, Carlin JB, Altman DG, Sterne JAC. Models for potentially biased evidence in meta-analysis using empirically based priors. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2009;172:119-136
- (133) Fischhoff B, Slovic P, Lichtenstein S. Knowing what you want: Measuring labile values. *Decision Making: Descriptive, Normative and Prescriptive Interactions,* Cambridge University Press, Cambridge 1988;398-421

- (134) Bala MV, Wood LL, Zarkin GA, Norton EC, Gafni A, O'Brien B. Valuing outcomes in health care: a comparison of willingness to pay and quality-adjusted life-years. *Journal of Clinical Epidemiology* 1998;51:667-676
- (135) Bravata DM, Nelson LM, Garber AM, Goldstein MK. Invariance and inconsistency in utility ratings. *Medical Decision Making* 2005;25:158-167
- (136) Chaloner K, Rhame FS. Quantifying and documenting prior beliefs in clinical trials. *Statistics in Medicine* 2001;20:581-600
- (137) Frew EJ, Whynes DK, Wolstenholme JL. Eliciting willingness to pay: comparing closed-ended with open-ended and payment scale formats. *Medical Decision Making* 2003;23:150-159
- (138) Garthwaite PH, Kadane JB, O'Hagan A. Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association* 2005;100:680-701
- (139) Johnson SR, Tomlinson GA, Hawker GA, Granton JT, Grosbein HA, Feldman BM. A valid and reliable belief elicitation method for Bayesian priors. *Journal of Clinical Epidemiology* 2010;63:370-383
- (140) Kerstholt JH, van der Zwaard F, Bart H, Cremers A. Construction of health preferences: a comparison of direct value assessment and personal narratives. *Medical Decision Making* 2009;29:513-520
- (141) O'Hagan A, Buck CE, Daneshkhah A et al. *Uncertain judgements: eliciting experts' probabilities*. John Wiley & Sons, 2006.
- (142) Soares MO, Bojke L, Dumville J, Iglesias C, Cullum N, Claxton K. Methods to elicit experts' beliefs over uncertain quantities: application to a cost effectiveness transition model of negative pressure wound therapy for severe pressure ulceration. *Statistics in Medicine* 2011;30:2363-2380
- (143) White IR, Pocock SJ, Wang D. Eliciting and using expert opinions about influence of patient characteristics on treatment effects: a Bayesian analysis of the CHARM trials. *Statistics in Medicine* 2005;24:3805-3821
- (144) Koopman J. Modeling infection transmission. *Annual Review of Public Health* 2004;25:303-326
- (145) Drummond M, Barbieri M, Cook J et al. Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force report. *Value Health* 2009;12:409-418
- (146) Drummond M, Manca A, Sculpher M. Increasing the generalizability of economic evaluations: recommendations for the design, analysis, and reporting of studies. 2005;21:165-171

- (147) Bareinboim E, Pearl J. A general algorithm for deciding transportability of experimental results. *Journal of Causal Inference* 2013;1:107-134
- (148) Meta-Transportability of Causal Effects: A Formal Approach.: 2013.
- (149) Causal transportability with limited experiments.: 2013.
- (150) Transportability of causal and statistical relations: A formal approach.: IEEE, 2011.
- (151) Pearl J, Bareinboim E. External validity: from do-calculus to transportability across populations. 2012. DTIC Document.

Ref Type: Report

- (152) Pressler TR, Kaizar EE. The use of propensity scores and observational data to estimate randomized controlled trial generalizability bias. *Statistics in Medicine* 2013;32:3552-3568
- (153) Groot KB, Weinstein MC, Stijnen T, Heijnenbroek-Kal MH, Hunink MG. Uncertainty and patient heterogeneity in medical decision models. *Medical Decision Making* 2010;30:194-205
- (154) Groot KB, Stijnen T, Weinstein MC, Hunink MG. The combined analysis of uncertainty and patient heterogeneity in medical decision models.[Erratum appears in Med Decis Making. 2013 Feb;33(2):307]. *Medical Decision Making* 2011;31:650-661
- (155) Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine* 2004;23:1663-1682
- (156) Higgins JP. Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. *International Journal of Epidemiology* 2008;37:1158-1160
- (157) Higgins J, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2009;172:137-159
- (158) Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine* 2002;21:1559-1573
- (159) Greenland S. Invited commentary: a critical look at some popular meta-analytic methods. *American Journal of Epidemiology* 1994;140:290-296
- (160) Poole C, Greenland S. Random-effects meta-analyses are not always conservative. *American Journal of Epidemiology* 1999;150:469-475
- (161) Berlin JA, Santanna J, Schmid CH, Szczech LA, Feldman HI, Anti-Lymphocyte Antibody Induction Therapy Study Group. Individual patient- versus group-level data

- meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Statistics in Medicine* 2002;21:371-387
- (162) Donegan S, Williamson P, D'Alessandro U, Garner P, Smith CT. Combining individual patient data and aggregate data in mixed treatment comparison meta-analysis: Individual patient data may be beneficial if only for a subset of trials. *Statistics in Medicine* 2013;32:914-930
- (163) Koopman L, van der Heijden GJ, Hoes AW, Grobbee DE, Rovers MM. Empirical comparison of subgroup effects in conventional and individual patient data meta-analyses. *International Journal of Technology Assessment in Health Care* 2008;24:358-361
- (164) Riley RD, Simmonds MC, Look MP. Evidence synthesis combining individual patient data and aggregate data: a systematic review identified current practice and possible methods. *Journal of Clinical Epidemiology* 2007;60:431-439
- (165) Riley RD, Dodd SR, Craig JV, Thompson JR, Williamson PR. Meta-analysis of diagnostic test studies using individual patient data and aggregate data. *Statistics in Medicine* 2008;27:6111-6136
- (166) Schmid CH, Stark PC, Berlin JA, Landais P, Lau J. Meta-regression detected associations between heterogeneous treatment effects and study-level, but not patient-level, factors. *Journal of Clinical Epidemiology* 2004;57:683-697
- (167) Simmonds MC, Higgins JP. Covariate heterogeneity in meta-analysis: criteria for deciding between meta-regression and individual patient data. *Statistics in Medicine* 2007;26:2982-2999
- (168) Sutton AJ, Kendrick D, Coupland CA. Meta-analysis of individual- and aggregate-level data. *Statistics in Medicine* 2008;27:651-669
- (169) Kovalchik SA. Survey finds that most meta-analysts do not attempt to collect individual patient data. *Journal of Clinical Epidemiology* 2012;65:1296-1299
- (170) Briggs AH, Gray AM. Handling uncertainty when performing economic evaluation of healthcare interventions. *Health technology assessment* 1999;3:1-134. PMID: <http://www.ncbi.nlm.nih.gov/pubmed/10448202>
- (171) Briggs AH. Handling uncertainty in cost-effectiveness models. 2000;17:479-500
- (172) Briggs AH, O'Brien BJ, Blackhouse G. Thinking outside the box: recent advances in the analysis and presentation of uncertainty in cost-effectiveness studies. *Annual Review of Public Health* 2002;23:377-401
- (173) Claxton K, Sculpher M, McCabe C et al. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. 2005;14:339-347

- (174) Critchfield GC, Willard KE. Probabilistic analysis of decision trees using Monte Carlo simulation. *Medical Decision Making* 1986;6:85-92
- (175) Critchfield GC, Willard KE, Connelly DP. Probabilistic sensitivity analysis methods for general decision models. *Computers & Biomedical Research* 1986;19:254-265
- (176) Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation. A practical approach. *Medical Decision Making* 1985;5:157-177
- (177) Lord J, Asante MA. Estimating uncertainty ranges for costs by the bootstrap procedure combined with probabilistic sensitivity analysis. *Health Economics* 1999;8:323-333
- (178) O'Hagan A, McCabe C, Akehurst R et al. Incorporation of uncertainty in health economic modelling studies. *PharmacoEconomics* 2005;23:529-536. PMID: <http://www.ncbi.nlm.nih.gov/pubmed/15960550>
- (179) Pasta DJ, Taylor JL, Henning JM. Probabilistic sensitivity analysis incorporating the bootstrap: an example comparing treatments for the eradication of *Helicobacter pylori*. *Medical Decision Making* 1999;19:353-363
- (180) Ades AE, Claxton K, Sculpher M. Evidence synthesis, parameter correlation and probabilistic sensitivity analysis. 2006;15:373-381
- (181) Baio G, Dawid AP. Probabilistic sensitivity analysis in health economics. 2011. PMID: 21930515
- (182) Briggs AH, Ades AE, Price MJ. Probabilistic sensitivity analysis for decision trees with multiple branches: use of the Dirichlet distribution in a Bayesian framework. *Medical Decision Making* 2003;23:341-350
- (183) Brisson M, Edmunds WJ. Impact of model, methodological, and parameter uncertainty in the economic analysis of vaccination programs. *Medical Decision Making* 2006;26:434-446
- (184) Claxton K. Exploring uncertainty in cost-effectiveness analysis. *PharmacoEconomics* 2008;26:781-798
- (185) Griffin S, Claxton K, Hawkins N, Sculpher M. Probabilistic analysis and computationally expensive models: Necessary and required? 2006;9:244-252
- (186) O'Hagan A, Stevenson M, Madan J. Monte Carlo probabilistic sensitivity analysis for patient level simulation models: efficient estimation of mean and variance using ANOVA. *Health Economics* 2007;16:1009-1023
- (187) Chilcott J, Tappenden P, Rawdin A et al. Avoiding and identifying errors in health technology assessment models: qualitative study and methodological review. *Health Technol Assess* 2010;14:iii-xii, 1

