





Assessing the Use of Agent-Based Models for Tobacco Regulation

ISBN
978-0-309-31722-1

280 pages
6 x 9
PAPERBACK (2015)

Robert Wallace, Amy Geller, V. Ayano Ogawa, Editors; Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation; Board on Population Health and Public Health Practice; Institute of Medicine

 Add book to cart

 Find similar titles

 Share this PDF



Visit the National Academies Press online and register for...

- ✓ Instant access to free PDF downloads of titles from the
 - NATIONAL ACADEMY OF SCIENCES
 - NATIONAL ACADEMY OF ENGINEERING
 - INSTITUTE OF MEDICINE
 - NATIONAL RESEARCH COUNCIL
- ✓ 10% off print titles
- ✓ Custom notification of new releases in your field of interest
- ✓ Special offers and discounts

Distribution, posting, or copying of this PDF is strictly prohibited without written permission of the National Academies Press. Unless otherwise indicated, all materials in this PDF are copyrighted by the National Academy of Sciences. Request reprint permission for this book

4

An Evaluation Framework for Policy-Relevant Agent-Based Models

Policy-relevant agent-based models (ABMs) are resource-intensive, complex technical activities that are developed by large groups of people with varying areas of expertise. The results of these models need to be translated and communicated to various stakeholders in order to affect policy and improve health. Policy-relevant ABMs need to be built carefully using appropriate data and social science theories, rigorously tested, and clearly communicated. These requirements for ABMs are the same as for other types of computational models and simulations used to inform policy decisions (e.g., aggregate models, system dynamics, and econometric forecasting models, to name a few).

Given the amount of time, effort, and money required to build an effective policy-relevant model, it is critical to evaluate the process, its outcomes, and its overall value. Was the model built such that its results represent what the modelers intended? Did the model address important and timely policy questions? Were the results useful in guiding subsequent policy and regulatory decision making and research? And, in the end, were the results worth the cost? These types of questions can be answered by *evaluating* the model building process, the model outcomes, and the return on investment.

The goal of this chapter is to present an evaluation framework for assessing ABMs for tobacco control policy and regulation. The committee found that no such framework exists for tobacco control and that such a framework is needed to assess complex computational modeling projects in a wide variety of public health policy and regulatory contexts. The committee developed this evaluation framework both to guide the committee in its review of the model developed for the U.S. Food and Drug Administration

(FDA) (see Chapter 5), and to provide FDA with guidance for future model development and evaluation.

An evaluation framework for ABMs can provide answers to two broad types of questions:

- a. *Process*—How was the model informed (by subject-matter experts, by data, and other inputs), planned, developed, and tested?
- b. *Outcomes*—In what ways did the modeling produce results that were useful for guiding future policy and regulatory efforts?

The remainder of this chapter provides a grammar for describing policy-relevant ABMs, presents an evaluation framework, identifies high-priority categories for evaluation questions, and illustrates some of the evaluation concepts through two case-study descriptions of existing policy-relevant ABMs.

MODEL DESCRIPTION

Before establishing a framework for evaluating how ABMs inform public health or tobacco control policy and regulation, a consistent way to describe and talk about them is needed. Although there have been some attempts at classifying ABMs to aid in model description (e.g., Marietto et al., 2003), these have tended to be too broad to capture the diversity of types of models that can be helpful for advancing policy and regulation or else too technical and not applicable to policy-relevant models (e.g., Grimm et al., 2006).

Table 4-1 presents a set of model descriptors that can be thought of as a *grammar* for describing in detail the structure and purpose of a policy-relevant ABM. This set of descriptors is not meant to replace a complete technical description of the model (sometimes called a “design document”). Instead, this gives a formal way to provide a rich description of the important elements of the model to be evaluated. Model evaluation requires a concise but thorough description of the model and what it was designed to accomplish.

The descriptors in Table 4-1 fall into seven broad categories: basic model description, model agents, use of data and theories, model context, model outcomes, policy aspects, and communication aspects. Within each of these categories is a small set of individual descriptors. For example, *physical space* is the indicator under *context* that describes whether and how the ABM depicts the physical space within which agents are allowed to move. Consistent use and application of these terms during model development will lead to better communication among the model developers and users of the model and will maximize the chance that the model meets the

TABLE 4-1 Grammar for Describing Agent-Based Models Relevant for Policy and Regulation

Descriptors	Definition
Basic Description	
Purpose (goal)	What is the main scientific, policy, or regulatory question that the model is addressing?
Breadth	What is the scope of the model? Is it designed to focus narrowly on one or a small number of social system components or processes, or is it designed to broadly encompass most or all parts of a complex system?
Abstraction	Is the model designed to be highly abstract, with the agents, rules, and context (i.e., physical and social spaces) not meant to precisely match real world settings, behavior, and processes, or is it designed with realism in mind?
Agents	
Agent type	Does the model include one type of agent, or multiple types (i.e., a multi-agent model)?
Agent definition	What are the agents in the model? For example, are the agents people or some other type of social agent (e.g., tobacco retailer)?
Data and Theories	
Data—rules	Are empirical data (quantitative or qualitative) used to inform the agent rules (e.g., smoking prevalence used to shape smoking initiation decision by an agent)?
Data—characteristics	Are empirical data used to inform the characteristics of the agents and environment?
Data—validation	Are empirical data used to validate model results?
Theories	What are the primary social science and behavioral theories used in the model design?
Context	
Physical space	Does the model include an explicit depiction of the physical space (e.g., built environment, geography) within which agents are allowed to move?
Social space	Does the model include an explicit depiction of the social space (i.e., connections or relationships among social entities such as people, communities, and organizations) that influences agent behavior or structures flow of information or other resources?
Physical dynamics	Is the physical space static or allowed to change as part of the model?
Social dynamics	Is the social space static or allowed to change as part of the model?

continued

TABLE 4-1 Continued

Descriptors	Definition
Outcomes	
Primary outcome	What is the primary outcome that is being modeled?
Proximal/distal outcome	Is the primary outcome a proximal or distal behavioral indicator? For example, reduction of smoking prevalence may be the ultimate goal of a policy that is being modeled, but the model may focus on addictive properties of new products or new restrictions on marketing. In these cases, these would be considered proximal outcomes.
Policy	
Policy definition	Description of the policy or policies that are being examined in the models.
Policy realism	How realistic are the policies being examined in the model? Are they reflective of actual policies that are being implemented, or do they reflect more abstract policy mechanisms or classes?
Policy tests	Does the model include formal tests of policy effects?
Communications	
Model sharing	What aspects of the model are (or will be) publicly available?

NOTE: The grammar in this table is meant to offer guidance on how to describe an ABM and is not meant to provide an evaluation of the quality of the model, which is something that is done later in the model development process.

needs of the model sponsor (Kuntz et al., 2013). See the chapter annex (see Table 4-2) for examples of how the descriptive grammar can be applied to three different policy-relevant ABMs.

EVALUATION FRAMEWORK

Fundamentally, systematic evaluations of policy-relevant ABMs are important because they can lead to better and more effective models in the future. A comprehensive evaluation provides useful information to at least four different groups involved with models:

- a. It helps the model *developers* improve their modeling efforts;
- b. It helps the *funders* understand better how to use model results and how to guide future funding of modeling work;
- c. It helps *policy makers* understand how to translate model results into more effective policies and increases their trust in the analysis; and
- d. It helps *modelers* and *scientists* by suggesting new avenues for research, modeling, and data collection.

Figure 4-1 presents an evaluation framework for policy-relevant ABMs that can be used specifically to evaluate models designed to inform tobacco control policy and regulation. The framework uses a logic-model approach, following the U.S. Centers for Disease Control and Prevention's program evaluation framework (CDC, 1999, 2007), and is based on best practices identified in a number of modeling fields. Although logic models have been used primarily to guide program evaluation, they are also useful for designing evaluations of policy development and implementation (Jordan, 2010; Langer et al., 2011) and larger systems evaluations (CDC, 2011; CORE, 2009). A logic model helps to ensure that all important aspects of the model building process are included in a systematic evaluation, and it guides the prioritization of evaluation questions. Logic models define the domains that are important for understanding the relevant processes and outcomes; however, actual evaluations based on the logic model will typically focus on a subset of domains (and associated evaluation questions) (CDC, 2007). The evaluation framework is not meant to be used as a checklist—each area deemed relevant to a particular model requires consideration on how each decision point will affect the model in the end.

