Methods Guide
for Comparative Effectiveness Reviews

Guidance for the Conduct and Reporting of Modeling and Simulation Studies in the Context of Health Technology Assessment

This report is based on research conducted by the Tufts Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-2007-10055-I). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, 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.

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

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report is made available to the public under the terms of a licensing agreement between the author and the Agency for Healthcare Research and Quality. This report may be used and reprinted without permission except those copyrighted materials that are clearly noted in the report. Further reproduction of those copyrighted materials is prohibited without the express permission of copyright holders.

AHRQ or U.S. Department of Health and Human Services endorsement of any derivative products that may be developed from this report, such as clinical practice guidelines, other quality enhancement tools, or reimbursement or coverage policies, may not be stated or implied.

Persons using assistive technology may not be able to fully access information in this report. For assistance contact EffectiveHealthCare@ahrq.hhs.gov.

Suggested citation: Dahabreh IJ, Trikalinos TA, Balk EM, Wong JB. Guidance for the Conduct and Reporting of Modeling and Simulation Studies in the Context of Health Technology Assessment. Methods Guide for Comparative Effectiveness Reviews. (Prepared by the Tufts Evidence-based Practice Center under Contract No. 290-2007-10055-I.) AHRQ Publication No. 16(17)-EHC025-EF. Rockville, MD: Agency for Healthcare Research and Quality; October 2016. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Prepared by:
Tufts Evidence-based Practice Center
Boston, MA

Investigators:
Issa J. Dahabreh, M.D., M.S.[1,2,3]
Thomas A. Trikalinos, M.D. [1,2]
 Ethan M. Balk, M.D., M.P.H. [1,2]
John B. Wong, M.D.[4]

[1]Center for Evidence-based Medicine, School of Public Health, Brown University, Providence, RI
[2]Department of Health Services, Policy & Practice, School of Public Health, Brown University, Providence, RI
[3]Department of Epidemiology, School of Public Health, Brown University, Providence, RI
[4]Division of Clinical Decision Making, Tufts Medical Center, Boston, MA

Editors' Foreword

These guidelines encourage Evidence-based Practice Centers to quantify biases. The idea is not new—in fact, it was proposed and debated at length in the late 1980s and early 1990s.[a] Most systematic reviews take a less transparent approach. In the example in Box 1, the investigators have pooled five studies and propose a "fair" rating for the overall quality of this evidence. Judging by the effect size, they seem to be saying that an intervention decreases mortality, but they also seem to be saying not to trust that estimate. When the investigators have a belief about the magnitude and direction of bias, the default approach (depicted in the box) is to put forward a numerical estimate they believe to be wrong and then qualify it—for example, "We think the estimate may be high because these are the first trials of this intervention, and early trials tend to have exaggerated effect sizes." This approach makes it difficult to justify putting out a numerical estimate.

Box 1. Example

Outcome	Pooled Effect Size	Overall Quality of Evidence	Number of Studies
Mortality	0.85 ± 0.13	Fair	5

Explicit adjustment for bias means that the pooled effect size estimate would take into account concerns about quality and be interpretable as the authors' best estimate of the effect.
While the explicit approach is appealing in many ways, most systematic reviewers fear that bias adjustments will introduce subjectivity and error rather than improve transparency. They note correctly that there is no reliable reproducible approach to making these estimates, and that the magnitude and direction of bias are often unpredictable.
Given this concern, it is important to note the following:
1. The basic recommendation is to use quantitative bias adjustments to integrate the reported effect sizes with the assessment of risk of bias or quality when meta-analysis is used alongside decision modeling or simulation.
2. Evidence-based Practice Centers are not required to use quantitative bias adjustments in systematic reviews and meta-analyses when decision or simulation modeling is not done.
3. If the investigators' true belief is that no adjustment is needed (e.g., that the adjustment factor should be 1.0), it is important to convey that judgment. There is no requirement to use a factor different from the investigators' true belief. What is important is to be transparent about our confidence in the estimate instead of using vague qualitative statements. If, in the example shown, the investigators are so unsure about the likely magnitude and direction of bias that they are unwilling to use an adjustment factor other than 1.0, this conveys important information to the reader about the emptiness of the "fair" rating.

Mark Helfand, M.D., M.P.H.
Stephanie Chang, M.D., M.P.H.

[a]See, for example, the article Eddy DM, Hasselblad V, Shachter R. An introduction to a Bayesian method for meta-analysis: the confidence profile method. Med Decis Making. 1990 Jan-Mar;10(1):15-23. PMID: 2182960.

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.

Strong methodological approaches to systematic review improve the transparency, consistency, and scientific rigor of these reports. Through a collaborative effort of the Effective Health Care (EHC) Program, the Agency for Healthcare Research and Quality (AHRQ), the EHC Program Scientific Resource Center, and the AHRQ Evidence-based Practice Centers have developed a Methods Guide for Comparative Effectiveness Reviews. This Guide presents issues key to the development of Systematic Reviews and describes recommended approaches for addressing difficult, frequently encountered methodological issues.

The Methods Guide for Comparative Effectiveness Reviews is a living document, and will be updated as further empiric evidence develops and our understanding of better methods improves.

If you have comments on this Methods Guide paper, they may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857, or by email to epc@ahrq.hhs.gov.

Andrew Bindman, M.D.
Director
Agency for Healthcare Research and Quality

Arlene S. Bierman, M.D., M.S.
Director
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Task Order Officer
Director
Evidence-based Practice Center Program
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Acknowledgments

M. Hassan Murad, M.D., M.P.H.
Director
Mayo Clinic Evidence-based Practice Center
Professor of Medicine
Mayo Clinic
Rochester, MN

Paul Shekelle, Ph.D.
Codirector
Southern California Evidence-based Practice Center
RAND Corporation
Santa Monica, CA

Jonathan R. Treadwell, Ph.D.
Associate Director
ECRI Evidence-based Practice Center
ECRI Institute
Plymouth Meeting, PA

C. Michael White, Pharm.D., FCP, FCCP
Director
University of Connecticut Evidence-based Practice Center
Professor & Head
Department of Pharmacy Practice
University of Connecticut School of Pharmacy
Storrs, CT

Technical Expert Panel

In designing the study questions and methodology at the outset of this report, the EPC consulted several technical and content experts. Broad expertise and perspectives were sought. Divergent and conflicted opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore, in the end, study questions, design, methodologic approaches, and/or conclusions do not necessarily represent the views of individual technical and content experts.

Technical Experts must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

The list of Technical Experts who provided input to this report follows:

Eric B. Bass, M.D., M.P.H.
Director
Johns Hopkins Evidence-based Practice Center
John Hopkins University
Baltimore, MD

Andrew Briggs, D.Phil.
William R. Lindsay Chair in Health Economics
University of Glasgow
Glasgow, Scotland, UK

Robert M. Golub, M.D.
Deputy Editor
Journal of the American Medical Association
Professor of Medicine
Northwestern University Feinberg School of Medicine
Chicago, IL

Douglas Owens, M.D., M.S.
Henry J. Kaiser, Jr., Professor
Director
VA Postdoctoral Medical Informatics Fellowship
VA Palo Alto Health Care System
Director
Program on Clinical Decision Making and Guideline Development
Center for Primary Care and Outcomes Research
Stanford University
Palo Alto, CA

J. Sanford Schwartz, M.D.
Leon Hess Professor in Internal Medicine
University of Pennsylvania
Philadelphia, PA

Milton C. Weinstein, Ph.D.
Henry J. Kaiser Professor of Health Policy and Management
Harvard Medical School
Director
Program on the Economic Evaluation of Medical Technology
Harvard T.H. Chan School of Public Health
Boston, MA

Peer Reviewers

Prior to publication of the final version of this chapter, the EPC sought input from independent Peer Reviewers. However, the conclusions and synthesis of the scientific literature presented in this report do not necessarily represent the views of individual reviewers.

Peer Reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential nonfinancial conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential nonfinancial conflicts of interest identified.

The list of Peer Reviewers follows.

Andrew Briggs, D.Phil.
William R. Lindsay Chair in Health Economics
University of Glasgow
Glasgow, Scotland, UK

Scott Cantor, Ph.D.
Professor
Department of Health Services Research
Division of Cancer Prevention and Population Sciences
The University of Texas MD Anderson Cancer Center
Houston, TX

Ashish Deshmukh, Ph.D.
Postdoctoral Student
Department of Health Services Research
Division of Cancer Prevention and Population Sciences
The University of Texas MD Anderson Cancer Center
Houston, TX

David B. Matchar, M.D.
Professor and Director
Program in Health Services and Systems Research
Duke-NUS Medical School
Singapore
Professor (Internal Medicine and Pathology)
Duke University Medical Center
Durham, NC

Steven Teutsch, M.D., M.P.H.
Senior Fellow
Leonard D. Schaeffer Center for Health Policy and Economics
University of Southern California
Senior Fellow
Public Health Institute
Adjunct Professor
UCLA Fielding School of Public Health
Los Angeles, CA

Guidance for the Conduct and Reporting of Modeling and Simulation Studies in the Context of Health Technology Assessment

Structured Abstract

Objectives. The U.S. Agency for Healthcare Research and Quality (AHRQ) solicited the development of guidance for modeling and simulation studies conducted in the context of health technology assessment and systematic comparative effectiveness reviews.

Guidance development process. We updated and expanded existing systematic reviews of recommendations for the conduct and reporting of modeling and simulation with input from a multidisciplinary team of clinical, policy, and decision analysis experts. The results of the systematic review were discussed in person with a panel of 28 stakeholders, including patient representatives, providers and purchasers of care, payers, policymakers, and principal investigators. Stakeholders commented on existing recommendations from various sources and identified gaps, limitations, and areas for elaboration. We subsequently reviewed the Web sites of 126 health technology assessment organizations providing guidance on the conduct and reporting of modeling and simulation. We also solicited input from senior researchers with experience in modeling and simulation from AHRQ and its Evidence-based Practice Centers, and from external reviewers.

Results. We developed principles and good-practice recommendations for modeling and simulation studies conducted to enhance and contextualize the findings of systematic reviews. The guidance applies to structural mathematical models, including declarative, functional, and spatial models. The recommendations address the identification, estimation, verification, and validation of such models, as well as the use of sensitivity, stability, and uncertainty analyses in model development and assessment. We organized recommendations by whether they pertain to the model conceptualization and structure, data, model assessment and consistency, or the interpretation and reporting of results. We provide the rationale for each recommendation, supportive evidence, or best judgment where adequate evidence was lacking.

Conclusions. We hope that this work will contribute to increased use and better conduct and reporting of modeling and simulation studies in health technology assessment.

Contents

Introduction

Understanding the effects of interventions and using evidence to inform decisions are difficult tasks. Systematic reviews generally do not fully address uncertainty, tradeoffs among alternative outcomes, and differences among individuals in their preferences (values) for the alternative outcomes.[1] Uncertainty may remain when clinical studies provide evidence only for surrogate outcomes; have small sample sizes, limited followup durations, or deficiencies in their design and conduct; or provide insufficient information on relevant patient subgroups. Tradeoffs among patient-relevant outcomes are common; for example, effective treatments may be associated with adverse effects (e.g., drug reactions), and informative diagnostic tests may result in overdiagnosis and overtreatment. Patients' preferences for different outcomes (along with those of their families and other caregivers) need to be considered when assessing the consequences of alternative actions.

Models and simulations are valuable tools for inference and decisionmaking in the presence of uncertainty, tradeoffs, and varying preferences. Models can also be used to structure investigators' thinking, facilitate the communication of assumptions and results, synthesize data from disparate sources, make predictions, and examine and understand the impact of (possibly counterfactual) interventions. These goals are highly relevant to the evidence syntheses prepared by the Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Centers (EPCs).

This document provides general guidance in the form of good-practice principles and recommendations for EPCs preparing modeling and simulation studies. Currently, EPCs have variable expertise and experience in modeling, but the growing complexity of the health care questions addressed in EPC reports suggests that use of modeling may increase in the future. AHRQ has recognized the need for an overview of good practices for modeling and simulation to guide these efforts within the EPC Program.[2,3]

The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and the Society for Medical Decision Making (SMDM) have recently published detailed modeling recommendations (including discussion of many applied examples reviewed in the context of those recommendations). The availability of this information makes the development and use of models in conjunction with systematic reviews prepared by EPCs more feasible.[4-17] In view of this, we sought to produce general guidance for modeling and simulation in the context of health technology assessment[*] on the basis of a systematic review of existing guidance documents, discussions with technical experts, and extensive deliberation within the EPC Program. The guidance aims to (1) encourage the use of good modeling and reporting practices in conjunction with systematic reviews without being too prescriptive about how to develop specific models and (2) describe how systematic reviews can increase the transparency of the modeling process and contribute to the development of useful models. We believe that this guidance applies generally, but to maintain focus, we emphasize models that could accompany systematic reviews produced by the EPC Program. We aim to establish a baseline understanding of modeling and simulation for EPC reports. We also provide an extensive list of references that can be a source of detailed information (and numerous examples) about specific modeling and simulation methods. We

---

[*]Readers who wish to review the recommendations without going over this introductory material can turn to the section titled "Good-Practice Recommendations for Modeling and Simulation in the Context of Health Technology Assessment."

deemphasize issues specific to economic modeling, because economic assessments are not a priority of the EPC Program.

Model Types and Scope of the Guidance

A model is a representation of select aspects of reality in a simplified way. There exist physical (e.g., a scaled-down airplane wing tested in a wind tunnel), analog (e.g., a DNA molecule is like a staircase), or theoretical (formal) models (e.g., a mathematical description of the flow of air around an airplane wing).[18,19] Models that can be prepared in conjunction with systematic reviews are exclusively theoretical. The starting point for most theoretical models is a conceptual model, a simplified natural language or pictorial representation of reality. A sound conceptual model is a prerequisite for the development of mathematical (quantitative) models. The analytic frameworks that are used to guide the conduct of systematic reviews prepared by the EPCs can form the basis of conceptual models representing the underlying decision or care process; for example, analytic frameworks[20-24] often resemble the gist of decision trees or influence diagrams.[25,26] Background information on analytic frameworks is provided elsewhere in the Methods Guide for Comparative Effectiveness Reviews.[27,28]

Mathematical models are a large and diverse group of formal models that use variables, together with mathematical symbols that represent relationships between the variables. Perhaps the most common mathematical models encountered in practice are multivariable regression models (e.g., ordinary least-squares regression, logistic regression). 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 (also referred to as "models of data").[29,30] They describe how the response varies over covariates, without necessarily referring to assumptions about the underlying mechanisms.[†] The literature addressing these models is vast (e.g., in Statistics and Computer Science) and is not covered in this guidance document. Instead, we address mathematical models that attempt to capture "true" (structural) relationships among their components (also referred to as "models of phenomena") and combine information from multiple sources.[19,31-36] These models include declarative (e.g., Markov models), functional (e.g., compartmental models), and spatial (e.g., geographic information systems) models. In applied work, elements of these model subtypes are commonly combined (multimodels).[31] Thus, the models covered by this guidance include those considered by the National Research Council[37] ("replicable, objective sequences of computations used for generating estimates of quantities of concern") and the 2003 ISPOR principles of good practice[38] ("analytic methodology that accounts for events over time and across populations, that is based on data drawn from primary and/or secondary sources"). Simulation studies operate ("run") fully specified models to understand the phenomenon or process of interest, to predict its behavior, or to obtain insight into how its course can be modified by an intervention.

Goals of Modeling and Simulation in EPC Reports

We briefly consider the potential goals of modeling when performed in conjunction with systematic reviews.[39-53]

- **To structure investigators' thinking and to facilitate the communication of data, assumptions, and results:** Modeling can help investigators organize knowledge about a

---

[†]In some cases behavioral models (e.g., regression) are used to estimate the parameters of structural models.

topic area, formalize the research questions, and communicate assumptions and results to peers (e.g., topic or methodological experts) and other stakeholders (e.g., patients or decisionmakers).[40]

- **To synthesize data from disparate sources:** Evidence on a specific research question may be available from multiple sources, and a single source may contribute information to the estimation of more than one model parameter (or functional combinations of parameters). Modeling provides mathematical tools for evidence synthesis and the assessment of consistency among data sources. For example, models can be used to combine information from clinical trials on the effect of treatment on intermediate outcomes with information from cohort studies on 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"), the future (forecasts), or other populations or outcomes (sometimes referred to as "extrapolations"). They can also pertain to the prioritization and planning of future research.[54] These predictions may be useful in themselves, even without reference to the anticipated effects of interventions.

- **To support causal explanations and infer the impact of interventions:** Modeling can be used to assess the effects of possible (hypothetical) interventions and to provide mechanistic explanations for observed phenomena.[55-57] When used this way, models are taken to encode structural causal mechanisms or to approximate such mechanisms sufficiently well. This allows models to examine counterfactual scenarios ("what if" analyses).

- **To inform decisionmaking:** The decisions that can be informed by modeling, even in the relatively narrow context of systematic reviews, are extremely diverse.[39,40] They include decisions about patient-level care (accounting for treatment heterogeneity and variation in patient preferences), drug or device licensing, health care policy, and the need to conduct additional research.

Communicating assumptions, synthesizing evidence, and informing decisions are probably the most common goals of the models and simulations that would be developed in conjunction with systematic reviews. That said, the listed goals are not mutually exclusive, and typically the same model is used to achieve multiple goals.

When Is Modeling Worth the Effort?

Because issues related to the appropriateness of modeling in EPC reports are addressed by existing guidance,[2,58] this document does not provide detailed recommendations to help investigators decide whether modeling and simulation should be undertaken. However, we briefly describe the typical conditions under which modeling is worth the extra effort.

