Agent-Based Modeling of Noncommunicable Diseases: A Systematic Review

Roch A. Nianogo, MD, MPH, and Onyebuchi A. Arah, MD, DSc, PhD, MPH, MSc

We reviewed the use of agent-based modeling (ABM), a systems science method, in understanding noncommunicable diseases (NCDs) and their public health risk factors.

We systematically reviewed studies in PubMed, ScienceDirect, and Web of Sciences published from January 2003 to July 2014. We retrieved 22 relevant articles; each had an observational or interventional design. Physical activity and diet were the most-studied outcomes. Often, single agent types were modeled, and the environment was usually irrelevant to the studied outcome. Predictive validation and sensitivity analyses were most used to validate models.

Although increasingly used to study NCDs, ABM remains underutilized and, where used, is suboptimally reported in public health studies. Its use in studying NCDs will benefit from clarified best practices and improved rigor to establish its usefulness and facilitate replication, interpretation, and application. (*Am J Public Health.* 2015;105:e20–e31. doi: 10.2105/AJPH.2014.302426)

THERE HAS BEEN AN INCREASING interest in using systems science approaches such as agent-based modeling (ABM) to investigate and understand complex public health problems.[1–4] Complex systems are systems that are not fully explained by just understanding the individual elements of the system.[4] In other words, these systems cannot be reduced to their component parts because of the interactions among the parts.[5]

Complex systems are made of heterogeneous elements or agents (e.g., individuals, organizations) whose interactions with one another yield an unpredictable yet organized emerging behavior that can persist over time.[5–7] When agents are capable of adapting to changing circumstances, the systems are said to be adaptive and thus called complex adaptive systems (CAS).[7,8] Examples of such complex systems include stock markets, insect colonies, immune systems, social systems, traffic jams, epidemics, and pandemics. All these phenomena have been studied in various fields such as economy, ecology, molecular biology, sociology, and epidemiology.[5,9]

Noncommunicable diseases (NCDs) are by far the leading cause of mortality in the world, killing 36 million people in 2008 worldwide, which accounted for about 63% of all deaths.[10] Cardiovascular diseases, cancer, chronic respiratory diseases, and diabetes represent about 80% of all NCD deaths.[10] These diseases constitute a huge health and economic burden across the world. Four main behavioral risk factors—tobacco use, physical inactivity, unhealthy diet, and harmful use of alcohol—are responsible for most NCDs.[10] Noncommunicable diseases are diseases that are not passed from person to person[10] and can have a chronic or acute progression.[11] They differ from chronic diseases in that the latter can be communicable or not and they require a long-term management.[11]

The study of NCDs can be recast as one of complex systems. Noncommunicable diseases are caused by factors that are influenced by one's individual behaviors as well as interaction with the physical, social, or economic environment.[12–14] Researchers have described obesity as a health problem that exhibits attributes that are characteristic of a CAS and have argued that techniques used to model such systems can and should be used to model obesity.[15] Importantly, obesity involves substantial diversity and heterogeneity in relevant actors at many different levels of scales (e.g., individuals, communities, policy), with a multiplicity of mechanisms in which actors interact with one another with dynamic feedback loops and changes over time.[2,15,16]

To study complex systems, traditional analysis (e.g., multivariate analyses) will often not suffice. The latter often assumes linearity (at least on some scale), normality, homogeneity, and independence between individuals and over time, and is concerned with variables often representing a single-level system.[4] This type of analysis is said to be reductionist or top-down.[4] In contrast, complex systems are often nonlinear, nonnormal, and involve heterogeneous actors or agents that interact at different levels with possibility of dynamic feedback loops. These systems approaches are said to be holistic and, in particular, bottom-up in the case of ABM.[17]

Besides ABM, other key systems science approaches have been developed to study complex systems and include systems dynamics and network analysis, as well as discrete event simulation.[4,18] Briefly, system dynamics uses computer simulation models to uncover and understand endogenous sources of complex system behavior.[4] They are based on the premise that complex behaviors of a system result from the interplay of feedback loops, stocks, and flows that all occur within the bounded endogenous system.[4,19] Unlike ABM, which is an individual-based modeling technique, systems dynamics is an aggregate-level modeling type. Network analysis, on the other hand, focuses on the measurement and analysis of relationships and flows among a set of actors.[4] Discrete event simulation is a type of modeling simulation that models the system as a sequence of discrete events over time. It is most known for being used in clinical care settings to determine patient flow through the system.[20]

These systems science approaches have been used for decades in different fields but have only been recently introduced to public health, with the exception of infectious diseases and epidemics.[2] In fact, ABM is most known to public health for its use in the study of epidemics and infectious disease dynamics.[4,21] Unfortunately, the use of ABM in behavioral health and NCDs is relatively new and perhaps lagging.[2,22] Among the few NCDs and related risk factors being explored, physical activity; diet, smoking, and drinking behaviors; and obesity have taken the spotlight.[4,23–25] Increasingly, researchers are advocating the use of such systems science approaches—namely, ABM—in understanding the complexities of NCDs.[2,3,15,16,22]

To fill the gap on whether and how ABM is being used in studying NCDs in public health, we conducted a systematic review examining the use of ABM in understanding various NCDs and their risk factors.

OVERVIEW OF AGENT-BASED MODELING

Agent-based modeling entails computer representations of systems consisting of a collection of microentities (referred to as agents) interacting and changing over time and whose interactions give rise to macrosystems.[4,7,17,26,27] Also called individual-based modeling in ecology,[26] and extensively described by many authors,[8,9,17] ABM consists of 3 key elements. The first element is a set of agents that composes the CAS. Each agent is characterized by specific attributes (e.g., age, gender) and behaviors (e.g., going to school). An agent can be of different natures (e.g., individuals, communities, organizations). They can be autonomous, goal-directed, intelligent (i.e., capable of adapting and learning), heterogeneous, dynamic, and interacting, and their internal states can vary over time. The second element is a set of agent relationships and methods (also known as decision rules or conceptual model) of interaction outlining how agents interact with each other and with their environment and how their internal states evolve over time. The third element is the agent's environment or topology (e.g., spatial location, lattice). Nevertheless, the environment is not always taken into account in building an ABM because it may not be relevant to the process being studied. When considered, the environment can be passive or active with its own dynamic properties and behavioral rules.

Agent-based modeling is particularly useful and attractive when the system being modeled is a CAS; one that involves agents that are autonomous, heterogeneous, and intelligent (e.g., individuals, organizations); whose environment (e.g., spatial location) is crucial and not fixed; and whose dynamic interactions between agents or with their environment give rise to an emergent phenomenon that is complex and nonlinear with feedback loops.[4,5] The ability and flexibility of ABM to capture emergent phenomena, to describe systems as a whole from the bottom up,[17] give it advantages over other analytical techniques.[5,28] Nonetheless, ABM is not without flaws[5,28]:

(1) it can require large amounts of data,
(2) it can be highly computational,
(3) it can be difficult to calibrate and justify its rules,
(4) the verification and valida tion of the model can be difficult to achieve, and
(5) it may have limited scope for reuse in different contexts.

