Comparative Effectiveness Research/Health Technology Assessment (HTA)

Adding Decision Models to Systematic Reviews: Informing a Framework for Deciding When and How to Do So

François Sainfort, PhD1, Karen M. Kuntz, ScD1, Sean Gregory, MBA1, Mary Butler, PhD, MBA1, Brent C. Taylor, PhD2,3,4, Shalini Kulasingam, PhD2, Robert L. Kane, MD1

1Division of Health Policy and Management, School of Public Health, University of Minnesota, Minneapolis, MN, USA; 2Veterans Affairs Health Care System, Minneapolis, MN, USA; 3Department of Medicine, University of Minnesota, Minneapolis, MN, USA; 4Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN, USA

ABSTRACT

Objective: Decision models are sometimes used alongside systematic reviews to synthesize evidence. Clarity, however, is lacking about when and how to conduct modeling studies in tandem with systematic reviews, as well as about how to evaluate and present model results. The objective of this study was to collect and analyze information from various sources to inform the development of a framework for deciding when and how a decision model should be added to a systematic review. Methods: We collected data through 1) review and analysis of evidence reports that used decision models; 2) review and synthesis of current best practices for the development of decision models; 3) interviews of Evidence-Based Practice Center directors and selected staff, United States Preventive Services Task Force members, and decision modelers who developed models used by the United States Preventive Services Task Force; and 4) a focus group of expert modelers. Results: Models are well suited to address gaps in the literature, better suited for certain types of research questions, and essential for determining the value of information relating to future research. Opinions differ regarding whether model outputs constitute evidence, but interviewees expressed concern over the lack of standards and directions in grading and reporting such “evidence.” Interviews of stakeholders and modelers revealed the importance of communication and presentation of model results as well as the importance of model literacy and involvement of stakeholders. Conclusions: The study demonstrates the need for a framework for deciding when and how to use models alongside systematic reviews and provides information to develop such a framework. Keywords: decision model, evidence, guidelines, systematic review.

Introduction

Systematic reviews are widely used to evaluate and synthesize the scientific literature on a particular topic. Results from systematic reviews are meant to help public and private organizations develop guidelines and strategies that improve the quality of health care and decision making. In the United States, the Agency for Healthcare Research and Quality funds projects through its Effective Health Care Program [1] to

- “Review and synthesize published and unpublished scientific evidence.
- Generate new scientific evidence and analytic tools.
- Compile research findings that are synthesized and/or generated and translates them into useful formats for various audiences.” [1]

Often, however, existing evidence cannot fully address the relevant questions being asked. Decision models may enhance the value of systematic reviews by adding a formal structure that can be informed by, and also extrapolate beyond, the evidence to produce additional outcomes relevant to decision makers. Since the first application reported by Henschke and Flehinger [2] in 1967, decision-analytic models have been increasingly used to evaluate and compare competing public health and medical interventions. While decision models may provide added value, the underlying philosophical approach—creating additional knowledge through modeling—is not necessarily congruent with the underlying philosophical approach of systematic reviews —critical assessment and evaluation of all research studies. At issue are the nature and strength of the “evidence” generated by a model as opposed to the nature and strength of evidence observed in a collection of studies carefully screened and systematically evaluated. Historically, a “hierarchy of evidence” has been used that ranks randomized clinical trials higher than other types of studies [3]. It is unclear where decision models “fit” within this hierarchy.