The development of models, especially models that can be used to understand complex phenomena and to inform difficult decisions, is a demanding process. 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 weaknesses. Although this document and the references cited herein provide information on methods for modeling and simulation, defining the circumstances under which modeling is worth the investment of time and resources beyond those required for a systematic review is challenging. In general, modeling is most useful

when the research question is complex, data sources have limitations (e.g., sparse or conflicting evidence, high risk of bias, or followup durations shorter than the time horizon of interest), outcomes involve complex tradeoffs, and choices are preference laden. Consideration should also be given to whether modeling is likely to produce results that the review's intended audience will deem credible and useful. 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 health care policy should also be considered when deciding whether modeling is worth the effort.

Modeling and Simulation Process

The specifics of model development and assessment vary across specific applications because modeling and simulation studies are used to address diverse research questions. Nonetheless, the key activities for the development of quantitative models, within the scope of this guidance, can be identified:[31,59-69]

1. **Specifying the research question and setting the modeling goals:** Specifying an answerable research question that addresses the needs of the relevant stakeholder(s)
2. **Conceptualizing the model and specifying its mathematical structure:** Determining which components of a disease or process need to be represented in the model to address the research question and describing their relationships
3. **Assembling information:** Identifying data sources, eliciting expert opinions, and processing all information that will be used as an input to the model
4. **Implementing and running the model:** Running the model using mathematical or numerical analysis methods
5. **Assessing the model:** Examining whether the model attains its stated goals; detecting model shortcomings by examining the model, as well as by comparing its output with prior beliefs, data, and other similar models
6. **Interpreting and reporting results:** Presenting the model findings in a way that addresses the research question

Model development is an iterative and dynamic process.[12,62] Multiple iterations are typically needed between the phases outlined here 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 arise that require further data collection. Similarly, data deficiencies that are detected during model assessment often require restructuring of the model, supplemental data collection, or other modifications of the modeling strategy.

Guidance Development Process

This guidance document is the culmination of a multistep process of summarizing existing recommendations and soliciting stakeholder input. Earlier steps of the process are described in a companion report ("Modeling and Simulation in the Context of Health Technology Assessment: Review of Existing Guidance, Future Research Needs, and Validity Assessment"; publication details to be provided by OCKT). Briefly, with input from a multidisciplinary team of clinical, policymaking, modeling, and decision analysis experts, we updated and expanded two systematic

reviews of recommendations for the conduct and reporting of modeling and simulation studies,[2,70] as described in detail in the companion report. To assess the relevance of published methodological recommendations for modeling and simulation, we discussed the results of our systematic review in person with a panel of 28 stakeholders, including patient representatives, care providers, purchasers of care, payers, policymakers (including research funders and professional societies), and principal investigators. To examine worldwide health technology assessment procedures and practices, we reviewed the Web sites of 126 international health technology assessment agencies and institutes for their guidance or standards for whether to conduct modeling and simulation studies, and if so, how to perform and report such studies. We solicited feedback on a set of draft recommendations from senior researchers at EPCs and AHRQ with experience in modeling and simulation methods. The draft recommendations were presented for comment and affirmation or dissent at the 2014 Annual Meeting of SMDM[71] and the 2015 Annual Meeting of ISPOR.[72] In these oral presentations, participants in the audience were invited to express disagreement (by raising their hands) after the presenter read out each recommendation statement. These reactions and additional comments from the audience were incorporated in revised versions of the report; however, this should not be taken to imply endorsement of these recommendations by either SMDM or ISPOR. External peer reviewers and AHRQ personnel further vetted a draft version of this report, and we incorporated their input into the final version; however, this should not be taken to imply endorsement of the recommendations by the reviewers. Lastly, EPC Directors were invited to consider the final report for adoption as guidance for the EPC Program, and the majority voted in favor of including this report in the EPC Methods Guide for Comparative Effectiveness Reviews.

Based on the gathered systematic review evidence on modeling recommendations and the processes described here, we prepared the final version of good-practice principles and recommendations for developing models in conjunction with systematic reviews. We categorized the modeling recommendations by whether they pertain to the conceptualization and structure of the model, data, consistency, or interpretation and reporting of results, as was done in earlier related work.[3,70,73-75]

Our approach in crafting the guidance was consistent with the framework for making methodological recommendations that we have described in an earlier publication.[76] We made a concerted effort to review the relevant literature and consulted with recognized experts in the field. Although the final version of the guidance reflects the authors' best judgment and is not the product of consensus among the stakeholders involved in the process, we have explained the rationale for each recommendation and, when available, have provided evidence that the recommendation should be preferred.

Terminology and Definitions

Table 1 defines terms used in this document.

Table 1. Definitions of terms

Term	Explanation and Elaboration	Comments
Model	A representation of select aspects of a phenomenon or system in simplified form.	Models are constructed following a process of abstraction and idealization. We focus on models that represent reality by means of mathematical relationships (using mathematical operations, symbols, and variables).
Decision model	A model of the choice between two or more alternative options (i.e., alternative actions or rules for making decisions over time).	These models are used to explore the consequences of different choices (e.g., the decision to use a particular treatment when two or more alternatives are available).
Simulation	The operation ("running") of a model.	"Simulation" is sometimes defined more restrictively as using numerical methods (especially methods that use computer-generated pseudorandom numbers) when analytic solutions are cumbersome or intractable. We use a broader definition in this document.[77-79]
Computer simulation	A simulation carried out using a computer.	Almost all simulations relevant to health technology assessment are computer simulations.
Model component	An element of a model (e.g., variables, health states, agents, and processes).	The term is purposely generic to encompass all model types.
Uncertainty	Lack of certainty; a state of limited knowledge.	Many typologies of uncertainty have been proposed in the context of modeling and simulation.[9,16] Such typologies are useful to the extent that they contribute to the success of applied modeling work; a complete theoretical treatment of the concept of uncertainty is beyond the scope of this guidance but is central to modeling practice (and addressed by a vast literature to which we provide a few selective links[80-89]). In applied work we find it useful to consider scenario, structural, and parametric uncertainty (defined in this table). In some cases it is also useful to consider residual (predictive) uncertainty (i.e., in nondeterministic models), to represent inherent stochasticity not captured by scenario, parametric, or structural sources.[90-92]
Scenario uncertainty	Uncertainty in the scenarios for which the model is run.	Exogenous forces not under the control of the investigators can be a source of scenario uncertainty. For example, in a model of a hospital that uses data on the number of patients attending the emergency room, running the model under an extreme scenario (e.g., to simulate a doubling of the number of patients attending) is subject to scenario uncertainty.
Structural uncertainty	Uncertainty due to incomplete understanding of the modeled phenomenon. Typically, this pertains to functional forms of relationships between model variables. At a more fundamental level, structural uncertainty always exists because the "true" relationship between variables in the real world cannot be uncovered from data.	In some cases, the distinction between structural and parametric uncertainty is a matter of definition. Some structural uncertainty would become parametric uncertainty if relevant data were available to the modeler. Structural uncertainty is conditional on the modeled scenario.
Parametric uncertainty	Uncertainty about the values of model parameters.	Parametric uncertainty is sometimes referred to as "aleatory" or "stochastic" uncertainty. It is conditional on the model structure and the scenario being modeled.

Term	Explanation and Elaboration	Comments
Uncertainty analysis	Analysis that addresses lack of knowledge regarding the model structure, parameter values, and scenarios, or any inherently nondeterministic aspect of the model.	The most common goal of uncertainty analyses is the propagation of uncertainty from model inputs to model outputs (e.g., accounting for lack of certainty when estimating treatment effects or event rates).
Propagation of uncertainty	The process of assessing uncertainty in model outputs by incorporating all sources of uncertainty in the model inputs. This derives from the uncertainty in model inputs. Uncertainty can be propagated analytically (exactly or up to an approximation) or numerically (e.g., with forward Monte Carlo simulations or with Markov Chain Monte Carlo [MCMC] methods).	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 the term PSA in this work to avoid confusion. We draw a distinction between the propagation of uncertainty and sensitivity analysis. Propagation of uncertainty is important for obtaining valid results (especially in nonlinear models) and for correctly assessing the value of obtaining additional information. However, simply propagating uncertainty from inputs to outputs does not fulfill the goals of sensitivity analysis (assess the influence of inputs on outputs).
Sensitivity analysis	The process of varying model variables over a set of values that are of interest (e.g., because they are deemed plausible) and examining impact on results.	Sensitivity analysis can be used to evaluate the impact of different inputs on model outputs or to examine the implications of different values of unidentifiable model parameters on model results. Sensitivity analysis often has a "continuous" character (e.g., the magnitude of a treatment effect is varied over multiple finely spaced values within the plausible range).
Stability analysis	Performing discrete actions and evaluating their impact on results. Examples include changing the structure of the model, such as using alternative specifications (e.g., different functional forms for relationships between variables) and systematically excluding input data (e.g., leave one study out in a meta-analysis).	Stability analysis involves discrete analytical decisions (e.g., the summary treatment effect can be obtained via a random effects or common effect meta-analysis model). Stability analyses assess the impact of alternative analytic approaches on model results (e.g., using a weekly or monthly cycle length in a discrete time Markov chain, or assessing robustness of model results to estimate treatment effects after excluding studies one at a time). Such analyses are often described as explorations of "methodological uncertainty;" however, the term "stability analyses" is more appropriate because their purpose is to examine robustness to alternative methodological choices, not to reflect uncertainty about the modeled phenomenon.
Model verification	The assessment of the correctness of the mathematical structure (e.g., absence of mistakes in the logic) and of the implementation of the computational model (e.g., absence of software bugs, suitability of numerical algorithms).	Model verification includes the identification and correction of mistakes in the model logic and software bugs, and an assessment of the suitability of numerical algorithms used in the model.
Model validation	Validation is the comparison of the model and its output with expert beliefs (conceptual validity and face validity), data (operational and predictive validity), and other models (cross-model validity).	Validation includes various checks of the face, operational, and external validity of the model. It is closely related to the concepts of representational fidelity (is the model a good representation of the actual system or process?) and behavioral fidelity (is the model output similar to the behavior of the actual system or process?). Because complete model validity cannot be established in the affirmative, a model can only be evaluated with respect to a specific purpose. The examination of model consistency (model assessment) includes attempts to verify and validate attributes of the model and establish its credibility.
Preferences (values)	We use the term in a broad sense, to denote how desirable a given outcome is for an individual or a group.[93]	Preferences are used in the valuation of outcomes. Sometimes "utility" is defined as a "cardinal measure of the strength of one's preference."[94-96]

Principles for Good Practice in Modeling and Simulation

We begin by outlining general principles for the conduct and reporting of modeling and simulation studies (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.

At the start of the modeling work, investigators should consider (1) the goals of the modeling application; (2) the nature of the modeled phenomena; (3) the available expertise; and (4) objective constraints in terms of available time, data, or other resources. Further, when developing, implementing, and running models, there are many methodological decisions to be made.[97-99] These decisions should be recorded, justified, and subjected to stability analysis.[98,100,101] This can be done most conveniently by specifying the modeling methods in a protocol (while the study is being planned) and by generating detailed documentation (while the study is conducted and after it is completed).

Table 2. Principles for good practice in modeling and simulation

The research question, modeling goals, and the scope of the model 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 the model's scope; data sources should be described.
The model should reflect uncertainty in inputs.
Sensitivity analyses (to assess the influence of model inputs) and stability analyses (to evaluate the impact of modeling decisions) should be undertaken and reported.
Models should be assessed for their ability to address the research question within the stated scope.
Modeling methods should be transparent. Adequate details about the structure, data, and assessment methods should be reported so that the modeling process is replicable.

We provide good-practice recommendations for modeling and simulation in conjunction with systematic reviews, organized by conceptualization and structure, data, assessment and consistency, and interpretation and reporting.[2,3,70,73,74] Briefly, structure and data comprise 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 of these areas of modeling.

This guidance is provided to facilitate the use of modeling and simulation 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 assessing models and simulations. In general, all recommendations should be viewed pragmatically when embarking on a specific project. (This is sometimes referred to as the "rule of reason."[102]) Systematic reviewers and modelers should exercise judgment when deciding whether specific recommendations are likely to have an appreciable impact on the review conclusions and should balance feasibility with the desire to conduct extensive analyses.

It is not possible to provide detailed recommendations about which structures to use in which cases or instructions about the implementation and manipulation of various model types. Interested readers should consult some of the numerous books, technical reports, and papers available on this topic (several of which are cited in the section on good-practice recommendations), including systematic overviews of existing guidance,[2,3,70,74,103] the recent ISPOR-SMDM recommendations,[4-17] other sources of guidance (including methodological appraisals),[38,47,73,98,100,102,104-162] expository works with a focus on biomedical modeling,[61,62,163-183] and the vast literature on modeling and simulation in other subject areas.[184-210]

Table 3. Operational definitions for the conceptualization and structure, data, model assessment and consistency, and interpretation and reporting framework

Recommendation Areas	Description of What Is Encompassed	Examples
Conceptualization and structure	Conceptualization pertains to the decision to use modeling, and the delineation of the perspective and scope. Structure pertains to variables, health states, and other components of the model, as well as how they relate to each other (i.e., the mathematical scaffold of the model).	In a discrete-time Markov model the disease states, variables informing transition probabilities, mathematical relationships among the variables, and time horizon of the model characterize the model's structure.
Data	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).
Model assessment and consistency	A model can only be evaluated with respect to the specific goals of modeling (the goals are determined by the research question). Model assessment examines the extent to which the model achieves the stated goals of representing the phenomenon of interest and the effects of alternative actions on pertinent outcomes. Model assessment activities occur throughout the process of model development. Model consistency includes attempts to verify and validate the model and establish its credibility.	Determination of whether the model has logical errors and whether the model output is consistent with expert opinion, observed data, or other models.
Interpretation and reporting of results	Summarizing model output to achieve the goals of modeling (e.g., to further understanding the topic or to inform decisionmaking).	Risk diagrams (to represent model-based risk analyses) and tornado diagrams (to summarize sensitivity analyses using ordered bar charts).

Sensitivity, Stability, and Uncertainty Analyses

Many of the recommendations in this guidance emphasize the need to perform sensitivity and stability analyses. By sensitivity analysis we mean the process of varying model variables over a set of values that are of interest and examining impact on results. Such analyses can be used to evaluate the impact of different inputs on model outputs or to examine the implications of different values of unidentifiable model parameters on model results. Sensitivity analysis can be local (e.g., examining changes in output in response to infinitesimal perturbations of the inputs) or global (e.g., examining changes over a broader range of input values). Many methods for sensitivity analysis, both stochastic and deterministic, have been proposed; the choice among available methods should be dictated by the goals for the modeling effort.[69,203,211-224]

Stability analyses are assessments of the impact of alternative analytic approaches on model results (e.g., whether to use a monthly or weekly cycle length in a Markov model). Often such analyses are described as explorations of "methodological uncertainty;" however, the term "stability analyses" is more appropriate because their purpose is not to reflect uncertainty about the modeled phenomenon but simply to examine robustness to alternative methodological choices.[101]

Handling of uncertainty is related to, but distinct from, sensitivity analysis. We organize uncertainty into three types: scenario, structural, and parametric uncertainty. A fourth type, "residual" (predictive) uncertainty, may also be important to consider. (For definitions, see Table 1.) Structural uncertainty is perhaps the most challenging to address because empirical observations are always compatible with a large number of alternative model structures. Methods for handling structural uncertainty include stability analyses (i.e., building models with alternative structures),[183] model expansion by "parameterizing" alternative structures, and formal model averaging.[92,225-228] Proper handling of parametric uncertainty is necessary for valid inference using on models and simulations.[90-92,225-227,229,230] Although uncertainty in model outputs should have no impact on decisionmaking under a Bayesian decision theoretic view,[231,232] we believe that decisionmakers are often interested in the degree of certainty around model outputs and (heuristically) consider that information when making decisions. In addition, proper handling of uncertainty is critical for using models to determine the need for future research, prioritize specific research activities, and plan future studies.[233]

We draw a distinction between uncertainty analysis (i.e., the propagation of stochastic uncertainty from model inputs to outputs) and sensitivity analysis. Propagation of uncertainty (analytically or via various Monte Carlo methods) is important for obtaining valid results (especially in nonlinear models) and for correctly assessing the value of obtaining additional information. However, simply propagating uncertainty from inputs to outputs cannot fulfill the goals of sensitivity analysis (i.e., to assess the influence of model inputs). Confusingly, the term "probabilistic sensitivity analysis" is often used in the literature to describe uncertainty propagation that is not coupled with attempts to identify influential inputs. We propose that the term "uncertainty analysis" or "uncertainty propagation" should be used to describe such analyses, and that the term "sensitivity analysis" should be reserved for analyses that aim to assess the influence of model inputs.[234-240] In many cases, sensitivity analysis will be conducted in models that also propagate uncertainty, but the two activities have different goals.

Model Assessment and Consistency

Model assessment is meaningful only with respect to the goals of modeling. Assessment activities occur throughout the life cycle of model development.[65,241-243] The examination of model consistency includes attempts to verify and validate the model and establish its credibility; this is a critical issue in modeling and has been extensively considered in the literature (both in health care contexts and beyond).[65,67,68,128,241,242,244-270] Verification (internal validity) is the assessment of the correctness of the mathematical structure (e.g., absence of mistakes in the logic) and of the implementation of the computational model (e.g., absence of software bugs, suitability of numerical algorithms). Validation is the comparison of the model and its output with expert beliefs, data, and other models. In practice, validation includes various "checks" of face, external, predictive, and cross-model validity. It is closely related to the concepts of representational fidelity (i.e., whether the model is a good representation of the modeled system or process) and behavioral fidelity (i.e., whether the model output is similar to the behavior of

the modeled system or process, sometimes referred to as dynamic fidelity), which arise in all types of modeling.[271] Importantly, a model can be evaluated only with respect to a specific purpose; complete model validity cannot be established in the affirmative; in fact, it has been suggested that a model that successfully passes an evaluation should be considered corroborated, not validated.[272,273]

Good-Practice Recommendations for Modeling and Simulation in the Context of Health Technology Assessment

Conceptualization and Structure:

Explicitly state the research question and the modeling goals. Describe and justify the decision to use modeling to address the research question. Use a conceptual model to guide the development of the mathematical model.

