INVITED ARTICLE

Modeling Infectious Diseases in Healthcare Network (MInD-Healthcare) Framework for Describing and Reporting Multidrug-resistant Organism and Healthcare-Associated Infections Agent-based Modeling Methods

Rachel B. Slayton,[1] Justin J. O'Hagan,[1] Sean Barnes,[2] Sarah Rhea,[3] Rainer Hilscher,[3] Michael Rubin,[4] Eric Lofgren,[5] Brajendra Singh,[1] Alberto Segre,[6] and Prabasaj Paul[1]; for the Centers for Disease Control and Prevention MInD-Healthcare Program

[1]Division of Healthcare Quality Promotion, Centers for Disease Control and Prevention, Atlanta, Georgia, USA, [2]Robert H. Smith School of Business, University of Maryland, College Park, Maryland, USA, [3]RTI International, Research Triangle Park, North Carolina, USA, [4]Division of Epidemiology, University of Utah School Medicine, Salt Lake City, Utah, USA, [5]Washington State University, Pullman, Washington, USA, and [6]Department of Computer Science, University of Iowa, Iowa City, Iowa, USA

Mathematical modeling of healthcare-associated infections and multidrug-resistant organisms improves our understanding of pathogen transmission dynamics and provides a framework for evaluating prevention strategies. One way of improving the communication among modelers is by providing a standardized way of describing and reporting models, thereby instilling confidence in the reproducibility and generalizability of such models. We updated the Overview, Design concepts, and Details protocol developed by Grimm et al [11] for describing agent-based models (ABMs) to better align with elements commonly included in healthcare-related ABMs. The Modeling Infectious Diseases in Healthcare Network (MInD-Healthcare) framework includes the following 9 key elements: (1) Purpose and scope; (2) Entities, state variables, and scales; (3) Initialization; (4) Process overview and scheduling; (5) Input data; (6) Agent interactions and organism transmission; (7) Stochasticity; (8) Submodels; and (9) Model verification, calibration, and validation. Our objective is that this framework will improve the quality of evidence generated utilizing these models.

Keywords. healthcare-associated infections; mathematical modeling; antimicrobial resistance.

FRAMEWORK PURPOSE AND AUDIENCE

Healthcare-associated infections (HAIs) are associated with substantial morbidity, mortality, and costs [1, 2]. Mathematical modeling transmission of HAI pathogens and multidrug-resistant organisms (MDROs) improves our understanding of the dynamics of the spread of pathogens, provides a framework for evaluating prevention strategies, and can accelerate prevention efforts [3]. Instilling confidence in the reproducibility and generalizability of models is a challenge that may be addressed through clear and thorough communication of the underlying methodology.

The International Committee of Medical Journal Editors has compiled recommendations for best practices for conducting and reporting high-quality scholarly work. Reporting guidelines utilized for different types of study designs include STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) for observational studies and Consolidated Standards of Reporting Trials (CONSORT) for randomized trials [4, 5]. However, there is a paucity of best practices for reporting HAI and MDRO transmission modeling studies to biomedical journals. As transmission modeling becomes increasingly prominent and findings are used to inform public health practice, a common framework for reporting would strengthen the knowledge base and acceptability of findings from such studies [6, 7], analogous to recent efforts to improve modeling of human papillomavirus–related cancer control [8].

In 2017, the Centers for Disease Control and Prevention created the Modeling Infectious Diseases in Healthcare Network (MInD-Healthcare), a consortium of investigators collaboratively developing and using mathematical models to investigate the spread and prevention of HAIs and MDROs. This document describes the MInD-Healthcare Framework and aims to improve communication among mathematical modelers about their respective agent-based models (ABMs). We specifically focus on transmission of HAI pathogens and MDROs, although there is substantial overlap with describing other infectious diseases. With an increasing proportion of ABMs among all HAI transmission models, this framework is focused on ABMs in contrast to other models (eg, compartmental models) for which

complete mathematical specification is provided for methodological transparency [9]. However, the guiding principles are generalizable. This framework can be utilized by epidemiologists to better understand the methods used by mathematical modelers and appropriately interpret modeling studies' results. This framework is an adaptation of the Overview, Design concepts, and Details (ODD) developed by Grimm et al [10, 11] to increase the transparency and reproducibility of modeling methods in the context of ecological modeling. The ODD framework has infrequently been used to describe HAI ABMs, although it has improved methodological clarity when employed [12–14]. The MInD-Healthcare Framework aims to align the traditional ODD with elements commonly included in healthcare-related ABMs and adopt terminology commonly used by epidemiologists to describe infectious disease processes.

