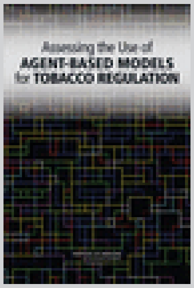


# Assessing the Use of Agent-Based Models for Tobacco Regulation

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

# 6

# Data and Implementation Needs for Computational Modeling for Tobacco Control

The committee noted in the previous chapter that many types of data can be used to inform tobacco research and modeling. Three uses of data were highlighted: qualitative data and "stylized facts" that offer qualitative benchmarks; individual-level data on personal characteristics; and quantitative aggregated data. Empirical data, such as the results from cross-sectional studies, can be useful for indicating patterns of tobacco use in specific settings. Other types of model inputs, such as theoretical models and grounded theories that conceptualize social patterns and structures, qualitative data, and heuristics are also important to consider. In this chapter, the committee provides a high-level overview of existing tobacco-use data sources and identifies data gaps. The committee then discusses inputs and data sources for agent-based models (ABMs) that build off of the discussion of data use in the SnapDragon model (Moore et al., in press a,b) in Chapter 5, including types of data sources, types of network data, and future data collection needs. Different types of agents that could be included in individual-level models, from molecules and cells to individuals and institutions, are also discussed. The chapter finishes with recommendations for the future implementation of computational models at the U.S. Food and Drug Administration's (FDA's) Center for Tobacco Products (CTP).

## EXISTING TOBACCO USE DATA SOURCES

Data are often critical at many, if not all, stages of model development. This section outlines existing data sources and identifies data gaps for tobacco control that could be filled to better understand the evolving tobacco landscape and to inform tobacco control models.

National, state, and local surveillance and evaluation systems primarily collect data on tobacco use behaviors and may also gather information on knowledge and attitudes about pro-tobacco and anti-tobacco influences, effects of tobacco use, and other important risk factors and health outcomes (CDC, 2014). These surveillance resources often vary in their timing, sampling methods, data collection modes, participation rates, and operational definitions and questions regarding tobacco use, initiation, and cessation. For example, the National Youth Tobacco Survey developed by the U.S. Centers for Disease Control and Prevention (CDC) uses self-administered surveys in classrooms to collect nationally representative data biennially on middle and high school youth's tobacco-related beliefs, attitudes, behaviors, social norms, and exposure to pro- and anti-tobacco influences. Every 3 years the Tobacco Use Supplement to the Current Population Survey (TUS–CPS) uses household interviews and telephone follow-ups to capture both national and state data on the age of initiation, secondhand smoke exposure, attitudes toward smoke-free policies, and cessation behavior among young adults and adults. (Box 6-1 provides an overview of national and state survey tools that include information on tobacco use; for more detailed information see CDC, 2014.) Although these surveys rely on different methodologies and are cross-sectional, they have provided general evidence on tobacco use and have offered insight for use in the planning, implementation, and evaluation of tobacco control programs as well as in policy making over the past few decades.

The National Longitudinal Study of Adolescent Health (Add Health) collects data from a nationally representative sample of U.S. adolescents who were in grades 7–12 during the 1994–1995 school year and includes, among many other topics, survey questions on tobacco use (Harris et al., 2009).[1] In 1994, Add Health collected nationally representative behavioral and network data on a baseline "core" sample of more than 90,000 students, including an "in-home" subsample drawn from the core who received more extensive interviews (n = 12,105); of these in-home respondents, 3,702 attended 1 of 16 "saturation schools" where a near-complete social network could be mapped out using answers to the questionnaires (Harris, 2013). Add Health also includes data on family, neighborhood, community, and schools; it is one of the few sources to provide data on both social networks and tobacco use.

[1] Only one cohort of adolescents was selected and followed into adulthood. The study collected follow-up data in 1994–1995, 1996, 2001–2002, and 2007–2008 using in-home interviews. Both public-use and restricted-use datasets are available. See http://www.cpc.unc.edu/projects/addhealth/data for more information (accessed March 2, 2015).

**BOX 6-1**
**Tobacco Data Sources and Tools**

**National and State Surveys and Tools**

Adult Tobacco Survey
Alaska Native Adult Tobacco Survey
American Indian Adult Tobacco Survey
Behavioral Risk Factor Surveillance System
Hispanic/Latino Adult Tobacco Survey Guide
Monitoring the Future
National Adult Tobacco Survey
National Health and Nutrition Examination Survey
National Health Interview Survey
National Survey on Drug Use and Health
National Youth Tobacco Survey
Pregnancy Risk Assessment Monitoring System
School Health Policies and Practices Study
School Health Profiles
Smoking-Attributable Morbidity, Mortality, and Economic Costs
State Tobacco Activities Tracking and Evaluation System
Tobacco Use Supplement to the Current Population Survey
Youth Risk Behavior Surveillance System
Youth Tobacco Survey

**Registries and Vital Statistics**

National Program of Cancer Registries
National Vital Statistics System

**Health Systems and Clinical Settings**

Healthcare Effectiveness Data and Information Set
Hospital discharge data
National Ambulatory Medical Care Survey
National Hospital Ambulatory Medical Care Survey
National Mental Health Services Survey
National Quitline Data Warehouse
National Survey of Substance Abuse Treatment Services

**Sales Data**

Information Resources, Inc.
Scanner data
Substance Abuse and Mental Health Services Administration
Tax revenue data
U.S. Food and Drug Administration compliance checks