Choose a perspective depending on the research question and the relevant stakeholders. There is no a priori preferred modeling perspective.

Specify the model scope to be consistent with the research question and modeling perspective. Describe and justify the model scope.

Specify and implement the structure of the mathematical model to correspond to the research question, the model's scope, and the modeling perspective. Provide the rationale for the chosen mathematical structure; explain and justify structural assumptions and computational approximations.

Allow for comparisons among all interventions that are relevant to the research question and the model's scope.

Use a time horizon long enough to allow all relevant outcomes to be fully evaluated.

When deciding how to handle time, spatial location, interactions among agents, and health states, consider the nature of the modeled phenomenon and the convenience of (and approximation errors associated with) alternative choices.

Determine the targeted level of complexity (or parsimony) based on the research question and the model scope. It is often preferable to build a simpler model first and progressively increase the degree of complexity.

Data:

Describe the methods for identifying and analyzing data. Make data choices based on the research question and the model's scope and structure. Report all data sources clearly and provide explicit references. Obtain values for model inputs following epidemiological and statistical principles. Use a "best evidence approach" when selecting data sources for model parameters. Obtain estimates for influential parameters using systematic review methods.

Assess the risk of bias of the available evidence and account for sources of bias when estimating values for model parameters.

Use formal elicitation methods to quantify expert opinion and its associated uncertainty. Use appropriate methods to quantify preferences for different outcomes.

Describe and justify the assumptions required for extrapolating beyond observed data and transporting information from various data sources to a common (target) setting. Subject these assumptions to stability and sensitivity analyses.

In statistical analyses, account for heterogeneity.

Use modeling methods that propagate uncertainty from inputs to outputs.

Model Assessment and Consistency:

Evaluate the model with respect to the specified modeling goals.

Anticipate, detect, and correct errors in the model's logic and implementation.

Invite topic experts to review the model's structure and outputs and to judge whether these seem consistent with their expectations. Verify, describe, and explain counterintuitive model results.

Assess the consistency between model outputs and the data on which the model was based.

Do not withhold data from model development for the sole purpose of assessing model validity.

Decide whether using future observations to assess a model is appropriate based on the research question and the modeling goals.

Update the model as new data become available, new interventions are added, and the understanding of the investigated phenomenon improves.

If models addressing the same research question are available, compare their results to the new model and explain any discrepancies.

Interpretation and Reporting of Results:

Be transparent about the model structure, computational implementation, and data. Report results clearly and in a way that addresses users' needs.

Interpret and report results in a way that communicates uncertainty in model outputs.

Fully disclose any potential conflicts of interest.

Explanation and Elaboration of Recommendations for Modeling and Simulation

Conceptualization and Structure

Explicitly state the research question and the modeling goals. Describe and justify the decision to use modeling to address the research question. Use a conceptual model to guide the development of the mathematical model.
Modeling is useful for addressing many research questions, especially questions that are not directly answerable using existing empirical data. A well-specified and explicitly stated research question is critical for modeling and simulation.[274,275] Models prepared in conjunction with systematic reviews should be based on a clear conceptual model.[5,12,276-279] Defining the question and objective of the analysis may require using literature-based information, expert knowledge, and input from stakeholders (e.g., the Key Informants and Technical Experts who provide input during the preparation of EPC reviews).[12] If a model is to be prepared in conjunction with a systematic review, issues related to the model should be considered during the planning of the review (e.g., possible model structures, anticipated data). Conversely, decisions related to the construction and use of models should be informed by the design of the systematic review (e.g., regarding the various populations, interventions, and outcomes that could be considered in the model).

Choose a perspective depending on the research question and the relevant stakeholders. There is no a priori preferred modeling perspective.
The modeling perspective determines the methods for choosing and handling consequences, preferences, and, if examined, costs in the model; thus it should depend on the research question and the relevant stakeholders (e.g., decisionmakers).[280] For example, when modeling aims to address the clinical options faced by an identifiable patient, the appropriate perspective is that of the individual patient. In contrast, when the goal is to inform health policy decisionmaking of a public payer or a Federal agency, one should prefer a payer or societal perspective.[281] The societal perspective (which considers impact on sectors beyond health care and includes time costs, opportunity costs, and community preferences) may allow for a comprehensive accounting of benefits, harms, and costs and can serve as a "base case," facilitating comparability of the results across health policy analyses.[282] For this reason, it has been recommended as an appropriate "default" perspective.[102] However, obtaining appropriate data for modeling from a societal perspective can be challenging (e.g., accommodating equity concerns).[282-284]

Specify the model scope to be consistent with the research question and modeling perspective. Describe and justify the model scope.
The scope of a mathematical model includes the condition or disease of interest, populations, risk factors, and diagnostic or therapeutic interventions. For decision models, the scope also includes alternative strategies, decision-relevant outcome quantities (e.g., life-years gained, quality-adjusted life-years, disability-adjusted life-years), the decision (optimality) criteria, the time horizon, and the decisionmaking perspective. Determining the scope of the model is akin to defining a systematic review's study selection criteria (e.g., population, intervention, comparator,

outcomes, timing, and setting). By necessity, a model represents only some aspects of the phenomenon or process under study. The research question defines how complex the model should be and what aspects of reality are represented or omitted (for parsimony). For example, many research questions in health care pertain to length of life; thus, mortality outcomes should be within the scope of models answering these questions.

Specify and implement the structure of the mathematical model to correspond to the research question, the model's scope, and the modeling perspective. Provide the rationale for the chosen mathematical structure; explain and justify structural assumptions and computational approximations.

The preferred model structure depends on the research question and the model's 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). Health states, transitions between health states, and functional relationships between parameters should reflect the understanding of the course of the disease. Detailed guidance on choosing among alternative mathematical structures and on implementing them in computational models is beyond the scope of this document. Readers are referred to the extensive technical literature in health care[4-17,38,46,61,163,165,167,181,285-307] and other fields.[31,187,192,308-310] Of note, relatively simple models (e.g., decision trees, time-homogeneous Markov state transition models) may be appropriate for use in the setting of many EPC evidence reports, particularly when the goal of modeling is to contextualize the evidence and extend review findings.

Allow for comparisons among all interventions that are relevant to the research question and the model's scope.

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 a 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 justified.

Use a time horizon long enough to allow all relevant outcomes to be fully evaluated.

When comparing alternative interventions, the time horizon should be long enough to allow the manifestation of differences in relevant 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 lifetime horizon is needed, particularly when modeling the effects of long-term treatment of chronic disease. The time horizon choice has implications for the data used to populate models; for example, lifetime horizons almost always require the extrapolation of treatment effects well beyond the followup duration of available clinical trials.

When deciding how to handle time, spatial location, interactions among agents, and health states, consider the nature of the modeled phenomenon and the convenience of (and approximation errors associated with) alternative choices.
For example, when deciding how to deal with time, we have three options: (1) do not model it explicitly (as in simple 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 models). 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).[311] 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 spatial location, interagent interactions (e.g., interactions between modeled individuals), and health states in various degrees of granularity (e.g., disease severity).

Determine the targeted level of complexity (or parsimony) based on the research question and the model scope. It is often preferable to build a simpler model first and progressively increase the degree of complexity.
Models should be as complex as needed to capture all pertinent aspects of the system being modeled, but not more ("rule of reason").[102,312-314] At the same time, models should be as simple as possible to facilitate timely development, error checking, and validation. Simple models are generally more accessible to nontechnical stakeholders, and results from such models 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.[5,12,313,315-317] In general, it is preferable to first build a simpler model and progressively increase the degree of complexity in order to facilitate error checking and ultimately obtain a reliable model that satisfies the goals of the modeling effort.

Data

Describe the methods for identifying and analyzing data. Make data choices based on the research question and the model's scope and structure. Report all data sources clearly and provide explicit references. Obtain values for model inputs following epidemiological and statistical principles. Use a "best evidence approach" when selecting data sources for model parameters. Obtain estimates for influential parameters using systematic review methods.
To enhance transparency and face validity, the source of each data element should be identified fully. This applies both to the base case data and to the range of values examined in sensitivity analyses for each data element. Particularly for data that are not derived from systematic review and meta-analysis, the rationale for why the given value was chosen should be provided.

All major assumptions and methodological choices for determining model inputs should be reported and justified. Modelers must select, appraise, and synthesize appropriate study types for

each model parameter.[9,16,179,318] A recent EPC Methods Research Report provides general guidance on "best evidence" strategies in systematic reviews.[319] Data from randomized trials cannot be used to inform all model parameters because (1) some parameters are best estimated from other study designs (e.g., the prevalence of a risk factor is best estimated from a sampling survey of a representative population; the 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 at all. In all these cases, evidence from other study designs will have to be included in the model.

For modeling and simulation studies prepared jointly with a systematic review of studies of interventions, estimates of treatment effects and other inputs (together with corresponding measures of sampling variability) should be used to inform the relevant model parameters. In particular, model parameters likely to have a large influence on model results should be informed by a systematic and replicable process that aims to minimize bias.[320-324] 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 include and the choice of synthesis methods.[158,318,324-330]

Data on other model inputs (e.g., prevalence, incidence, resource use or costs, and utilities) may be obtained through processes other than systematic review. Appropriate sources of such data can include de novo analyses of registries and other large observational studies, completed studies found through a nonsystematic approach, stakeholder panel opinions, and domain expert judgments. When retrieving and processing data, modelers often make decisions that may appreciably impact results (e.g., use of operational selection criteria to determine the relevance of published studies or use of approximate calculations when extracting data from published studies). All such decisions should be recorded, justified, and reported in the model's documentation. Supplementary material describing detailed methods and data sources can be made available electronically.

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 methods (meta-analysis, network meta-analysis, or generalized evidence synthesis).[321,322] When data from multiple sources are combined to estimate model parameters, the examination of consistency among sources is an important task. For example, inconsistency in network meta-analyses can indicate the presence of effect measure modification or bias in the evidence base. This guidance does not provide detailed information on the conduct of quantitative synthesis for different types of data structures; both EPC guidance and many other sources can be consulted for detailed descriptions of meta-analysis and evidence synthesis methods.[93,326,331-358]

Related to the idea of estimation of model parameters is the concept of model calibration, the tweaking of (typically unidentifiable or weakly identifiable) model parameters to improve the "closeness" of the model outputs with empirical data.[270,359-364] We do not distinguish sharply between processes for calibration and estimation because the analytic goals and the methods to achieve them are similar.[365,366] This becomes particularly clear when considering Bayesian simulation models.[367-370] A more detailed discussion of model calibration is provided in a companion report ("Modeling and Simulation in the Context of Health Technology Assessment: Review of Existing Guidance, Future Research Needs, and Validity Assessment").

Assess the risk of bias of the available evidence and account for sources of bias when estimating values for model parameters.

Models typically are specified with respect to "true" parameters, but empirical studies provide parameter estimates that are subject to bias. Consequently, model inputs should be adjusted ("corrected") for biases. In general, modelers should avoid using unadjusted, incompletely adjusted, or inappropriately adjusted results simply because no other information is available.[371,372] When the available evidence base is large, this may be possible by obtaining information from studies free of such problems. However, in many cases, studies free of bias are unavailable or represent a very small fraction of the available evidence. In such cases, modelers should adjust study results to account for bias and associated uncertainties (i.e., multiple bias modeling) and should undertake sensitivity analyses.[373,374]

Because the factors that determine the direction and magnitude of bias depend on the modeling context and the design, conduct, and analysis of the studies under consideration, bias assessment has to be tailored on a case-by-case basis.[375,376]

The direction and magnitude of bias introduced by different factors, uncertainty about bias, and the relationship between biasing factors should be incorporated into the analyses. In most cases, "bias parameters" cannot be identified from study data; thus, modelers have to use methods that incorporate external information (empirical and 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).[371,373,377-388]

Use formal elicitation methods to quantify expert opinion and its associated uncertainty. Use appropriate methods to quantify preferences for different outcomes.

When no empirical evidence is available for parameters of interest, modelers have to rely on expert opinion (e.g., to estimate probabilities of event occurrence). Preferences for different outcomes can also be elicited using specialized methods. The literature on methods for eliciting expert opinions and for determining preferences is extensive and is not covered in this report; the measurement of preferences is a contentious topic.[389-398]

Current technical expert and stakeholder engagement processes in systematic reviews can incorporate formal methods for eliciting expert opinions and quantifying preferences for different outcomes (e.g., by expanding the roles of Key Informants and Technical Experts involved in the development, refinement, and conduct of systematic reviews). 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 quantities of interest.[399] When elicitation of preferences cannot be performed *de novo*, the literature can be used as a source of information.

Describe and justify the assumptions required for extrapolating beyond observed data and transporting information from various data sources to a common (target) setting. Subject these assumptions to stability and sensitivity analyses.

A particular challenge arises when there is a need to extrapolate beyond the observed data (e.g., to longer followup periods or to other populations). Such extrapolations are based on untestable assumptions that 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).

Models often use data obtained from diverse sources.[45] 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 mechanisms underlying the modeled phenomenon.[121,400,401] Consideration should be given to formal (causal) methods for assessing the transportability of results across domains.[402-407]

In statistical analyses, account for heterogeneity.

As a general principle, models and simulations should account for heterogeneity, defined as nonrandom (systematic) variation.[408-412] Attempts should be made to explain heterogeneity by incorporating information on determinants of variability via appropriate statistical methods (e.g., subgroup or regression analyses). Because scientific understanding of any topic is likely to be incomplete (e.g., important modifiers of effect may be unknown) and because lack of data 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 is common in meta-analyses of treatment effects that use published (aggregate) level data. In such cases, efforts to explain heterogeneity rely primarily on metaregression methods, and residual heterogeneity is accounted for by using random-effects models.[330,413-416] 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.[417,418] Person-level data can allow models and simulations to meaningfully incorporate heterogeneity;[419-426] however, such data are rarely available in systematic reviews prepared by EPCs or in meta-analyses published in peer-reviewed journals.[427]

Use modeling methods that propagate uncertainty from inputs to outputs.

Appropriate data analysis methods should be used to obtain valid parameter estimates and to propagate uncertainty from inputs to outputs.[62,235-238,408,409,428-434] Sometimes this can be done analytically, either exactly or by approximating up to an order of error (e.g., with the delta method[435]). In most cases, it is computationally convenient to propagate uncertainty with numerical methods, typically with a forward Monte Carlo approach; in the medical modeling literature, this is often, and somewhat inappropriately, termed "probabilistic sensitivity analysis."[238,305,306]

Detailed descriptions of methods for conducting probabilistic analyses are available elsewhere in the literature on modeling in health care and other fields.[229,230,232,238,239,408,409,436-440] Of note, probabilistic methods for incorporating and propagating uncertainty in models do not eliminate the need for stability and sensitivity analyses. For example, the choice of the distribution is rarely unique. Thus, it may be important to assess the impact of using alternative probability distributions (stability analysis) or to assess the impact of varying the parameters determining the distribution (e.g., location, scale, as applicable) over a range (sensitivity analysis).

In rare cases, it may be unnecessary to perform analyses that propagate uncertainty, based on the goals of the model. 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 performing probabilistic analyses.

Model Assessment and Consistency

Evaluate the model with respect to the specified modeling goals.

A model can only be evaluated with respect to the specific goals of modeling (as determined by the research question). The preferred model assessment methods and criteria depend on the intended use of the model.

Model Verification

Anticipate, detect, and correct errors in the model's logic and implementation.

Errors are unavoidable when developing any nontrivial model.[166] Mistakes in research question formulation, 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.[10,441] Errors in logic and implementation can be challenging to detect and can have important consequences. The risk of mistakes in question formulation and model structure can be reduced by adhering to some of the principles outlined previously in this document (e.g., consulting with topic experts, using a conceptual model to guide the implementation of the mathematical model), together with transparent reporting of methods and results and the use of teams with sufficient expertise. Several checking techniques have been advocated for health care–related models (e.g., sensitivity analysis, extreme value analysis, dimensional analysis).[166] In addition, software production techniques, such as unit testing, code review (review of one programmer's work by another team member), and paired programming (i.e., one programmer's coding being monitored by another in real time), can be considered. Duplicate implementation of the same model by an independent team or implementation of the same model in a different software package can also be used to identify errors in coding. Because these strategies can substantially increase the time and resources required for model development, their use should be balanced against the modeling goals, model complexity, and anticipated frequency and impact of errors.

Model Validation

Face Validation

Invite topic experts to review the model's structure and outputs and to judge whether these seem consistent with their expectations. Verify, describe, and explain counterintuitive model results.
An examination of the model and its results by a group of topic experts can alert modelers to the presence of deficiencies in the model's structure or data.[10] For example, a formal version of this examination involves providing model-generated output (e.g., incidence rates, mortality rates, distributions of patients across stages at diagnosis) and empirical data on the same quantities to users of the model. The experts are then asked to identify which results are "real" and which are model generated. This procedure can be used to assess the credibility of a given model (and relates to the Turing test in artificial intelligence research).[442-444]

Counterintuitive model results ("paradoxical findings") may indicate "bugs" or errors, so such results should be examined carefully. If an error has been ruled out, the results should be described and explained with reference to model structure, available data, and current understanding of the modeled phenomena.