FRAMEWORK ELEMENTS

The framework is composed of 9 elements: (1) Purpose and scope; (2) Entities, state variables, and scales; (3) Initialization; (4) Process overview and scheduling; (5) Input data; (6) Agent interactions and organism transmission; (7) Stochasticity; (8) Submodels; and (9) Model verification, calibration, and validation. (See Table 1 for a checklist for users of this framework.) Each element should be described clearly and concisely utilizing tables and diagrams, where appropriate.

Purpose and Scope

The purpose of a model is grounded by the primary objective(s) of the study and the potential impact of the model-driven insights on clinical practice. The scope provides boundaries for which dynamics will be included and therefore the external validity of the model. The purpose and scope should align with each other and drive the description of the remaining framework elements. The purpose and scope of a model are typically described in the introduction of an article because they provide critical context for understanding why the model was developed and provide a guide for what to expect in the following model description. This element contains a summary description and justification for the model's level of complexity and its intended use.

Entities, State Variables, and Scales

An entity is a person or object that may interact with other entities or be affected by external environmental conditions. The entities of an ABM (eg, patients, healthcare personnel [HCP]) are characterized by a set of state variables. A state variable is

Table 1. Framework Elements

Element	Checklist Item and Brief Definition	Reported on Page No.
1. Purpose and scope	Purpose: Specify the primary problem under consideration or the objective(s) of the study. Scope: Specify the boundaries for which dynamics will included and which are ignored.	_____
2. Entities, state variables, and scales	Describe each entity, state variable, and scale. Entity: A distinct or separate object or actor that behaves as a unit; may interact with other entities or be affected by external environmental factors. Each entity's current state is characterized by its state variables. State variable: Attribute which performs at least 1 of the following functions: distinguishes an entity from other entities of the same type or category, or traces how the entity changes over time. Scale: Temporal and spatial resolutions and extents of the model.	_____
3. Initialization	Describe the initial states of the model entities and environment, (ie, at time $t = 0$) including how many entities of each type are present initially and the exact values of their state variables (or how they were set stochastically).	_____
4. Process overview and scheduling	Process overview: Specify who (ie, what entity) does what (ie, what actions are executed) and who is affected (ie, which entities and state variables). Scheduling: The order in which actions are taken and the order in which the effects of those actions are realized.	_____
5. Input data	Describe whether the model uses input from external sources such as data files or other models to represent processes that change over time.	_____
6. Agent interactions and organism transmission	Agent interactions: Specify both direct interactions in which individuals encounter and affect others and any indirect interactions. Organism transmission: Describe how pathogen transmission is specified, including which interaction types result in changes to health states (eg, incident colonization, infection transmission).	_____
7. Stochasticity	Describe if stochasticity is part of the model and the major underlying reasons including which elements are modeled as fully stochastic, partly stochastic, or deterministic.	_____
8. Submodels	Describe the equations and algorithms used in the submodels. Include tables of parameter definitions, units, and values used.	_____
9. Model verification, calibration, and validation	Describe the processes of verification, calibration, and validation. Verification: Process of ensuring the model was implemented correctly and meets the specifications of the model design. Calibration: Process of tuning the model parameters so that the model output matches a selected set of statistics from the real-world system or from simulated data. Validation: Process of evaluating how well the model represents the underlying truth of the real-world process it aims to represent.	_____

We strongly encourage using this checklist in conjunction with the current article "Modeling Infectious Diseases in Healthcare Network (MInD-Healthcare) Framework for Describing and Reporting Multidrug Resistant Organism (MDRO) and Healthcare-Associated Infections (HAIs) Agent-Based Modeling (ABM) Methods" for important clarifications on all included elements. Additional extensions are forthcoming: For those and for up-to-date references relevant to this checklist, see https://www.cdc.gov/hai/research/MIND-Healthcare.html.

an attribute that performs at least 1 of the following functions: it distinguishes an entity from other entities of the same type or traces how the entity changes over time. State variables may also be global and accessible to all entities. Scales refer to the temporal and spatial resolutions of the model and how they influence the representation of transmission (eg, minimum duration for contact) [15]. Modelers should describe what kinds of entities are in the model and by what state variables or attributes they are characterized.