The development of the evaluation framework was based on a review of relevant literature and on committee members' experience in building and assessing ABMs and in developing public health policy evaluations. The evaluation framework is designed to cover the important aspects of designing, implementing, testing, and disseminating policy-relevant ABMs, especially for tobacco control regulatory and policy efforts. It can be used to assess the model development processes as well as its outcomes. The framework has five major sections—resources, activities, outputs, outcomes, and environment—any of which can be the focus of a systematic evaluation.

Model development is an iterative process, and there are no clear divisions between the various evaluation steps (Berk, 2008), so there will be some overlap between the domains of the major sections (resources, activities, outputs, outcomes, and environment) of the logic model. It is often the case that a logic model will have boxes with the same or similar names across the columns (CDC, 2011). For example, there are policy activities that lead to policy outputs, which in turn influence policy outcomes. One important reason for using this structure is that the evaluation questions and timeline are quite different for early activities versus long-term outcomes.

Resources

A successful policy-relevant ABM is made up of a wide variety of critical ingredients. The domains listed in the resources section of the framework reflect the most important of these individual elements, which are the people, knowledge, infrastructure, and financial support neces-

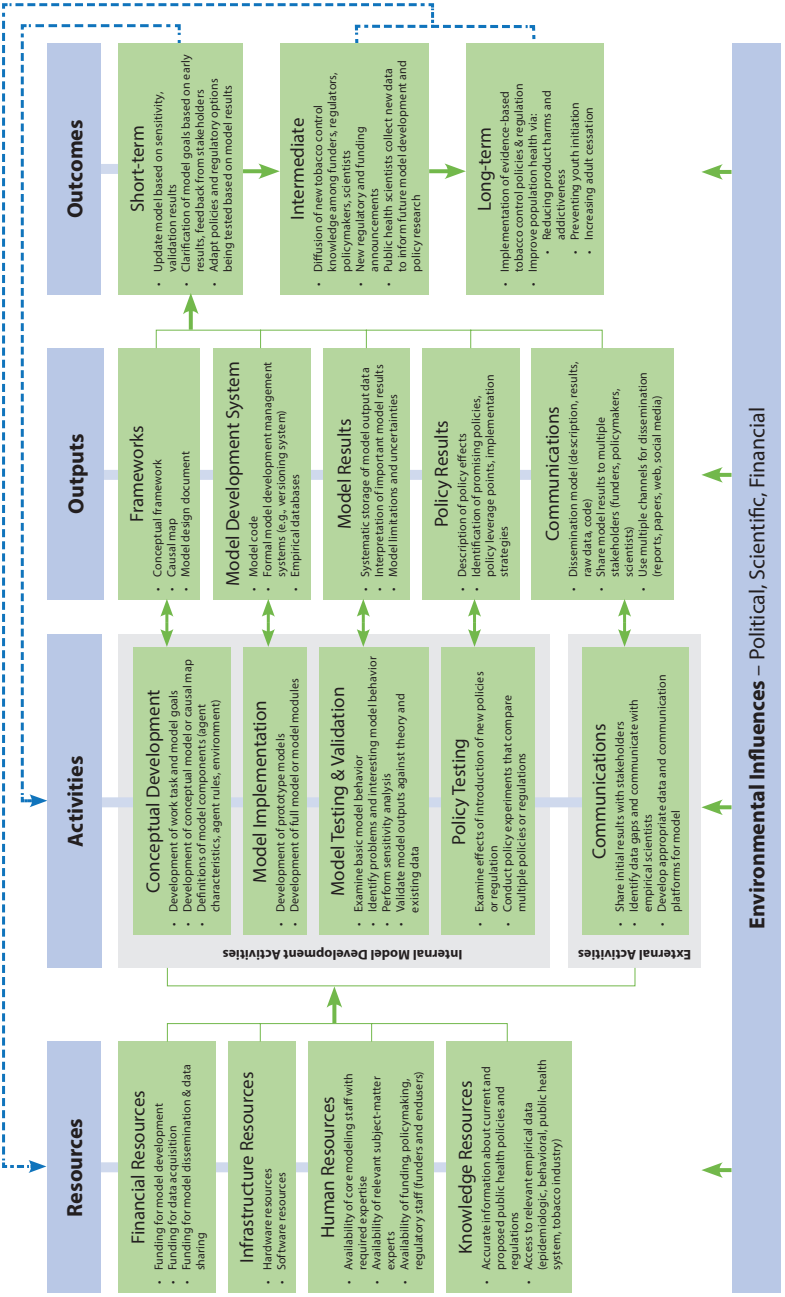


FIGURE 4-1 An evaluation framework for policy-relevant agent-based models.

sary for the successful development of an ABM. Typically, for example, a sophisticated modeling effort would not be started without the necessary funding and infrastructure already in place, but additional resources are often needed for data acquisition, model alterations, or an expanded dissemination of model findings, without any of which the model effectiveness could be diminished. Even more important for the successful development of a model are the necessary people and knowledge. Before model development begins, discussions with modelers from varied backgrounds (such as economics, engineering, and social sciences, among others) can help identify the best modeling approach (or combination of approaches) for the policy question under consideration (Roberts et al., 2012). These conversations can help ensure that an appropriate modeling strategy is chosen from the outset. In addition to the core modeling staff, a model must be informed by input from relevant subject-matter experts (Kuntz et al., 2013). For tobacco control policy and regulatory models, these may include experts from the social, clinical, and basic sciences. Another important group of people consists of the model funders and other policy makers who may use the results of the model to inform their policy and regulatory work. Finally, a modeling team will need access to a wide variety of knowledge resources, including relevant data, empirical findings, and current and proposed policies and regulatory options that may be addressed by the ABM.

Activities

The actions required to develop a policy-relevant ABM are listed under the activities section of the logic model. The first three boxes contain those activities that are common to any ABM, or almost any type of computational model. Model development starts with a conceptual phase, followed by implementing the model in code and then performing a series of validation tests. It is particularly important that the initial conceptual phase is included in an evaluation, yet this does not always happen:

Some of the most important model choices are made at the conceptual stage, yet most model evaluation activities tend to avoid a critical evaluation at this stage. Often a peer review panel will begin its efforts with the implicit acceptance of all the key assumptions made to establish the conceptual model and then devote all of its attention to the model building and model application stages. (NRC, 2007, p. 115)

The technical requirements for strong model development are many and varied, but the modeling community has developed a number of systematic approaches for describing, managing, and monitoring this development (Badham, 2010; Berk, 2008; Caro et al., 2012; CREM, 2009; Gurcan et al., 2011; Helbing and Balietti, 2011; NRC, 2007, 2012; Salamon, 2011;

Weinstein et al., 2003).¹ (See Appendix A for a broad overview of some of these requirements as well as the PARTE framework, which provides a helpful framework for defining model agents and context.) Some of these technical assessments can be useful for an evaluation of the model development process, but it is important to distinguish a more technical assessment of model validation, verification, and uncertainty quantification from a more general evaluation of the modeling processes and model outcomes. As also noted in a 2007 report by the National Research Council (NRC), the process of model evaluation is more than just a strict validation or verification procedure; it is a process that builds confidence in model applications and increases the understanding of model strengths and limitations. “Regulatory model evaluation must consider how accurately a particular model application represents the system of interest while being reproducible, transparent, and useful for the regulatory decision at hand” (NRC, 2007, p. 3). In essence, models should be evaluated with regard to their suitability as tools to address the question or process under study.

Iterative data collection throughout the model development process is often crucial in the development of effective policy-relevant models. In many cases, it can be difficult to predict what data may be needed to parameterize the model before—and even during—model development. When the model is implemented and outputs are generated, data gaps may be identified, signifying areas where data (whether available or not) is needed to inform a critical component of the model. Having identified these gaps, the model developers may revamp their model with newly integrated available data, or they may acknowledge the limitations of the data and encourage further data collection so that future models addressing a similar purpose can be more useful for the regulatory decision at hand. (See Case Examples later in this chapter for an illustration of this point.) Close and ongoing communication with subject-matter experts (detailed on the next page) can facilitate identifying these gaps early in the process.