Developing a decision model for a particular question requires a synthesis of relevant literature pertaining to the natural history...
(or risk) of disease, effectiveness and risks of alternative interventions, and health-related quality of life. Thus, modeling endeavors rely on much of the same information provided by systematic reviews. Models, however, must typically be supplemented by additional data, as well as by clinically reasonable assumptions where data are limited or nonexistent. As a result, decision analysis methods are an obvious companion to systematic reviews. However, when a systematic review is conducted, it is unclear whether a decision model should be developed and used in tandem with the systematic review and whether the decision model results add value to the nature and strength of the evidence summarized in the review. There is a lack of clarity about how to simultaneously evaluate and present results from both the systematic review and the model. Prior work has raised (and in some cases, addressed) important issues including 1) which input estimates should be used as a result of a systematic review [4] and 2) how to incorporate quality or strength of evidence into decision models [5]. Thus, we conducted a study to help inform a framework for deciding when and how a decision model and its results should be added to a systematic review. As part of this study, we collected information from all key stakeholders involved in conducting and using systematic reviews in the United States: 1) the Evidence-Based Practice Centers (EPCs) as they are the largest producers of systematic reviews; 2) the United States Preventive Services Task Force (USPSTF), a key user of a large number of systematic reviews; and 3) expert modelers. Working in conjunction with the Agency for Healthcare Research and Quality and a technical expert panel established to provide guidance on all aspects of the study, we identified the following main questions:

1. How and when have decision models been used alongside systematic reviews?
2. How do key individuals who routinely perform systematic reviews view the use and potential value of decision models?
3. What is the recent experience of the USPSTF, a key user of systematic reviews, with respect to using decision models alongside systematic reviews?
4. What were the experiences of the modeling teams who developed decision models alongside the systematic reviews used by the USPSTF?
5. What can be learned from expert modelers frequently involved in the development and use of decision models alongside systematic reviews?

Prior to answering these questions, the concept of a “decision model” needed to be defined. Decision models vary from very simple “back of the envelope” calculations to complex computer-based microsimulation and optimization. While the term “model” conveys different meanings in different clinical settings [6], we followed the taxonomy developed by Brennan et al. [7] and focused on decision-analytic model structures (Table 1). Decision-analytic models are typically used in medical decision-making applications: decision trees, Markov (cohort) models, microsimulation (individual) models, dynamic models, and discrete event simulation models.

### Table 1: Summary of types of decision model structures.

<table>
<thead>
<tr>
<th>Model type</th>
<th>General description</th>
<th>Type of decision best suited for</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Decision tree</strong></td>
<td>Diagrams the risk of events and states of nature over a fixed time horizon</td>
<td>Interventions for which the relevant time horizon is short and fixed</td>
</tr>
<tr>
<td><strong>Markov (cohort) model</strong></td>
<td>Simulates a hypothetical cohort of individuals through a set of health states over time</td>
<td>Modeling interventions for diseases or conditions that involve risk over a long time horizon and/or recurrent events</td>
</tr>
<tr>
<td><strong>Microsimulation (individual) model</strong></td>
<td>Simulates one individual at a time; tracks the past health states of individual and models risk of future events stochastically</td>
<td>Modeling complex disease processes, when Markov models are too limiting</td>
</tr>
<tr>
<td><strong>Dynamic model</strong></td>
<td>System of differential equations that simulates the interactions between individuals and the spread of disease</td>
<td>Modeling interventions for communicable diseases, such as vaccinations</td>
</tr>
<tr>
<td><strong>Discrete event simulation model</strong></td>
<td>Simulates time to an event and subsequent events, one individual at a time as well as interactions among individuals or within a health care system</td>
<td>Evaluating alternative health care systems (e.g., workflow, staffing) though flexible enough to address questions in several different areas</td>
</tr>
</tbody>
</table>

**Methods**

To answer the key questions and inform the development of a framework, we used a mixed-method qualitative approach. First, we summarized the use of decision models in all EPCs’ evidence reports to identify when and why decision models were developed alongside systematic reviews. As noted earlier, EPCs are the largest producers of systematic reviews in the United States. Another literature review summarized current best practices for the development of decision models. Next, we conducted a series of elite interviews with three groups of key informants: EPC directors and their staff, USPSTF members who participated in recent recommendations based on the results from decision models in addition to systematic reviews, and the corresponding group of expert modelers who developed the models used by the USPSTF. Last, a focus group of expert modelers offered views regarding best practices for model development. We briefly describe each of these methods below; the full report [8] provides complete detail for the methods.