At any point in time, the state variables provide a snapshot of the model, with sufficient information to restore and restart the model from that instant. State variables can change over time (eg, use of an antibiotic) or remain constant (eg, sex, physical location of a hospital). Patient-level state variables should be epidemiologically meaningful and comparable to data available in electronic medical records (eg, patients' treatment statuses with each class of antibiotic could be included as state variables [eg, neither, only penicillin, only fluoroquinolone, both], but "receipt of 2 antibiotic classes" would not be considered a state variable). Health states are a set of state variables that represent the underlying natural history of disease processes. These health states may include, but are not limited to, susceptible, asymptomatically colonized, infected and infectious, and recovered. The health states that are modeled may vary by agent type (eg, HCP may only become transiently colonized, whereas patients progress to infection).

Many ABMs include the following types of entities and their associated state variables:

Agents: A model can have different types of agents; individuals can be considered agents and may be further differentiated in categories (eg, patients, HCP). There may even be different subtypes within a category (eg, nurses, physicians). Agents are not limited to humans (eg, pathogens or facilities). Example state variables of agents include: identity number (ie, even if all other state variables would be the same, the agent would still maintain a unique identity), demographics (eg, age, sex), location (eg, assigned unit within a hospital), and comorbidities (eg, previous antibiotic use).

Spatial units: Spatial units are often used to model the state of specific spatial locations within a hospital (eg, patient room, unit) and are referred to as grid cells or patches. Example state variables of spatial units include: location, size, list of agents within the spatial unit, and descriptors of physical environmental conditions (eg, surface type) within the spatial unit. Some overlap of roles can occur; for example, a spatial unit used to model a patient room may be an entity with its own state variables (eg, likelihood of surfaces to become and remain contaminated) but may also function as a location, and hence a state variable, of patients.

Collectives and cohorts: We group entities with common state variables together. These groups of entities can have distinct behaviors, so it may make sense to distinguish them as a defined group (eg, entire healthcare facilities, groups of HCP within a unit, or households). A collective is a group that is usually characterized by the list of its agents, and by specific actions that are only performed by the collective, not their constitutive entities (eg, a hospital ward made up of spatial unit subcomponents in which certain agent to agent interactions can only occur). In contrast, a group of agents that are only considered as a unit of analysis (and do not have distinct behavior) are referred to as a cohort (eg, intensive care unit patients to aid in model verification, calibration, and validation).

In describing spatial and temporal scales it is important to specify what the model's units represent in reality, both within and among healthcare facilities. For temporal scale: describe the time period being modeled (eg, 5-year period or specific calendar years 2011–2015). Describe time as discrete steps (eg, day) or as a continuum over which both continuous processes and discrete events can occur. Describe if time scales vary by agent class and at which time step events occur. For spatial scale: describe if agents operate and interact within a single ward, across an entire healthcare facility, or across a network of healthcare facilities and the community. Describe if the model explicitly represents a geographic area.

Initialization

Results generated from ABMs can depend significantly on the initial conditions of the model. Initialization describes the starting states of the model entities and environment (ie, at time $t = 0$ of a simulation run), including how many entities of each type are present initially and the exact values of their state variables or how they were set stochastically. Describe if initialization is always the same or varies between simulation runs. Are the initial values chosen arbitrarily or based on data? If the latter, references to those data should be provided. It may not be possible to accurately replicate model results unless the initial conditions are known. Different models, and different analyses using the same model, can of course depend quite differently on initial conditions. Sometimes the purpose of a model is to analyze consequences of its initial state, while other times modelers aim to minimize the effect of initial conditions on results (eg, excluding the output from a predetermined warm-up period). These considerations should be described explicitly and align with the overall objective of the study. If a random number seed is used for stochastic elements, this section is the appropriate place to specify the value of that seed.