External and Predictive Validation

Assess the consistency between model outputs and the data on which the model was based.
A combination of graphical and statistical methods should be used to compare model outputs with expected results.[138,360,445-450] For parameters that are identifiable using available data, model validation is essentially an assessment of model fit. As such, comparisons of observed versus model-predicted values (graphical or statistical) can be used to identify potential areas of improvement in model structure, assumptions, and data.

Do not withhold data from model development for the sole purpose of assessing model validity.
Generally, data should not be withheld during model development for the purpose of using them for model validation. Using all of the available data during model development improves the efficiency of parameter estimation, facilitates the appropriate handling of correlated inputs, and allows an assessment of consistency across all available sources of evidence.[359] For example, problems may exist when model predictions do not agree well with observations. Model validation, in terms of agreement of model predictions with the corresponding data, can be formalized with metrics of model fit. Resampling methods (cross-fold sampling, bootstrap) can be used to assess model fit and to detect outlying or influential observations that may guide further explorations. Additional model validation methods are available in a Bayesian framework (e.g., posterior predictive checks).[368,451] Validation assessments that use ideas of model fit require careful application and interpretation in over-parameterized models that have parameters that cannot be fully identified from the data (e.g., parameters related to the unobservable rate of tumor cell growth in cancer microsimulation models). Even when such parameters (e.g., tumor growth) are not identifiable by available data, withholding data on identifiable parameters (or on

functional combinations of identifiable and nonidentifiable parameters) is, in general, less efficient than joint modeling.

Decide whether using future observations to assess a model is appropriate based on the research question and the modeling goals.
Predictive validation is an important component of the assessment of models intended as forecasting tools. However, a comparison of model output with empirical results unavailable at the time of model development is not an appropriate method of assessment for models intended to guide decisionmaking using the best available data at a specific point in time.[48,73,452] Models developed in conjunction with systematic reviews are likely to belong in this category.

Update the model as new data become available, new interventions are added, and the understanding of the investigated phenomenon improves.
Models should be updated when new data about important parameters become available (e.g., updated systematic reviews with new or different effect estimates). In addition, as the understanding of disease mechanisms (causal agents, natural history) and interventions and their consequences evolve, model updating should be considered. The model structure and its software implementation must be flexible enough to accommodate this updating process.

Cross-Model Validation

If models addressing the same research question are available, compare their results to the new model and explain any discrepancies.
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).[97,453-456] If such independent models are available (known to the modelers or identified through literature review), then their outputs should be compared as part of cross-model validation, and any discrepancies should be explained with reference to the structure and data inputs of each model.

Interpretation and Reporting of Results

Be transparent about the model structure, computational implementation, and data. Report results clearly and in a way that addresses users' needs.
The implementation of the model structure and data used to populate it should meet the standards of reproducible research.[10,17,106,107,122,457-459] This is particularly important for models that are supported by public funds (e.g., models created in conjunction with EPC evidence reports) or models used to inform decisions that affect health care 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), and a detailed tabular presentation of model inputs (e.g., probability distributions and their parameters) together with the data sources used to estimate these parameters.[10] 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).[156,460] Using the best analytic approach might make complete reporting more challenging; however, accessibility should not be pursued at the expense of model performance (i.e., models should not be oversimplified in order to make their operation understandable to users).[461] The recent ISPOR–SMDM good research practices report provides detailed guidance regarding appropriate elements for technical and nontechnical documentation for modeling studies.[10,17]

Reporting of the results of modeling studies should be tailored to the goals of the relevant stakeholders while remaining faithful to the model structure and assumptions, and communicating uncertainty in the results.[73,86,462-464] Every effort should be made to present the model findings and analyses in a manner that will be most useful to the stakeholders who would be expected to use them.[180,465] For models prepared in conjunction with EPC reports, stakeholders (e.g., Key Informants and Technical Experts) can provide useful suggestions for presenting the results of modeling efforts.[466] It is impossible to give specific guidance to address all model types and uses of modeling covered by this document. Interested readers are referred to the many available texts on health care modeling, the reporting of statistical and simulation analyses, and graphing quantitative information.[93,304-306,352,355,467-472]

Interpret and report results in a way that communicates uncertainty in model outputs.

Results should be reported in a way that effectively communicates uncertainty in model outputs.[429,430] This may include the use of graphical and statistical summaries that describe 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.

Fully disclose any potential conflicts of interest.

All persons who developed the model, conducted and analyzed simulations, or interpreted model results, and those who provided input during any stage of the modeling process should fully disclose any potential conflicts of interest. Both financial and nonfinancial conflicts of interest should be reported.[473-478] For models produced for the AHRQ EPC Program and many other health technology assessment groups, it is necessary that conflicts of interest be avoided. Modelers should adhere to established guidance for avoiding and managing conflicts of interest for EPC products (e.g., Institute of Medicine recommendations and existing EPC guidance).[479,480]

Concluding Remarks

This report provides guidance in the form of widely accepted principles and good-practice recommendations for the conduct and reporting of modeling and simulation studies in the context of health technology assessments. Development of the guidance was based on a systematic review; input from clinical, policy, and decision analysis experts; and stakeholder

discussions. Leadership within the EPC Program, AHRQ personnel, and external reviewers provided extensive feedback. The principles and recommendations are applicable to the class of structural mathematical models that can be developed and used in conjunction with systematic reviews. Because of this broad scope, the guidance does not prescribe specific modeling approaches. We hope that this work will contribute to increased use and better conduct and reporting of modeling and simulation studies in health technology assessment.

Bibliographic Note

Table 4 organizes the references cited throughout the report into categories by (1) modeling topics covered and (2) whether the exposition of the methods was primarily targeting applications in medical, epidemiological, or health services research versus other research fields. The categories are not exclusive, and some references are cited under more than one category. The list is by no means exhaustive of the vast literature on modeling and simulation; it is meant only as a starting point for readers who wish to further explore this literature. We obtained guidance documents on mathematical and simulation modeling in medical, epidemiological, and health services research through systematic review; we obtained all other references from our personal bibliographies or through recommendations by stakeholders, topic experts, or peer reviewers.

Table 4. References cited in the report

Topic	Primarily Target Medical, Epidemiological, or Health Services Research	Primarily Target Other Research Fields
Developing methodological guidance	76	None
Structural modeling, representation; goals of modeling, modeling process; model complexity	19,50,56,66	18,32-36,55,57,77-79,271,274,275,311,313-317,466
General modeling tutorials, overviews of modeling practices, expository papers	1,39-42,44,46,51,53,54,61,62,73,163-167,179-182,25575,168,170-176,178,183,200,285-303,351	37,52,169,177,185-191,193-196,198,199,201,203,204,211,310,312
Guidance for modeling (the decision to conduct modeling, the methods for conducting the modeling, and issues related to reporting); empirical assessments of published modeling studies	2-17,38,47-49,58,70,74,102-116117-119,121-147,400148-162,170,359-361,445,463-465	184,197,202,207,447,459
Books on decision analysis, economic and mathematical modeling	93,304-307,352,470,471	31,59,60,63,64,192,205,206,209,210,308,309,460
Analytic frameworks, influence diagrams, conceptual modeling; choice of model perspective	20-28,280-282,284	276-279
Sources of information for obtaining values for model inputs; methods for systematic reviews and meta-analyses, evidence synthesis	45,421-427 93,239,318-336,338-345,347,348,350,352-357,413-420	337,346,349
Elicitation of probabilities and preferences	94-96,389-392,394,395,397,398	393,396,399
Risk-of-bias assessment and bias adjustment	340-343,371-381,383-388	382
Statistical modeling (behavioral)	30	29
Visualization of quantitative information	None	450,467-469
Heterogeneity in modeling	408-412	
Concepts of uncertainty; methods for uncertainty, stability, and sensitivity analysis; value of information analysis; estimation, calibration, and identifiability	43,97-101,178,183,212,213,220,221,225-240,326,361,362,367-370,408,429-432,432-434,436-438,463	80-92,208,214-219,222-224,270,364-366,435,439,440,451
Transporting study and modeling results across settings; generalizing results	111,121,137,153,400,407,428	401-406
Model assessment (verification, validation)	65,138,446,448,453-456,461	67,68,241-254,256-269,271-275,441-444,447,449-452,462
Conflict of interest; potential for bias in the modeling process	472-480	
Reproducible research	458	457

Citations to the two abstracts presenting preliminary results from this report are not included in the table.

References

1. Owens DK. Analytic tools for public health decision making. Med Decis Making. 2002;22:S3-10. PMID: 12369229.

2. Kuntz, K, Sainfort, F, Butler, M, et al. Decision and simulation modeling in systematic reviews. AHRQ Publication No. 11(13)-EHC037-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2013. PMID: 23534078.

3. 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 Jan;16(1):133-39. PMID: 23337224.

4. Caro JJ, Briggs AH, Siebert U, et al. Modeling good research practices--overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. Med Decis Making. 2012 Sep;32(5):667-77. PMID: 22990082.

5. Roberts M, Russell LB, Paltiel AD, et al. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-2. Med Decis Making. 2012 Sep;32(5):678-89. PMID: 22990083.

6. Siebert U, Alagoz O, Bayoumi AM, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. Med Decis Making. 2012 Sep;32(5):690-700. PMID: 22990084.

7. Karnon J, Stahl J, Brennan A, et al. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. Med Decis Making. 2012 Sep;32(5):701-11. PMID: 22990085.

8. Pitman R, Fisman D, Zaric GS, et al. Dynamic transmission modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-5. Med Decis Making. 2012 Sep;32(5):712-21. PMID: 22990086.

9. Briggs AH, Weinstein MC, Fenwick EA, et al. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. Med Decis Making. 2012 Sep;32(5):722-32. PMID: 22990087.

10. Eddy DM, Hollingworth W, Caro JJ, et al. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. Med Decis Making. 2012 Sep;32(5):733-43. PMID: 22990088.

11. Caro JJ, Briggs AH, Siebert U, et al. Modeling good research practices--overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--1. Value Health. 2012 Sep;15(6):796-803. PMID: 22999128.

12. Roberts M, Russell LB, Paltiel AD, et al. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--2. Value Health. 2012 Sep;15(6):804-11. PMID: 22999129.

13. Siebert U, Alagoz O, Bayoumi AM, et al. State-Transition Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. Value Health. 2012;15(6):812-20. PMID: 22999130.

14. Karnon J, Stahl J, Alan B, et al. Modeling using Discrete Event Simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. Value Health. 2012;15:821-27. PMID: 22999131.

15. Pitman R, Fisman D, Zaric GS, et al. Dynamic Transmission Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-5. Value Health. 2012;15(6):828-34. PMID: 22999132.

16. Briggs A, Weinstein MC, Fenwick E, et al. Model Parameter Estimation and Uncertainty: A Report of the ISPOR-SMDM

Modeling Good Research Practices Task Force-6. Value Health. 2012;15:range. PMID: 22999133.

17. Eddy DM, Hollingworth W, Caro JJ, et al. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--7. Value Health. 2012 Sep;15(6):843-50. PMID: 22999134.

18. Giere RN, Bilcke J, Mauldin R. Understanding scientific reasoning. Wadsworth; 2005.

19. Weisberg M. Simulation and similarity: Using models to understand the world. Oxford University Press; 2012.

20. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. Am J Prev Med. 2001;20:21-35. PMID: 11306229.

21. U.S.Preventive Services Task Force. U.S. Preventive Services Task Force (USPSTF) Procedure Manual: Section 3. Topic Work Plan Development. 2011.

22. Whitlock EP, Orleans CT, Pender N, et al. Evaluating primary care behavioral counseling interventions: an evidence-based approach. Am J Prev Med. 2002 May;22(4):267-84. PMID: 11988383.

23. Woolf SH, DiGuiseppi CG, Atkins D, et al. Developing evidence-based clinical practice guidelines: lessons learned by the US Preventive Services Task Force. Annu Rev Public Health. 1996;17:511-38. PMID: 8724238.

24. Samson, D, Schoelles, KM. Developing the Topic and Structuring Systematic Reviews of Medical Tests: Utility of PICOTS, Analytic Frameworks, Decision Trees, and Other Frameworks. Methods Guide for Medical Test Reviews [Internet] Rockville (MD): Agency for Healthcare Research and Quality (US). 2012. PMID: 22834028.

25. Owens DK, Shachter RD, Nease RF, Jr. Representation and analysis of medical decision problems with influence diagrams.

Med Decis Making. 1997 Jul;17(3):241-62. PMID: 9219185.

26. Nease RF, Jr., Owens DK. Use of influence diagrams to structure medical decisions. Med Decis Making. 1997 Jul;17(3):263-75. PMID: 9219186.

27. Methods guide for effectiveness and comparative effectiveness reviews AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2014.

28. Buckley, DI, Ansari, M, Butler, M, et al. The refinement of topics for systematic reviews: lessons and recommendations from the Effective Health Care Program Report No.: 13-EHC023-EF. [Internet]. Rockville, MD: Agency for Healthcare Research and Quality; 2013. PMID: 23409302.

29. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning. Berlin, Germany: Springer; 2001.

30. Harrell FE. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. Springer Science & Business Media; 2013.

31. Fishwick PA. Simulation model design and execution: building digital worlds. Prentice Hall PTR; 1995.

32. Neyman J. Current problems of mathematical statistics. Report at the International Congress of Mathematicians, Amsterdam. 1954

33. Koopmans TC, Reiersol O. The identification of structural characteristics. Ann Math Stat. 1950;21(2):165-81.

34. Haavelmo T. The probability approach in econometrics. Econometrica. 1944;12(Supplement):iii-115.

35. Wright S. Correlation and causation. J Agric Res. 1921;20(7):557-85.

36. Reiss PC, Wolak FA. Structural econometric modeling: Rationales and examples from industrial organization. Handbook of econometrics. 2007;6:4277-415.

37. Improving information for social policy decisions: the uses of microsimulation modeling; Volume 1: Review and recommendations. Washington, DC: National Academy Press; 1991.

38. 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 Jan;6(1):9-17. PMID: 12535234.

39. Albert DA. Decision theory in medicine: a review and critique. Milbank Mem Fund Q Health Soc. 1978;56(3):362-401. PMID: 100721.

40. Brennan A, Akehurst R. Modelling in health economic evaluation. What is its place? What is its value? Pharmacoeconomics. 2000;17(5):445-59. PMID: 10977387.

41. Buxton MJ, Drummond MF, Van Hout BA, et al. Modelling in economic evaluation: an unavoidable fact of life. Health Econ. 1997 May;6(3):217-27. PMID: 9226140.

42. Eddy D. Technology assessment: the role of mathematical modeling. Assessing medical technologies. Washington, DC: National Academies Press; 1985. p. 144-75.

43. Halpern EF, Weinstein MC, Hunink MG, et al. Representing both first- and second-order uncertainties by Monte Carlo simulation for groups of patients. Med Decis Making. 2000;20(3):314-22. PMID: 10929854.

44. Hodges JS. Six (or so) things you can do with a bad model. Oper Res. 1991;39(3):355-65.

45. Mulrow C, Langhorne P, Grimshaw J. Integrating heterogeneous pieces of evidence in systematic reviews. Ann Intern Med. 1997 Dec 1;127(11):989-95. PMID: 9412305.

46. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Medical Decis Making. 1993 Oct;13(4):322-38. PMID: 8246705.

47. 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(1):94-111. PMID: 11987445.

48. Weinstein MC, Toy EL, Sandberg EA, et al. Modeling for health care and other policy decisions: uses, roles, and validity. Value Health. 2001;4(5):348-61. PMID: 11705125.

49. Siebert U. When should decision-analytic modeling be used in the economic evaluation of health care? Eur J Health Econ. 2003;4(3):143-50.

50. Massoud TF, Hademenos GJ, Young WL, et al. Principles and philosophy of modeling in biomedical research. FASEB J. 1998 Mar;12(3):275-85. PMID: 9580086.

51. Sterman JD. Learning from evidence in a complex world. Am J Public Health. 2006 Mar;96(3):505-14. PMID: 16449579.

52. Rescigno A, Beck JS. The use and abuse of models. J Pharmacokinet Biopharm. 1987 Jun;15(3):327-44.

53. Ness RB, Koopman JS, Roberts MS. Causal system modeling in chronic disease epidemiology: a proposal. Ann Epidemiol. 2007 Jul;17(7):564-68. PMID: 17329122.

54. Chilcott J, Brennan A, Booth A, et al. The role of modelling in prioritising and planning clinical trials. Health Technol Assess. 2003;7(23):iii-125. PMID: 14499052.

55. Berk RA. Causal inference as a prediction problem. Crime & Just. 1987;9:183.

56. Greenland S. Causal inference as a prediction problem: assumptions, identification, and evidence synthesis. Causal Inference: Statistical Perspectives and Applications. New York, NY: Wiley; 2012. p. 43-58.

57. Pearl J. Causality: models, reasoning and inference. New York, NY: Cambridge Univ Press; 2009.

58. Trikalinos TA, Kulasingam S, Lawrence WF. Deciding whether to complement a systematic review of medical tests with decision modeling. Methods Guide for Medical Test Reviews [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US). Methods Guide for Medical Test Reviews [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US) ed. 2012. PMID: 22834021.

59. Anderson MP, Woessner WW. Applied groundwater modeling: simulation of flow and advective transport. San Diego, CA: Academic Press; 1992.

60. Bender EA. An introduction to mathematical modeling. Mineola, NY: Courier Dover Publications; 2012.

61. Detsky AS, Naglie G, Krahn MD, et al. Primer on medical decision analysis: Part 2--Building a tree. Med Decis Making. 1997 Apr;17(2):126-35. PMID: 9107607.

62. Habbema JD, Bossuyt PM, Dippel DW, et al. Analysing clinical decision analyses. Stat Med. 1990 Nov;9(11):1229-42. PMID: 2277875.