**National, State, and Local Policy Tracking**

ACTIVE Life Tobacco Free Worksite Survey
American College Health Association College Campus Tobacco Cessation and Prevention Survey

*continued*

**BOX 6-1 Continued**

- American Lung Association's State Legislated Actions on Tobacco Issues
- American Nonsmokers' Rights Foundation: U.S. Tobacco Control Laws Database
- California Student Tobacco Survey
- California Tobacco Use Prevention Education Evaluation Teacher Survey
- CDC School Health Profiles
- Worksite and Restaurant Smoking Policy Questionnaires and Guide

**Media Tracking**

- Adobe SiteCatalyst
- Arbitron
- Cision
- Clicktracks Optimizer
- DataSift
- Facebook Insights
- Gnip
- Google Analytics
- HootSuite
- Legacy Media Tracking Survey and Legacy Media Tracking Online
- LexisNexis
- Nielsen
- Pinterest
- Radian6
- Sysomos
- Topsy
- Webalyzer
- YouTube Analytics

**Global Survey Tools**

- Global Adult Tobacco Survey
- Global Health Professions Student Survey
- Global School Personnel Survey
- Global School-Based Student Health Survey
- Global Youth Tobacco Survey

**Tobacco Industry Monitoring**

- Network of the National Cancer Institute
- New Product Watch, funded by Tobacco Surveillance, Epidemiology, and Evaluation
- Project SMART Money of California State Department of Public Health
- Retail Advertising Tobacco Survey
- University of California at San Francisco Tobacco Control Archives

SOURCE: Adapted from CDC, 2014, which contains details on each of these sources.

The smoking and social network data available in Add Health have been used in numerous studies. For example, Pollard and colleagues found that membership in "higher-use" trajectories of tobacco smoking, as these adolescents moved into adulthood, were predicted by the number of perceived best friends who smoked and by changes in the numbers of these friends (Pollard et al., 2010). Several analyses of Add Health have employed the stochastic actor-based model SIENA developed by Snijders and colleagues (2010), which simultaneously models the social network change process and the peer influence process. Schaefer and colleagues (2012) found that students in a single Add Health saturation school smoked more frequently if their peers smoked, and they were also more likely to choose peers who smoked if they themselves smoked. Using the Add Health saturation schools, Lakon and colleagues (2014) considered the effects of parental influences as well and found that smoking by parents and peers increased the probability of an adolescent's smoking. The SIENA model is amenable to simulation and could serve as a basis for computational experiments for smoking prevention, as was done by Schaefer and colleagues (Haas and Schaefer, 2013, 2014).[2] Other studies using network and smoking data include the six-country European Smoking Prevention Framework studies by Mercken and colleagues (2009), a study of online social networks supporting tobacco cessation (Cobb et al., 2010), and studies by Valente and colleagues (2006, 2013).

Data from the Population Assessment of Tobacco and Health (PATH) Study could provide useful data in the future. PATH is a national-cohort longitudinal study of tobacco use and how it affects health in the United States. Sponsored by the National Institutes of Health and FDA, PATH began in 2011 and is a prospective study that will follow an estimated 46,000 U.S. household residents age 12 years and older (PATH, 2015a). The study's goals include explaining various aspects of tobacco use patterns and characterizing the natural history of tobacco dependence, cessation, and relapse. However, the PATH study is still in the early phases, and data from it are not yet available.[3] PATH will collect some data that are not routinely collected in other data sources. For example, PATH will identify trends in tobacco use patterns, including the use of new products, dual use,

[2] It is important to note some of the limitations of the SIENA model. For example, the model is limited to network change and peer influences on behavior, which is just one component of the model. Furthermore, SIENA is designed to fit a simulation model to data, and thus results (parameter estimates for network or behavior change) may not be generalizable to out-of-data scenarios. Finally, the model requires an initial network configuration.

[3] "The field test for the PATH Study took place between November 2012 and February 2013. Baseline data collection, which will last for 15 months, began in September 2013; the second annual data collection begins mid-October 2014 and will be followed by at least one additional data-collection wave" (PATH, 2015b).

poly use and switching; it will monitor changes in risk perceptions and other attitudes, such as social acceptability and individual preferences; and it will assess differences among and within critical subgroups, including youth, young adults, daily users, racial/ethnic minority groups, and users of new tobacco products, among others.

In addition to national and state surveys, tobacco-related information can be gathered from a variety of other sources, even if many of these sources are dedicated to other topics. Cancer registries, vital statistics, and medical records[4] offer data on health status and outcomes, such as incidence data on smoking-related morbidity and mortality. Quitline data warehouses collect information on the use and success of quitlines and identify knowledge gaps in order to inform the design of new strategies that can improve cessation services. Mass media and social media trackers can gather data on the level of influence of both anti- and pro-tobacco advertisements and campaigns as well as on tobacco-related beliefs, attitudes, social norms, and behaviors, particularly among youth (CDC, 2014). Finally, consumer purchase data have been collected and analyzed to assess trends in purchasing in order to identify patterns relevant to specific geographic locations and demographics characteristics of consumers and also to assess the impact of specific marketing strategies. One source of these data are the Nielsen data (2014), which can be purchased to understand better how and where specific products are selling[5] (see, for example, Amerson et al., 2014; NYSDOH, 2011; Terry-McElrath et al., 2011).