Not all ABMs are intended to test the effects of different policies, but such policy testing is the *raison d’être* for policy-relevant models. Develop-

¹For example, Berk provides a six-stage evaluation process based on the work of Bayarri and colleagues (2007). These six stages are model specification, including the interactions between agents; the determination of model evaluation criteria, including calibration (fit) and use of visualizations; data collection of model inputs (for calibration) and “ground truth” test data (for model outputs), which refers to any data that capture the empirical process under investigation; construction of model approximations, e.g., statistical summaries of inputs and outputs and data reduction; an analysis of model output, e.g., search for obvious discrepancies between ground truth and model output; and the feedback of information to model development, while avoiding turning test data into calibration data.

ing informative policy-testing ABMs is challenging,² but once such an ABM has been developed, it can be used to explore the hypothesized effects of specific policy or regulatory interventions to reveal the possible mechanisms by which these policies operate or even to perform *in silico* policy experiments where different policy options are compared to one another (Auchincloss and Diez Roux, 2008; Hammond, 2015; Lempert, 2002). Finally, throughout all model building phases, it is necessary to perform a number of communication tasks. Funders and sponsors need to be kept apprised of progress, and content experts need to talk frequently with the model building staff to avoid making major errors related to essential model implementation decisions (Kuntz et al., 2013), such as programming agent behaviors, and to identify data gaps throughout model development. Because complex problems such as tobacco control require collaborative, interdisciplinary efforts—and thus varying backgrounds among the model-building stakeholders—communication will inevitably be interdisciplinary in nature. It can be a difficult process to get all team members on the same page early in the modeling process; there can be differences in the use of terminology, approaches, and strategy (Hovelynck et al., 2010; Nicolson et al., 2002). However, this process results in a stronger model in the end. It is often helpful to have as part of the team a translator or “knowledge broker” (Bammer, 2012; Bammer et al., 2010; Bielak et al., 2008; Dobbins et al., 2004; Meyer, 2010) who has a solid understanding both of the policy issues at hand and of modeling (while likely not actually being a modeler) and who can ensure that the information from the subject-matter experts is effectively translated into the model.

Outputs

Each of the model building activity domains has an associated set of products and outputs. For example, the outputs from the conceptual development stage may include causal maps, conceptual frameworks, and the general model design document. These outputs are often the primary subjects of process-focused evaluations. For example, if a model evaluation is focused on the validity of the agent behavioral rules incorporated in the model, the model design document will be an important source of information on how the agent behaviors were constructed. To ensure that the end users understand the scope of the model and properly apply the modeling results during development and upon completion of the model, the documentation of model’s limitations and uncertainties is an important output (Eddy et al., 2012).

²See Appendixes B and C for a comprehensive discussion on the challenges of developing informative policy-testing ABMs.

Outcomes

A variety of short- and longer-term outcomes is typically of interest when evaluating a policy-relevant ABM. The short-term outcomes include the immediate results of the model both as it is being developed and right after the model is finalized. The longer-term effects of a model include the diffusion of its results across a variety of stakeholder audiences (e.g., regulators, policy scientists, policy makers, tobacco content experts, and other modelers). Ultimately, models may result in changes in the policy and regulatory environments, shifts in funding priorities, changes in the types of data collected, implementation of new policies and regulations, and subsequent changes in the behavior of individuals, of the public health sector, and of organizations (e.g., tobacco companies). Although these longer-term effects are of obvious interest, by their very nature they take a long time to manifest themselves. In addition, whether model results are used by policy makers is largely out of the control of the model development team. Even when a model has useful outputs that are effectively quantified and communicated, the policy maker might not understand the value that the model has to offer (Kuntz et al., 2013), or unexpected changes in the environment can make the model results outdated (NRC, 2007). Thus, even in an outcome-focused ABM evaluation, long-term policy outcomes may not be explicitly included. However, this highlights the need for the policy maker to be involved with model development from its conception and the importance of translating the model results properly. Although this involvement will not guarantee that the model is used, it will increase the likelihood that the model addresses the current questions the policy makers are faced with and that they have a deep understanding of the value that the model offers (Wagner et al., 2010).

Environment

A variety of external environmental characteristics and forces might influence model development, either positively or negatively. For example, a shift in governmental policy priorities may make certain modeling efforts of greater interest to stakeholders. Alternatively, a change to federal, state, or local tobacco laws might affect tobacco use patterns, which would then need to be accounted for in a model. The rapid introduction of non-combustible tobacco products (e.g., e-cigarettes) in recent years, which is having dramatic effects on tobacco product purchasing and consumption as well as on industry and retailer behavior, is a good example of the sort of environmental change that can influence policy-relevant model development. Environmental factors are typically not the focus of an evaluation, but the entity conducting the evaluation needs to be aware of these influ-

ences so that the evaluation results and conclusions can be put into the appropriate context (Weinstein et al., 2001).

In addition to the individual domains listed in the logic model, the framework also focuses on some of the relationships between the domains, which illustrate the dynamic nature of ABM development. In particular, three types of feedback loops are present throughout model development and dissemination. First, the direct outputs of model development (i.e., conceptual frameworks and causal maps, model code, model testing and validation results, and policy testing results) provide feedback directly to the modeling team. In particular, the validation results invariably lead to modeling changes and improvements. (This feedback is depicted via double-headed arrows connecting the activities and outputs boxes.) Second, modeling results are typically disseminated in a number of ways, including via meetings with funders, reports, conference presentations, and peer-reviewed scientific papers. Immediate reactions to this dissemination can lead to further data collection, model development, or model expansion. For example, funders may ask the modeling team to consider new types of policy experiments or questions based on initial model results. Or new data may be made available that could be used to improve or expand the initial model. (This is depicted by the inner feedback loop connecting the short-term outcomes box to model activities.) Finally, intermediate and long-term outcomes constitute the types of policy, regulatory, and public health changes that were the goals of the modeling in the first place. These major changes in the policy and health landscapes will lead to completely new modeling efforts. (This feedback is depicted by the outer line connecting the intermediate and long-term outcomes boxes to future model resources.) And, of course, model development does not end here, as it is an iterative process. In particular, it is useful to think about taking a life-cycle approach to model development and testing (NRC, 2007).

Identifying High-Priority Evaluation Questions

The evaluation framework for policy-relevant ABMs provides a guide for designing an evaluation of a specific ABM project or a broader modeling initiative. Evaluations can be used to answer many questions, but they are most effective when there is a clearly stated purpose for the evaluation. For example, the main purpose of an ABM outcome evaluation may be to identify how the modeling results influenced subsequent policy and regulatory research. Once this overall purpose is decided on, the next task is to identify the set of specific questions that will be addressed in the evaluation. Despite the broad nature of the framework presented in this chapter, it would not be feasible to have an evaluation focus on every single domain. Instead, a short list of specific evaluation questions should be identified that

are most important to implement and to derive from the overall purpose of the evaluation. Typically, this prioritization process starts with a longer list of potential questions that can be linked to the logic model, and the list is then shortened by deciding which questions are most important. In this chapter's annex, the committee provides a list of example questions based on the evaluation framework that could be appropriate for evaluating an ABM project, especially in the context of tobacco control policy and regulatory science.

Fundamental Evaluation Categories

As the committee developed its evaluation framework, five fundamental evaluation categories for policy-relevant ABMs emerged. These are broader categories of relevant evaluation and assessment domains that the committee believes need to be included in most policy-relevant ABM evaluations. The five categories are listed below, as well as some sample questions for consideration (with a longer list of evaluation questions available in the chapter annex).