**Literature Reviews**

Use of decision models in evidence reports

To identify reports for which a decision model was developed alongside a systematic review, we reviewed all 193 evidence reports generated by the EPCs available, but not archived, on the Agency for Healthcare Research and Quality EPC Evidence.
Best practices in decision modeling
A second systematic review identified articles that recommended best practices for developing decision models. We used two search strategies: 1) a keyword search algorithm designed to locate both model and “best practices” constructs from database inception to March 2010 and 2) an update of the search strategy used by Phillips et al. [10] for their review of good practice guidelines. To complement the review, we also searched the gray literature via Web search engines for published guidelines from professional societies, governmental bodies, and other health-related organizations. We reviewed articles and extracted information that could further inform a framework for developing a decision model alongside a systematic review.

Key Informants Interviews
EPC directors and staff
To complement the review of EPC reports and to better understand if, when, and why decision models are developed alongside systematic reviews, we interviewed EPC directors and staff members. We sought to 1) discuss whether and why EPCs have been involved in decision modeling activities and whether the modeling results were incorporated into reports; and 2) determine the factors instrumental in considering, developing, and completing/abandoning modeling activities. We interviewed members from all EPCs rather than only those EPCs that have incorporated models because we wanted to incorporate lessons and gather reasoning and perspectives from EPCs who 1) have considered and attempted to incorporate modeling but decided not to or 2) have not considered developing or incorporating models at all or 3) have no familiarity with modeling. We developed a semi-structured interview guide, identified 20 potential respondents, and conducted interviews from December 15, 2009, through March 2010, either by phone or face to face. Interviews were recorded and analyzed.

USPSTF and expert modelers
To evaluate strengths and weaknesses of current approaches to jointly conducting a systematic review and modeling exercise, we conducted key informant interviews with members of the USPSTF and modeling groups involved with the recent colorectal cancer, breast cancer, and cervical cancer modeling projects [11–13]. The interviews were designed to 1) evaluate all stakeholder experiences with the process including their perceived needs and whether they were met, 2) understand the impact of the modeling exercises on USPSTF decision making, and 3) make recommendations for the process of conducting similar projects. The final sample of respondents consisted of the leaders of the three USPSTF cancer modeling projects, members of each modeling team, and USPSTF members who were involved in either or both the development of the models and voting on recommendations (the evidence for which included modeling). The interview guide focused on strengths and weaknesses of current approaches, perceived needs, degree to which needs are met, lessons learned from the cancer screening modeling projects, and perceived impact of these projects on USPSTF decision making. Interviews lasted about an hour via telephone between April 5, 2010, and May 25, 2010.

Expert Modeler Focus Group
We convened a focus group of expert modelers to discuss, characterize, and qualify best practices in decision and simulation modeling in the context of systematic reviews. Discussions explored model development and construction, handling and presentation of modeling assumptions, and presentation and communication of results. Focus group participants included expert modelers who work frequently with the United Kingdom National Institute for Health and Clinical Excellence, an organization that uses decision or simulation modeling to provide guidance to the National Health Service on the clinical effectiveness and cost-effectiveness of selected medical technologies. We provided participants with advance summaries of preliminary findings from interviews with EPC members and with a selection of articles on best practices. The focus group was conducted in May 2010 and (with prior consent of participants) was recorded, analyzed, and summarized.

Results
The EPCs’ Perspective(s)
Evidence reports
Out of 192 evidence reports reviewed, 10 reports (5.2%) used decision analysis. All but two developed new models as part of the process. One evidence report [14] adapted a previously published model while the second report [15] further refined the model. Six reports modeled diagnostic tests or screening strategies along with subsequent treatments [16–21], while three reports modeled treatments only [14,15,22]. Only three reports used models as the prime methodology to answer key questions [14,15] or address the main research aim [18]. The remaining seven reports used models to augment systematic review results in cases where preliminary searches suggested that the literature would be unable to address the key question directly.

The reason most often given for incorporating models into the evidence reports was to provide a link between intermediate outcomes and clinical, or patient-centered, outcomes. Other reasons included simulating head-to-head comparisons otherwise unavailable in the literature, examining cost-effectiveness, and determining the effectiveness of screening under specific scenarios by modeling a novel hypothesis for disease progression not previously mentioned in the literature. The reports, however, infrequently stated clear purposes for incorporating decision-analytic models.