A number of general toolkits can and have been used to perform ABM such as Microsoft Excel (Microsoft, Redmond, WA), the statistical package R (R Foundation for Statistical Computing, Vienna, Austria), MATLAB (MathWorks, Natick, MA), Mathematica (Wolfram Research, Champaign, IL), Java (Oracle Corporation, Redwood City, CA), C++ (Bell Laboratories, Murray Hill, NJ), and Python (Python Software Foundation, Wilmington, DE). Other more specific toolkits include software that are free such as NetLogo (The Center for Connected Learning and Computer-Based Modeling, Northwestern University, Chicago, IL), Repast (Argonne National Laboratory, Argonne, IL), MASON (George Mason University's Evolutionary Computation Laboratory and the George Mason University Center for Social Complexity, Fairfax County, VA), StarLogo (Massachusetts Institute of Technology Media Lab and Massachusetts Institute of Technology Teacher Education Program, Cambridge, MA), and Swarm (Swarm Development Group, Santa Fe, NM), and proprietary ones such as AnyLogic (The AnyLogic Company, St Petersburg, Russian Federation).[8,18]

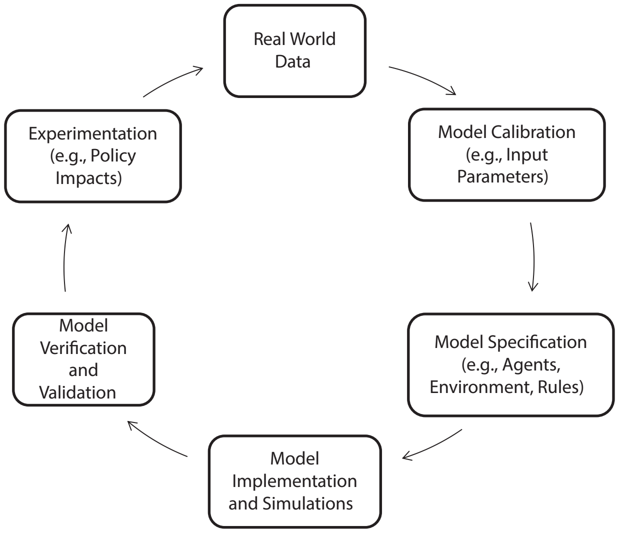
To implement an ABM, an investigator usually follows a number of key steps from abstracting parameters from real-world data to implementing them and testing hypothetical scenarios. In addition, a series of iteratively implemented steps are used to calibrate, verify, and validate the model. These steps are illustrated in Figure 1, adapted from Sargent.[29] We briefly mention the key ones in the next paragraphs.[9,29]

Typically, the real world data set is divided into 2: a training data set used for model development, calibration, and evaluation, and a test data set for validation.[30]

Calibration is concerned with assigning baseline or trend characteristics (also known as input parameters) to the virtual neosystem, by using the training data set. Calibration is part of the internal quantitative credibility.[30] In addition, model evaluation, which is also part of the internal quantitative credibility, is concerned with checking whether the model is well calibrated—that is, how well the projected output using the training data set corresponds with the observed data from the training data set.[30] The input parameters may come from empirical studies, regional surveys, or longitudinal data, for instance. They can be stochastically or deterministically obtained.

Verification is concerned with debugging the model, checking for errors in coding, and making sure that the model does what it is intended to do. Baseline output operation of the codes can be compared with the expectations stated in the design documents. It can be achieved via structured code walk-throughs, debugging walk-throughs, or unit testing, and so on.

Validation is concerned with how accurately the virtual system reflects the real-world system. From a practical standpoint, this refers to how well the model output using inputs from the training data set accurately predicts the observed data in the test data set. This is also referred to as external quantitative credibility.[30] This is done, for instance, in predictive validation. In addition, one could compare the predicted output to historical data whenever available (also known as historical data validation) to assess whether the model behaves as the system does. As a result, one can better calibrate and validate a model by using historical cases. Sensitivity analyses can also be used to assess how stable or sensitive the model is to small changes in the input parameters. In face validity, experts in the field are consulted to assess whether the conceptual model and



Note. Adapted from Sargent.[29]

FIGURE 1—Steps for building an agent-based model.

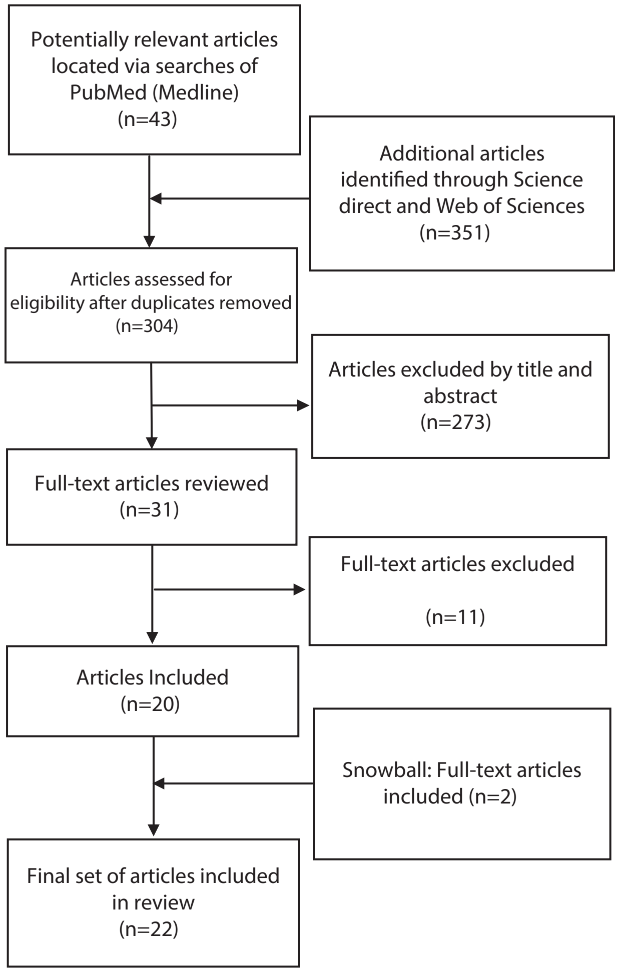


FIGURE 2—Flow diagram showing the process of inclusion and exclusion of articles.

the input–output relationship seem reasonable. Process validation, another important step, but often omitted, is another type of validation that is concerned with whether the steps in the model and the internal flows (e.g., as exemplified in state charts) reflect the real-world processes and behaviors.[9]

METHODS

We conducted a systematic review of studies that used ABM to study NCDs as well as their risk factors in the context of public health. We searched 3 scientific and medical electronic databases—PubMed (Medline), ScienceDirect, and Web of Sciences—to capture articles most relevant to public health. We did not search journals that focused on simulation exclusively. We employed the widely used evidence-based protocol PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) in our review.[31]

In particular, we used the following search keywords and categories: (1) agent-based model, agent-based modeling, or individual-based modeling; (2) noncommunicable diseases, noninfectious diseases, nontransmissible diseases, or chronic diseases; (3) heart disease, diabetes, stroke, cancer, chronic respiratory disease, or neurodegenerative diseases; and (4) unhealthy diet, physical activity, exercise, smoking, alcohol, hyperlipidemia, high cholesterol, high triglyceride, overweight, or obesity. (A detailed search strategy is available in Appendix A, available as a supplement to the online version of this article at http://www.ajph.org).