As the tobacco landscape has evolved in recent years, the need for different types of data has grown. After the enactment of the Tobacco Control Act—and in response to emerging trends in tobacco use—FDA and CDC began including detailed questions on nonconventional tobacco products in the 2012 National Youth Tobacco Survey (Apelberg et al., 2014). However, most surveys still focus on cigarettes, and the data sources are not available for every state (CDC, 2014). The surveys that have included questions on other tobacco products still lack the quality, depth, and breadth to capture data on the effects of multiple product use, substitution, and branding on initiation, cessation, addiction, and tobacco-related disparities among population groups (Delnevo, 2014; Mermelstein, 2014). There also continues to be a gap in the data on the interacting effects of multiple tobacco control

[4] Some tobacco studies have used medical records as data sources. For example, to study the relationship between passing smokefree indoor policies and incidence of myocardial infarction, Hurt and colleagues (2012) used medical records from the Mayo Clinic. Potentially, larger health care datasets could be used for future tobacco research.

[5] There are restrictions on how those data can be disseminated, but government agencies now purchase these data to assess local and national sales trends because of their relevance to a variety of outcomes (e.g., increases in sales of tobacco in a location might be linked with increased health care costs in that same area) (Amerson et al., 2014).

policies (Farrelly, 2009). Finally, network data for tobacco use—that is, the salient social connections between potential or current users and their peers, family members, and others who may influence tobacco use—are almost completely lacking, including data on special populations such as minority groups and high-risk groups such as those with mental illness. Such data could provide a better understanding of the influence of social networks and social context on tobacco use and on the behavior change process involved. Given the changing tobacco landscape, it is likely there will be an increasing need for detailed yet timely and accurate data for informing tobacco control efforts nationwide. Data needs for ABMs are discussed in more detail later in this chapter.

## DATA NEEDS FOR FUTURE MODELING EFFORTS

Although various types of existing data sources related to tobacco use can be used to inform and strengthen ABMs, these sources do not contain all of the relevant agent attributes, behaviors, and social and spatial interactions related to tobacco use. As noted above, Add Health data are commonly used to study peer influences on smoking behavior. However, the Add Health baseline data lack detailed information on the mechanisms underlying peer selection. Also, its tobacco-related questions, which are concerned only with smoking and chewing/snuff, capture limited information (CPC, 1998). Furthermore, the biological and clinical data collected by Add Health are not comprehensive, especially in the first two waves of the study. The lack of data on networks and smoking can make modeling the social interactions that influence tobacco use a challenge.

Other existing data sources could also be used to inform computational models but pose some challenges as well (North and Macal, 2007, p. 240). The tobacco industry has collected much data on the uptake of smoking and the effectiveness of marketing (see Cummings et al., 2002, as well as the Legacy Tobacco Documents Library[6] for historical industry documents), which could be used to inform computational models. Sifting through these documents and finding those most relevant to inform computational models used to guide regulatory efforts could be difficult, however (Bero, 2003; Cruz, 2009). Another approach would be to try to maximize the use of available administrative data from states and regions, but, except in unusual circumstances, this information is not likely to contain many of the behaviors and interactions wanted. Alternatively, one could combine

[6]The Legacy Tobacco Documents Library is a digital archive of tobacco industry documents, containing more than 14 million documents, which was developed by a variety of tobacco companies and which relates to their advertising, marketing, manufacturing, sales, and research activities. For more information, see http://legacy.library.ucsf.edu (accessed March 2, 2015).

data from various sources, such as large-area administrative information and small-area detailed surveys. However, using such combinations would require considerable care. These challenges are compounded by the fact that, in general, "the data most useful for modeling is often among the most jealously guarded resources in many organizations" (North and Macal, 2007, p. 240).

A longer-term approach is to try to anticipate critical data needs and either fund or otherwise encourage the collection of data that best suit ABMs or other modeling approaches. Similarly, encouraging the standardization of data collection items and methods might improve model quality. Even for administrative data that are "routinely" collected, such as tobacco marketing and sales information or population smoking prevalence estimates, it could be possible to evaluate those data periodically for validity and consistency. It may also be possible to substitute existing or newly developed biomarkers of certain smoking behaviors for other forms of data collection, and in selected instances, information from other countries with similar populations may be of value.

Network data, which are thought to require the elucidation of an entire social network, are particularly difficult to collect. For example, many network measures, particularly centrality measures, are prone to biases (Costenbader and Valente, 2003; Kossinets, 2006; Smith and Moody, 2013). However, there may be some modeling efforts for which whole-network (sociometric) data are not required. As noted by the Statnet[7] Development Team, egocentric data was "long regarded as the poor country cousin in the network data family," yet such data "contain a remarkable amount of information" (Butts et al., 2014). Adding in egocentric (sampled) network questions would add information that is relevant to ABMs. In collecting network data of this type, one employs traditional survey methods to assemble representative samples of the population. Respondents could be asked questions about important contacts. For example: How many of your five best friends smoke? What are their relevant attributes? Do they know one another? How many of your family members smoke? Such questions could be added, for example, to the Behavioral Risk Factor Surveillance System or the TUS–CPS. Novel developments in sampled networks permit the simulation of disease outbreaks, which could be applied to behavioral "epidemics" as well as to infectious ones.[8]

The use of network-based ABMs in epidemiology has increased over the past decade. There are two issues in such modeling. One, ground truth

[7]Statnet is a statistical modeling package for the R platform. See http://statnet.org for more information (accessed March 2, 2015).