63. Meyer WJ. Concepts of mathematical modeling. Mineola, NY: Courier Dover Publications; 2012.

64. Complex Systems Modelling Group. Modelling in healthcare. Providence, RI: American Mathematical Society; 2010.

65. Cobelli C, Carson ER, Finkelstein L, et al. Validation of simple and complex models in physiology and medicine. Am J Physiol. 1984 Feb;246(2 Pt 2):R259-R266. PMID: 6696149.

66. Withers BD, Pritsker AA, Withers DH. A structured definition of the modeling process. Proceedings of the 25th Winter Simulation Conference. 1993:1109-17.

67. U.S.General Accounting Office. Guidelines for model evaluation (PAD-79-17). 1979.

68. Gass SI. Decision-aiding models: validation, assessment, and related issues for policy analysis. Oper Res. 1983;31(4):603-31.

69. French S. Modelling, making inferences and making decisions: the roles of sensitivity analysis. Top. 2003;11(2):229-51.

70. 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 Sep;8(36):iii-158. PMID: 15361314.

71. Dahabreh IJ, Balk E, Wong JB, et al. Systematically developed guidance for the conduct and reporting of decision and simulation models. Med Decis Making. 2015;35(1):E1-E167.

72. Dahabreh IJ, Trikalinos TA, Balk E, et al. Guidance for the conduct and reporting of modeling and simulation in the context of health technology assessment. Value Health. 2015;18:A1-A307.

73. Sculpher M, Fenwick E, Claxton K. Assessing quality in decision analytic cost-effectiveness models. A suggested framework and example of application. Pharmacoeconomics. 2000;17:461-77. PMID: 10977388.

74. Philips Z, Bojke L, Sculpher M, et al. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics. 2006;24(4):355-71. PMID: PM:16605282.

75. 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 Jul;31(4):530-39. PMID: 21673186.

76. Trikalinos, TA, Dahabreh, IJ, Wallace, BC, et al. Towards a framework for communicating confidence in methodological recommendations for systematic reviews and meta-analyses. 13-EHC119-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2013. PMID: 24156117.

77. Oren T. A critical review of definitions and about 400 types of modeling and simulation. SCS M&S Magazine. 2011;2(3):142-51.

78. Oren T. The many facets of simulation through a collection of about 100 definitions. SCS M&S Magazine. 2011;2(2):82-92.

79. Padilla JJ, Diallo SY, Tolk A. Do We Need M&S Science? SCS M&S Magazine. 2011;8:161-66.

80. De Finetti B. Foresight: its logical laws, its subjective sources (1937); translated from the French by Henry E. Kyberg, Jf. In: Kotz S, Johnson NL (eds). Breakthroughs in statistics, Volume 1. New York, NY: Springer Science+Business Medi; 1980.

81. De Finetti B. Theory of Probability, Volume I. New York, NY: Wiley; 1974.

82. Savage LJ. The foundations of statistics. New York: Courier Dover Corporation; 1972.

83. Lindley DV. Understanding uncertainty. Hoboken, N.J.: John Wiley & Sons; 2006.

84. Lindley DV, Lindley DV. Making decisions. London, New York: Wiley-Interscience; 1985.

85. Halpern JY. Reasoning about uncertainty. Cambridge, M.A.: MIT Press; 2003.

86. Morgan MG, Henrion M, Small M. Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge, New York: Cambridge University Press; 1992.

87. Knight FH. Risk, uncertainty and profit. Newburyport, M.A.: Courier Dover Corporation; 2012.

88. Arrow KJ. Alternative approaches to the theory of choice in risk-taking situations. Econometrica. 1951;19(4):404-37.

89. Matheron G. Estimating and choosing: an essay on probability in practice. Springer Science & Business Media; 2012.

90. Draper D. Model uncertainty in stochastic and deterministic systems. Proceedings of the 12th International Workshop on Statistical Modeling. 1997;5:43-59.

91. Draper D, Pereira A, Prado P, et al. Scenario and parametric uncertainty in GESAMAC: a methodological study in nuclear waste disposal risk assessment. Comput Phys Commun. 1999;117(1):142-55.

92. Draper, D, Hodges, JS, Leamer, EE, et al. A research agenda for assessment and propagation of model uncertainty. Document No. N-2683-RC. RAND Corporation; 1987.

93. Petitti DB. Meta-analysis, decision analysis, and cost-effectiveness analysis: methods for quantitative synthesis in medicine. New York: Oxford University Press; 1999.

94. Torrance GW. Health status index models: a unified mathematical view. Management Science. 1976;22(9):990-1001.

95. Torrance GW. Utility approach to measuring health-related quality of life. Journal of chronic diseases. 1987;40(6):593-600.

96. Pliskin JS, Shepard DS, Weinstein MC. Utility functions for life years and health status. Operations Research. 1980;28(1):206-24.

97. Drummond MF, Barbieri M, Wong JB. Analytic choices in economic models of treatments for rheumatoid arthritis: What makes a difference? Med Decis Making. 2005;25(5):520-33. PMID: 16160208.

98. Karnon J, Brennan A, Akehurst R. A critique and impact analysis of decision modeling assumptions. Med Decis Making. 2007 Jul;27(4):491-99. PMID: 17761961.

99. Mauskopf J. Modelling technique, structural assumptions, input parameter values: which has the most impact on the results of a cost-effectiveness analysis? Pharmacoeconomics. 2014 Jun;32(6):521-23. PMID: 24743914.

100. 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 Jun;13(29):iii, ix-61. PMID: 19500484.

101. Rosenbaum PR. Observational studies. New York: Springer; 2002.

102. Weinstein MC, Siegel JE, Gold MR, et al. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. JAMA. 1996 Oct 16;276(15):1253-58. PMID: 8849754.

103. Penaloza Ramos MC, Barton P, Jowett S, et al. A Systematic Review of Research Guidelines in Decision-Analytic Modeling. Value Health. 2015 Jun;18(4):512-29. PMID: 26091606.

104. Canadian Agency for Drugs and Technologies in Health. Guidelines for the economic evaluation of health technologies: Canada. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.

105. Hay J, Jackson J. Panel 2: methodological issues in conducting pharmacoeconomic evaluations--modeling studies. Value Health. 1999 Mar;2(2):78-81. PMID: 16674337.

106. Husereau D, Drummond M, Petrou S, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. BMJ. 2013;346:f1049. PMID: 23529982.

107. Nuijten MJ, Pronk MH, Brorens MJ, et al. Reporting format for economic evaluation. Part II: Focus on modelling studies. Pharmacoeconomics. 1998 Sep;14(3):259-68. PMID: 10186465.

108. Russell LB, Gold MR, Siegel JE, et al. The role of cost-effectiveness analysis in health and medicine. Panel on Cost-Effectiveness in Health and Medicine. JAMA. 1996 Oct 9;276(14):1172-77. PMID: 8827972.

109. Decision analytic modelling in the economic evaluation of health technologies. a consensus statement. consensus conference on guidelines on economic modelling in health technology assessment. Pharmacoeconomics. 2000 May;17(5):443-44. PMID: 10977386.

110. Bae EY, Lee EK. Pharmacoeconomic guidelines and their implementation in the positive list system in South Korea. Value Health. 2009 Nov;12 Suppl 3:S36-S41. PMID: 20586979.

111. Boulenger S, Nixon J, Drummond M, et al. Can economic evaluations be made more transferable? Eur J Health Econ. 2005 Dec;6(4):334-46. PMID: 16249933.

112. Chiou CF, Hay JW, Wallace JF, et al. Development and validation of a grading system for the quality of cost-effectiveness studies. Med Care. 2003 Jan;41(1):32-44. PMID: 12544542.

113. Cleemput I, van WP, Huybrechts M, et al. Belgian methodological guidelines for pharmacoeconomic evaluations: toward standardization of drug reimbursement requests. Value Health. 2009 Jun;12(4):441-49. PMID: 19900251.

114. Clemens K, Townsend R, Luscombe F, et al. Methodological and conduct principles for pharmacoeconomic research. Pharmaceutical Research and Manufacturers of America. Pharmacoeconomics. 1995 Aug;8(2):169-74. PMID: 10155611.

115. Colmenero F, Sullivan SD, Palmer JA, et al. Quality of clinical and economic evidence in dossier formulary submissions. Am J Manag Care. 2007 Jul;13(7):401-07. PMID: 17620035.

116. Davalos ME, French MT, Burdick AE, et al. Economic evaluation of telemedicine: review of the literature and research guidelines for benefit-cost analysis. Telemed J E Health. 2009 Dec;15(10):933-48. PMID: 19954346.

117. Detsky AS. Guidelines for economic analysis of pharmaceutical products: a draft document for Ontario and Canada. Pharmacoeconomics. 1993 May;3(5):354-61. PMID: 10146886.

118. Drummond M, Brandt A, Luce B, et al. Standardizing methodologies for economic evaluation in health care. Practice, problems, and potential. Int J Technol Assess Health Care. 1993;9(1):26-36. PMID: 8423113.

119. Drummond M, Sculpher M. Common methodological flaws in economic

evaluations. Med Care. 2005 Jul;43(7 Suppl):5-14. PMID: 16056003.

120. Drummond M, Manca A, Sculpher M. Increasing the generalizability of economic evaluations: recommendations for the design, analysis, and reporting of studies. Int J Technol Assess Health Care. 2005;21(2):165-71. PMID: 15921055.

121. Drummond M, Barbieri M, Cook J, et al. Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force report. Value Health. 2009 Jun;12(4):409-18. PMID: 19900249.

122. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ. 1996 Aug 3;313(7052):275-83. PMID: 8704542.

123. Evers S, Goossens M, de VH, et al. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. Int J Technol Assess Health Care. 2005;21(2):240-45. PMID: 15921065.

124. Fry RN, Avey SG, Sullivan SD. The Academy of Managed Care Pharmacy Format for Formulary Submissions: an evolving standard--a Foundation for Managed Care Pharmacy Task Force report. Value Health. 2003 Sep;6(5):505-21. PMID: 14627057.

125. Garattini L, Grilli R, Scopelliti D, et al. A proposal for Italian guidelines in pharmacoeconomics The Mario Negri Institute Centre for Health Economics. Pharmacoeconomics. 1995 Jan;7(1):1-6. PMID: 10155289.

126. Gartlehner G, West SL, Mansfield AJ, et al. Clinical heterogeneity in systematic reviews and health technology assessments: synthesis of guidance documents and the literature. Int J Technol Assess Health Care. 2012 Jan;28(1):36-43. PMID: 22217016.

127. Glennie JL, Torrance GW, Baladi JF, et al. The revised Canadian Guidelines for the Economic Evaluation of Pharmaceuticals.

Pharmacoeconomics. 1999 May;15(5):459-68. PMID: 10537963.

128. Goldhaber-Fiebert JD, Stout NK, Goldie SJ. Empirically evaluating decision-analytic models. Value Health. 2010 Aug;13(5):667-74. PMID: 20230547.

129. Graf von der Schulenburg JM, Greiner W, Jost F, et al. German recommendations on health economic evaluation: third and updated version of the Hanover Consensus. Value Health. 2008 Jul;11(4):539-44. PMID: 18194408.

130. Grutters JP, Seferina SC, Tjan-Heijnen VC, et al. Bridging trial and decision: a checklist to frame health technology assessments for resource allocation decisions. Value Health. 2011 Jul;14(5):777-84. PMID: 21839418.

131. Hoomans T, Severens JL, van der RN, et al. Methodological quality of economic evaluations of new pharmaceuticals in The Netherlands. Pharmacoeconomics. 2012 Mar;30(3):219-27. PMID: 22074610.

132. Kolasa K, Dziomdziora M, Fajutrao L. What aspects of the health technology assessment process recommended by international health technology assessment agencies received the most attention in Poland in 2008? Int J Technol Assess Health Care. 2011 Jan;27(1):84-94. PMID: 21262087.

133. Liberati A, Sheldon TA, Banta HD. EUR-ASSESS Project Subgroup report on Methodology. Methodological guidance for the conduct of health technology assessment. Int J Technol Assess Health Care. 1997;13(2):186-219. PMID: 9194352.

134. Lopez-Bastida J, Oliva J, Antonanzas F, et al. Spanish recommendations on economic evaluation of health technologies. Eur J Health Econ. 2010 Oct;11(5):513-20. PMID: 20405159.

135. Lovatt B. The United Kingdom guidelines for the economic evaluation of medicines. Med Care. 1996 Dec;34(12 Suppl):DS179-DS181. PMID: 8969324.

136. Luce BR, Simpson K. Methods of cost-effectiveness analysis: areas of consensus

and debate. Clin Ther. 1995 Jan;17(1):109-25. PMID: 7758053.

137. Mason J. The generalisability of pharmacoeconomic studies. Pharmacoeconomics. 1997 Jun;11(6):503-14. PMID: 10168092.

138. McCabe C, Dixon S. Testing the validity of cost-effectiveness models. Pharmacoeconomics. 2000 May;17(5):501-13. PMID: 10977390.

139. McGhan WF, Al M, Doshi JA, et al. The ISPOR Good Practices for Quality Improvement of Cost-Effectiveness Research Task Force Report. Value Health. 2009 Nov;12(8):1086-99. PMID: 19744291.

140. Menon D, Schubert F, Torrance GW. Canada's new guidelines for the economic evaluation of pharmaceuticals. Med Care. 1996 Dec;34(12 Suppl):DS77-DS86. PMID: 8969316.

141. Karnon J, Goyder E, Tappenden P, et al. A review and critique of modelling in prioritising and designing screening programmes. Health Technol Assess. 2007 Dec;11(52):iii-xi, 1. PMID: 18031651.

142. Mullahy J. What you don't know can't hurt you? Statistical issues and standards for medical technology evaluation. Med Care. 1996 Dec;34(12 Suppl):DS124-DS135. PMID: 8969321.

143. Mullins CD, Ogilvie S. Emerging standardization in pharmacoeconomics. Clin Ther. 1998 Nov;20(6):1194-202. PMID: 9916612.

144. Mullins CD, Wang J. Pharmacy benefit management: enhancing the applicability of pharmacoeconomics for optimal decision making. Pharmacoeconomics. 2002;20(1):9-21. PMID: 11817989.

145. Blackmore CC, Magid DJ. Methodologic evaluation of the radiology cost-effectiveness literature. Radiology. 1997 Apr;203(1):87-91. PMID: 9122421.

146. Murray CJ, Evans DB, Acharya A, et al. Development of WHO guidelines on generalized cost-effectiveness analysis. Health Econ. 2000 Apr;9(3):235-51. PMID: 10790702.

147. Canadian Agency for Drugs and Technologies in Health. Guidelines for economic evaluation of pharmaceuticals: Canada. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.

148. Neumann PJ, Stone PW, Chapman RH, et al. The quality of reporting in published cost-utility analyses, 1976-1997. Ann Intern Med. 2000 Jun 20;132(12):964-72. PMID: 10858180.

149. Olson BM, Armstrong EP, Grizzle AJ, et al. Industry's perception of presenting pharmacoeconomic models to managed care organizations. J Manag Care Pharm. 2003 Mar;9(2):159-67. PMID: 14613345.

150. Paisley S. Classification of evidence in decision-analytic models of cost-effectiveness: a content analysis of published reports. Int J Technol Assess Health Care. 2010 Oct;26(4):458-62. PMID: 20923588.

151. Ramsey S, Willke R, Briggs A, et al. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. Value Health. 2005 Sep;8(5):521-33. PMID: 16176491.

152. Sassi F, McKee M, Roberts JA. Economic evaluation of diagnostic technology. Methodological challenges and viable solutions. Int J Technol Assess Health Care. 1997;13(4):613-30. PMID: 9489253.

153. Sculpher MJ, Pang FS, Manca A, et al. Generalisability in economic evaluation studies in healthcare: a review and case studies. Health Technol Assess. 2004 Dec;8(49):iii-192. PMID: 15544708.

154. Severens JL, van der Wilt GJ. Economic evaluation of diagnostic tests. A review of published studies. Int J Technol Assess Health Care. 1999;15(3):480-96. PMID: 10874376.

155. Siegel JE, Torrance GW, Russell LB, et al. Guidelines for pharmacoeconomic studies. Recommendations from the panel on cost effectiveness in health and medicine. Panel on cost Effectiveness in Health and Medicine. Pharmacoeconomics. 1997 Feb;11(2):159-68. PMID: 10172935.

156. Sonnenberg FA, Roberts MS, Tsevat J, et al. Toward a peer review process for medical decision analysis models. Med Care. 1994 Jul;32(7 Suppl):JS52-JS64. PMID: 8028413.

157. Taylor RS, Elston J. The use of surrogate outcomes in model-based cost-effectiveness analyses: a survey of UK Health Technology Assessment reports. Health Technol Assess. 2009 Jan;13(8):iii, ix-50. PMID: 19203465.

158. Trikalinos TA, Siebert U, Lau J. Decision-analytic modeling to evaluate benefits and harms of medical tests: uses and limitations. Med Decis Making. 2009 Sep;29(5):E22-E29. PMID: 19734441.

159. Udvarhelyi IS, Colditz GA, Rai A, et al. Cost-effectiveness and cost-benefit analyses in the medical literature. Are the methods being used correctly? Ann Intern Med. 1992 Feb 1;116(3):238-44. PMID: 1530808.

160. Ungar WJ, Santos MT. The Pediatric Quality Appraisal Questionnaire: an instrument for evaluation of the pediatric health economics literature. Value Health. 2003 Sep;6(5):584-94. PMID: 14627065.

161. Vegter S, Boersma C, Rozenbaum M, et al. Pharmacoeconomic evaluations of pharmacogenetic and genomic screening programmes: a systematic review on content and adherence to guidelines. Pharmacoeconomics. 2008;26(7):569-87. PMID: 18563949.