We included studies published in the English language between January 2003 and July 2014 in our review. Furthermore, we selected only studies that were actual applications (not mere illustrations) of ABM to public health problems. We excluded studies if they pertained to animal research, infectious diseases, molecular biology, immunology, health care management, or ecology, or if they were not freely available. In addition, we reviewed the references of our included studies for additional studies.

After we reviewed the abstracts of all articles that met the inclusion criteria, we extracted the following data from the final set: (1) the first author and year of publication, (2) the study design (goal, outcome studied, exposure or intervention tested), (3) model specification (agents, environment, and conceptual model), and (4) model calibration and validation (platform, calibration, and validation).

RESULTS

Our initial literature search identified 304 citations matching the search criteria. We screened the abstracts of these publications and excluded 273 articles, as they did not meet the inclusion criteria (Figure 2). We retrieved the full texts of 31 articles. We excluded a further 11 from these, leaving

20 articles. We identified 2 additional articles through "snowball searches"—that is, through our knowledge and academic networks.

Study Design

Among the 22 studies included in this review, 8 were observational, 12 were interventional, and the remaining 2 incorporated components of both study designs. The exposures of interest and interventions tested pertained to agents' behaviors, their environment, or the interaction between agents and between agents and their environment. Furthermore, 19 of the studies assessed the effect of multiple exposures or interventions on a particular outcome. Six articles studied were interested in physical activity (e.g., walking behavior) as an outcome, 4 in diet, and 4 in diabetes and related complications or management.

Model Specifications

Of the 22 studies, 5 modeled more than 1 single agent type, and in 17 articles, the agents were individuals (e.g., patients, residents, students; Tables 1 and 2). The physical environment or spatial location was irrelevant to the process under study in half of the studies. When relevant, the spatial location of agents was a virtual representation of a known city. In the latter case, the environment was active, with inherent properties and rules. In addition to textual description of the conceptual model, 7 studies presented diagrams and 13 gave the sets of equations used. (A summary of the results is presented in Appendix B, available as a supplement to the online version of this article at http://www.ajph.org).

Platform, Calibration and Validation

Six studies implemented their model by using the software Repast, 4 used AnyLogic, and 4 used NetLogo. Fifteen studies conducted a predictive validation and 5 studies conducted sensitivity analyses to validate models. One study also performed face and process validation (Table 3).

DISCUSSION

In our systematic review of the use of ABM to study NCDs and their risk factors in public health during the period 2003 to mid-2014, we found 22 studies. A higher proportion of the studies was intended to test the impacts of hypothetical interventions on specified outcomes (i.e., interventional studies). Physical activity, diet, and diabetes-related complications and management were the most-studied outcomes. Often, single agent types were modeled and, in most cases, the environment was irrelevant to the outcome under study. Predictive validation and sensitivity analyses were most used to validate models. Almost all of the studies in this review incorporated some version of feedback loop in their modeling.

Model Design

Most of the studies reviewed here had an observational,[24,23,32–37] interventional,[38–49] or hybrid[40,51] (or mixed) design. In the former type of ABM studies, the authors investigated the contribution of known causal processes,[37,40] and examined emerging phenomena (e.g., drinking behaviors) as a result of agent–environment interactions.[23,33,35] In the interventional or experimental ABM designs, however, the investigators have sought to predict "What would happen to a population's outcome level if they could implement a certain intervention in this population?" In such studies included in this review, the authors examined the impact or effect of hypothetical interventions or policies on health behaviors and NCDs. For instance, Yang et al. examined the impact of policies aimed at changing people's attitude toward walking as well as policies improving the contextual safety level on residents' walking behavior.[40] In this case and others, certain interventions cannot be implemented in real life because of ethical considerations and cost in time and money.

Agent-based modeling offers researchers the added benefit of conducting virtual randomized controlled trials wherein dynamic interactions are accounted for. The hypothetical intervention is then implemented as a counterfactual "what if" scenario and the potential outcome under intervention is contrasted with the actual outcome under no intervention. This makes ABM very appealing especially to epidemiologists who are most concerned with causality but have to resort to observational studies.

Model Specification

Modeling a single agent type as individuals with inherent properties who are placed in a physical environment half the time was the most common instance observed in this review. These agents interacted among one another[38,40,41,44,46,47] or with their environment[23,24,34–36] to yield a dynamic and emergent behavior. As seen in this review, the environment or spatial location wherein the agent was placed was not always described.[33,39,43,50] This can occur, for instance, when the environment is irrelevant to the experiment and conclusions. In particular, the modeling team may not be interested in investigating the interactions between agents and their environment but rather is interested in the interactions among agents only or their evolvement through time.

Moreover, the conceptual model (i.e., rules governing the agents' behaviors and their interactions) was explained by using textual description in all the studies included in this review. Less often was the conceptual model also described with diagrams or equations.[23,34,44,45] This is unfortunate as these visual aids and equations can help with better visualizing the problem under study and understanding the logic behind the modeling process. It is often said that "a picture is worth 1000 words," yet this view has been poorly used when studies conveyed complex phenomena in ABM in the context of public health. One could even augment the graph by adding equations, parameters, or transitional probabilities from state to state, for instance, using state charts as done in Day et al.[32]

Likewise, this approach was done in the modeling of drinking behavior by Gorman et al. in which the underlying rules were borrowed from the well-known Susceptible–Infected–Recovered model widely used in infectious diseases and epidemics.[23] Others have also used diagrams when representing the modeling structure in NCDs.[51–53] In light of the previous modeling structures, we illustrate an example of a conceptual model of diabetes progression (Figure 3). In addition to conveying well the conceptual model, one has to make sure to provide basis and justification for the decision rules, especially the equations underlying the relationships between and within agents. To conclude, it is critically important that authors explicitly specify the equations underlying their system to improve transparency of what was done and aid replication if necessary.