- a. **Resources:** A modeling team needs access to adequate financial, infrastructure, human, and knowledge resources to successfully design, build, and test its model.
 - To what extent were relevant staff available (e.g., funders, policy makers, end users) as the model was being built, especially in the conceptual development phase?
- b. **Technical best practices:** Model implementation, testing, and validation phases need to be reviewed throughout model development.
 - What kinds of analyses were done to quantify uncertainty?
 - How do the results compare to the results of other models addressing similar policy questions or having similar purpose?
- c. **Model suitability:** Models need to be developed in a manner that makes them suitable for their intended purpose and that will allow for exploration or testing of specific policy options or conditions. Some models could be developed for answering very narrow questions related to tobacco use, others as broader tools to look at a larger range of tobacco policies.
 - Does the goal of the model match the methods used and the assumptions made?
 - To what extent does the model capture the fundamental dynamics thought to be operating in the real world?
 - Does the model provide information that is helpful to making tobacco control policy?

- d. **Communication and translation:** Communication and translation strategies are essential during every stage of model development for enhancing the model building process and ensuring that the model is focused on the key issues that will affect policy outcomes. Modeling requirements, descriptions, and results need to be communicated effectively to a variety of audiences, including agency staff, regulators, politicians, and the general public.
- Does the model documentation include a write-up of model uncertainties, interpretations of results, and considerations for maintenance of the model?
 - How were preliminary results fed back into subsequent model improvements?
 - Were model processes and results communicated in a manner that allows for reproducibility?
 - If proprietary issues and requirements limited the communication of modeling information, were the costs and benefits of those limitations assessed or articulated?
- e. **Policy outcomes:** Ultimately, policy-relevant models will be used to inform policy and regulatory action or to advance scientific progress. Many of the likely evaluation questions in this category are not in the control of the model development team, and policy-related evaluation results do not necessarily reflect the quality of the model, but this reinforces the need for collaboration with policy makers from the onset.
- Was the model used to inform policy decisions? Did policies and regulatory options change in response to the model results?
 - How flexible is the model (i.e., capacity for the model to be modified or revised and applied to situations as new data arise or alternative objectives are specified)? What factors might trigger the need for major revisions, or what circumstances might prompt users to seek an alternative model?
 - How has the sponsor (e.g., FDA) used model results to inform its own regulatory activities?
 - How relevant are the modeling results to the tobacco control field? Have the results informed tobacco control knowledge and influenced decisions among funders, regulators, policy makers, scientists?

Recommendation 4-1: The Center for Tobacco Products should adopt an evaluation framework for its modeling work, either the one presented in this chapter or one similar to it. Key dimensions of the evaluation framework should include considerations of resources, technical best practices, model suitability, communication and translation, and

policy outcomes. The evaluation plan should be designed early in the model development process and should be carried out throughout model development.

This evaluation framework would apply to all efforts funded by the Center for Tobacco Products (CTP)—internal model development, inter-agency agreements, contracts, and grants. In addition to internal CTP reviewers, external experts need to be part of the evaluation process (see NRC, 2007, for guidance on the peer review of models).³ If CTP chooses to adopt the evaluation framework developed by the committee, the framework should be used as a guideline and not as a mechanical exercise or checklist, because different ABMs will require differing evaluation strategies based on intended use, modeling approach, and other aspects of model development.

CASE EXAMPLES

In this section, the committee explores published models that illustrate many of the areas outlined in the evaluation framework. These examples cover subjects from two different areas: transportation and illicit drugs. The committee chose these examples because they illustrate several of the important aspects of model development discussed in this chapter; however, the committee did not formally review or assess the overall strengths or weaknesses of the models. It is difficult to provide examples of all of the

³Guidance on peer review can be found in NRC, 2007. Options for receiving external review include contracts, special government appointees, and advisory panels.

Peer review should be considered, but not necessarily performed, at each stage in a model's life cycle. Some simple, uncontroversial models might not require any peer review, whereas others might merit peer review at several stages. Appropriate peer review requires an effort commensurate with the complexity and significance of the model application. When a model peer review is undertaken, EPA should allow sufficient time, resources, and structure to assure an adequate review. Reviewers should receive not only copies of the model and its documentation but also documentation of its origin and history. Peer review for some regulatory models should involve comparing the model results with known test cases, reviewing the model code and documentation, and running the model for several types of problems for which the model might be used. Reviewing model documentation and results is not sufficient peer review for many regulatory models. Because many stakeholders and others interested in the regulatory process do not have the capability or resources for a scientific peer review, they need to be able to have confidence in the evaluation process. This need requires a transparent peer review process and continued adherence to criteria provided in EPA's guidance on peer review. Documentation of all peer reviews, as well as evidence of the agency's consideration of comments in developing revisions, should be part of the model origin and history. (NRC, 2007, pp. 5–6)

elements from the evaluation framework because these activities are often not documented when a model is published (such as those that fall in the resources category). The two examples below illustrate a range of the elements in the framework.

Agent-Based Model of Potential Plug-in Hybrid Electric Vehicles (PHEVs) Market Adoption

Eppstein et al. (2011) describes an ABM of the potential market adoption of plug-in hybrid electric vehicles (PHEVs) that features spatial, social, and media influences. The model's purpose is to inform manufacturers and policy makers on the prioritization of investments toward potential leverage points and to identify combinations of policies that may be the most effective for PHEV market penetration.⁴ In developing the ABM, however, Eppstein et al. recognized the need for additional data to inform the model. Thus, to strengthen the model, the developers conducted an extensive survey to gather and integrate data to the ABM (Krupa et al., 2014). As the researchers reworked the model with the new data, they generated results that could provide better insights for policy makers and manufacturers into the factors influencing the potential for PHEV market penetration (Eppstein et al., under review). Below is a detailed description of the process the researchers used in developing their ABM.

The original model by Eppstein et al. (2011) included agents who are individual vehicle consumers restricted to certain attributes.⁵ When agents make decisions to buy a car in the model, they compare the relative costs and benefits of all pairs of vehicles and fuel types and then choose the most desirable vehicle. While agents think about their decisions, they may be susceptible to media and social influences. To put these agent attributes and decision rules into the model and determine their cross-correlations, Eppstein et al. (2011) drew on available data as well as on social science theories⁶ and relevant literature.

The original model generated several findings relevant to PHEV

⁴Specifically, Eppstein et al. (2011) examined the effects of the following: gas prices; the ability of agents to consider fuel costs, PHEV purchase price, and rebates; PHEV all-electric battery range; consumer values regarding financial versus nonfinancial concerns in vehicle purchase; agent comfort threshold with the PHEV technology; social and media influence on PHEV market penetration; and fuel efficiency of the resulting fleet after 25 years.

⁵Agent attributes included age, annual salary, residential location, typical years of car ownership, annual vehicle miles traveled, vehicle age, fuel type, and fuel economy of current vehicle.

⁶The theories they used included threshold effects (Granovetter, 1978), homophily (McPherson et al., 2001), and conformity (Axelrod, 1997).

adoption.⁷ Eppstein et al. (2011) also discussed a number of limitations of the ABM, including the lack of data for accurate parameterization and model realism. The modelers used data where possible to initialize agent attributes and simulations; where they did not have data, they tried to make reasonable yet simplifying assumptions. For example, the developers made many assumptions on spatial and inter-attribute cross-correlations and distributions, such as estimating the mean and standard deviation of the threshold distribution for new PHEV technology consumer adoption. These assumptions may not have necessarily been realistic, and they could have significantly affected model outcomes. Eppstein and colleagues did not claim that their model provided accurate quantitative predictions, but they stated that the findings offered preliminary insights into the combinations of policies and procedures that may be most effective for PHEV market penetration.