Models contributed to conclusions in several ways. Seven evidence reports used model results to conclude more optimal practices or no clinically important distinguishable differences [12,13,16,17,19,20,22,23]. One model demonstrated that outputs were sensitive to changes in key input variables and thus contributed to concluding that the current evidence was insufficient to support broad implementation of the treatment [15]. Models that relied on low-quality evidence were reported as exploratory [22]. Two modeling exercises were performed to promote understanding of the interactions between the variables of an analytic framework rather than to provide a basis for clinical recommendations [18,20].

One evidence report described a modeling effort intended to evaluate the usefulness of diagnostic modalities that differentiate epileptic seizures from seizures commonly mistaken for epilepsy [24]. The effort would have required diagnostic performance data from multiple sources to accurately model the
clinical differential diagnosis. Lack of available evidence, however, prevented the model from being developed.

**EPC directors and staff**

Nineteen of the 20 EPC directors and designated staff contacted were interviewed, representing 12 of the 13 EPCs (92.3%). Seven main themes emerged from the discussions (Table 2). Most themes were addressed across all EPC discussions, depending on the interviewees’ experience with modeling. Of the 19 interviewees, 15 had either personal modeling experience/expertise or were members of an EPC with modeling experience. Interviewees with experience tended to respond more similarly than did those without experience. Table 2 summarizes the key differences between interviewees with and without modeling experience with respect to the seven major themes that emerged from the interviews.

Interviewees with modeling experience were unanimously positive about how models can augment the evidence from systematic reviews. They stated that models are well suited to address gaps in the literature and to synthesize literature from differing sources and contexts into a single representation of the empirical evidence. They identified certain research questions as lending themselves to modeling, including those regarding screening as well as those involving harms or benefits measured with intermediate outcomes instead of the terminal outcome of interest, such as survival or disease prevention. They also noted comparison of testing, prevention, and diagnostic strategies as areas of great benefit for modeling. They pointed out that decision models can best quantify net benefit (benefits less harms). More generally, respondents viewed models as well suited for research questions in which there is a high degree of uncertainty in assumptions or input parameters, or in situations in which there is a great amount of discordance between estimates in empirical studies. They felt that models offer great benefit over alternative methods (such as randomized controlled trials or observational methods) to derive findings for small specific subpopulations of interest. Last, they noted that models are essential for determining the value of information as related to future research priorities and directions. Interviewees identified the lack of defined standards and methods as a major problem for evaluating models and presentation of outputs; they expressed the need for a framework to fill this gap. Many

### Table 2 – Differences between responses of interviewees with and without modeling experience.

<table>
<thead>
<tr>
<th>Interview theme</th>
<th>Interviewees with experience (n = 15)</th>
<th>Interviewees without experience (n = 4)</th>
</tr>
</thead>
</table>
| **Attitudes Toward Models and Appropriateness of Modeling in Systematic Reviews** | - Important set of techniques and strategies for analysis and should be incorporated into systematic reviews  
- Natural extension of the systematic review by addressing gaps in the literature and extending information about intermediate benefits and harms to terminal outcomes | - Systematic review should be limited to synthesis and meta-analysis of all available empirical and observational evidence  
- Models are outside the scope and purpose of the systematic review |
| **Research Questions and Contexts Best Suited for Decision and Simulation Modeling** | - Comparison of testing strategies (start, stop, and interval)  
- Determination of complicated net benefit calculations by linking intermediate to terminal benefits and harms with additional data sources  
- Questions with high degree of uncertainty  
- Application of findings to subpopulations not included in the original study | - Situations with high degree of uncertainty  
- Difficulty enumerating, but agreed with the “with experience” examples when prompted |
| **Definition of Decision and Simulation Models** | - Mathematical representation of a decision based on empirical input parameters, supported by a framework, and subject to a set of identifiable assumptions | - Confusion on where modeling is defined differently from statistical inference |
| **Evaluation of Models and Assessment of Model Outcomes** | - Quality and expertise of the modeler(s)  
- Lack of defined standards  
- Inspection of assumptions and theoretical framework (natural history of disease representation) | - Focus on the quality and “believability” of the output parameters, and whether multiple models generated similar results  
- Lacked familiarity with any empirical measures of model quality |
| **Decision and Simulation Models Results as Evidence** | - Outputs generated from models merit inclusion in systematic reviews as evidence  
- Modeling offers access to parameters that might not otherwise be available (e.g., subpopulations) | - Model evidence is “manufactured” or “model produced” and thus must be kept separate from empirical evidence (RCT or observational)  
- There is no evidence grading for model-based parameters |
| **Impact of Decision and Simulation Modeling on Systematic Reviews** | - Models require additional time and expense, and are not always able to be anticipated at the initiation of a project  
- Likely to add 20%–40% to the time and expense of a typical systematic review  
- Need a mechanism to include a model after the question refinement phase has been completed! | - Would require expertise that some EPCs do not have in-house, and thus must contract for externally  
- Need to have guidelines from the Methods Manual |
| **Training Needs** | - Increase training opportunities for doctoral and postdoctoral positions to train modelers | - Need for seminars and programs to train existing EPC staff  
- Identify modeling groups with specific expertise to contract with for model components of systematic reviews |