TABLE 1—Study Design and Model Specification

	Study Design			Model Specification		
Study, Year	Goal of the Model	Exposure or Intervention Tested	Outcome	Agents (Types and Properties)	Environment	Conceptual Model
Day et al., 2013[32]	To develop a template for a cohort of patients with or at risk for diabetic retinopathy	None (observation of the cohort over time as patients age)	Progression to diabetic retinopathy	Patients (age, diabetes, BMI, HbA1c, hypertension, state of diabetic retinopathy, smoking)	ND	Diagram Equations Text
Yang et al., 2011[24]	To investigate the contributions of built and social environments to SES differences in people's walking behavior within a city	Spatial distribution of nonhousehold locations Spatial distribution of safety SES residential segregation	Walking behavior	Residents (gender, age, SES, family size, friends, workplace, walking ability)	Ann Arbor, MI	Equations Text
Gorman et al., 2006[23]	To examine agent–environment interactions that support the development and maintenance of drinking behavior at the population level	Bar topology; agent motion Probability of a former drinker to resume drinking, and a current drinker to stop drinking	Drinking behavior	Individuals (location on the lattice, drinking status)	NS	Diagram Equations Text
Veloso, 2013[50]	To predict individual long-term disability (i.e., prognosis) and examine the effect of treatment among patients with multiple sclerosis	Treatment (e.g., Avonex [biogen idec, Cambridge, MA]) Prognostic factors (e.g., age at disease onset, gender, sphincter onset, pure motor onset, sequel after onset)	Progression to multiple sclerosis	Patients (gender, age at disease onset, BREMS scores, EDSS states, QALY)	ND	Text
Yang and Diez Roux, 2013[38]	To explore how various policies may influence children's active travel to school	Traffic-level safety Catchment area definition School location School size; population density	Walking behavior	Households (concern about traffic safety) Child (attitude toward walking)	NS	Equation Text
Subramanian et al., 2009[39]	To assess the most cost-effective approach to improve outcomes from colorectal cancer screening at the population level	Patient's compliance (initial and at diagnostic follow-up) Level of screening Patient's preference	Cost-effectiveness of colorectal cancer screening tests	Patients (colorectal cancer: presence, average risk, stage; screening: types, preference)	ND	Diagram Text
Verella and Patek, 2009[33]	To investigate the effects of the interactions between patient and physician on the adoption of continuous glucose monitoring	Regulatory approval for decision-making Favorable publication Affordability, effectiveness of continuous glucose monitoring Insurance coverage	Adoption of continuous glucose monitoring in diabetes-type 1 management	Patients and physicians (interest in continuous glucose monitoring) Device manufacturer (cost, affordability)	ND	Equation Text
Yang et al., 2012[40]	To examine the possible impact of interventions on socioeconomic differences in walking	People's attitude toward walking Safety level Modification of mixed land use	Physical activity or walking behavior	Residents (gender, age, SES, family size, friends, workplace, walking ability)	Ann Arbor, MI	Text
Auchincloss et al., 2011[25]	To explore the role of economic segregation in creating income disparities in diet, and policy levers that may be appropriate for countering them	Spatial segregation Healthy food preference and relative price	Income differential in diet	Households (income, food preference) Food stores (price and types of food)	NS	Text

Continued

TABLE 1—*Continued*

Schaefer et al., 2013[41]	To evaluate alternative intervention scenarios in the smoking behaviors	Peer influence on smoking Popularity of smokers relative to nonsmokers	Smoking behavior	Students (smoking rate, gender, age, parent smoking, grade point average, delinquency, alcohol)	ND	Text
Jin and White, 2012[34]	To explore the influences of neighborhood design on trip and traffic patterns with an emphasis on pedestrian movements	Neighborhood design types (e.g., traditional grid, suburban) Specific design features (e.g., pedestrian-only routes, locations of facilities)	Trip and traffic patterns Walking behavior	Residents (individual perception of utility; knowledge about road or traffic)	Ottawa, Ontario	Diagram Equations Text
Widener et al., 2013[42]	To simulate the impact of various policy interventions on low-income households' consumption of fruits and vegetables	Introduction of mobile markets, farmers market; shopping frequency Percentage of convenience stores that sell fresh fruits and vegetables	Consumption of fruits and vegetables	Households (size, location, fruits and vegetables shopping frequency and probability)	Buffalo, NY	Equations Text
Yin, 2013[35]	To explore the patterns of walkability arising from microlevel interactions between people and their built environment across the city	Interactions between people and the social environment Interactions between people and their built environment	Walking behavior	Residents (SES, walking speed)	Buffalo, NY	Equations Text
Rein et al., 2007[43]	To determine the cost-effectiveness of vitamin therapy (antioxidants plus zinc) for all indicated patients diagnosed with age-related macular degeneration	Vitamin therapy (antioxidants plus zinc)	Cost-effectiveness of vitamin therapy	Patients (demographics, longevity, use of ophthalmologic care)	ND	Diagram Text
Garrison and Babcock, 2009[36]	To model student drinking behavior	None	Drinking behavior	Students (attitude toward drinking, shyness, class rank, residence, friends)	ND	Equations Text
Zhang et al., 2014[44]	To examine the impact of different policies on unhealthy eating behaviors	Tax on unhealthy food Subsidies for healthy food Promotion of healthy norms Regulation of local food environment	Unhealthy eating behavior	Individuals (health belief, demographics) Food outlets (types of food sold)	Pasadena, CA	Diagram Equations Text
Day et al., 2014[45]	To examine the effect of changes to screening interval on the incidence of vision loss in a simulated cohort of veterans with diabetic retinopathy	Screening interval	Vision loss incidence	Patients (age, diabetes, BMI, HbA1c, hypertension, state of diabetic retinopathy, smoking)	ND	Diagram Equations Text
Yang et al., 2014[46]	To examine the impact of the walking school bus on children's active travel to school	Walking school bus plus Education campaign Walking speed and waiting time Bus route placements	Walking behavior	Households (concern about traffic safety) Child (attitude toward walking)	NS	Text
Orr et al., 2014[47]	To explore the efficacy of a policy that improved the quality of neighborhood schools in reducing racial disparities in obesity-related behavior	School quality policy Social network effect Social norm type	Diet behavior	Residents (age, race, school attendance, education, diet, neighborhood attributes)	Large US metropolitan areas	Diagram Text Equations

Continued

TABLE 1—Continued

Hammond and Ornstein, 2014[37]	To test whether social influence through body weight norms (i.e., "follow the average") can independently support the development and persistence of obesity	Social influence	BMI	Individuals (age, sex, network body image, actual, ideal body image, BMI)	ND	Text
Zhang et al., 2014[48]	To gain insights into what network mechanisms are salient for obesity and which obesity-related approaches might leverage social networks	Social influence (e.g., peer selection, strength of peer influence, targeted weight loss in the overweight population)	Overall overweight or obesity prevalence	Individuals (network parameters, age, grade, sex, household income, BMI)	ND	Diagram Text Equations
Li et al., 2014[49]	To evaluate the effect of a lifestyle program on short- and long-term health outcomes in a primary care practice serving a Medicare-age population	Population health management (e.g., diet and exercise improvement and weight reduction)	Diabetes Hypertension High cholesterol	Medicare patients (age, sex, smoking status, BMI, physical activity, diet)	ND	Text

Note. BMI = body mass index; BREMS = Bayesian risk estimate for multiple sclerosis; EDSS = Expanded Disability Status Scale; HbA1c = hemoglobin A1C; ND = not described; NS = nonspecific; QALY = quality-adjusted life year; SES = socioeconomic status; S-I-R = Susceptible, Infected, Recovered.