162. von der Schulenburg J, Vauth C, Mittendorf T, et al. Methods for determining cost-benefit ratios for pharmaceuticals in Germany. Eur J Health Econ. 2007 Sep;8 Suppl 1:S5-31. PMID: 17582539.

163. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. J Health Serv Res Policy. 2004 Apr;9(2):110-18. PMID: 15099459.

164. Hjelmgren J, Berggren F, Andersson F. Health economic guidelines--similarities, differences and some implications. Value Health. 2001 May;4(3):225-50. PMID: 11705185.

165. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. Health Econ. 2006 Dec;15(12):1295-310. PMID: 16941543.

166. 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 May;14(25):iii-107. PMID: 20501062.

167. Stahl JE. Modelling methods for pharmacoeconomics and health technology assessment: an overview and guide. Pharmacoeconomics. 2008;26(2):131-48. PMID: 18198933.

168. Goeree R, O'Brien BJ, Blackhouse G. Principles of good modeling practice in healthcare cost-effectiveness studies. Expert Rev Pharmacoecon Outcomes Res. 2004;4(2):189-98. PMID: 19807523.

169. Sheiner LB, Steimer JL. Pharmacokinetic/pharmacodynamic modeling in drug development. Annu Rev Pharmacol Toxicol. 2000;40:67-95.:67-95. PMID: 10836128.

170. Brandeau ML, McCoy JH, Hupert N, et al. Recommendations for modeling disaster responses in public health and medicine: a position paper of the society for medical decision making. Med Decis Making. 2009 Jul;29(4):438-60. PMID: 19605887.

171. Holford NH, Kimko HC, Monteleone JP, et al. Simulation of clinical trials. Annu Rev Pharmacol Toxicol. 2000;40:209-34.:209-34. PMID: 10836134.

172. Ip EH, Rahmandad H, Shoham DA, et al. Reconciling statistical and systems science approaches to public health. Health Educ

Behav. 2013 Oct;40(1 Suppl):123S-31S. PMID: 24084395.

173. Girard P, Cucherat M, Guez D. Clinical trial simulation in drug development. Therapie. 2004 May;59(3):287-304. PMID: 15559184.

174. Luke DA, Stamatakis KA. Systems science methods in public health: dynamics, networks, and agents. Annu Rev Public Health. 2012 Apr;33:357-76. PMID: 22224885.

175. Gunal MM. A guide for building hospital simulation models. Health Systems. 2012;1(1):17-25.

176. Bonate PL. Clinical trial simulation in drug development. Pharm Res. 2000;17(3):252-56. PMID: 10801212.

177. Barton HA, Chiu WA, Woodrow SR, et al. Characterizing uncertainty and variability in physiologically based pharmacokinetic models: state of the science and needs for research and implementation. Toxicol Sci. 2007 Oct;99(2):395-402. PMID: 17483121.

178. Cox LA, Jr. Confronting deep uncertainties in risk analysis. Risk Anal. 2012 Oct;32(10):1607-29. PMID: 22489541.

179. Naglie G, Krahn MD, Naimark D, et al. Primer on medical decision analysis: Part 3-- Estimating probabilities and utilities. Med Decis Making. 1997 Apr;17(2):136-41. PMID: 9107608.

180. Krahn MD, Naglie G, Naimark D, et al. Primer on medical decision analysis: Part 4-- Analyzing the model and interpreting the results. Med Decis Making. 1997 Apr;17(2):142-51. PMID: 9107609.

181. Naimark D, Krahn MD, Naglie G, et al. Primer on medical decision analysis: Part 5-- Working with Markov processes. Med Decis Making. 1997 Apr;17(2):152-59. PMID: 9107610.

182. Detsky AS, Naglie G, Krahn MD, et al. Primer on medical decision analysis: Part 1-- Getting started. Med Decis Making. 1997 Apr;17(2):123-25. PMID: 9107606.

183. Koopman J. Modeling infection transmission. Annu Rev Public Health. 2004;25:303-26. PMID: 15015922.

184. Martinez-Moyano IJ, Richardson GP. Best practices in system dynamics modeling. System Dynamics Review. 2013;29(2):102-23.

185. Sanchez PJ. Fundamentals of simulation modeling. Proceedings of the 39th Winter Simulation Conference. 2007:54-62.

186. Shannon RE. Introduction to the art and science of simulation. Proceedings of the 30th Winter Simulation Conference. 1998:7-14.

187. Macal CM, North MJ. Introductory tutorial: Agent-based modeling and simulation. Proceedings of the 43rd Winter Simulation Conference. 2011:1451-64.

188. Sturrock DT. Tutorial: tips for successful practice of simulation. Proceedings of the 44th Winter Simulation Conference. 2012:1-8.

189. Barton RR. Designing simulation experiments. Proceedings of the 45th Winter Simulation Conference. 2013:342-53.

190. Currie CS, Cheng RC. A practical introduction to analysis of simulation output data. Proceedings of the 45th Winter Simulation Conference. 2013:328-41.

191. Sanchez SM, Wan H. Work smarter, not harder: A tutorial on designing and conducting simulation experiments. Proceedings of the 44th Winter Simulation Conference. 2012:1-15.

192. Law A. Simulation Modeling and Analysis. New York: McGraw-Hill Science/Engineering/Math; 2006.

193. Uhrmacher AM. Seven pitfalls in modeling and simulation research. Proceedings of the 44th Winter Simulation Conference. 2012:318.

194. Jurishica C, Zupick N. Tutorial: Tools and methodologies for executing successful simulation consulting projects. Proceedings

of the 44th Winter Simulation Conference. 2012:1-13.

195. Loizou G, Spendiff M, Barton HA, et al. Development of good modelling practice for physiologically based pharmacokinetic models for use in risk assessment: the first steps. Regul Toxicol Pharmacol. 2008 Apr;50(3):400-11. PMID: 18331772.

196. Andersen ME, Clewell HJ, III, Frederick CB. Applying simulation modeling to problems in toxicology and risk assessment--a short perspective. Toxicol Appl Pharmacol. 1995 Aug;133(2):181-87. PMID: 7645013.

197. Byon W, Smith MK, Chan P, et al. Establishing best practices and guidance in population modeling: an experience with an internal population pharmacokinetic analysis guidance. CPT Pharmacometrics Syst Pharmacol. 2013 Jul 3;2:e51:e51. PMID: 23836283.

198. Jones H, Rowland-Yeo K. Basic concepts in physiologically based pharmacokinetic modeling in drug discovery and development. CPT Pharmacometrics Syst Pharmacol. 2013 Aug 14;2:e63 PMID: 23945604.

199. Clark LH, Setzer RW, Barton HA. Framework for evaluation of physiologically-based pharmacokinetic models for use in safety or risk assessment. Risk Anal. 2004 Dec;24(6):1697-717. PMID: 15660623.

200. Auchincloss AH, Diez Roux AV. A new tool for epidemiology: the usefulness of dynamic-agent models in understanding place effects on health. Am J Epidemiol. 2008 Jul 1;168(1):1-8. PMID: 18480064.

201. Yates FE. Good manners in good modeling: mathematical models and computer simulations of physiological systems. Am J Physiol. 1978 May;234(5):R159-R160. PMID: 645933.

202. Richiardi MG, Leombruni R, Saam NJ, et al. A common protocol for agent-based social simulation. JASSS. 2006;9(1)

203. Young P. Data-based mechanistic modelling, generalised sensitivity and dominant mode analysis. Computer Physics Communications. 1999;117(1):113-29.

204. Young P. The data-based mechanistic approach to the modelling, forecasting and control of environmental systems. Annual Reviews in Control. 2006;30(2):169-82.

205. Tolk A. Engineering principles of combat modeling and distributed simulation. Wiley Online Library; 2012.

206. Banks J. Handbook of simulation. New York: Wiley Online Library; 1998.

207. Refsgaard JC, Henriksen HJ. Modelling guidelines - terminology and guiding principles. Advances in Water Resources. 2004;27(1):71-82.

208. Beven K. Prophecy, reality and uncertainty in distributed hydrological modelling. Advances in Water Resources. 1993;16(1):41-51.

209. Pratt JW, Raiffa H, Schlaifer R. Introduction to statistical decision theory. Cambridge, MA: MIT press; 1996.

210. Von Neumann J, Morgenstern O. Theory of games and economic behavior. Princeton, N.J.: Princeton University Press; 2007.

211. Rabitz H. Systems analysis at the molecular scale. Science. 1989 Oct 13;246(4927):221-26. PMID: 17839016.

212. Briggs A, Sculpher M, Buxton M. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. Health Econ. 1994 Mar;3(2):95-104. PMID: 8044216.

213. Felli JC, Hazen GB. Sensitivity analysis and the expected value of perfect information.[Erratum appears in Med Decis Making 2001 May-Jun;21(3):254], [Erratum appears in Med Decis Making. 2003 Jan-Feb;23(1):97.]. Med Decis Making. 1998;18(1):95-109. PMID: 9456214.

214. Saltelli A, Tarantola S, Campolongo F. Sensitivity analysis as an ingredient of modeling. Statistical Science. 2000:377-95.

215. Saltelli A. Sensitivity analysis for importance assessment. Risk Anal. 2002 Jun;22(3):579-90. PMID: 12088235.

216. Hofer E. Sensitivity analysis in the context of uncertainty analysis for computationally intensive models. Comput Phys Commun. 1999;117(1):21-34.

217. Kleijnen JP. Sensitivity analysis versus uncertainty analysis: When to use what? Predictability and nonlinear modelling in natural sciences and economics. Springer; 1994. p. 322-33.

218. Iman RL, Helton JC. An investigation of uncertainty and sensitivity analysis techniques for computer models. Risk Anal. 1988;8(1):71-90.

219. Cukier RI, Levine HB, Shuler KE. Nonlinear sensitivity analysis of multiparameter model systems. J Comput Phys. 1978;26(1):1-42.

220. Jain R, Grabner M, Onukwugha E. Sensitivity analysis in cost-effectiveness studies: from guidelines to practice. Pharmacoeconomics. 2011 Apr;29(4):297-314. PMID: 21395350.

221. Campbell JD, McQueen RB, Libby AM, et al. Cost-effectiveness Uncertainty Analysis Methods: A Comparison of One-way Sensitivity, Analysis of Covariance, and Expected Value of Partial Perfect Information. Med Decis Making. 2014 Oct 27 PMID: 25349188.

222. Frey HC, Patil SR. Identification and review of sensitivity analysis methods. Risk Anal. 2002 Jun;22(3):553-78. PMID: 12088234.

223. Saltelli A, Tarantola S, Campolongo F, et al. Sensitivity analysis in practice: a guide to assessing scientific models. Chichester, West Sussex, England: John Wiley & Sons; 2004.

224. Saltelli A, Ratto M, Andres T, et al. Global sensitivity analysis: the primer. Chichester, West Sussex, England: John Wiley & Sons; 2008.

225. Bojke L, Claxton K, Sculpher M, et al. Characterizing structural uncertainty in decision analytic models: a review and application of methods. Value Health. 2009;12(5):739-49. PMID: 19508655.

226. Bilcke J, Beutels P, Brisson M, et al. Accounting for methodological, structural, and parameter uncertainty in decision-analytic models a practical guide. Medical Decision Making. 2011;31(4):675-92. PMID: 21653805.

227. Jackson CH, Bojke L, Thompson SG, et al. A framework for addressing structural uncertainty in decision models. Medical Decision Making. 2011;31(4):662-74. PMID: 21602487.

228. Jackson CH, Thompson SG, Sharples LD. Accounting for uncertainty in health economic decision models by using model averaging. Journal of the Royal Statistical Society: Series A (Statistics in Society). 2009;172(2):383-404. PMID: 19381329.

229. Claxton K. Exploring uncertainty in cost-effectiveness analysis. Pharmacoeconomics. 2008;26(9):781-98. PMID: 18767898.

230. Brisson M, Edmunds WJ. Impact of model, methodological, and parameter uncertainty in the economic analysis of vaccination programs. Med Decis Making. 2006 Sep;26(5):434-46. PMID: 16997923.

231. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ. 1999 Jun;18(3):341-64. PMID: 10537899.

232. Baio G, Dawid AP. Probabilistic sensitivity analysis in health economics. Stat Methods Med Res. 2011 Sep 18 PMID: 21930515.

233. Ades AE, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. Med Decis Making. 2004;24:207-27. PMID: 15090106.

234. Whang W, Sisk JE, Heitjan DF, et al. Probabilistic sensitivity analysis in cost-effectiveness. An application from a study of vaccination against pneumococcal bacteremia in the elderly. Int J Technol Assess Health Care. 1999;15(3):563-72. PMID: 10874382.

235. Critchfield GC, Willard KE. Probabilistic analysis of decision trees using Monte Carlo simulation. Med Decis Making. 1986 Apr;6(2):85-92. PMID: 3702625.

236. Critchfield GC, Willard KE, Connelly DP. Probabilistic sensitivity analysis methods for general decision models. Comput Biomed Res. 1986 Jun;19(3):254-65. PMID: 3709122.

237. Doubilet P, Begg CB, Weinstein MC, et al. Probabilistic sensitivity analysis using Monte Carlo simulation. A practical approach. Med Decis Making. 1985;5(2):157-77. PMID: 3831638.

238. Claxton K, Sculpher M, McCabe C, et al. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. Health Econ. 2005;14(4):339-47. PMID: 15736142.

239. Ades AE, Claxton K, Sculpher M. Evidence synthesis, parameter correlation and probabilistic sensitivity analysis. Health Econ. 2006;15(4):373-81. PMID: 16389628.

240. Oakley JE, O'Hagan A. Probabilistic sensitivity analysis of complex models: a Bayesian approach. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2004;66(3):751-69.

241. Robinson S. Simulation model verification and validation: increasing the users' confidence. Proceedings of the 29th Winter Simulation Conference. 1997:53-59.

242. Balci O. Validation, verification, and testing techniques throughout the life cycle of a simulation study. Ann Oper Res. 1994;53(1):121-73.

243. Windrum P, Fagiolo G, Moneta A. Empirical validation of agent-based models:

Alternatives and prospects. J Artif Soc S. 2007;10(2):8.

244. Committee on Mathematical and Statistical Foundations of Verification Validation and Uncertainty Qualification. Assessing the Realiability of Complex Models: Mathematical and Statistical Foundations of Verification,Validation,and Uncertainty Qualification. Washington, D.C.: National Academies Press; 2012.

245. Kleijnen JP. Verification and validation of simulation models. Eur J Oper Res. 1995;82(1):145-62.

246. Sargent RG. Verification and validation of simulation models. Proceedings of the 37th Winter Simulation Conference. 2005:130-43.

247. Sargent RG. An introduction to verification and validation of simulation models. Proceedings of the 45th Winter Simulation Conference. 2013:321-27.

248. Balci O. Verification, validation, and certification of modeling and simulation applications: verification, validation, and certification of modeling and simulation applications. Proceedings of the 35th Winter Simulation Conference. 2003:150-58.

249. Balci O. Quality assessment, verification, and validation of modeling and simulation applications. Proceedings of the 36th Winter Simulation Model. 2004;1

250. Law AM. How to build valid and credible simulation models. Proceedings of the 41st Winter Simulation Conference. 2009:24-33.

251. Beck MB, Ravetz JR, Mulkey LA, et al. On the problem of model validation for predictive exposure assessments. Stoch Hydrol Hydraul. 1997;11(3):229-54.

252. Caswell H. The validation problem. Systems analysis and simulation in ecology. 1976;4:313-25.

253. Dery R, Landry M, Banville C. Revisiting the issue of model validation in OR: an epistemological view. Eur J Oper Res. 1993;66(2):168-83.

254. Landry M, Oral M. In search of a valid view of model validation for operations research. Eur J Oper Res. 1993;66(2):161-67.

255. Caro JJ, Moller J. Decision-analytic models: current methodological challenges. Pharmacoeconomics. 2014 Oct;32(10):943-50. PMID: 24986039.

256. Kopec JA, Fines P, Manuel DG, et al. Validation of population-based disease simulation models: a review of concepts and methods. BMC Public Health. 2010;10:710. PMID: 21087466.

257. Schlesinger S, Crosbie RE, Gagn+¬ RE, et al. Terminology for model credibility. Simulation. 1979;32(3):103-04.

258. Sargent RG. Verification and validation of simulation models. Journal of Simulation. 2013;7(1):12-24.

259. Naylor TH, Finger JM. Verification of computer simulation models. Management Science. 1967;14(2):B-92.

260. Gass SI, Thompson BW. Letter to the Editor – Guidelines for Model Evaluation: An Abridged Version of the US General Accounting Office Exposure Draft. Oper Res. 1980;28(2):431-39.

261. Gass SI, Joel LS. Concepts of model confidence. Comput Oper Res. 1981;8(4):341-46.

262. Oren TI. Concepts and criteria to assess acceptability of simulation studies: a frame of reference. Communications of the ACM. 1981;24(4):180-89.

263. Thacker, BH, Doebling, SW, Hemez, FM, et al. Concepts of model verification and validation. Los Alamos National Lab., Los Alamos, NM (United States). Funding organisation: US DOE (United States); 2004.

264. Oberkampf WL, Trucano TG, Hirsch C. Verification, validation, and predictive capability in computational engineering and physics. Appl Mech Rev. 2004;57(5):345-84.

265. Roy CJ, Oberkampf WL. A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. Comput Method Appl M. 2011;200(25):2131-44.

266. Oberkampf WL, Roy CJ. Verification and validation in scientific computing. New York, N.Y.: Cambridge University Press; 2010.

267. Rykiel EJ. Testing ecological models: the meaning of validation. Ecol Model. 1996;90(3):229-44.

268. Mihram GA. Some practical aspects of the verification and validation of simulation models. Oper Res Quart. 1972:17-29.

269. Kleijen JPC. Validation of models: statistical techniques and data availability. Proceedings of the 31st Winter Simulation Conference. 1999;1:647-54.