EPC, Evidence-Based Practice Center; RCT, randomized controlled trial.
respondents stressed the need to determine the opportunity or need for a model and/or simulation in advance, specifically before the completion of the question refinement phase or the early stage literature review.

Interviewees without experience believed that developing a model was beyond the scope of a systematic review of the literature and questioned whether published models and related output constitute valid information that could be included in systematic reviews. They questioned incorporating and evaluating the "engineered" evidence provided by such modeling studies. While opinions differed regarding whether outputs of models should be presented along with other evidence in systematic reviews, both groups of interviewees recognized the lack of standards and directions in grading and reporting such "evidence" and indicated that a framework should address such issues.

Users’ and modelers’ perspectives

Three main themes emerged from the interviews with cancer modelers and USPSTF members about conducting decision and simulation models alongside systematic reviews to inform USPSTF recommendations: 1) communication and presentation of model results, 2) modeling literacy, and 3) recommendations for future projects. Although the modeling efforts were conducted differently across the three cancer projects, interviewees expressed consistent views about the difficulty of communicating model results and the challenge of communicating model results, particularly to decision makers who were not knowledgeable about modeling. They stressed that these issues must be addressed to improve the success, acceptance, and use of models.

Respondents offered recommendations for future projects along five basic categories: 1) goals and objectives for the project, 2) outputs and results, 3) USPSTF interactions with modelers and/or reviewers, 4) project lead on the USPSTF, and 5) interactions between modeling and systematic review teams.

USPSTF’s goals and objectives for the modeling effort must be explicit and understood by both modelers and USPSTF leads. Project members must identify the areas within each question where modeling is anticipated to have the most impact and benefit, so that modelers can tune the analysis accordingly. Models have the greatest potential impact in determining the start, stop, and interval for different testing strategies, an essential USPSTF objective. Furthermore, the USPSTF can use models in key questions to assess the net benefit and magnitude of the effect/benefit. Only when these goals and objectives are specific, clear, and aligned, can the modeling effort deliver the necessary results and associated impact. Lack of clarity has been a problem. Modelers need to be very specific with USPSTF leads and the systematic review team as to the questions (or components thereof) that modeling can likely impact, and whether the evidence for a specific issue is sufficient to develop a valid model.

Outputs from a model must be designed in a purposeful and careful manner. From a model user’s perspective as opposed to a modeler’s perspective, the “outputs are the model,” and as such, outputs must be constructed to answer questions that inform and support decision making. One interviewee suggested that designing tables and figures before the start of the project would clarify goals and expectations while ensuring results of value for decision making. According to this thinking, not only should the modeling effort confirm an existing conclusion, but its outputs should also be directly usable by decision makers to inform and aid in the specific decision or recommendation of interest. Some respondents expressed concern that modeling efforts provide too much or too little information, and in some cases, fail to provide the information that decision makers need, thus leaving them to interpret or interpolate the results to address the recommendations.