Platform and Programs

As seen in this review, Repast, AnyLogic, and NetLogo were the platforms that most researchers used to implement their ABM. The choice of these specific software packages was possibly motivated by a number of factors such as cost, ease of use, power, and preference. Some of the differences observed in the use of these platforms have been extensively outlined in the literature.[6,8,18,29,54–57] For instance, NetLogo is freely available and offers a relative ease of learning and use for beginning modelers. However, it does not scale well to larger and more complex models. Repast, on the other hand, is also freely available, offers an added flexibility, and tends to scale more effectively than NetLogo. However, Repast requires a little more knowledge to use than NetLogo.[6] AnyLogic is a proprietary software that is characterized by its added flexibility, its use of informative visual aids, its tendency to scale more effectively, and its ability to incorporate hybrid modeling (e.g., ABM and systems dynamics, network analysis, and discrete event simulation).[58]

Although some platforms may offer apparent advantages compared with others, they may not always be accessible to all and may require learning a new language. This can be a hindrance for researchers with a long experience in analytical research but who are new to ABM and wish to incorporate it into their work. Admittedly, there exist today appropriate software tailored just for ABM, yet it might be easier for the long-time SAS users and public health professionals to implement an ABM within their general statistical packages if and when possible.

Calibration and Validation

In our review, with the exception of a few models that were not calibrated to real-world data and used abstract or arbitrary scenarios instead,[33,36] almost all models abstracted their initial parameters from real-world data. The use of such arbitrary parameters has led some to call complex systems dynamic approaches a "fact-free science."[59] In some cases in this review, the parameters were pulled from empirical studies,[23,50] whereas others have estimated them through multivariate regression modeling.[32,45] Most included studies used multiple data sources such as national surveys and longitudinal data.[43,47–49] This finding is corroborated by some authors, including Auchincloss and Diez Roux, who have also suggested using a diversity of sources (whenever possible) to better calibrate an ABM.[3] In addition, as previously noted by Grimm et al., the use of stochasticity in the input parameters should be specified in the description of an ABM.[60,61] More importantly, reporting input parameters such as random numbers (and seeds) as well as the probability distributions used can be vital for full model reproducibility.[60,61]

In this review, most authors have used predictive validation[24,32,38,40,50] or sensitivity analyses[25,39,47] to validate their models. This is encouraging as model validation is a paramount step to implementing a credible agent-based model and failure to do so may prevent readers from trusting the validity of the results. As a consequence, transparency and making assumptions explicit should guide future ABM-based research. To better achieve this transparency, we propose a set of checklists or guidelines to assess

TABLE 2—Platform and Model Calibration and Validation

Study, Year	Platform[a]	Parameters and Calibration	Model Validation
Day et al., 2013[32]	AnyLogic	Agent deaths (life table by the SSA)	Predictive validation (comparison of the simulated data with a real-world test cohort)
		Prevalence of diabetic retinopathy (empirical studies)	
		Patient records and predicted probability of diabetic retinopathy progression (from a retrospective cohort of current patients of the VA St Louis Healthcare System eye clinic)	
Yang et al., 2011[24]	Java and Repast	Ann Arbor features, agent's properties (census data)	Predictive validation (comparison of the simulated data with the NHTS)
		Daily probabilities of performing the activity and maximum walking distances for different activities (NHTS, 2001)	
Gorman et al., 2006[23]	Matlab	Agent motion, natural tendency of a former drinker to resume drinking, natural tendency of a current drinker to stop drinking (from empirical studies)	Not specified
		Bar locations and number (arbitrarily chosen)	
Veloso, 2013[50]	NetLogo	Disease progression data: years from disease onset to EDSS scores 4, 5, 6 (Lyon database)	Predictive validation (comparison of the simulation data with the real value presented by the patients observed by the authors, for 10- and 20-y estimations)
		Relapse-free and progression-free at 2 y of treatment (empirical studies)	
Yang and Diez Roux, 2013[38]	Java and Repast	Distance decay parameter; % of children who walk to school, total number of people who walk (NHTS, empirical studies)	Predictive validation (comparison of the model output with the data on the percentages of children who walk to school)
		Traffic safety level; household concern toward traffic	
Subramanian et al., 2009[39]	AnyLogic	Progression rates through cancer stages and probability of presenting colorectal cancer symptoms (from SEER, 1988); deaths (National Center for Health Statistics)	Sensitivity analyses done to assess the effect of varying test performance on model results
		Baseline compliance and preference (NHIS, 2003, 2005)	
		Effectiveness and cost parameters (from empirical studies)	
Verella and Patek, 2009[33]	NetLogo	Not calibrated to a specific population (purely experimental)	No validation
Yang et al., 2012[40]	Java and Repast	Ann Arbor features, agent's properties (census data)	Predictive validation (comparison of the simulated data at baseline with the NHTS)
		Daily probabilities of performing the activity and maximum walking distances for different activities (NHTS, 2001)	
Auchincloss et al., 2011[25]	Repast (code supplied)	Behaviors were guided with econometrics literature and observational and survey data	Face validity; process validation (comparison of agent behaviors with available data to reflect intuitive and known behaviors)
		Attributes at baseline were randomly chosen (stochastic)	Sensitivity analyses
Schaefer et al., 2013[41]	R-SIENA (R package)	Smoking function and network function (using data from 1 school in the National Longitudinal Study of Adolescent Health); obtained via simulations	Predictive validation (comparison of the model baseline parameters with observed data)
Jin and White, 2012[34]	Repast OpenMap GIS	Trip survey data (from 7 Ottawa traffic analysis zones)	Predictive validation (comparison of the simulated model output with observed data)
		Probability of choosing a particular route (using a maximum likelihood approach)	
		SES factors (census data)	
		Route characteristics for driving and for walking	
Widener et al., 2013[42]	Not specified	Percentage of households participating in SNAP; percentage and probability of purchasing fresh fruits and vegetables at a supermarket (USDA)	Not specified
		Location of food vendors (Hoover's business directory, NY)	
		Household size and location (city of Buffalo, NY, 2010); probability of shopping for food	
Yin, 2013[35]	ArchObjects	Agent and environment attributes (from the city of Buffalo, NY, US Census, NY state GIS clearinghouse, Department of Education, Greater Buffalo–Niagara Regional Transportation Council)	Predictive validation (comparison of model output: pedestrian count data with real-world walkability data)