[8]For example, see details on the EpiModel at http://cran.r-project.org/web/packages/EpiModel/index.html for more information (accessed March 2, 2015).

(i.e., any data that capture the empirical process under investigation) is often limited to stylized facts and theoretical models with little empirical data. Two, confirmation rarely moves beyond internal validation and calibration. Networks provide structure—who interacts with whom—and are incorporated in a number of ways. The two most common approaches are to generate a stylized network or to input an actual network. In the case of generating a stylized network, the choice of which stylized network has implications for diffusion processes, including the time course and peak of an epidemic (Rahmandad and Sterman, 2008).

One of the key data needs for ABM is data that inform agent interactions, either with other agents or with the agent's environment. Such interactions are difficult or impossible to capture empirically. Traditional data from survey methods may not always provide the detailed data required for reproducing the relevant interactions, motives, sequence of events, or decision processes associated with the behavior of an agent. Alternative data collection methodologies could include qualitative methods (such as ethnography) that tap into the experience of social interactions (Falkin et al., 2007; Rothwell and Lamarque, 2011), experiential or situational sampling (e.g., ecological momentary assessment [EMA]; see Shiffman et al., 2008), and time-use data (e.g., the American Time Use Survey, or ATUS) that capture "with whom," "where," and "when" types of questions. The use of EMA has been particularly enlightening for understanding the context of tobacco cravings (Chandra et al., 2011). Time-use data capture a representative slice of daily activities; the ATUS sample is drawn from the Current Population Survey (CPS), which means it can be linked to the TUS–CPS (NCI, 2014). Such linkage provides a rich source of daily activities within a geographic context. Experimental or quasi-experimental data may also be relevant, such as those from random roommate assignments (Eisenberg et al., 2014), and data from randomized controlled trials of smoking cessation programs (Bullen et al., 2013; Strecher et al., 2008). Such studies, especially individual-level trials, would be useful in parameterizing empirically based rules for agent behavior.

Online platforms may offer yet another way to collect data. While tobacco companies are making extensive use of online social media to market their products, the tobacco control community is using online platforms to counter the marketing of tobacco products (Legacy, 2012), provide cessation support services and forums (Gutierrez and Newcombe, 2012), and mobilize advocates to strengthen tobacco-control efforts (Hefler et al., 2013). Because various stakeholders of the tobacco environment use online social media (see Box 6-1 for other online platforms and related trackers), enormous amounts of data have been generated, including the social connections and interactions among individuals online. Such data may be mined to better understand the diffusion dynamics of and the role of social network structure in tobacco use (Centola, 2013; Cobb et al., 2010). Content analysis is now possible on a massive scale, a development that could help enrich the understanding of the mechanisms that drive tobacco addiction (Myslín et al., 2013). It is important to keep in mind, however, that online and face-to-face networks are distinct and potentially interact with one another (Huang et al., 2014), so it will continue to be necessary to use a range of research methods.

There are also data needs at the aggregate (state/national/local) level. It will be necessary to remain vigilant in collecting both qualitative and quantitative observations of American tobacco use habits over time. Changes will likely occur in the types of tobacco user groups and their general characteristics, such as age, gender, race/ethnicity, socioeconomic status, cultural beliefs, health characteristics, the types of tobacco delivery devices used, and the use of other relevant substances. Of course, the policy questions themselves may change over time, which will also affect the nature of data collection. Not all of these changes can be easily predicted, making ongoing population tobacco surveillance necessary, if only for basic data needs and to identify more targeted surveys for policy promulgation. Such data would be useful as ground truth against which simulation results could be compared.

### Other Types of Agents for Application in Agent-Based Models

Another area for future data collection is capturing information on the many agents that could be modeled in ABMs developed for tobacco control policy. The agents in SnapDragon (the central ABM evaluated in this report, see Chapter 5) are people and media, but other types of agents are possible. These agents could, for example, include state and local legislators, policy makers, and health departments if one wished to better understand how they approach tobacco control and regulation at the local level. Social networks, particularly the ways in which information and resources are shared among stakeholder groups in the tobacco control regulatory landscape, have been described by Luke and colleagues (Harris et al., 2008; Luke et al., 2010). Organizational collaborations among public health agencies, advocacy groups, and funders, among others, are critical in the dissemination of tobacco control research and evidence-based best practices (Luke et al., 2013). ABMs could also be used to consider the tobacco industry's behavior, with tobacco companies being the agents in the model. For example, an ABM could examine the role of current cigarette manufacturers in the alternative nicotine delivery market. ABMs that aim to capture industry behavior could complement other models and research in illuminating the implications of tobacco product use and could provide guidance on the type of industry data that is needed for policy evaluation. Agents may also include state excise tax collectors, private area-wide commerce organizations, and other organizations and government agencies that do not have tobacco regulation as their fundamental mission but whose policies impact tobacco use. For example, agents might include housing and environmental agencies, chambers of commerce, commercial trade organizations, or police organizations that may become involved in contraband tobacco products. Health systems and health professionals might also be considered as agents in some policy models.