In order to provide more accurate parameterization and model realism, Eppstein et al. collected relevant quantitative data by administering an extensive consumer survey (Krupa et al., 2014). Each survey respondent corresponded one-to-one with an agent in the model so that each agent's attitudes and attributes, such as demographic information and susceptibility to social and media influences, were based on a real person. In this way, Eppstein et al. could populate the model with realistic (instead of assumed) distributions and cross-correlations of agent attributes. The survey included questions on different aspects of potential PHEV adoption barriers and attempted to fill in the holes left from the original model. Based on the analyses of the survey questions (Krupa et al., 2014), Eppstein et al. inserted agent vehicle purchasing decision rules in the model (Eppstein et al., under review). Data from the surveys revealed that many of the cross-correlations and estimates used in the original model, such as the standard deviation of the threshold distribution, were not accurate. The model developers continued to use some assumptions (e.g., rules for social network updates) in the modified model, but now they were equipped with more data, which resulted in different implementation decisions. Consequently, the updated model generated results that differed slightly from those of the original model.⁸

⁷Some of the findings included if there are sufficient potential early adopters, readily accessible estimates of lifetime vehicle fuel costs could be important for promoting PHEV market penetration; increasing gas costs could help people choose PHEV over traditional vehicles; temporal incentive programs like tax credits are not likely to have lasting effects on long-term fuel efficiency unless manufacturers are able to lower sticker prices after the rebates are discontinued; and increasing PHEV battery range may be an important leverage point.

⁸The results of the modified model indicated, among other things, that consumer uneasiness with the new PHEV technology was the biggest barrier to potential PHEV market penetration; that manufacturers and policy makers may need to take more action to help consumers feel

This model illustrates several of the elements outlined in the evaluation framework. Although the committee does not have information about the human or infrastructure resources for this model, the authors did strive to develop a model with the intended users—i.e., policy makers—in mind. The developers grounded their assumptions in theories during the conceptual phase of development. Although, as discussed, the model still contains assumptions, Eppstein and colleagues quantified and communicated the uncertainties and limitations of the model, provided additional data to better ground the model after the initial iteration of the model was completed, and incrementally developed the model, taking a key step in providing better insight into factors influencing the potential for PHEV market penetration. Although the initial model design did not properly represent the agents' behaviors, the authors made needed adjustments to improve the model for its intended purpose (exemplifying the necessary feedbacks in the evaluation framework presented in this chapter). The authors were clear on how the results of the model could be interpreted by policy makers, and where more information was needed.

SimAmph

A group of Australian researchers developed an ABM to study how individual perceptions, peer influence, and subcultural settings shape the use of psychostimulants and related harms among young Australians (Dray et al., 2012; Moore et al., 2009; Perez et al., 2012). The team was composed of modelers as well as experts in epidemiology, anthropology, economics, and drug policy. Within an interdisciplinary team, the researchers focused on collective design and incremental development of the model to address the study question. The team developed an ABM called SimAmph that iteratively integrated ethno-epidemiological data (Moore et al., 2009).

The researchers conducted both ethnographic and epidemiological studies simultaneously in three research sites, led by the appropriate experts on the team.⁹ When developing the model, the ethnographers and epidemi-

more at ease with the new technology, whether it is through advertisements or well-publicized incentives; that many consumers choose used cars instead of new cars, whereas PHEVs are not likely to become part of an extensive used-car market anytime soon; that consumers may not feel limited when PHEVs are offered as only compact cars; that increases in gasoline prices may lead to small effects on PHEV market penetration (a finding that was contrary to the results from the first model); and that governmental and manufacturer rebates may allow PHEVs to be more competitive, but because many consumers may not know of the rebates, the rebates need to be more available until the prices of PHEVs decrease.

⁹The findings of the ethno-epidemiological data drawn from participant observation and in-depth interviews, and two surveys have been published (Green and Moore, 2009; Jenkinson et al., 2014; Siokou and Moore, 2008; Siokou et al., 2010).

ologists advised on the input data and the conceptual underpinnings of the ABM based on the findings of their studies, and the modelers asked questions, reworked the model, and conducted partial verification at each stage in the process. In addition, the team used secondary sources from national drug surveys as well as other qualitative research on similar populations to complement the findings of the ethno-epidemiological research and to further develop the model.

From these various sources, the researchers found that the use of psychostimulants among young Australians occurred mostly in the context of weekend partying and poly-drug use at licensed and other leisure venues. The researchers also learned that many young Australians were influenced by social relationships and the settings in which drug use took place. Using these findings, the researchers developed a model that included agents (young people) with particular attributes (e.g., socio-demographic characteristics, peer relationships) in various social settings who are able to access different types of drugs, have a set of friends whom they can exchange information with, such as drug experiences, and use drugs variably, depending on time and circumstance. The researchers set up rules, specifically concerning peer influence and health experience, that were designed to capture the dynamic process of the agents' use of psychostimulants (see Moore et al., 2009, for more details about the model). Over many iterations of model development, the researchers produced an ABM that could run such policy scenarios as the impact of pill testing (Moore et al., 2009) and the use of drug detection dogs by police and the dissemination of mass media prevention campaign (Dray et al., 2012).

SimAmph provides a good example of several of the criteria laid out in the evaluation framework. Having an interdisciplinary team in place from the outset allowed the researchers to explore many angles of the research question. Although SimAmph is simple and has several limitations,¹⁰ the researchers integrated (or considered) concepts and data from relevant disciplines to capture and adequately justify the conceptual basis and inputs of the model while acknowledging the model's shortcomings. The team faced tensions brought on by the existence of multiple epistemologies rooted in different disciplines, but with ongoing, open dialogue throughout the project, the team was able to produce a model that integrates triangulated data and that begins to encapsulate and promote discussions concerning the complexity of drug use and policy (Moore et al., 2009). This type of interaction, which is highlighted in the evaluation framework, can help build a

¹⁰For example, the simulation was in a closed system that simplified a more complex reality of transient movements among individuals in drug scenes. For a comprehensive list, refer to page 70 of Perez et al. (2012).

strong conceptual framework for a model and increase the likelihood that the model will meet its intended purpose.

Several of the authors of SimAmph are part of the Drug Policy Modeling Program (DPMP), which created a series of models, including four ABMs (SimARC, SimDrug, SimDrugPolicing, and SimHero), that were designed to examine the effects of drug policies.¹¹ DPMP is tasked with generating new research evidence, translating evidence for policy makers, and studying how policy is made with teams that span many disciplines.¹² These goals are incorporated in the ABMs they have created. The team consults with policy makers to improve their use of the models and research. Although the model documentation does not include information on the financial resources available to DPMP, it is evident that input from an array of disciplines was considered and that the researchers sought critical human and knowledge resources during the course of model development. Because of the policy focus of DPMP, the researchers work with policy makers to ensure that the model is suitable for their purposes, and they regularly assign a “knowledge broker” to translate model findings into policy language and communicate the limitations of the modeled scenarios as well as the predictive ability of the model to the policy makers (MacDonald, 2012).¹³ Because DPMP aims to ensure that modelers understand the needs of the model they are developing and to make certain that the models are used properly by policy makers, communication and translation strategies are considered throughout model development.

¹¹ See <http://dpmp.unsw.edu.au/resource/models>.

¹² These disciplines include complex systems science, criminology, economics, epidemiology, integration and implementation sciences, law, medicine, political science, psychology, public health, public policy, sociology, and systems thinking.

¹³ Personal communication, P. Perez, A. Ritter, and Institute of Medicine staff, April 15, 2014.

Chapter 4 Annex

EXAMPLE OF APPLYING GRAMMAR TO DESCRIBE AGENT-BASED MODELS

The following Table 4-2 illustrates how the descriptive grammar presented in Table 4-1 could be applied to existing policy-relevant models. The grammar is meant to be descriptive only—it is not an evaluation of a model, but rather a systematic way to describe ABMs early in the model development process. The use of the grammar will improve communication between the model development team and the policy makers and help ensure that they are all in agreement about the goals and intended uses of the model. The models listed in Table 4-2 are described more fully in Chapter 4 (PHEV Market Adoption and SimAmph) and Chapter 5 (SnapDragon).