Continued

TABLE 2—*Continued*

Rein et al., 2007[43]	AnyLogic	Prevalence of AMD in the United States (Eye Diseases Prevalence Research Group)	Predictive validation (comparison of the baseline output with the National Eye Institute data)
		Annual incidence of AMD; ages and initial state of AMD	
		Use of ophthalmic services (2002 National Ambulatory Medical Care)	Sensitivity analysis
		Stochastic	
Garrison and Babcock, 2009[36]	C++	Most equations and numbers are arbitrarily chosen	Predictive validation (comparison of model outputs with existing literature)
Zhang et al., 2014[44]	NetLogo	Factors influencing individual decision-making such as demographics, taste preference, health beliefs, food price index, price sensitivity, food accessibility (from 2000 US Census, 2007 Food Attitudes and Behaviors Survey, and empirical studies)	Predictive validation (comparison of the model baseline output with the Los Angeles County Survey 2007)
		Environment: Pasadena, CA (2010 US Census)	Sensitivity analyses
Day, 2014[45]	AnyLogic	Agent deaths (life table by the SSA)	Predictive validation (comparison of the simulated data with a real-world test cohort)
		Prevalence of diabetic retinopathy (empirical studies)	
		Patient records and predicted probability of diabetic retinopathy progression (from a retrospective cohort of current patients of the VA St Louis Healthcare System eye clinic)	Area under the curve
			Hosmer–Lemeshow partition
Yang et al., 2014[46]	Repast and Java	Same as in Yang and Diez Roux, 2013[38]	Same as in Yang and Diez Roux, 2013[38]
		Distance decay parameter; percentage of children who walk to school, total number of people who walk (NHTS, empirical studies)	Predictive validation (comparison of the model output with the data on the percentages of children who walk to school)
		Traffic safety level; household concern toward traffic	
Orr et al., 2014[47]	Not specified	Racial/ethnic and economic distribution (empirical studies, 2011 US Census)	Sensitivity analyses (varying the size of various influences)
		Food availability, activity level (2007 NHIS), smoking pattern (2010 NSDUH)	
		BMI levels (2007 NHIS), cardiovascular health, death rates (2007 NHIS, Health USA)	
		Healthy Eating Index, influence of activity (empirical studies)	
Hammond and Ornstein, 2014[37]	Not specified	Satisficing rule, networks (Quebec En Forme)	Conceptual model validation (empirical analyses)
		Physiological equations (e.g., resting energy expenditure) from empirical studies	Predictive validation (comparison of the model output with real-world longitudinal data)
		Simulations calibrated to longitudinal data (NLSY97)	
		Agents' sex, age, height at baseline randomly assigned	
Zhang et al., 2014[48]	R-SIENA NetLogo	Network characteristics abstracted from the National Longitudinal Study of Adolescent Health	Predictive validation (comparing simulated results at the end of the model run with empirical data)
Li et al., 2014[49]	Not specified	Agents' attributes abstracted from the BRFSS	Predictive validation (comparing simulated and actual health outcomes using the BRFSS)

Note. AMD = age-related macular degeneration; BMI = body mass index; BRFSS = Behavioral Risk Factor Surveillance System; EDSS = Expanded Disability Status Scale; GIS = geographic information system; NCHS = National Center for Health Statistics; NHIS = National Health Interview Survey; NHTS = National Household Travel Survey; NLSY = National Longitudinal Survey of Youth; NSDUH = National Survey on Drug Use and Health; SEER = Surveillance, Epidemiology, and End Results; SNAP = Supplemental Nutrition Assistance Program; SSA = Social Security Administration; USDA = United States Department of Agriculture; VA = Veterans Affairs.
[a]"Codes supplied" only mentioned when it is the case.

TABLE 3—Properties of Complex Adaptive Systems Addressed in the Included Studies

	Agents[a]		Environment[b]		Dynamic Interactions		
Study, Year	More Than 1 Type	Varying States	Spatial Location	Active	Interactions Among Agents	Interactions Between Agents and Environment	Feedback Loops
Interventional studies (n = 12)							
Yang and Diez Roux,[38] 2013	✓		✓	✓	✓	✓	✓
Subramanian et al.,[39] 2009		✓					✓
Yang et al.,[40] 2012			✓	✓	✓	✓	✓
Schaefer et al.,[41] 2013					✓		✓
Widener et al.,[42] 2013			✓	✓		✓	✓
Rein et al.,[43] 2006		✓					✓
Zhang et al.,[44] 2014	✓		✓	✓	✓	✓	✓
Day et al.,[45] 2014		✓					✓
Yang et al.,[46] 2014	✓		✓	✓	✓	✓	✓
Orr et al.,[47] 2014		✓	✓	✓	✓	✓	✓
Zhang et al.,[48] 2014					✓		✓
Li et al.,[49] 2014							✓
Observational studies (n = 8)							
Day et al.,[32] 2013		✓					✓
Yang et al.,[24] 2011			✓	✓	✓	✓	✓
Gorman et al.,[23] 2006		✓	✓	✓	✓	✓	✓
Verella and Patek,[33] 2009	✓				✓		✓
Jin and White,[34] 2012			✓	✓	✓	✓	✓
Yin et al.,[35] 2013			✓	✓	✓	✓	✓
Garrison and Babcock,[36] 2008		✓			✓	✓	✓
Hammond and Ornstein,[37] 2014					✓		✓
Mixed (n = 2)							
Veloso et al.,[50] 2013		✓					
Auchincloss et al.,[25] 2011	✓		✓	✓	✓	✓	✓
Total studies with attributes, no. (%)	5 (23)	8 (36)	11 (50)	11 (50)	15 (68)	12 (55)	21 (95)

[a]All agents in the included studies are heterogeneous with respect to sociodemographics and other properties.
[b]The environment refers to the physical environment.

whether a model is valid or reliable. These guidelines can be found in Appendix C and D (available as supplements to the online version of this article at http://www.ajph.org) and have been inspired by the works of Sargent as well as North and Macal.[9,29]

Other templates or protocols for describing ABMs, such as the Overview Design Details protocol, have been previously described by Grimm et al. and widely implemented.[60,61] Briefly, the 7 elements of the updated Overview Design Details protocol are as follows: (1) overview—(i) purpose; (ii) entities, state variables, and scales; and (iii) process overview and scheduling; (2) design concepts—(iv) design concepts and basic principles (emergence, adaptation, objectives, learning, prediction, sensing, interaction, stochasticity, collectives, observation); (3) details—(v) initialization, (vi) input data, and (vii) submodels. More in-depth discussion and definitions of the different elements of the Overview Design Details protocol are given in Grimm et al.[60,61]

Strengths and Limitations

The strengths of this review include its thorough systematic nature, the moderate to large number of studies found, and the detailed and structured information extracted from the studies retrieved.

To our knowledge, this is one of the first comprehensive reviews of ABM of NCDs in public health, but a few limitations should be considered. First, we limited our review to studies published from 2003 onward as ABM is relatively new to public health and especially to NCD epidemiology. Second, by limiting our review to the English language, we may have missed relevant articles published in other languages. Third, we only searched public health and clinical research journals, and might have missed few yet relevant articles in other journals. Finally, we only searched terms that were synonyms to agent-based and individual-based modeling intentionally omitting from the final search query other nonspecific terms such as computer simulation, in silico, nonlinear dynamics, chaos theory, and complex adaptive systems because such searches yielded very large unfocused numbers of articles not dealing with NCDs.