270. Trucano TG, Swiler LP, Igusa T, et al. Calibration, validation, and sensitivity analysis: What's what. Reliab Eng Syst Safe. 2006;91(10):1331-57.

271. Weisberg M. Who is a modeler? Brit J Phil Sci. 2007;58(2):207-33.

272. Oreskes N, Shrader-Frechette K, Belitz K. Verification, validation, and confirmation of numerical models in the Earth sciences. Science. 1994 Feb 4;263(5147):641-46. PMID: 17747657.

273. Oreskes N. Evaluation (not validation) of quantitative models. Environ Health Perspect. 1998 Dec;106 Suppl 6:1453-60.:1453-60. PMID: 9860904.

274. Balci O, Ormsby WF. Well-defined intended uses: an explicit requirement for accreditation of modeling and simulation applications. Proceedings of the 32nd Winter Simulation Conference. Society for Computer Simulation International; 2000. p. 849-54.

275. Balci O, Nance RE. Formulated problem verification as an explicit requirement of model credibility. Simulation. 1985;45(2):76-86.

276. Robinson S. Conceptual modelling for simulation Part I: definition and requirements. J Oper Res Soc. 2008;59(3):278-90.

277. Robinson S. Conceptual modelling for simulation Part II: a framework for conceptual modelling. J Oper Res Soc. 2008;59(3):291-304.

278. Robinson S. Conceptual modeling for simulation: issues and research requirements. Proceedings of the 38th Winter Simulation Conference. 2006:792-800.

279. Robinson S. Conceptual modeling for simulation. Proceedings of the 45th Winter Simulation Conference. 2013:377-88.

280. Russell LB, Fryback DG, Sonnenberg FA. Is the societal perspective in cost-effectiveness analysis useful for decision makers? Jt Comm J Qual Improv. 1999 Sep;25(9):447-54. PMID: 10481813.

281. Roy S, Madhavan SS. Making a case for employing a societal perspective in the evaluation of Medicaid prescription drug interventions. Pharmacoeconomics. 2008;26(4):281-96. PMID: 18370564.

282. Garrison LP, Jr., Mansley EC, Abbott TA, III, et al. 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 Jan;13(1):8-13. PMID: 19883405.

283. Drummond M, Sculpher M, Torrance G, et al. Methods for the Economic Evaluation of Health Care Programmes. Oxford, England: Oxford University Press; 2005.

284. Neumann PJ. Costing and perspective in published cost-effectiveness analysis. Med Care. 2009 Jul;47(7 Suppl 1):S28-S32. PMID: 19536023.

285. Ademi Z, Kim H, Zomer E, et al. Overview of pharmacoeconomic modelling methods. Brit J Clin Pharmacol. 2013 Apr;75(4):944-50. PMID: 22882459.

286. Beck JR, Pauker SG. The Markov process in medical prognosis. Med Decis Making. 1983;3(4):419-58. PMID: 6668990.

287. Briggs A, Sculpher M. An introduction to Markov modelling for economic evaluation. Pharmacoeconomics. 1998 Apr;13(4):397-409. PMID: 10178664.

288. Caro JJ, Moller J, Getsios D. Discrete event simulation: the preferred technique for health economic evaluations? Value Health. 2010 Dec;13(8):1056-60. PMID: 20825626.

289. Jun JB, Jacobson SH, Swisher JR. Application of discrete-event simulation in health care clinics. J Oper Res Soc. 1999;50(2):109-23.

290. Karnon J, Brown J. Selecting a decision model for economic evaluation: a case study and review. Health Care Manag Sci. 1998 Oct;1(2):133-40. PMID: 10916592.

291. Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. Health Econ. 2003 Oct;12(10):837-48. PMID: 14508868.

292. 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. Pharmacoeconomics. 2014 Mar 14;32(6):547-58. PMID: 24627341.

293. Soares MO, Canto E Castro. Continuous time simulation and discretized models for cost-effectiveness analysis. Pharmacoeconomics. 2012 Dec 1;30(12):1101-17. PMID: 23116289.

294. Standfield L, Comans T, Scuffham P. Markov modeling and discrete event simulation in health care: a systematic comparison. Int J Technol Assess Health Care. 2014 Apr 28;30(2):1-8. PMID: 24774101.

295. van Rosmalen J, Toy M, O'Mahony JF. A mathematical approach for evaluating Markov models in continuous time without discrete-event simulation. Med Dec Making. 2013 Aug;33(6):767-79. PMID: 23715464.

296.	Ethgen O, Standaert B. Population- versus cohort-based modelling approaches. Pharmacoeconomics. 2012 Mar;30(3):171-81. PMID: 22283692.

297.	Marshall BD, Galea S. Formalizing the role of agent-based modeling in causal inference and epidemiology. Am J Epidemiol. 2015 Jan 15;181(2):92-99. PMID: 25480821.

298.	Spielauer M. Dynamic microsimulation of health care demand, health care finance and the economic impact of health behaviours: survey and review. International Journal of Microsimulation. 2007;1(1):35-53.

299.	Rutter CM, Zaslavsky AM, Feuer EJ. Dynamic microsimulation models for health outcomes: a review. Med Decis Making. 2011 Jan;31(1):10-18. PMID: 20484091.

300.	Edmunds WJ, Medley GF, Nokes DJ. Evaluating the cost-effectiveness of vaccination programmes: a dynamic perspective. Stat Med. 1999 Dec 15;18(23):3263-82. PMID: 10602150.

301.	Daun S, Rubin J, Vodovotz Y, et al. Equation-based models of dynamic biological systems. J Crit Care. 2008 Dec;23(4):585-94. PMID: 19056027.

302.	Jit M, Brisson M. Modelling the epidemiology of infectious diseases for decision analysis: a primer. Pharmacoeconomics. 2011 May;29(5):371-86. PMID: 21504239.

303.	Caro JJ. Pharmacoeconomic analyses using discrete event simulation. Pharmacoeconomics. 2005;23(4):323-32. PMID: 15853433.

304.	Hunink MM. Decision Making in Health and Medicine with CD-ROM: Integrating Evidence and Values. Cambridge University Press; 2001.

305.	Hunink MM, Weinstein MC, Wittenberg E, et al. Decision making in health and medicine: integrating evidence and values. Cambridge, England: Cambridge University Press; 2014.

306.	Briggs AH, Claxton K, Sculpher MJ. Decision modelling for health economic evaluation. Oxford, England: Oxford University Press; 2006.

307.	Torrance GW, Drummond MF. Methods for the economic evaluation of health care programmes. Oxford university press; 2005.

308.	Banks J, Carson II J, Nelson B, et al. Discrete-Event System Simulation. Upper Saddle River, N.J.: Prentice Hall; 2010.

309.	Zeigler BP, Praehofer H, Kim TG. Theory of modeling and simulation. San Diego, CA: Academic Press; 2000.

310.	Macal CM, North MJ. Tutorial on agent-based modeling and simulation. Proceedings of the 37th Winter Simulation Conference. 2005:2-15.

311.	Nance RE. The time and state relationships in simulation modeling. Communications of the ACM. 1981;24(4):173-79.

312.	Pidd M. Five simple principles of modelling. Proceedings of the 28th Winter Simulation Conference. 1996:721-28.

313.	Chwif L, Barretto MRP, Paul RJ. On simulation model complexity. Proceedings of the 32nd Winter Simulation Conference. 2000:449-55.

314.	Ward SC. Arguments for constructively simple models. J Oper Res Soc. 1989:141-53.

315.	Law AM. Simulation model's level of detail determines effectiveness. Ind Eng. 1991;23(10):16.

316.	Evans MR, Grimm V, Johst K, et al. Do simple models lead to generality in ecology? Trends in ecology & evolution. 2013;28(10):578-83.

317.	Brooks RJ, Tobias AM. Choosing the best model: Level of detail, complexity, and model performance. Math Comput Model. 1996;24(4):1-14.

318.	Zechmeister-Koss I, Schnell-Inderst P, Zauner G. Appropriate evidence sources for

populating decision analytic models within health technology assessment (HTA): a systematic review of HTA manuals and health economic guidelines. Med Decis Making. 2013 Oct 17;34(3):288-99. PMID: 24135150.

319.   Treadwell, JR, Singh, S, Talati, R, et al. A framework for "best evidence" approaches in systematic reviews. Rockville, MD: Agency for Healthcare Research and Quality (US); 2011. PMID: 21834173.

320.   Brettle AJ, Long AF, Grant MJ, et al. Searching for information on outcomes: do you need to be comprehensive? Qual Health Care. 1998 Sep;7(3):163-67. PMID: 10185143.

321.   Cooper N, Coyle D, Abrams K, et al. Use of evidence in decision models: an appraisal of health technology assessments in the UK since 1997. J Health Serv Res Policy. 2005 Oct;10(4):245-50. PMID: 16259692.

322.   Novielli N, Cooper NJ, Abrams KR, et al. 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 Health. 2010 Dec;13(8):952-57. PMID: 21029247.

323.   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. Health Technol Assess. 2003;7:iii, ix-x, 1-51. PMID: 14609481.

324.   Sutton AJ, Cooper NJ, Goodacre S, et al. Integration of meta-analysis and economic decision modeling for evaluating diagnostic tests. Med Decis Making. 2008 Sep;28(5):650-67. PMID: 18753686.

325.   Ades AE, Lu G, Higgins JP. The interpretation of random-effects meta-analysis in decision models. Med Decis Making. 2005 Nov;25(6):646-54. PMID: 16282215.

326.   Dias S, Sutton AJ, Welton NJ, et al. Evidence synthesis for decision making 6: embedding evidence synthesis in

probabilistic cost-effectiveness analysis. Med Decis Making. 2013 Jul;33(5):671-78. PMID: 23084510.

327.   Graham PL, Moran JL. Robust meta-analytic conclusions mandate the provision of prediction intervals in meta-analysis summaries. J Clin Epidemiol. 2012 May;65(5):503-10. PMID: 22265586.

328.   Oppe M, Al M, Rutten-van MM. Comparing methods of data synthesis: re-estimating parameters of an existing probabilistic cost-effectiveness model. Pharmacoeconomics. 2011 Mar;29(3):239-50. PMID: 21142288.

329.   Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. BMJ. 2011;342:d549. PMID: 21310794.

330.   Vemer P, Al MJ, Oppe M, et al. A choice that matters? Simulation study on the impact of direct meta-analysis methods on health economic outcomes. Pharmacoeconomics. 2013 Aug;31(8):719-30. PMID: 23736971.

331.   Ades AE, Cliffe S. Markov chain Monte Carlo estimation of a multiparameter decision model: consistency of evidence and the accurate assessment of uncertainty. Med Decis Making. 2002 Jul;22(4):359-71. PMID: 12150601.

332.   Ades AE. A chain of evidence with mixed comparisons: models for multi-parameter synthesis and consistency of evidence. Stat Med. 2003 Oct 15;22(19):2995-3016. PMID: 12973783.

333.   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(1):5-35. PMID: 18806188.

334.   Ades AE, Sculpher M, Sutton A, et al. Bayesian methods for evidence synthesis in cost-effectiveness analysis. Pharmacoeconomics. 2006;24(1):1-19. PMID: 16445299.

335.   Ades AE, Welton NJ, Caldwell D, et al. Multiparameter evidence synthesis in

epidemiology and medical decision-making. J Health Serv Res Policy. 2008 Oct;13 Suppl 3:12-22. PMID: 18806188.

336.   Borenstein M, Hedges LV, Higgins JP, et al. Introduction to meta-analysis. John Wiley & Sons; 2011.

337.   Cooper H, Hedges LV, Valentine JC. The handbook of research synthesis and meta-analysis. Russell Sage Foundation; 2009.

338.   Dias S, Sutton AJ, Ades AE, et al. 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 Jul;33(5):607-17. PMID: 23104435.

339.   Dias S, Welton NJ, Sutton AJ, et al. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. Med Decis Making. 2013 Jul;33(5):641-56. PMID: 23804508.

340.   Dias S, Sutton AJ, Welton NJ, et al. Evidence synthesis for decision making 3: heterogeneity--subgroups, meta-regression, bias, and bias-adjustment. Med Decis Making. 2013 Jul;33(5):618-40. PMID: 23804507.

341.   Eddy DM. The confidence profile method: a Bayesian method for assessing health technologies. Oper Res. 1989 Mar;37(2):210-28. PMID: 10292450.

342.   Eddy DM, Hasselblad V, Shachter R. An introduction to a Bayesian method for meta-analysis: The confidence profile method. Med Decis Making. 1990 Jan;10(1):15-23. PMID: 2182960.

343.   Eddy DM, Hasselblad V, Shachter R. A Bayesian method for synthesizing evidence. The confidence profile method. Int J Technol Assess Health Care. 1990;6(1):31-55. PMID: 2361818.

344.   Egger M, Smith GD, Altman D. Systematic reviews in health care: meta-analysis in context. Oxford, England: John Wiley & Sons; 2008.

345.   Fu R, Gartlehner G, Grant M, et al. Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. AHRQ Methods for Effective Health Care. 2008 PMID: 21433407.

346.   Hartung J, Knapp G, Sinha BK. Statistical meta-analysis with applications. New York, N.Y.: John Wiley & Sons; 2011.

347.   Kaizar EE. Estimating treatment effect via simple cross design synthesis. Stat Med. 2011 Nov 10;30(25):2986-3009. PMID: 21898521.

348.   Lau J, Terrin N, Fu R. Expanded guidance on selected quantitative synthesis topics. Methods Guide for Effectiveness and Comparative Effectiveness Reviews [Internet] Rockville (MD): Agency for Healthcare Research and Quality (US); 2008- AHRQ Methods for Effective Health Care. 2008 PMID: 23596640.

349.   Hedges L, Olkin I. Statistical methods for meta-analysis. San Diego, CA: Academic. 1985

350.   Schmid CH. Using Bayesian inference to perform meta-analysis. Eval Health Prof. 2001 Jun;24(2):165-89. PMID: 11523385.

351.   Spiegelhalter DJ, Best NG. Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. Stat Med. 2003 Dec 15;22(23):3687-709. PMID: 14652869.

352.   Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian approaches to clinical trials and health-care evaluation. Chichester, England: John Wiley & Sons; 2004.

353.   Stangl D, Berry DA. Meta-analysis in medicine and health policy. Boca Raton, FL: CRC Press; 2009.

354.   Sutton AJ, Abrams KR, Jones DR, et al. Methods for meta-analysis in medical research. New York: John Wiley & Sons; 2000.

355.   Sutton AJ, Cooper NJ, Abrams KR, et al. Evidence synthesis for decision making in

healthcare. Hoboken, N.J.: John Wiley & Sons; 2012.

356. Jansen JP, Fleurence R, Devine B, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. Value Health. 2011 Jun;14(4):417-28. PMID: 21669366.

357. Hoaglin DC, Hawkins N, Jansen JP, et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 2. Value Health. 2011 Jun;14(4):429-37. PMID: 21669367.

358. Kaizar EE. Incorporating Both Randomized and Observational Data into a Single Analysis. Annual Review of Statistics and Its Application. 2015;2:49-72.

359. Vanni T, Karnon J, Madan J, et al. Calibrating models in economic evaluation: a seven-step approach. Pharmacoeconomics. 2011 Jan;29(1):35-49. PMID: 21142277.

360. Stout NK, Knudsen AB, Kong CY, et al. Calibration methods used in cancer simulation models and suggested reporting guidelines. Pharmacoeconomics. 2009;27(7):533-45. PMID: 19663525.

361. Karnon J, Vanni T. Calibrating models in economic evaluation: a comparison of alternative measures of goodness of fit, parameter search strategies and convergence criteria. Pharmacoeconomics. 2011 Jan;29(1):51-62. PMID: 21142278.

362. Weinstein MC. Recent developments in decision-analytic modelling for economic evaluation. Pharmacoeconomics. 2006;24(11):1043-53. PMID: 17067190.

363. Taylor DC, Pawar V, Kruzikas D, et al. Methods of model calibration: observations from a mathematical model of cervical cancer. Pharmacoeconomics. 2010;28(11):995-1000. PMID: 20936883.

364. Hill MC. Methods and guidelines for effective model calibration. US Geological Survey Denver, CO, USA; 1998.

365. Hansen LP, Heckman JJ. The empirical foundations of calibration. J Econ Perspect. 1996:87-104.

366. Dawkins C, Srinivasan TN, Whalley J. Calibration. Handbook of econometrics. 2001;5:3653-703.

367. Jackson CH, Jit M, Sharples LD, et al. Calibration of complex models through bayesian evidence synthesis: a demonstration and tutorial. Med Decis Making. 2013 Jul 25 PMID: 23886677.

368. Rutter CM, Miglioretti DL, Savarino JE. Bayesian Calibration of Microsimulation Models. J Am Stat Assoc. 2009 Dec 1;104(488):1338-50. PMID: 20076767.

369. Whyte S, Walsh C, Chilcott J. Bayesian calibration of a natural history model with application to a population model for colorectal cancer. Med Decis Making. 2011 Jul;31(4):625-41. PMID: 21127321.

370. Welton NJ, Ades AE. Estimation of markov chain transition probabilities and rates from fully and partially observed data: uncertainty propagation, evidence synthesis, and model calibration. Med Decis Making. 2005 Nov;25(6):633-45. PMID: 16282214.

371. Greenland S. Basic methods for sensitivity analysis of biases. Int J Epidemiol. 1996 Dec;25(6):1107-16. PMID: 9027513.

372. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiology. 2009 Jul;20(4):488-95. PMID: 19525685.

373. Braithwaite RS, Roberts MS, Justice AC. Incorporating quality of evidence into decision analytic modeling. Ann Intern Med. 2007 Jan 16;146(2):133-41. PMID: 17227937.