Agents may also be “below the skin,” as components of complex biological systems—for example, neurons, nicotine, nicotinic acetylcholine receptors, and cytochrome P450 enzymes could all be considered types of agents. Neural pathways have been identified as key components in the addiction process, which entails the activation of reward-learning circuits (Hyman et al., 2006; Koob and Le Moal, 2001). According to Hyman and colleagues, “Humans and animals rapidly learn cues and contexts that predict the availability of these ‘addictive drugs’; once learned, these cues motivate drug seeking in humans and animal models” (Hyman et al., 2006, p. 567). Because addiction plays such a central role in tobacco use, modeling the process of addiction and the resulting difficult-to-change behavior could help strengthen ABMs.

The mathematical and computational modeling of biological systems has been helpful in understanding other disease processes, including hepatitis clearance and infectivity (Dahari et al., 2009), host–pathogen interactions (Stern et al., 2013), and inflammation and multiple organ failure (An, 2004, 2006). Relevant “below the skin” factors—that is, various elements that constitute and act on an organism’s biological systems—have not been largely used in ABM of tobacco use. Biological factors are not necessarily required if they can be represented by simply using proxies or if the model is concerned primarily with above-the-skin factors. In other words, the level at which agents are specified will depend on the questions being asked of the model. Such low-level detail may be necessary if individual responses to nicotine (e.g., half-life) and the toxicity of a tobacco product are important, but they may be unnecessary or undesired if the model concerns diffusion of information or norms. Following the recommendations of a National Research Council report (2008), models should strive for parsimony and avoid “kitchen sink” approaches. The report noted that “models can become unwieldy when weighed down by a proliferation of features and variables” (p. 347). Nevertheless, given that tobacco use initiation and cessation are at least in part based on human physiology, the modeling of relevant biological mechanisms would need to at least be considered.

## DATA COLLECTION AND MODEL DEVELOPMENT AT THE CENTER FOR TOBACCO PRODUCTS

In this report, the committee discussed the importance and challenges of incorporating data into ABMs that are intended to inform tobacco control policy. Models that use minimal amounts of data can be used to guide data collection and the development of future models. However, when the goal is to guide policies, data can help ensure that the agents capture realistic representations of actual entities to the extent possible, and data can also confirm the degree to which the model replicates or predicts real-world patterns, such as initiation and cessation processes among individuals or populations. In the tobacco control field, an assortment of data is available (as presented earlier in the chapter as well as in the tobacco use behavior section of Chapter 2), and much of these data could help strengthen ABMs developed to guide tobacco control policy. These data could be used creatively to inform models, and more data could be collected from efforts that go beyond traditional survey methods, such as gathering information from online social media platforms. Data collected with behavioral mechanisms in mind would allow agent-based modelers to capture more realistic agent characteristics as well as more realistic agent–agent and agent–environment interactions. It is important that these characteristics and interactions be captured meaningfully because they tend to be central elements of ABMs that aim to inform policy decisions, especially if the goal of the model is to understand how interdependent agent behavior will shape the outcomes experienced under a given policy (as discussed in Chapter 3). Because ABMs and other individual-level modeling techniques are promising tools to further our understanding of tobacco use behavior, it is worthwhile to collect such data. As a major funder and user of tobacco data (including for the modeling of tobacco use), CTP can help shape the tobacco data environment in the future.

*Conclusion 6-1: The committee concludes that agent-based models designed to inform policy decisions require data on the underlying mechanisms governing behavior and on agent-to-agent and agent-to-environment interactions. Currently, these data are not commonly collected.*

**Recommendation 6-1: The Center for Tobacco Products should identify and help develop data sources relevant to the questions it is trying to address using agent-based models and other modeling approaches.**

The use of data already being collected (either by CTP or other sources) could be incorporated into the modeling process. CTP could consider coordinating with other activities, such as the Tobacco Centers of Regulatory Science, to gather these data. As noted elsewhere in the report, models can help researchers identify data gaps and combine data from various sources (while recognizing the limitations of each), further guiding data collection and enhancing models used to inform policy.

To ensure that the processes of collecting the necessary data and of identifying agent attributes based on those data are done successfully, it is crucial to address implementation issues. Having the appropriate individuals overseeing these processes and ensuring that the models have broad input to inform them will both be important to the success of the models. Many different types of models have been developed by federal agencies; some of them developed within the agencies and others through contracts or grants. Regardless of where the models are developed, funders for policy-relevant models require access to expertise if they are to issue effective funding opportunity announcements or contracts; to make informed decisions about which modeling approaches are appropriate for the question at hand; to work effectively with the modeling team(s) throughout model development; to appropriately evaluate model inputs, processes, and outputs; and to interpret or translate model results appropriately to decision makers.

FDA is regularly confronted with uncertainty within the complex tobacco environment. Because of this, it will remain necessary to have models that represent potential tobacco policies to organize data, elucidate specific uncertainties, and forecast future scenarios. Because the use of models at CTP has the potential to affect regulatory decision making, it is essential that the development of these models be overseen by individuals who have the expertise and experience needed to maximize the benefit and reliability of the models. Subject-matter experts (that is, scientists and researchers who have a deep understanding of the tobacco literature and work in that field) could be essential partners in future CTP modeling endeavors (see Chapter 4 for more discussion on this topic).