Process Overview and Scheduling

The simulation of an ABM relies extensively on an explicitly defined schedule for how model processes are executed. Processes can be defined by who (ie, what entity) does what (ie, what actions are executed) and who is affected (ie, which entities and state variables). Scheduling determines the order

in which actions are taken, as well as the order in which the effects of those actions are realized. Many ABMs represent time in discrete steps, while others treat it as a continuous variable [16].

Describe if the model is a hybrid or combined model (ie, including compartmental components). In this element only the self-explanatory names of the model's processes should be listed: "move," "update plots," etc. These names are then the titles of the submodels that are described in the section entitled "Submodels." As Grimm et al defined, processes are performed either by one of the model's entities (eg, move) or by a higher-level controller (ie, an observer) that performs actions such as updating plots or writing output to files. To handle such higher-level processes, ABM software platforms like Repast [17] and NetLogo [18] include the concept of the "Model" or "Observer." By "in what order?" we refer to both the order in which the different processes are executed and the order in which a process is performed by a set of agents. For example, for HCP visiting a patient we specify the order in which HCP attempt to wash their hands, wear personal protective equipment, and touch the patient and environment. We specify, relative to other processes modeled, whether HCP perform hand hygiene in a random order, a fixed order, or a HCP type-sorted order. Differences in such ordering (eg, HCP visits to patients) can have a very large effect on model outputs [19, 20]. When processing a particular action, an entity's state can be updated immediately (asynchronous updating) or the update can be stored until all entities have executed the process, at which point all entity states are updated simultaneously (synchronous updating).

For this element, only an overview is required to provide an understanding of how the ABM is simulated (eg, visit patient, update environment). The specific details of these processes are reserved for the submodels element. Except for very simple models, authors should provide the full model code or, at a minimum, pseudo-code to fully describe the schedule, so the model can be replicated. Ideally, the pseudo-code corresponds fully to the actual ABM code.

Input Data

Describe whether the model uses input from external sources such as data files or other models to represent processes that change over time. In models of real systems, dynamics may be driven in part by a time series of environmental variables, sometimes called external forcings; for example, seasonality of disease or associations with environmental factors that could affect the hospital infrastructure (eg, legionellosis and rainy weather seasons, daily shift schedules of HCP). "Driven" means that 1 or more state variables or processes are affected by how these environmental variables change over time. However, these environmental variables are inputs and are not, themselves, affected by the internal variables of the model. Often, it makes

sense to use observed time series of environmental variables so that their statistical qualities (eg, mean, variability, temporal autocorrelation) are realistic. Alternatively, external models can be used to generate input (eg, a set of seasonal environmental conditions) [21]. Obviously, to replicate an ABM, any such input has to be specified and the data or models provided, if possible. The publication of input data for some simulations may be constrained by confidentiality considerations. Where such concerns exist, inclusion of pseudo-data ought to be considered. If these input data are obtained through an application program interface (API) to another model or data provider, details on how this API can be accessed should be provided. If a model does not use external data, this element should nevertheless be included, using the statement, "The model does not use input data to represent time-varying processes." Note that "Input data" does not refer to parameter values or initial values of state variables.

Agent Interactions and Organism Transmission

Describe the kinds of interactions among agents that are assumed. Include a description of both direct interactions (eg, in which individuals encounter and affect others) and indirect interactions (eg, patients sharing the same care team). Provide details on how transmission is specified, including which interactions result in changes to health states (eg, in a simple flow chart). How do agents react to the interactions, or lack thereof, with other agents and the processes that they undergo? How do the pathogen characteristics influence transmission? Is feedback from previous processes incorporated into future processes? Are dynamic responses based on exceeding some minimum threshold?