Modelers prefer an iterative process that allows interim readouts of results with the USPSTF lead. One modeler commented that an “iterative process is a much better discipline for modelers, especially with complex questions ... interaction with the lead would have served us well, and allowed us to develop a better model.” Unclear communication could result in modeling results that are less informative and require additional analyses. Such an iterative process will “give the task force members, or at least the lead, more confidence in the model and a better ability to accurately use and communicate the model results.”

Two of the three projects reviewed mentioned that informed, model-literate leadership within the USPSTF was an essential component of success. In both cases, the modelers and the USPSTF lead reported a modeling project that impacted the USPSTF recommendations and allowed the task force to make either a “more detailed recommendation” or “to increase the certainty and/or the magnitude of the effects.” Modelers noted that these USPSTF leads were familiar with models and had used them in their professional experiences, and thus were able to “be much more specific and answer detailed questions about their request ... also they were able to challenge us on some of our logic.”

Many recommendations emerged from interactions among modeling groups and with systematic review teams. Interviewees consistently cited the Cancer Intervention and Surveillance Modeling Network as providing the right balance between frequent interactions and collaboration among the modeling teams and the need to maintain distinct and separate models demonstrating disparate representations of the disease. They also saw the Cancer Intervention and Surveillance Modeling Network structure as an advantage for building repositories of expertise in specific diseases.

Best Practices

Literature review

Of the 39 articles that provided guidance on what constitutes a good decision model, 7 discuss good modeling practices; 4 discuss the roles, uses, or value of modeling in general; 20 focus on specific aspects of modeling; 3 propose comprehensive guidelines for modeling in a specific clinical domain; and 5 review and compare models in specific clinical areas. Although the guidance for model users is extensive and fairly consistent, it can be vague. For example, model structure is recognized as important, yet we have no explicit guidelines for how to judge this. Furthermore, existing guidance focuses on the technical aspects of models, not the process or expertise required to conduct a modeling study. Nor does guidance address how best to illustrate and present models and modeling results or develop capacity to understand decision models and overcome the black box problem. In addition, modeling guidelines focus primarily on Markov models and less often on other types such as dynamic or discrete event models. Nor does much guidance exist for the optimal approach to choosing the type of model for a particular problem.

Modelers face many challenges as they seek to assist decision makers and improve the quality of decision making. In the context of medical tests, Trikalinos et al. [25] summarize these challenges and the various situations in which they surface. The challenges include 1) insufficient data on key input quantities (such as prevalence, test performance, and effectiveness); 2) the potential nontransferrability of performance across studies; 3) the choice of modeling outcomes (e.g., event-free survival, survival, and quality-adjusted life-years); 4) the methods for meta-
analysis; and 5) challenges in the parameterization and appraisal of complex models. This list echoes Tavakoli et al. [26], who also emphasize the task of identifying data as a major difficulty in developing decision models, specifically 1) epidemiological data on the risk of subsequent outcomes, 2) effectiveness data essential to estimate treatment benefits and harms as well as the probabilities of various outcomes given specific decisions over clinical pathways, and 3) health state valuation data necessary to estimate the utilities to be attached to specific outcomes. By definition, models are simplified representations of a real problem; thus, they are incomplete and inherently limited. They are useful, however, precisely for that reason. They promote transparency by pinpointing the influential constituents of each problem and by providing systematic uncertainty analysis to fully appreciate the impact of parameter estimates. To capitalize on the potential value of models, it is necessary to clearly identify and communicate their assumptions, challenges, and limitations.

**Expert modeler focus group**

Focus group experts structured the discussion of modeling within the context of a decision-making framework. They reiterated the results from the literature review regarding the quality of models, they further developed the need for interaction between the model and the decision maker(s) the model is intended to inform, and they elaborated on the importance of communicating the model and its results. The focus group described an essential need to discuss modeling within the proper context of a decision-making framework. They stressed that this clarifies the modeling’s main goal: To generate an unbiased synthesis of available evidence on the basis of clearly stated assumptions, and thus to make the decisions of individual(s) who must present complex but well-defined recommendations. The focus group noted that most individuals tasked with complex decisions, such as members of the USPSTF, value both the availability of a decision analytic framework and help from decision models. In the view of the focus group, the main issues regarding the acceptance of models and their results generally stem from stakeholders and broader communities who may or may not welcome or accept the resulting decision and/or recommendations.