**Recommendation 6-2: The Center for Tobacco Products (CTP) should ensure that it has staff with, or access to, the necessary expertise to inform CTP's research, contracting, and evaluation efforts and to translate model results for various stakeholders.**

FDA could also consider obtaining input on the development of its models from tobacco stakeholders, including representatives from local, state, and federal public health agencies; scientists and other members of academia; other modelers; and end users, among others. CTP could acquire feedback in a number of ways, ranging from developing a standing expert panel to provide regular feedback on modeling initiatives, to using modeling networks or forums such as the Models of Infectious Disease Agent

Study (MIDAS),[9] the Cancer Intervention and Surveillance Modeling Network (CISNET),[10] the Drug Policy Modelling Program,[11] and the Energy Modeling Forum.[12,13]

Although individual models are a useful tool for informing policy decisions, having a range of modeling techniques will offer a fuller picture of the policy questions confronted by CTP—for example, by creating various models to approach the same question or process (e.g., multiple ABMs or ABMs and aggregate models), as is done by several modeling networks and forums.[14] The documentation of model inputs, activities, and outputs by the model developers (as discussed in Chapter 4) and a comparison of results with a rigorous discussion by the developers on why the results differ—or do not differ—will create a richer understanding of the models and the model results (Kuntz et al., 2013) and will help to address model uncertainty. Doing so will also help to increase policy makers' confidence in the model results, identify where assumptions need to be modified, and detect where further data are needed.[15]

> **Recommendation 6-3: The U.S. Food and Drug Administration should develop a range of models using various approaches. This would include agent-based models as well as other modeling approaches.**

It is important to note that the range of models FDA could use includes not only those that FDA commissions or develops but also those that others have already developed or will develop to help guide tobacco control policy.

---

[9] For more on MIDAS, see http://www.nigms.nih.gov/Research/SpecificAreas/MIDAS/Pages/default.aspx (accessed March 2, 2015).

[10] For more on CISNET, see http://cisnet.cancer.gov (accessed March 2, 2015).

[11] For more on Drug Policy Modelling Program, see https://dpmp.unsw.edu.au (accessed March 2, 2015).

[12] For more on Energy Modeling Forum, see https://emf.stanford.edu (accessed March 2, 2015).

[13] Modeling networks and forums can take several forms but generally consist of a collaborative network of researchers who develop various types of models to understand the topic at hand. These models are often intended for policy makers, public health officials, and other researchers to help them make better-informed decisions on the topic of study (see Appendix A).

[14] For example, the modeling done by CISNET is collaborative, and members address a common question using a common dataset. For a description of a specific instance of this, see the July 2012 supplement of *Risk Analysis*, which was devoted to the CISNET modeling of smoking and lung cancer, and various CISNET collaborative articles on breast, colon and prostate cancer.

[15] See Appendix A for a discussion on the benefits of using a multiple model approach, using MIDAS as an example.

## CONCLUSION

Although simulation modeling has been used for many years in tobacco control, CTP is still in the early stages of its efforts to use ABM to explore tobacco control policy and regulation. This report has illustrated many of the challenging and technical aspects surrounding ABMs. However, the committee believes that ABMs are a useful tool that could add to the understanding of tobacco use initiation, cessation, and relapse processes. The model developed for FDA (see Chapter 5) does not accurately represent many of the important characteristics of tobacco use, but there is much to be learned from its development that can be applied to future models of tobacco use, both agent-based and otherwise. There are some barriers to overcome, such as the collection of data to inform the development of ABMs and the elucidation of the empirical and theoretical challenges of specifying model inputs and appropriately interpreting model outputs (see Chapter 3). A strong evaluation framework (as described in Chapter 4) will be needed to track rigorous model development. As discussed in Chapters 3 and 4, it will be important to consult an interdisciplinary modeling team and subject-matter experts at the earliest stage of model conceptualization and then throughout the model development process in order to ensure that the model is grounded in the current state of tobacco science (that is, evidence-based research related to tobacco in the fields of epidemiology, social and behavioral sciences, biology, chemistry, and others), while carefully considering individual behavior. If the principles discussed in this report are followed, the value of ABMs for informing tobacco regulation will be greatly strengthened.

## REFERENCES

Amerson, N. L., B. S. Arbise, N. K. Kelly, and E. Traore. 2014. Use of market research data by state chronic disease programs, Illinois, 2012–2014. *Preventing Chronic Disease* 11:E165.

An, G. 2004. In silico experiments of existing and hypothetical cytokine-directed clinical trials using agent-based modeling. *Critical Care Medicine* 32(10):2050–2060.

———. 2006. Concepts for developing a collaborative in silico model of the acute inflammatory response using agent-based modeling. *Journal of Critical Care* 21(1):105–110.

Apelberg, B. J., C. L. Backinger, and S. J. Curry. 2014. Enhancing youth tobacco surveillance to inform tobacco product regulation: Findings from the 2012 National Youth Tobacco Survey. *American Journal of Preventive Medicine* 47(2 Suppl 1):S1–S3.

Bero, L. 2003. Implications of the tobacco industry documents for public health and policy. *Annual Review of Public Health* 24(1):267–288.

Bullen, C., C. Howe, M. Laugesen, H. McRobbie, V. Parag, J. Williman, and N. Walker. 2013. Electronic cigarettes for smoking cessation: A randomised controlled trial. *Lancet* 382(9905):1629–1637.

Butts, C. T., M. Morris, P. N. Krivitsky, Z. Almquist, M. S. Handcock, D. R. Hunter, S. M. Goodreau, and S. B. de-Moll. 2014. Introduction to exponential-family random graph (ERG or p*) modeling with ergm. http://statnet.csde.washington.edu/workshops/UNBELT/current/ergm/ergm_tutorial.pdf (accessed November 7, 2014).