- (188) Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Med Decis Making* 2012;32:733-743. PMID: 22990088
- (189) MacConnell S. *Code Complete : A Practical Handbook of Software Construction*. Microsoft Press, 2004.
- (190) Bennett C, Manuel DG. Reporting guidelines for modelling studies. *BMC Medical Research Methodology* 2012;12:168
- (191) Cooper BS. Confronting models with data. *J Hosp Infect* 2007;65 Suppl 2:88-92. PMID: 17540249
- (192) McCabe C, Dixon S. Testing the validity of cost-effectiveness models. *Pharmacoeconomics* 2000;17:501-513
- (193) Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans ASABE* 2007;50:885-900
- (194) Sendi PP, Craig BA, Pfluger D, Gafni A, Bucher HC. Systematic validation of disease models for pharmaco-economic evaluations. Swiss HIV Cohort Study. *Journal of Evaluation in Clinical Practice* 1999;5:283-295
- (195) Stout NK, Knudsen AB, Kong CY, McMahon PM, Gazelle GS. Calibration methods used in cancer simulation models and suggested reporting guidelines. *Pharmacoeconomics* 2009;27:533-545
- (196) Vanni T, Karnon J, Madan J et al. Calibrating models in economic evaluation: a seven-step approach. 2011;29:35-49. PMID: 21142277
- (197) Boer R, Plevritis S, Clarke L. Diversity of model approaches for breast cancer screening: a review of model assumptions by the Cancer Intervention and Surveillance Network (CISNET) Breast Cancer Groups. *Statistical Methods in Medical Research* 2004;13:525-538
- (198) Hodges JS, Dewar JA, Center A. *Is it you or your model talking?: A framework for model validation*. Rand Santa Monica, CA, 1992.
- (199) Peng RD. Reproducible research in computational science. 2011;334:1226-1227
- (200) National Research Council. *Models in Environmental Regulatory Decision Making*. Washington, DC: National Research Council, 2007.
- (201) Sonnenberg FA, Roberts MS, Tsevat J, Wong JB, Barry M, Kent DL. Toward a peer review process for medical decision analysis models. *Med Care* 1994;32:JS52-JS64

- (202) Krahn MD, Naglie G, Naimark D, Redelmeier DA, Detsky AS. Primer on medical decision analysis: Part 4--Analyzing the model and interpreting the results. *Medical Decision Making* 1997;17:142-151
- (203) Briggs AH, Claxton K, Sculpher MJ. *Decision modelling for health economic evaluation*. Oxford university press, 2006.
- (204) Cleveland WS. The elements of graphing data. Monterey. *Calif, Wadsworth* 1985;
- (205) Cleveland WS. *Visualizing data*. Hobart Press, 1993.
- (206) Drummond MF, McGuire A. *Economic evaluation in health care: merging theory with practice*. Oxford University Press, 2001.
- (207) Hunink MM. *Decision Making in Health and Medicine with CD-ROM: Integrating Evidence and Values*. 1 ed. Cambridge University Press, 2001.
- (208) Tufte ER, Graves-Morris PR. *The visual display of quantitative information*. 2 ed. Graphics press Cheshire, CT, 1983.
- (209) Baker CB, Johnsrud MT, Crismon ML, Rosenheck RA, Woods SW. Quantitative analysis of sponsorship bias in economic studies of antidepressants. *British Journal of Psychiatry* 2003;183:498-506
- (210) Barbieri M, Drummond MF. Conflict of interest in industry-sponsored economic evaluations: real or imagined? *Current Oncology Reports* 2001;3:410-413
- (211) Friedberg M, Saffran B, Stinson TJ, Nelson W, Bennett CL. Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *JAMA* 1999;282:1453-1457
- (212) Garattini L, Koleva D, Casadei G. Modeling in pharmacoeconomic studies: funding sources and outcomes. *International Journal of Technology Assessment in Health Care* 2010;26:330-333
- (213) Valachis A, Polyzos NP, Nearchou A, Lind P, Mauri D. Financial relationships in economic analyses of targeted therapies in oncology. *Journal of Clinical Oncology* 2012;30:1316-1320
- (214) Lo B, Field MJ. *Conflict of interest in medical research, education, and practice*. National Academies Press, 2009.
- (215) Viswanathan M, Carey TS, Belinson SE, Berliner E, Chang S, Graham E et al. Identifying and Managing Nonfinancial Conflicts of Interest for Systematic Reviews. 2013. Agency of Healthcare Research and Quality.

Ref Type: Report