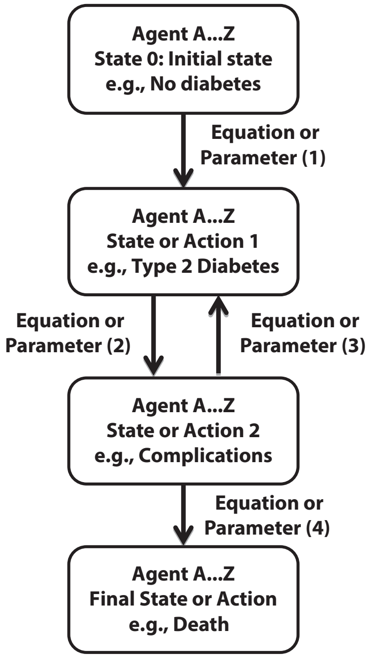


FIGURE 3—Illustration of a conceptual model underlying the rules governing the movements from state to state by using a state chart model.

Conclusion

The use of ABM in studying and understanding NCDs and risk factors is slowly growing. Much of the attention so far in this area has revolved around physical activity, diet, and diabetes-related complications and management, all of which exhibit attributes of CAS. Researchers have modeled individuals, and their interactions with one another and their physical environment to study the emergent phenomena arising from these interactions. In these ABMs, hypothetical interventions were implemented in a virtual randomized controlled study wherein dynamic interactions are accounted for. Systematic presentation of studies as well as use of diagrams and equations to better represent the modeling process was somewhat missing in a number of the studies included in this review. Admittedly, conducting an ABM study is not an easy task and researchers are dealing with the difficulty of creating relatively simple ABMs that capture complex phenomena. As a consequence, a systematic, rigorous, and transparent guideline for using this new tool is necessary to improve its usefulness, and for facilitating study replication and application. ■

About the Authors

Roch A. Nianogo and Onyebuchi A. Arah are with the Department of Epidemiology, Fielding School of Public Health, University of California, Los Angeles (UCLA). Onyebuchi A. Arah is also with the Center for Health Policy Research, UCLA, and the California Center for Population Research, UCLA, as well as the Academic Medical Center, University of Amsterdam, The Netherlands.

Correspondence should be sent to Roch A. Nianogo, Department of Epidemiology, UCLA Fielding School of Public Health, 650 Charles E. Young Dr S, Los Angeles, CA 90095-1772 (e-mail: niaroch@ucla.edu). Reprints can be ordered at http://www.ajph.org by clicking the "Reprints" link.

This article was accepted October 25, 2014.

Contributors

R. A. Nianogo contributed to the study design, conducted the data extraction and analysis, led the interpretation of results, drafted the initial article, and revised the article. O. A. Arah conceptualized and designed the study, contributed to data analysis and interpretation of results, edited the article, and supervised the study. Both authors critically edited the article for intellectual content.

Acknowledgments

R. A. N was supported by the UCLA Department of Epidemiology and the UCLA Fielding Fellowship for Epidemiology. O. A. A. was partially supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (grant 1R01HD072296-01A1) and by a Veni Career Grant (916.96.059) from the Netherlands Organization for Scientific Research.

Note. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of this article.

Human Participant Protection

Institutional review board approval was not required because this research did not involve human participants.

References

1. Nigel G, Klaus GT. *Simulation for the Social Scientist.* Buckingham, England: Open University Press; 1999.
2. Galea S, Riddle M, Kaplan GA. Causal thinking and complex system approaches in epidemiology. *Int J Epidemiol.* 2010;39(1):97–106.
3. Auchincloss AH, Diez Roux AV. A new tool for epidemiology: the usefulness of dynamic-agent models in understanding place effects on health. *Am J Epidemiol.* 2008;168(1):1–8.
4. Luke DA, Stamatakis KA. Systems science methods in public health: dynamics, networks, and agents. *Annu Rev Public Health.* 2012;33:357–376.
5. Bonabeau E. Agent-based modeling: methods and techniques for simulating human systems. *Proc Natl Acad Sci USA.* 2002;99(suppl 3):7280–7287.
6. Gilbert N. *Agent-Based Models: Quantitative Applications in the Social Science.* Los Angeles, CA: SAGE Publications; 2008.
7. Miller JH, Page SE. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life.* Princeton, NJ: Princeton University Press; 2007.
8. Macal CM, North MJ. Tutorial on agent-based modelling and simulation. *J Simul.* 2010;4(3):151–162.
9. North MJ, Macal CM. *Managing Business Complexity: Discovering Strategic Solutions With Agent-Based Modeling and Simulation.* New York, NY: Oxford University Press; 2007.
10. *Global Status Report on Noncommunicable Diseases 2010: Description of the Global Burden of NCDs: Their Risk Factors and Determinants.* Geneva, Switzerland: World Health Organization; 2011: 176.
11. Unwin N, Epping Jordan J, Bonita R, Ackland M, Choi BC, Puska P. Rethinking the terms non-communicable disease and chronic disease. *J Epidemiol Community Health.* 2004;58(9):801.
12. Barr VJ, Robinson S, Marin-Link B, et al. The expanded chronic care model: an integration of concepts and strategies from population health promotion and the chronic care model. *Hosp Q.* 2003;7(1):73–82.
13. Alamian A, Paradis G. Individual and social determinants of multiple chronic disease behavioral risk factors among youth. *BMC Public Health.* 2012;12(1):224.
14. Blas E, Kurup AS. *Equity, Social Determinants and Public Health Programmes.* Geneva, Switzerland: World Health Organization; 2010. Available at: http://whqlibdoc.who.int/publications/2010/9789241563970_eng.pdf. Accessed January 1, 2014.
15. Hammond RA. Complex systems modeling for obesity. *Prev Chronic Dis.* 2009;6(3):A97.
16. Huang TT, Drewnoski A, Kumanyika S, Glass TA. A systems-oriented multilevel framework for addressing obesity in the 21st century. *Prev Chronic Dis.* 2009;6(3):A82.
17. Epstein JM, Axtell R. *Growing Artificial Societies: Social Science From the Bottom Up.* Washington, DC: Brookings Institution Press; 1996:208.
18. Macal CM, North MJ. Agent-based modeling and simulation. In: *Proceedings of the 2009 Winter Simulation Conference.* Piscataway, NJ: Institute of Electrical and Electronics Engineers Press; 2009: 86–98.
19. Richardson GP. Reflections on the foundations of system dynamics. *Syst Dyn Rev.* 2011;27(3):219–243.

20. Connelly LG, Bair AE. Discrete event simulation of emergency department activity: a platform for system-level operations research. *Acad Emerg Med.* 2004;11(11):1177–1185.

21. Perez L, Dragicevic S. An agent-based approach for modeling dynamics of contagious disease spread. *Int J Health Geogr.* 2009;8:50.

22. Ness RB, Koopman JS, Roberts MS. Causal system modeling in chronic disease epidemiology: a proposal. *Ann Epidemiol.* 2007;17(7):564–568.

23. Gorman DM, Mezic J, Mezic I, Gruenewald PJ. Agent-based modeling of drinking behavior: a preliminary model and potential applications to theory and practice. *Am J Public Health.* 2006;96(11):2055–2060.