CDC (Centers for Disease Control and Prevention). 2014. *Surveillance and evaluation data resources for comprehensive tobacco control programs*. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

Centola, D. 2013. Social media and the science of health behavior. *Circulation* 127(21): 2135–2144.

Chandra, S., D. Scharf, and S. Shiffman. 2011. Within-day temporal patterns of smoking, withdrawal symptoms, and craving. *Drug and Alcohol Dependence* 117(2–3):118–125.

Cobb, N. K., A. L. Graham, and D. B. Abrams. 2010. Social network structure of a large online community for smoking cessation. *American Journal of Public Health* 100(7):1282–1289.

Costenbader, E., and T. W. Valente. 2003. The stability of centrality measures when networks are sampled. *Social Networks* 25(4):283–307.

CPC (Carolina Population Center, University of North Carolina at Chapel Hill). 1998. *National Longitudinal Study of Adolescent Health: Wave I*. http://www.cpc.unc.edu/projects/addhealth/codebooks/wave1 (accessed February 25, 2015).

Cruz, T. B. 2009. Monitoring the tobacco use epidemic IV. The vector: Tobacco industry data sources and recommendations for research and evaluation. *Preventive Medicine* 48(1 Suppl):S24–S34.

Cummings, K. M., C. P. Morley, J. K. Horan, C. Steger, and N. R. Leavell. 2002. Marketing to America's youth: Evidence from corporate documents. *Tobacco Control* 11(Suppl 1):I5–I17.

Dahari, H., E. Shudo, R. Ribeiro, and A. Perelson. 2009. Mathematical modeling of HCV infection and treatment. *Methods in Molecular Biology* 510:439–453.

Delnevo, C. D. 2014. Tobacco surveillance data for population modeling: Getting the inputs right. Paper presented at Tobacco Products Scientific Advisory Committee Meeting, Rockville, MD.

Eisenberg, D., E. Golberstein, and J. L. Whitlock. 2014. Peer effects on risky behaviors: New evidence from college roommate assignments. *Journal of Health Economics* 33:126–138.

Falkin, G. P., C. S. Fryer, and M. Mahadeo. 2007. Smoking cessation and stress among teenagers. *Qualitative Health Research* 17(6):812–823.

Farrelly, M. C. 2009. Monitoring the tobacco use epidemic V: The environment: Factors that influence tobacco use. *Preventive Medicine* 48(1 Suppl):S35–S43.

Gutierrez, K., and R. Newcombe. 2012. *Lessons learned globally: Tobacco control digital media campaigns*. Saint Paul, MN: Global Dialogue for Effective Stop Smoking Campaigns.

Haas, S., and D. Schaefer. 2014. With a little help from my friends? Asymmetrical social influence on adolescent smoking initiation and cessation. *Journal of Health and Social Behavior* 55(2):126–143.

Harris, J. K., D. A. Luke, R. C. Burke, and N. B. Mueller. 2008. Seeing the forest and the trees: Using network analysis to develop an organizational blueprint of state tobacco control systems. *Social Science and Medicine* 67(11):1669–1678.

Harris, K. M. 2013. *The Add Health study: Design and accomplishments*. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill.

Harris, K. M., C. T. Halpern, E. Whitsel, J. Hussey, J. Tabor, P. Entzel, and J. R. Udry. 2009. The National Logitudinal Study of Adolescent to Adult Health: Research design http://www.cpc.unc.edu/projects/addhealth/design (accessed March 2, 2015).

Hefler, M., B. Freeman, and S. Chapman. 2013. Tobacco control advocacy in the age of social media: Using Facebook, Twitter and change. *Tobacco Control* 22(3):210–214.

Huang, G. C., D. Soto, K. Fujimoto, and T. W. Valente. 2014. The interplay of friendship networks and social networking sites: Longitudinal analysis of selection and influence effects on adolescent smoking and alcohol use. *American Journal of Public Health* 104(8):e51–e59.

Hurt, R. D., S. A. Weston, J. O. Ebbert, S. M. McNallan, I. T. Croghan, D. R. Schroeder, and V. L. Roger. 2012. Myocardial infarction and sudden cardiac death in Olmsted County, Minnesota, before and after smoke-free workplace laws. *Archives of Internal Medicine* 172(21):1635–1641.

Hyman, S. E., R. C. Malenka, and E. J. Nestler. 2006. Neural mechanisms of addiction: The role of reward-related learning and memory. *Annual Review of Neuroscience* 29:565–598.

Koob, G. F., and M. Le Moal. 2001. Drug addiction, dysregulation of reward, and allostasis. *Neuropsychopharmacology* 24(2):97–129.

Kossinets, G. 2006. Effects of missing data in social networks. *Social Networks* 28(3):247–268.

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.

Lakon, C. M., C. Wang, C. T. Butts, R. Jose, D. S. Timberlake, and J. R. Hipp. 2014. A dynamic model of adolescent friendship networks, parental influences, and smoking. *Journal of Youth and Adolescents*.

Legacy. 2012. *Truth health fact sheet*. http://www.legacyforhealth.org/content/download/621/7337/file/truth_fact_sheet_January_2012.pdf (accessed February 18, 2015).