TABLE 4-2 Application of Descriptive Grammar to Three Policy-Relevant ABMs

	Models		
	PHEV Market Adoption ^a	SimAmph ^b	SnapDragon ^c
	Basic Description		
Purpose	To inform policies affecting plug-in hybrid vehicle market penetration.	To test policies that could influence drug use and experience among young Australians.	To study the effects of tobacco control policies in a single- or multiple-tobacco product environment.
Breadth	Moderately broad	Very broad	Moderately narrow
Abstraction	Moderately realistic	Moderately realistic	Moderately abstract
	Agents		
Type	Single type	Single type	Single type
Definition	Agents are consumers who make decisions about which vehicles to purchase	Agents are Australian youth who make decisions about drug use based on psychological and health status and social interactions	Agents are generic persons who have opinions about tobacco products and also tobacco use behaviors

TABLE 4-2 Continued

	Models		
	PHEV Market Adoption ^a	SimAmph ^b	SnapDragon ^c
	Data and Theories		
Data—rules	Yes	Yes	No
Data—characteristics	Yes	Yes	Yes
Data—validation	No—data for validation not available	Yes—validated with data from the 2004 National Drug Strategy Household Survey	Very simple validation using social network data ^d
Theories	Social threshold effects, social science theories (principles of homophily and conformity)	Broad set of social science theories; developed ethnographic framework, Stage of Social Engagement	Opinion dynamics
	Context		
Physical space	Abstract	Abstract	None
Social space	Simple	Simple	Simple
Physical dynamics	Static	Static	None
Social dynamics	Static	Static	Static (at the time of committee review)
	Outcomes		
Primary outcome	Fleet fuel efficiency resulting from agent vehicle purchase choices	Individual drug use and population prevalence of drug-related harm and of regular drug use	User or nonuser of tobacco products
Proximal/distal outcome	Distal	Proximal/Distal	Proximal/Distal
	Policy		
Policy definition	Effects of purchase rebates ^e	Effects of mass media drug prevention campaigns; effects of using drug-sniffing dogs	Introduction of non-specific communications campaign; introduction of new products
Policy realism	Realistic	Realistic	Abstract
Policy tests	Yes (although not a primary goal of study)	Yes	No (at the time of committee review)

continued

TABLE 4-2 Continued

	Models		
	PHEV Market Adoption ^a	SimAmph ^b	SnapDragon ^c
	Communications		
Model sharing	Collated results in the form of peer-reviewed papers and presentations.	Collated results in the form of peer-reviewed papers and presentations. Model code and documentation are available on a website. ^f	Some preliminary results have been presented at professional meetings; other aspects of the modeling process and outcomes have been presented to FDA; manuscripts have been submitted for publication.

^aSources: Eppstein et al., 2011; Krupa et al., 2014.

^bSources: Dray et al., 2012; Moore et al., 2009; Perez et al., 2012.

^cSources: Moore et al., in press a,b.

^dThese data were collected as part of NIH/NCI grant 3R01CA157577-02S1 (Extending a School-Based Cohort to Improve Longitudinal Modeling), Thomas W. Valente, principal investigator. This data collection was a follow-up to the Social Network Study cohort in 2010 through 2012 (Valente et al., 2013). The data are not yet published.

^eEppstein et al. (2011) did not identify specific policies to test from the beginning, but rather used the model to find key leverage points—that is, specific model parameters that, if changed, affected PHEV technology adoption—and then identified examples of potential government influence on the model parameters, through the form of a targeted policy. In addition to purchase rebates, other potential policy examples include gasoline taxes, tax breaks or other manufacturer incentives to keep PHEV sticker prices low, and public service announcements to educate consumers, among others.

^fA version of SimAmph (and relevant documentation) is available at: <http://cormas.cirad.fr/en/applica/simAmph.htm>.

EVALUATION QUESTIONS DERIVED FROM THE EVALUATION FRAMEWORK FOR POLICY RELEVANT AGENT-BASED MODELS

Based on the evaluation framework presented on page 92, this document contains sample questions for each of the categories outlined. Although many of these questions would be of interest to any modeling effort, some questions are specifically applicable for ABMs, and many are geared toward informing models specific to tobacco control policies. The questions are intended for modelers, subject-matter experts, funders, policy makers, and other relevant collaborative members involved with developing or using the model. Before modeling begins, it is suggested that these actors select a reasonable number (e.g., three to five) of high-priority evaluation categories from the framework, develop questions within each category (potentially

adapting the sample questions below), and build a tailored evaluation plan. If done properly, during and after model development, the evaluators (including independent third-party evaluators) would be able to understand the purpose of the model and apply the evaluation plan. Thus, the framework and associated questions are not meant to be used as a checklist but rather as a general guide that may help in determining if the model has fulfilled its objective. These questions do not reflect an evaluation of an actual ABM; however, many of these questions were considered by the committee as they reviewed the ABM developed for FDA (see Chapter 5). The questions are drawn from existing sources (ASPE, 2012; Caro et al., 2012; CREM, 2009; Grimm et al., 2006; Gurcan et al., 2011; Hammond, 2015; Kopec et al., 2010; Kuntz et al., 2013; NRC, 1991, 2007, 2012; Rochester, 2014; Šalamon, 2011; Wagner et al., 2010; Weinstein et al., 2001, 2003) as well as from committee expertise.

1. Resources

- a. Financial
 - i. Were the model development, data acquisition, and model dissemination and data sharing funded at a level commensurate with the scope of the model?
 - ii. Did the model developers have the required financial resources to reach the needs of the end users of the model?
- b. Infrastructure
 - i. What hardware resources did the model developers use?
 - ii. What software resources did the model developers use?
- c. Human
 - i. Did the modeling team use an interdisciplinary team or approach when building and testing the model?
 - ii. How were subject-matter experts involved (or not involved) in the model development?
 - iii. To what extent were relevant staff and stakeholders available (e.g., funders, policy makers, and end users) when building the model, especially in the conceptual development phase?
- d. Knowledge
 - i. Were the specific policy or regulatory goals of the modeling project clearly described before model development began?
 - ii. To what extent did the modeling team use (or at least take into account) the relevant studies and principles and frameworks in the area, not just knowledge of their own approach? That is, would all or some aspects of another approach be better suited to address the policy question or goal?

- iii. What types of decisions could the model support (e.g., strategic planning, compliance, enforcement)?
- iv. What kinds of data are available to support the model (e.g., epidemiologic, behavioral, public health system, tobacco industry)?

2. Activities

Internal Model Development Activities

- a. Conceptual
 - i. Why was the modeling method chosen (versus other approaches)? Were there other modeling methods that could have been used instead of or in tandem with this method?
 - ii. How did the particular theoretical framework enhance or weaken the validity of the model results?
 - iii. Was the level of abstraction employed in the model well justified, and did it match up well with the specific policies being examined?
 - iv. Was a rationale presented for the overall scope and timeline of the model?
 - v. What are the definitions of the major model components (e.g., agent characteristics, agent rules, environment, initiation, cessation, addiction, relapse)?
 - vi. Did the model developers use appropriate theories to inform agent characteristics and interactions?
- b. Model Implementation
 - i. Did the model developers make full use of existing, relevant datasets? When empirical data were lacking, how was this accounted for in the model?
 - ii. How are the assumptions supported (e.g., empirical evidence)?
 - iii. Are social networks important for the specific model application? If so, were the social network structures and processes too simple (or too complex) for the model?
 - iv. What kind of heterogeneity was captured? Did the model capture too little or too much?
 - v. What temporal and spatial scales were used in the model, and were they appropriate for the presumed behaviors of the policies and agents?
 - vi. What algorithms or mathematical methods are used in the model and how were they derived?
 - vii. Were various evolving environmental scenarios, not just the status quo and past trends, considered in the model? What features were held constant?

- viii. Is the model unreasonably complicated? (Are there, for example, too many parameters that increase model uncertainty? Did the modeling team consider trade-offs between the need for the model to be an accurate representation of the system of interest and the need for it to be reproducible, transparent, and useful for the regulatory decision?)
- c. Model Testing
 - i. What kinds of analyses were performed to quantify uncertainty?
 - ii. Was the model output compared to empirical outputs under some specified time frame to ensure that the model captures real-world dynamics?
 - iii. What problems and interesting or surprising model behaviors were identified, and how did the modeling team handle them?
 - iv. How do the results compare to the results of other models addressing similar policy questions or having similar purpose?
 - v. Do the results conform to or conflict with other relevant evidence and face validity?
 - vi. How appropriate are the verification, validation, and calibration techniques used in the model?
- d. Policy Testing
 - i. How were the specific policies or processes operationalized within the modeling framework?
 - ii. Were policies examined in isolation, or were multiple policies modeled and allowed to interact?