The focus group identified the interface between the model and the decision maker as a neglected aspect of best practice for conducting models alongside systematic reviews. They discussed four key issues: 1) nature of evidence produced by models; 2) nature and extent of a decision maker’s involvement in the modeling effort; 3) transparency versus trust in the model, and 4) communication and visualization.

Not surprisingly, expert modelers agreed that models constitute “inferential” or “carefully manufactured” evidence not otherwise available and that models need to be incorporated along with other evidence generated through systematic reviews. The nature of the evidence generated may differ and require viewing through different lenses, but it provides otherwise inaccessible information to support decisions. Furthermore, one could argue that an implicit “mental model” is applied in reviewing and evaluating evidence in systematic reviews and that this mental model should also be made more explicit.

One problem related to the acceptance (and therefore the use and usefulness) of models is that models are first developed by a technical team that passes on the results to decision makers without interacting with them in advance of and during the process. This practice may be ineffective and lead to the wrong model being developed, misunderstanding of the model and its results, and low acceptance and use. In a framework of decision support, the development of a model should be a multidisciplinary effort involving clinical experts, modeling experts, and decision makers from inception to completion of the project. A modeling report generally has multiple audiences and therefore the model and its results must be carefully explained and understood by the relevant stakeholders. Structured interaction and involvement between decision makers and the modeling team would greatly enhance the understanding, acceptance, and use of models.

Respondents perceived transparency as somewhat of a paradox. Transparency is essential if peer expert modelers are to review and evaluate models; however, stakeholders generally do not want transparency despite their stated wish for it. Rather, stakeholders want to “trust” models. Because the majority of stakeholders and users are unfamiliar with modeling, transparency into the intricacies of a model offers little value and may even detract from building trust and acceptance. The ultimate test of how good a model is resides in its usefulness, its actual use, and the degree to which stakeholders understand and accept the model outputs. Increased public acceptance of models and their results will hinge on the contribution of individuals who can clearly and simply explain to lay audiences what a model is and does. Focus group members offered no specific recommendations on how to find or train such individuals, but they did stress the importance of building such expertise.

**Conclusions**

We used a variety of approaches to address important issues regarding the development and use of decision models alongside systematic reviews. We aimed to generate useful information for developing a framework for determining if and when a decision-analytic model might add value to a systematic review, and how to best accomplish such an endeavor. Overall, the results demonstrated that models are well suited to address gaps in the literature, better suited for certain types of research questions, and essential for determining the value of information relating to future research. While opinions differed regarding whether model outputs constitute evidence, interviewees expressed concern over the lack of standards and directions in grading and reporting such “evidence.” Interviews of stakeholders and modelers revealed the importance of communication and presentation of model results as well as the importance of model literacy and involvement of decision makers. We organized the information that emerged from our inquiries into a proposed framework as described in the full report [8].

**Acknowledgments**

The authors thank William Lawrence, MD, and Ava John-Baptiste, PhD, for their excellent insights throughout this project; Joseph M. Anderson, Eric Mann, and Jonah Popp for their contribution to the literature review; Jeannine Ouellette for her outstanding assistance with editing; and Molly Moore Jeffery for assistance with data entry and formatting tables. We also thank Mary B. Barton, MD, MPP, Eric J. Feuer, PhD, Douglas K. Owens, MD, MS, Diana Pettiti, MD, MPH, Steven M. Teutsch, MD, MPH, and Thomas A. Trikalinos, MD, PhD, for serving on the Technical Expert Panel.

Source of financial support: This project was funded under Contract No. 290-2007-10064-I, Task Order No.8, from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality, the U.S. Department of Health and Human Services, or the U.S. Department of Veterans Affairs.
REFERENCES