374. Goldhaber-Fiebert JD. Accounting for biases when linking empirical studies and

simulation models. Med Decis Making. 2012 May;32(3):397-99. PMID: 22593033.

375. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928. PMID: 22008217.

376. Berger ML, Martin BC, Husereau D, et al. A questionnaire to assess the relevance and credibility of observational studies to inform health care decision making: an ISPOR-AMCP-NPC good practice task force report. Value Health. 2014;17(2):143-56. PMID: 24636373.

377. Dias S, Welton NJ, Marinho VCC, et al. 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(3):613-29.

378. 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]. Int J Epidemiol. 2009 Dec;38(6):1662-73. PMID: 19744933.

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

380. Gustafson P, McCandless LC. Probabilistic approaches to better quantifying the results of epidemiologic studies. Int J Environ Res Public Health. 2010 Apr;7(4):1520-39. PMID: 20617044.

381. Hofler M, Lieb R, Wittchen HU. Estimating causal effects from observational data with a model for multiple bias. Int J Methods Psychiatr Res. 2007;16(2):77-87. PMID: 17623387.

382. Kuroki, M, Pearl, J. Measurement bias and effect restoration in causal inference. DTIC Document; 2011.

383. Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data. New York, N.Y.: Springer; 2011.

384. Maldonado G. Adjusting a relative-risk estimate for study imperfections. J Epidemiol Community Health. 2008 Jul;62(7):655-63. PMID: 18559450.

385. Molitor N, Best N, Jackson C, et al. 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(3):615-37.

386. Thompson S, Ekelund U, Jebb S, et al. A proposed method of bias adjustment for meta-analyses of published observational studies. Int J Epidemiol. 2011 Jun;40(3):765-77. PMID: 21186183.

387. Turner RM, Spiegelhalter DJ, Smith G, et al. Bias modelling in evidence synthesis. J R Stat Soc Ser A Stat Soc. 2009;172(1):21-47. PMID: 19381328.

388. Welton NJ, Ades AE, Carlin JB, et al. 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(1):119-36.

389. Bala MV, Wood LL, Zarkin GA, et al. Valuing outcomes in health care: a comparison of willingness to pay and quality-adjusted life-years. J Clin Epidemiol. 1998 Aug;51(8):667-76. PMID: 9743315.

390. Bravata DM, Nelson LM, Garber AM, et al. Invariance and inconsistency in utility ratings. Med Decis Making. 2005 Mar;25(2):158-67. PMID: 15800300.

391. Chaloner K, Rhame FS. Quantifying and documenting prior beliefs in clinical trials. Stat Med. 2001 Feb 28;20(4):581-600. PMID: 11223902.

392. Frew EJ, Whynes DK, Wolstenholme JL. Eliciting willingness to pay: comparing closed-ended with open-ended and payment scale formats. Med Decis Making. 2003 Mar;23(2):150-59. PMID: 12693877.

393. Garthwaite PH, Kadane JB, O'Hagan A. Statistical methods for eliciting probability distributions. Journal of the American Statistical Association. 2005;100(470):680-701.

394. Johnson SR, Tomlinson GA, Hawker GA, et al. A valid and reliable belief elicitation method for Bayesian priors. J Clin Epidemiol. 2010 Apr;63(4):370-83. PMID: 19926253.

395. Kerstholt JH, van der Zwaard F, Bart H, et al. Construction of health preferences: a comparison of direct value assessment and personal narratives. Med Decis Making. 2009 Jul;29(4):513-20. PMID: 19237644.

396. O'Hagan A, Buck CE, Daneshkhah A, et al. Uncertain judgements: eliciting experts' probabilities. Hoboken, N.J.: John Wiley & Sons; 2006.

397. Soares MO, Bojke L, Dumville J, et al. Methods to elicit experts' beliefs over uncertain quantities: application to a cost effectiveness transition model of negative pressure wound therapy for severe pressure ulceration. Stat Med. 2011 Aug 30;30(19):2363-80. PMID: 21748773.

398. 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. Stat Med. 2005 Dec 30;24(24):3805-21. PMID: 16320265.

399. Fischhoff B, Slovic P, Lichtenstein S. Knowing what you want: Measuring labile values. Decision Making: Descriptive, Normative and Prescriptive Interactions. Cambridge, UK: Cambridge University Press; 1988. p. 398-421.

400. Drummond M, Manca A, Sculpher M. Increasing the generalizability of economic evaluations: recommendations for the design, analysis, and reporting of studies. Int J Technol Assess Health Care. 2005;21(2):165-71. PMID: 15921055.

401. Rosen R. Role of similarity principles in data extrapolation. Am J Physiol. 1983 May;244(5):R591-R599. PMID: 6846566.

402. Bareinboim E, Pearl J. A general algorithm for deciding transportability of experimental results. J Causal Inference. 2013;1(1):107-34.

403. Bareinboim E, Pearl J. Meta-transportability of causal effects: a formal approach. Proceedings of the 16th International Conference on Artificial Intelligence and Statistics (AISTATS). 2013:135-43.

404. Bareinboim E, Pearl J. Causal transportability with limited experiments. Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI 2013), Menlo Park, CA: AAAI Press. 2013

405. Pearl J, Bareinboim E. Transportability of causal and statistical relations: a formal approach. Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on. 2011:540-47.

406. Pearl, J, Bareinboim, E. External validity: from do-calculus to transportability across populations. DTIC Document; 2012.

407. Pressler TR, Kaizar EE. The use of propensity scores and observational data to estimate randomized controlled trial generalizability bias. Stat Med. 2013 Sep 10;32(20):3552-68. PMID: 23553373.

408. Groot Koerkamp B, Weinstein MC, Stijnen T, et al. Uncertainty and patient heterogeneity in medical decision models. Med Decis Making. 2010 Mar;30(2):194-205. PMID: 20190188.

409. Groot Koerkamp B, Stijnen T, Weinstein MC, et al. The combined analysis of uncertainty and patient heterogeneity in medical decision models.[Erratum appears in Med Decis Making. 2013 Feb;33(2):307]. Med Decis Making. 2011 Jul;31(4):650-61. PMID: 20974904.

410. Zaric GS. The impact of ignoring population heterogeneity when Markov models are used in cost-effectiveness analysis. Med Decis Making. 2003 Sep;23(5):379-96. PMID: 14570296.

411. Kuntz KM, Goldie SJ. Assessing the sensitivity of decision-analytic results to

unobserved markers of risk: defining the effects of heterogeneity bias. Med Decis Making. 2002 May;22(3):218-27. PMID: 12058779.

412. Ramaekers BL, Joore MA, Grutters JP. How should we deal with patient heterogeneity in economic evaluation: a systematic review of national pharmacoeconomic guidelines. Value Health. 2013 Jul;16(5):855-62. PMID: 23947981.

413. Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. Stat Med. 2004 Jun 15;23(11):1663-82. PMID: 15160401.

414. Higgins JP. Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. Int J Epidemiol. 2008 Oct;37(5):1158-60. PMID: 18832388.

415. 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(1):137-59. PMID: 19381330.

416. Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? Stat Med. 2002 Jun 15;21(11):1559-73. PMID: 12111920.

417. Greenland S. Invited commentary: a critical look at some popular meta-analytic methods. Am J Epidemiol. 1994 Aug 1;140(3):290-96. PMID: 8030632.

418. Poole C, Greenland S. Random-effects meta-analyses are not always conservative. Am J Epidemiol. 1999 Sep 1;150(5):469-75. PMID: 10472946.

419. Berlin JA, Santanna J, Schmid CH, et al. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. Stat Med. 2002 Feb 15;21(3):371-87. PMID: 11813224.

420. Donegan S, Williamson P, D'Alessandro U, et al. 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. Stat Med. 2013 Mar 15;32(6):914-30. PMID: 22987606.

421. Koopman L, van der Heijden GJ, Hoes AW, et al. Empirical comparison of subgroup effects in conventional and individual patient data meta-analyses. Int J Technol Assess Health Care. 2008;24(3):358-61. PMID: 18601805.

422. Riley RD, Simmonds MC, Look MP. Evidence synthesis combining individual patient data and aggregate data: a systematic review identified current practice and possible methods. J Clin Epidemiol. 2007 May;60(5):431-39. PMID: 17419953.

423. Riley RD, Dodd SR, Craig JV, et al. Meta−analysis of diagnostic test studies using individual patient data and aggregate data. Stat Med. 2008;27(29):6111-36. PMID: 18816508.

424. Schmid CH, Stark PC, Berlin JA, et al. Meta-regression detected associations between heterogeneous treatment effects and study-level, but not patient-level, factors. J Clin Epidemiol. 2004 Jul;57(7):683-97. PMID: 15358396.

425. Simmonds MC, Higgins JP. Covariate heterogeneity in meta-analysis: criteria for deciding between meta-regression and individual patient data. Stat Med. 2007 Jul 10;26(15):2982-99. PMID: 17195960.

426. Sutton AJ, Kendrick D, Coupland CA. Meta−analysis of individual− and aggregate− level data. Stat Med. 2008;27(5):651-69. PMID: 17514698.

427. Kovalchik SA. Survey finds that most meta-analysts do not attempt to collect individual patient data. J Clin Epidemiol. 2012 Dec;65(12):1296-99. PMID: 22981246.

428. Briggs AH, Gray AM. Handling uncertainty when performing economic evaluation of healthcare interventions. Health Technol Assess. 1999;3(2):1-134. PMID: 10448202.

429. Briggs AH. Handling uncertainty in cost-effectiveness models. Pharmacoeconomics. 2000;17(5):479-500. PMID: 10977389.

430. Briggs AH, O'Brien BJ, Blackhouse G. Thinking outside the box: recent advances in the analysis and presentation of uncertainty in cost-effectiveness studies. Annu Rev Public Health. 2002;23:377-401. PMID: 11910068.

431. Lord J, Asante MA. Estimating uncertainty ranges for costs by the bootstrap procedure combined with probabilistic sensitivity analysis. Health Econ. 1999 Jun;8(4):323-33. PMID: 10398525.

432. O'Hagan A, McCabe C, Akehurst R, et al. Incorporation of uncertainty in health economic modelling studies. Pharmacoeconomics. 2005;23:529-36. PMID: 15960550.

433. Pasta DJ, Taylor JL, Henning JM. Probabilistic sensitivity analysis incorporating the bootstrap: an example comparing treatments for the eradication of Helicobacter pylori. Med Decis Making. 1999 Jul;19(3):353-63. PMID: 10424842.

434. Parmigiani G. Measuring uncertainty in complex decision analysis models. Stat Method Med Res. 2002;11(6):513-37.

435. Oehlert GW. A note on the delta method. Am Stat. 1992;46(1):27-29.

436. Briggs AH, Ades AE, Price MJ. Probabilistic sensitivity analysis for decision trees with multiple branches: use of the Dirichlet distribution in a Bayesian framework. Med Decis Making. 2003 Jul;23(4):341-50. PMID: 12926584.

437. Griffin S, Claxton K, Hawkins N, et al. Probabilistic analysis and computationally expensive models: Necessary and required? Value Health. 2006;9(4):244-52. PMID: 16903994.

438. 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 Econ. 2007 Oct;16(10):1009-23. PMID: 17173339.

439. Law AM. A tutorial on how to select simulation input probability distributions.

Simulation Conference (WSC), Proceedings of the 2012 Winter. IEEE; 2012. p. 1-15.

440. Apostolakis G. The concept of probability in safety assessments of technological systems. Science. 1990;250(4986):1359-64. PMID: 2255906.

441. MacConnell S. Code Complete : A practical handbook of software construction. Redmond, Wash.: Microsoft Press; 2004.

442. Schruben LW. Establishing the credibility of simulations. Simulation. 1980;34(3):101-05.

443. Turing A. Computing machinery and intelligence. Mind-A Quarterly Review of Psychology and Philosophy. 1950;59(236):433-60.

444. Harel D. A Turing-like test for biological modeling. Nat Biotechnol. 2005 Apr;23(4):495-96. PMID: 15815679.

445. Bennett C, Manuel DG. Reporting guidelines for modelling studies. BMC Medical Research Methodology. 2012;12:168. PMID: 23134698.

446. Cooper BS. Confronting models with data. J Hosp Infect. 2007 Jun;65 Suppl 2:88-92.

447. Moriasi DN, Arnold JG, Van Liew MW, et al. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans ASABE. 2007;50(3):885-900.

448. Sendi PP, Craig BA, Pfluger D, et al. Systematic validation of disease models for pharmacoeconomic evaluations. Swiss HIV Cohort Study. J Eval Clin Pract. 1999 Aug;5(3):283-95. PMID: 10461580.

449. Roberts S, Pashler H. How persuasive is a good fit? A comment on theory testing. Psychol Rev. 2000 Apr;107(2):358-67. PMID: 10789200.

450. Sargent RG. Some subjective validation methods using graphical displays of data. Proceedings of the 28th Winter Simulation Conference. 1996. p. 345-51.

451. Garrett ES, Zeger SL. Latent class model diagnosis. Biometrics. 2000 Dec;56(4):1055-67. PMID: 11129461.

452. Hodges JS, Dewar JA, Center A. Is it you or your model talking?: A framework for model validation. Santa Monica, CA: Rand; 1992.

453. 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. Stat Methods Med Res. 2004 Dec;13(6):525-38. PMID: 15587437.

454. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, et al. Evaluating test strategies for colorectal cancer screening: a decision analysis for the U.S. Preventive Services Task Force. Ann Intern Med. 2008;149(9):659-69. PMID: 18838717.

455. Lansdorp-Vogelaar I, Gulati R, Mariotto AB, et al. Personalizing age of cancer screening cessation based on comorbid conditions: model estimates of harms and benefits. Ann Intern Med. 2014 Jul 15;161(2):104-12. PMID: 25023249.

456. Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med. 2005;353(17):1784-92. PMID: 16251534.

457. Peng RD. Reproducible research in computational science. Science. 2011;334:1226-27. PMID: 22144613.

458. Laine C, Goodman SN, Griswold ME, et al. Reproducible research: moving toward research the public can really trust. Ann Intern Med. 2007;146:450-53. PMID: 17339612.

459. Rahmandad H, Sterman JD. Reporting guidelines for simulation–based research in social sciences. System Dynamics Review. 2012;28(4):396-411.

460. National Research Council. Models in environmental regulatory decision making. Washington, DC: National Research Council; 2007.

461. Eddy DM. Accuracy versus transparency in pharmacoeconomic modelling: finding the right balance. Pharmacoeconomics. 2006;24(9):837-44. PMID: 1694119.

462. Landry M, Banville C, Oral M. Model legitimisation in operational research. Eur J Oper Res. 1996;92(3):443-57.

463. Thompson KM. Variability and uncertainty meet risk management and risk communication. Risk Anal. 2002 Jun;22(3):647-54. PMID: 12088239.

464. Politi MC, Han PK, Col NF. Communicating the uncertainty of harms and benefits of medical interventions. Med Decis Making. 2007 Sep;27(5):681-95. PMID: 17873256.

465. Redelmeier DA, Detsky AS, Krahn MD, et al. Guidelines for verbal presentations of medical decision analyses. Med Decis Making. 1997 Apr;17(2):228-30. PMID: 9107619.

466. Brailsford SC, Bolt T, Connell C, et al. Stakeholder engagement in health care simulation. Proceedings of the 41st Winter Simulation Conference. 2009. p. 1840-49.

467. Cleveland WS. The elements of graphing data. Monterey, CA.: Wadsworth; 1985.

468. Cleveland WS. Visualizing data. Summit, N.J.: Hobart Press; 1993.

469. Tufte ER, Graves-Morris PR. The visual display of quantitative information. Cheshire, CT: Graphics Press; 1983.

470. Drummond MF, McGuire A. Economic evaluation in health care: merging theory with practice. Oxford, England: Oxford University Press; 2001.

471. Sox HC, Owens D, Higgins MC. Medical decision making. Hoboken, N.J.: John Wiley & Sons; 2013.

472. Gray AM, Clarke PM, Wolstenholme JL, et al. Applied methods of cost-effectiveness analysis in healthcare. Oxord, England: Oxford University Press; 2010.

473. Baker CB, Johnsrud MT, Crismon ML, et al. Quantitative analysis of sponsorship bias in economic studies of antidepressants. Br J Psychiatry. 2003 Dec;183:498-506. PMID: 14645020.

474. Barbieri M, Drummond MF. Conflict of interest in industry-sponsored economic evaluations: real or imagined? Curr Oncol Rep. 2001 Sep;3(5):410-13. PMID: 11489241.

475. Friedberg M, Saffran B, Stinson TJ, et al. Evaluation of conflict of interest in economic analyses of new drugs used in oncology. JAMA. 1999 Oct 20;282(15):1453-57. PMID: 10535436.

476. Garattini L, Koleva D, Casadei G. Modeling in pharmacoeconomic studies: funding sources and outcomes. Int J Technol Assess Health Care. 2010 Jul;26(3):330-33. PMID: 20584363.

477. Valachis A, Polyzos NP, Nearchou A, et al. Financial relationships in economic analyses of targeted therapies in oncology. J Clin Oncol. 2012 Apr 20;30(12):1316-20. PMID: 22430267.

478. Bell CM, Urbach DR, Ray JG, et al. Bias in published cost effectiveness studies: systematic review. BMJ. 2006 Mar 25;332(7543):699-703. PMID: 16495332.

479. Lo B, Field MJ. Conflict of interest in medical research, education, and practice. Washington D.C.: National Academies Press; 2009.

480. Viswanathan, M, Carey, TS, Belinson, SE, et al. Identifying and managing nonfinancial conflicts of interest for systematic reviews. 13-EHC085-EF. Rockville, MD: Agency of Healthcare Research and Quality; 2013. PMID: 23844449.