External Activities

- e. Communications
 - i. Were relevant stakeholders included in all aspects of the model development, or just at the end?
 - ii. How were initial results shared with the stakeholders?
 - iii. Were appropriate data and communication platforms developed for the model?
 - iv. Were the model processes and results communicated in a manner that allows for reproducibility?
 - v. If proprietary issues and requirements limited the communication of modeling information, were the costs and benefits of those limitations assessed or articulated?

- f. Peer Review
 - i. At what stages of the model development did the modelers seek peer review? What did the peer review involve (e.g., reviewing the conceptual framework of the model, running the model several times, comparing the model's results with known test cases, reviewing the model code)? How did the modelers incorporate the feedback into the model? Is there documentation of this?
3. **Outputs**
- a. Frameworks
 - i. How does the model design documentation describe all of the important details of the model implementation and testing process?
 - ii. Does the model documentation include a write-up of model uncertainties, an interpretation of results, and considerations for maintenance of the model?
 - iii. Did the authors provide a conceptual framework and causal map (this would be developed during the conceptual phase of model development)?
 - iv. Did the authors clearly discuss the model's strengths and weaknesses and implications for tobacco control policy?
 - b. Development/Software Versioning System
 - i. How did the modeling team use a management system to enhance model development? ("Management systems" are needed when building a model that requires a complicated software program.) How was progress documented?
 - ii. Did the authors publish the model code and empirical databases?
 - c. Model Results
 - i. What kinds of results were generated (e.g., morbidity, mortality, prevalence, DALYs [disability-adjusted life years])?
 - ii. To what extent can the model address short-term, intermediate, and long-term effects?
 - iii. Did the authors provide for a systematic storage of model output data?
 - d. Policy Results
 - i. How useful are the model results for informing or setting priorities of future policy or regulatory activity (e.g., identification of promising policies, policy leverage points, implementation strategies)?
 - ii. Does the model fulfill its designated task (i.e., address the specified policy goal(s))?

- iii. How are the policy results translated and interpreted? Can the various audiences understand the model results, strengths, and limitations?
 - e. Communications
 - i. Were the relevant stakeholders included in dissemination activities?
 - ii. What kinds of multimedia platforms were used for dissemination?
 - iii. Did contract restrictions or proprietary concerns inhibit dissemination?
 - iv. Was a dissemination plan discussed with the funders?
 - v. Were the model details and results clearly described? How accessible is the model?
- 4. Outcomes**
- a. Short-term
 - i. How were preliminary results fed back into subsequent model improvements?
 - ii. Based on model results, did policies and regulatory options change?
 - iii. Who is going to use the model? How will it be applied? Do the end users have the expertise needed for using the model, or will they always need to partner with a contractor to use it?
 - iv. How can this type of model be used to inform other models—for example, aggregate (compartment) models?
 - v. How flexible is the model (i.e., capacity for the model to be modified or revised and applied to situations as new data arise or alternative objectives are specified)? What factors might trigger the need for major revisions, or what circumstances might prompt users to seek an alternative model?
 - b. Medium-term
 - i. What was the return on investment for the modeling efforts? Are the results justified, given the amount of money invested and the amount of time taken to develop, test, and disseminate the model?
 - ii. How has the sponsor (e.g., FDA) used the model's results to inform its own regulatory activities? Did the results help shape new regulatory and funding announcements?

- iii. How relevant are the modeling results to the tobacco control field? Have the results informed tobacco control knowledge and influenced decisions among funders, regulators, policy makers, scientists?
- iv. Have public health scientists collected new data to inform future model development and policy research?
- c. Long-term
 - i. How has the sponsor or other stakeholders used the model to implement evidence-based tobacco control policies and regulation?
 - ii. How has the sponsor or other stakeholders used the model to improve population health via reducing product harms and addictiveness, preventing youth initiation, or increasing adult cessation?
 - iii. Did the model inform new promising avenues of research, study, or exploration?

REFERENCES

- ASPE (Office of the Assistant Secretary for Planning and Evaluation). 2012. *Demystifying microsimulation meeting report*. Washington, DC: Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services.
- Auchincloss, A. H., and A. V. Diez Roux. 2008. A new tool for epidemiology: The usefulness of dynamic-agent models in understanding place effects on health. *American Journal of Epidemiology* 168(1):1–8.
- Axelrod, R. 1997. The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution* 41(2):203–226.
- Badham, J. 2010. A compendium of modelling techniques. *Integration Insights* #12. <http://i2s.anu.edu.au> (accessed May 5, 2014).
- Bammer, G. 2012. *Disciplining interdisciplinarity: Integration and implementation sciences for researching complex real-world problems*. Canberra, Australia: Australian National University Press.
- Bammer, G., A. Michaux, and A. Sanson. 2010. *Bridging the “know-do” gap: Knowledge brokering to improve child wellbeing*. Canberra, Australia: Australian National University Press.
- Bayarri, M. J., J. O. Berger, R. Paulo, J. Sacks, J. A. Cafeo, J. Cavendish, C.-H. Lin, and J. Tu. 2007. A framework for validation of computer models. *Technometrics* 49(2):138–154.
- Berk, R. 2008. How you can tell if the simulations in computational criminology are any good. *Journal of Experimental Criminology* 4(3):289–308.
- Bielak, A., A. Campbell, S. Pope, K. Schaefer, and L. Shaxson. 2008. From science communication to knowledge brokering: The shift from “science push” to “policy pull.” In *Communicating science in social contexts*, edited by D. Cheng, M. Claessens, T. Gascoigne, J. Metcalfe, B. Schiele and S. Shi. Netherlands: Springer. Pp. 201–226.
- Caro, J. J., A. H. Briggs, U. Siebert, and K. M. Kuntz. 2012. Modeling good research practices—overview: A report of the ISPOR-SMDM Modeling Good Research Practices Task Force—1. *Value in Health* 15(6):796–803.

- CDC (Centers for Disease Control and Prevention). 1999. Framework for program evaluation in public health. *Morbidity and Mortality Weekly Report: Recommendations and Reports* 48(RR11):1–40.
- . 2007. *Evaluation guide: Developing and using a logic model*. http://www.cdc.gov/dhdsp/programs/spha/evaluation_guides/logic_model.htm (accessed May 5, 2014)
- . 2011. *Introduction to program evaluation for public health programs: A self-study guide*. Atlanta, GA: Centers for Disease Control and Prevention.
- CORE (Cornell Office for Research on Evaluation). 2009. *The evaluation facilitator's guide to systems evaluation protocol*. Ithaca, NY: Cornell Digital Print Services.
- CREM (Council for Regulatory Environmental Modeling). 2009. *Guidance on the development, evaluation, and application of environmental models*. Washington, DC: Office of the Science Advisor, Council for Regulatory Environmental Modeling, U.S. Environmental Protection Agency.
- Dobbins, M., K. DeCorby, and T. Twiddy. 2004. A knowledge transfer strategy for public health decision makers. *Worldviews on Evidence-Based Nursing* 1(2):120–128.
- Dray, A., P. Perez, D. Moore, P. Dietze, G. Bammer, R. Jenkinson, C. Siokou, R. Green, S. L. Hudson, and L. Maher. 2012. Are drug detection dogs and mass-media campaigns likely to be effective policy responses to psychostimulant use and related harm? Results from an agent-based simulation model. *International Journal of Drug Policy* 23(2):148–153.
- Eddy, D. M., W. Hollingworth, J. J. Caro, J. Tsevat, K. M. McDonald, and J. B. Wong. 2012. Model transparency and validation: A report of the ISPOR-SMDM Modeling Good Research Practices Task Force—7. *Value Health* 15(6):843–853.
- Eppstein, M. J., D. K. Grover, J. S. Marshall, and D. M. Rizzo. 2011. An agent-based model to study market penetration of plug-in hybrid electric vehicles. *Energy Policy* 39(6):3789–3802.
- Eppstein, M. J., D. M. Rizzo, B. H. Y. Lee, J. S. Krupa, and N. Manukyan. Under review. National survey respondents as agents in a model of plug-in hybrid electric vehicle adoption.
- Granovetter, M. 1978. Threshold models of collective behavior. *American Journal of Sociology* 83(6):1420–1443.
- Green, R., and D. Moore. 2009. “Kiddie drugs” and controlled pleasure: Recreational use of dexamphetamine in a social network of young Australians. *International Journal of Drug Policy* 20(5):402–408.
- Grimm, V., U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. K. Heinz, G. Huse, A. Huth, J. U. Jepsen, C. Jørgensen, W. M. Mooij, B. Müller, G. Pe'er, C. Piou, S. F. Railsback, A. M. Robbins, M. M. Robbins, E. Rossmannith, N. Rüger, E. Strand, S. Souissi, R. A. Stillam, R. Vabø, U. Visser, and D. L. DeAngelis. 2006. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling* 198(1):115–126.
- Gurcan, O., O. Dikenelli, and C. Bernon. 2011. Towards a generic testing framework for agent-based simulation models. In *Proceedings of the Federated Conference on Computer Science and Information System*, edited by M. Ganzha, L. Maciaszek, and M. Paprzycki. Los Alamitos, CA: IEEE Computer Society Press. Pp. 635–642.
- Hammond, R. A. 2015. Considerations and best practices in agent-based modeling to inform policy. Paper commissioned by the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation (see Appendix A).
- Helbing, D., and S. Balietti. 2011. *How to do agent-based simulations in the future: From modeling social mechanisms to emergent phenomena and interactive systems design*. Tech. Rep. 11-06-024, Santa Fe, NM: Santa Fe Institute.
- Hovelynck, J., A. Dewulf, G. François, and T. Taillieu. 2010. Interdisciplinary knowledge integration through group model building: Recognizing dualities and triadizing the conversation. *Environmental Science & Policy* 13(7):582–591.