24. Yang Y, Diez Roux AV, Auchincloss AH, Rodriguez DA, Brown DG. A spatial agent-based model for the simulation of adults' daily walking within a city. *Am J Prev Med.* 2011;40(3):353–361.

25. Auchincloss AH, Riolo RL, Brown DG, Cook J, Diez Roux AV. An agent-based model of income inequalities in diet in the context of residential segregation. *Am J Prev Med.* 2011;40(3):303–311.

26. Railsback SF, Grimm V. *Agent-Based and Individual-Based Modeling.* Princeton, NJ: Princeton University Press; 2012.

27. Page SE. Agent-based models. In: Durlauf SN, Blume LE, eds. *The New Palgrave Dictionary of Economics.* 2nd ed. New York, NY: Palgrave Macmillan; 2008.

28. Nilsson F, Darley V. On complex adaptive systems and agent-based modelling for improving decision-making in manufacturing and logistics settings: experiences from a packaging company. *Int J Oper Prod Manage.* 2006;26(12):1351–1373.

29. Sargent RG. Verification and validation of simulation models. In: *Proceedings of the 2007 Winter Simulation Conference.* Piscataway, NJ: Institute of Electrical and Electronics Engingeers; 2007: 124–137.

30. Berk R. How you can tell if the simulations in computational criminology are any good. *J Exp Criminol.* 2008;4(3):289–308.

31. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 2009;6(7):e1000097.

32. Day TE, Ravi N, Xian H, Brugh A. An agent-based modeling template for a cohort of veterans with diabetic retinopathy. *PLoS ONE.* 2013;8(6):e66812.

33. Verella JT, Patek SD. Toward an agent-based patient–physician model for the adoption of continuous glucose monitoring technology. *J Diabetes Sci Technol.* 2009;3(2):353–362.

34. Jin X, White R. An agent-based model of the influence of neighbourhood design on daily trip patterns. *Comput Environ Urban Syst.* 2012;36(5):398–411.

35. Yin L. Assessing walkability in the City of Buffalo: application of agent-based simulation. *J Urban Plann Dev.* 2013;139(3):166–175.

36. Garrison LA, Babcock DS. Alcohol consumption among college students: an agent-based computational simulation. *Complexity.* 2009;14(6):35–44.

37. Hammond RA, Ornstein JT. A model of social influence on body mass index. *Ann N Y Acad Sci.* 2014; 1331(1):34–42.

38. Yang Y, Diez Roux AV. Using an agent-based model to simulate children's active travel to school. *Int J Behav Nutr Phys Act.* 2013;10(1):67.

39. Subramanian S, Bobashev G, Morris RJ. Modeling the cost-effectiveness of colorectal cancer screening: policy guidance based on patient preferences and compliance. *Cancer Epidemiol Biomarkers Prev.* 2009;18(7):1971–1978.

40. Yang Y, Diez Roux AV, Auchincloss AH, Rodriguez DA, Brown DG. Exploring walking differences by socioeconomic status using a spatial agent-based model. *Health Place.* 2012;18(1):96–99.

41. Schaefer DR, Adams J, Haas SA. Social networks and smoking: exploring the effects of peer influence and smoker popularity through simulations. *Health Educ Behav.* 2013;40(1, suppl):24S–32S.

42. Widener MJ, Metcalf SS, Bar-Yam Y. Agent-based modeling of policies to improve urban food access for low-income populations. *Appl Geogr.* 2013;40:1–10.

43. Rein DB, Saaddine JB, Wittenborn JS, et al. Cost-effectiveness of vitamin therapy for age-related macular degeneration. *Ophthalmology.* 2007;114(7):1319–1326.

44. Zhang D, Giabbanelli PJ, Arah O, Zimmerman JF. Impact of different policies on unhealthy dietary behaviors in an urban adult population: an agent-based simulation model. *Am J Public Health.* 2014;104(7):1217–1222.

45. Day TE, Ravi N, Xian H, Brugh A. Sensitivity of diabetic retinopathy associated vision loss to screening interval in an agent-based/discrete event simulation model. *Comput Biol Med.* 2014;47:7–12.

46. Yang Y, Diez-Roux A, Evenson KR, Colabianchi N. Examining the impact of the walking school bus with an agent-based model. *Am J Public Health.* 2014;104(7):1196–1203.

47. Orr MG, Galea S, Riddle M, Kaplan GA. Reducing racial disparities in obesity: simulating the effects of improved education and social network influence on diet behavior. *Ann Epidemiol.* 2014;24(8):563–569.

48. Zhang J, Tong L, Lamberson PJ, Durazo-Arvizu RA, Luke A, Shoham DA. Leveraging social influence to address overweight and obesity using agent-based models: the role of adolescent social networks. *Soc Sci Med.* 2014: 1–11.

49. Li Y, Kong N, Lawley MA, Pagan JA. Using systems science for population health management in primary care. *J Prim Care Community Health.* 2014;5(4):242–246.

50. Veloso M. An agent-based simulation model for informed shared decision making in multiple sclerosis. *Mult Scler Relat Disord.* 2013;2(4):377–384.

51. Atanasijević-Kunc M, Drinovec J, Ručigaj S, Mrhar A. Simulation analysis of coronary heart disease, congestive heart failure and end-stage renal disease economic burden. *Math Comput Simul.* 2011;82(3):494–507.

52. Weinstein MC, Coxson PG, Williams LW, Pass TM, Stason WB, Goldman L. Forecasting coronary heart disease incidence, mortality, and cost: the Coronary Heart Disease Policy Model. *Am J Public Health.* 1987;77(11):1417–1426.

53. Barhak J, Isaman DJM, Ye W, Lee D. Chronic disease modeling and simulation software. *J Biomed Inform.* 2010;43(5):791–799.

54. Lytinen SL, Railsback SF. The evolution of agent-based simulation platforms: a review of NetLogo 5.0 and ReLogo. In: *European Meetings on Cybernetics and Systems Research.* Vienna, Austria: Bertalanffy Center for the Study of Systems Science; 2010.

55. Railsback SF, Lytinen SL, Jackson SK. Agent-based simulation platforms: review and development recommendations. *Simulation.* 2006;82(9):609–623.

56. North MJ, Collier NT, Ozik J, et al. Complex adaptive systems modeling with Repast Simphony. *Complex Adapt Syst Model.* 2013;1:3.

57. Macal CM, North MJ. Agent-based modeling and simulation: ABMS examples. In: *Proceedings of the 2008 Winter Simulation Conference.* Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2008: 101–112.

58. The AnyLogic company. AnyLogic. 2014. Available at: http://www.anylogic.com/overview. Accessed March 2, 2014.

59. Smith J. Review of "The Origins of Order." New York, NY: *New York Rev Books.* 1995;30:30–33.

60. Grimm V, Berger U, Deangelis DL, et al. The ODD protocol: a review and first update. *Ecol Modell.* 2010;221:2760–2768.

61. Grimm V, Berger U, Bastiansen F, et al. A standard protocol for describing individual-based and agent-based models. *Ecol Modell.* 2006;198(1-2):115–126.