Luke, D. A., J. K. Harris, S. Shelton, P. Allen, B. J. Carothers, and N. B. Mueller. 2010. Systems analysis of collaboration in 5 national tobacco control networks. *American Journal of Public Health* 100(7):1290–1297.

Luke, D. A., L. M. Wald, B. J. Carothers, L. E. Bach, and J. K. Harris. 2013. Network influences on dissemination of evidence-based guidelines in state tobacco control programs. *Health Education and Behavior* 40(1 Suppl):33S–42S.

Mercken, L., T. A. Snijders, C. Steglich, and H. de Vries. 2009. Dynamics of adolescent friendship networks and smoking behavior: Social network analyses in six European countries. *Social Science and Medicine* 69(10):1506–1514.

Mermelstein, R. J. 2014. Adapting to a changing tobacco landscape: Research implications for understanding and reducing youth tobacco use. *American Journal of Preventive Medicine* 47(2 Suppl 1):S87–S89.

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*.

Myslín, M., S. H. Zhu, W. Chapman, and M. Conway. 2013. Using Twitter to examine smoking behavior and perceptions of emerging tobacco products. *Journal of Medical Internet Research* 15(8):e174.

NCI (National Cancer Institute). 2014. *Tobacco use supplement—current population survey fact sheet*. http://appliedresearch.cancer.gov/tus-cps/TUS-CPS_fact_sheet.pdf (accessed November 8, 2014).

The Neilsen Company. 2014. *About us*. http://www.nielsen.com/us/en/about-us.html (accessed November 10, 2014).

North, M. J., and C. M. Macal. 2007. *Managing business complexity: Discovering strategic solutions with agent-based modeling and simulation*. Oxford, UK: Oxford University Press.

NRC (National Research Council). 2008. *Behavioral modeling and simulation: From individuals to societies*. Washington, DC: The National Academies Press.
NYSDOH (New York State Department of Health). 2011. *Key tobacco control outcome indicators*. New York State Department of Health. https://www.health.ny.gov/prevention/tobacco_control/docs/2011-09_key_tobacco_outcome_indicators.pdf (accessed October 28, 2014).
PATH (Population Assessment of Tobacco and Health). 2015a. *Study overview*. http://www.pathstudyinfo.nih.gov/UI/StudyOverviewMobile.aspx (accessed February 18, 2015).
———. 2015b. *FAQS for researchers*. http://www.pathstudyinfo.nih.gov/UI/FAQsResMobile.aspx (accessed November 21, 2014).
Pollard, M. S., J. S. Tucker, H. D. Green, D. Kennedy, and M. H. Go. 2010. Friendship networks and trajectories of adolescent tobacco use. *Addictive Behaviors* 35(7):678–685.
Rahmandad, H., and J. Sterman. 2008. Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. *Management Science* 54(5):998–1014.
Rothwell, E., and J. Lamarque. 2011. The use of focus groups to compare tobacco attitudes and behaviors between youth in urban and rural settings. *Health Promotion Practice* 12(4):551–560.
Schaefer, D. R., S. A. Haas, and N. J. Bishop. 2012. A dynamic model of U.S. adolescents' smoking and friendship networks. *American Journal of Public Health* 102(6):e12–e18.
Schaefer, D. R., J. Adams, and S. A. Haas. 2013. Social networks and smoking: Exploring the effects of peer influence and smoker popularity through simulations. *Health Education and Behavior* 40(1 Suppl):24S–32S.
Shiffman, S., A. A. Stone, and M. R. Hufford. 2008. Ecological momentary assessment. *Annual Review of Clinical Psychology* 4:1–32.
Smith, J. A., and J. Moody. 2013. Structural effects of network sampling coverage I: Nodes missing at random. *Social Networks* 35(4):652–668.
Snijders, T. A., G. G. Van de Bunt, and C. E. Steglich. 2010. Introduction to stochastic actor-based models for network dynamics. *Social Networks* 32(1):44–60.
Stern, J. R., A. D. Olivas, V. Valuckaite, O. Zaborina, J. C. Alverdy, and G. An. 2013. Agent-based model of epithelial host-pathogen interactions in anastomotic leak. *Journal of Surgical Research* 184(2):730–738.
Strecher, V. J., J. B. McClure, G. L. Alexander, B. Chakraborty, V. N. Nair, J. M. Konkel, S. M. Greene, L. M. Collins, C. C. Carlier, C. J. Wiese, R. J. Little, C. S. Pomerleau, and O. F. Pomerleau. 2008. Web-based smoking-cessation programs: Results of a randomized trial. *American Journal of Preventive Medicine* 34(5):373–381.
Terry-McElrath, Y. M., S. Emery, M. A. Wakefield, P. M. O'Malley, G. Szczypka, and L. D. Johnston. 2011. Effects of tobacco-related media campaigns on smoking among 20–30-year-old adults: Longitudinal data from the USA. *Tobacco Control* 22(1):38–45.
Valente, T. W., J. B. Unger, A. Ritt-Olson, S. Y. Cen, and C. Anderson Johnson. 2006. The interaction of curriculum type and implementation method on 1-year smoking outcomes in a school-based prevention program. *Health Education Research* 21(3):315–324.
Valente, T. W., K. Fujimoto, D. Soto, A. Ritt-Olson, and J. B. Unger. 2013. A comparison of peer influence measures as predictors of smoking among predominately Hispanic/Latino high school adolescents. *Journal of Adolescent Health* 52(3):358–364.