- Jenkinson, R., D. Jolley, and P. Dietze. 2014. "Weekend on the town": Discrete sessions of drug use for a sample of young psychostimulant users. *Drug and Alcohol Review* 33(4):428–435.
- Jordan, G. B. 2010. A theory-based logic model for innovation policy and evaluation. *Research Evaluation* 19(4):263–273.
- Kopec, J. A., P. Finès, D. G. Manuel, D. L. Buckeridge, W. M. Flanagan, J. Oderkirk, M. Abrahamowicz, S. Harper, B. Sharif, A. Okhmatovskaia, E. C. Sayre, M. M. Rahman and M. C. Wolfson. 2010. Validation of population-based disease simulation models: A review of concepts and methods. *BMC Public Health* 10(1):710.
- Krupa, J. S., D. M. Rizzo, M. J. Eppstein, D. Brad Lanute, D. E. Gaalema, K. Lakkaraju, and C. E. Warrender. 2014. Analysis of a consumer survey on plug-in hybrid electric vehicles. *Transportation Research Part A* 64:14–31.
- Kuntz, K., F. Sainfort, M. Butler, B. Taylor, S. Kulasingam, S. Gregory, E. Mann, J. M. Anderson, and R. L. Kane. 2013. *Decision and simulation modeling in systematic reviews*. Rockville, MD: Agency for Healthcare Research and Quality.
- Langer, E. M., A. L. Gifford, and K. Chan. 2011. Comparative logic modeling for policy analysis: The case of HIV testing policy change at the Department of Veterans Affairs. *Health Services Research* 46(5):1628–1645.
- Lempert, R. 2002. Agent-based modeling as organizational and public policy simulators. *Proceedings of the National Academy of Sciences of the United States of America* 99(Suppl 3):7195–7196.
- MacDonald, D. 2012. Assessing the influence on drug policy of a program of drug policy research: The Australian Drug Policy Modelling Program 2006–2011. Paper presented at sixth annual conference of the international society for the study of drug policy, University of Kent, Canterbury, United Kingdom.
- Marietto, M., N. David, J. Sichman, and H. Coelho. 2003. A classification of paradigmatic models for agent-based social simulation. In *Multi-agent-based simulation III*. Vol. 2927, Lecture notes in computer science, edited by D. Hales, B. Edmonds, E. Norling, and J. Rouchier. Germany: Springer Berlin Heidelberg. Pp. 193–208.
- McPherson, M., L. Smith-Lovin, and J. M. Cook. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27(1):415–444.
- Meyer, M. 2010. The rise of the knowledge broker. *Science Communication* 32(1):118–127.
- Moore, D., A. Dray, R. Green, S. L. Hudson, R. Jenkinson, C. Siokou, P. Perez, G. Bammer, L. Maher, and P. Dietze. 2009. Extending drug ethno-epidemiology using agent-based modelling. *Addiction* 104(12):1991–1997.
- Moore, T. W., P. D. Finley, N. S. Brodsky, T. J. Brown, B. Apelberg, B. Ambrose, R. J. Glass. In press a. Modeling education and advertising with opinion dynamics. *The Journal of Artificial Societies and Social Simulation*.
- Moore, T. W., P. D. Finley, B. J. Apelberg, B. Ambrose, N. S. Brodsky, T. J. Brown, C. Husten, R. J. Glass. In press b. An opinion-driven behavioral dynamics model for addictive behaviors. *European Physical Journal B*.
- Nicolson, C. R., A. M. Starfield, G. P. Kofinas, and J. A. Kruse. 2002. Ten heuristics for interdisciplinary modeling projects. *Ecosystems* 5(4):376–384.
- NRC (National Research Council). 1991. *Improving information for social policy decisions—the uses of microsimulation modeling: Volume 1, Review and recommendations*. Washington, DC: National Academy Press.
- . 2007. *Models in environmental regulatory decision making*. Washington, DC: The National Academies Press.
- . 2012. *Assessing the reliability of complex models: Mathematical and statistical foundations of verification, validation, and uncertainty quantification*. Washington, DC: The National Academies Press.

- Perez, P., A. Dray, D. Moore, P. Dietze, G. Bammer, R. Jenkinson, C. Siokou, R. Green, S. L. Hudson, and L. Maher. 2012. SimAmph: An agent-based simulation model for exploring the use of psychostimulants and related harm amongst young Australians. *International Journal of Drug Policy* 23(1):62–71.
- Roberts, M., L. B. Russell, A. D. Paltiel, M. Chambers, P. McEwan, and M. Krahn. 2012. Conceptualizing a model: A report of the ISPOR-SMDM modeling good research practices task force—2. *Value in Health* 15(6):804–811.
- Rochester, C. G. 2014. Developing simulation models for assessing effects of tobacco products. Paper presented at 2014 Tobacco Products Scientific Advisory Committee Meeting, Rockville, MD.
- Šalamon, T. 2011. *Design of agent-based models: Developing computer simulations for a better understanding of social processes*. Czech Republic: Tomáš Bruckner.
- Siokou, C., and D. Moore. 2008. “This is not a rave!”: Changes in the commercialised Melbourne rave/dance party scene. *Youth Studies Australia* 27(3):50–57.
- Siokou, C., D. Moore, and H. Lee. 2010. “Muzzas” and “old skool ravers”: Ethnicity, drugs and the changing face of Melbourne’s dance party/club scene. *Health Sociology Review* 19(2):192–204.
- Valente, T. W., K. Fujimoto, J. B. Unger, D. W. Soto, and D. Meeker. 2013. Variations in network boundary and type: A study of adolescent peer influences. *Social Networks* 35(3):309–316.
- Wagner, W., E. Fisher, and P. Pascual. 2010. Misunderstanding models in environmental and public health regulation. *NYU Law Environmental Law Journal* 18:293–356.
- Weinstein, M. C., E. L. Toy, E. A. Sandberg, P. J. Neumann, J. S. Evans, K. M. Kuntz, J. D. Graham, and J. K. Hammitt. 2001. Modeling for health care and other policy decisions: Uses, roles, and validity. *Value in Health* 4(5):348–361.
- Weinstein, M. C., B. O’Brien, J. Hornberger, J. Jackson, M. Johannesson, C. McCabe, and B. R. Luce. 2003. 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 in Health* 6(1):9–17.