Stochasticity

Often, we choose to model variability in a parameter or process without fully representing all the underlying mechanisms that account for that variability occurring. A simplified approach is to incorporate stochastic elements that randomly draw numbers to change parameters or processes. ABMs are inherently stochastic, although they may contain deterministic elements (eg, a module describing room contamination deterministically may output contamination level as a parameter value in an ABM more broadly). Demographic stochasticity, where only whole individuals, as opposed to fractions of individuals, can undergo changes (eg, birth, death, disease-state changes) is a mechanistic basis for including stochasticity in mathematical models and is especially relevant in small populations [17]. What elements of the submodels are modeled fully stochastically, partially stochastically, or deterministically? Why are elements modeled this way? Is stochasticity used, for example, to reproduce variability in processes for which it is important to model the actual causes of the variability? How many model replications are utilized and how was this number determined?

Submodels

All submodels should be presented completely and in detail (utilizing the applicable framework elements). Submodels are described at a high level in the process overview and scheduling, although specific details of these processes are described here. The factual description of the submodel (ie, equation[s] and algorithms) should be described. If parameterization is not discussed outside of the description based upon this framework, it can be included here. The parameter definitions, units, and values used should be presented in tables. The purpose of including this information is to prevent the description from seeming ad hoc and to strengthen model credibility. Justification can be very brief in the earlier sections, but the complete description of submodels is likely to include references to relevant literature, as well as independent implementation, testing, and analysis of submodels. Additionally, in most cases, it will be necessary to have simulation experiments or a model analysis section following the model description.

Model Verification, Calibration, and Validation

Verification is the process of ensuring the model was implemented correctly. Calibration is the process of tuning the model parameters so that model output matches a selected set of statistics from the real-world system or simulated data and is possibly conducted in tandem with validation (eg, incident infections hospitalwide). Validation is the process of evaluating how well the model represents the underlying truth of the process it aims to represent. What is the process of model verification? What is the process for model calibration? What are the calibration targets and on what information are they based? What is the process for model validation?

Sensitivity analysis may be a part of multiple processes. Sensitivity analysis is the process of determining which model parameters and structures have the most significant effect on model outputs. Uncertainty analysis is the process of evaluating how the uncertainty of model parameter values may affect the model's output [22]. The precision with which parameters can be measured in the real-world impacts the reliability of model results generated using those estimates [3]. Therefore, uncertainty analysis relates to validation and calibration. Sensitivity analysis may be performed for verification (eg, extreme value test) or validation (eg, analyzing the impact of increasing hand hygiene compliance on acquisition rates) or calibration (eg, choosing a wider range for certain parameters to be chosen from, fixing noncalibrated parameters to different plausible values, employing a different fitting method). What sensitivity analyses were conducted? What uncertainty analyses were conducted?

INTERPRETATION

The MInD-Healthcare Framework was developed as a set of best practices for conducting and reporting transmission modeling studies, in particular those utilizing ABMs. It primarily aims to improve communication among mathematical modelers about their respective ABMs and builds upon the ODD protocol [10]. By increasing methodological transparency, we hope this framework will improve quality and reproducibility of the evidence generated by these models, just as development of the CONSORT guideline improved completeness of reporting of randomized controlled trials [23]. Additionally, this framework can facilitate multimodel comparisons by aiding in harmonization, systematic exploration of variability, and pooling of results [24].

High-quality scientific literature provides a complete methodological description and increased certainty for decision makers. Mathematical models are often criticized as being "black boxes" because they provide a limited methodological description. Adoption of these standards will enable consumers of modeling studies to better understand the methods and accompanying strengthens and limitations of these studies. Increased transparency also ought to improve the interpretation of modeling studies and adoption of recommendations based upon these studies. The MInD-Healthcare Framework should be re-evaluated and revised in the future to consider evolution of the field, generation of new evidence, and incorporate feedback from those using this framework.

References

1. Centers for Disease Control and Prevention. Antibiotic resistance threats in the United States, 2013. Atlanta, GA: Centers for Disease Control and Prevention, **2013**.
2. Scott RD. The direct medical costs of healthcare-associated infections in U.S. hospitals and the benefits of prevention. **2009**.
3. Railsback SF, Grimm V. Agent-based and individual-based modeling: a practical introduction. Princeton, NJ: Princeton University Press, **2011**.
4. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet **2007**; 370:1453–7.
5. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ **2010**; 340:c869.
6. Tracy M, Cerdá M, Keyes KM. Agent-based modeling in public health: current applications and future directions. Annu Rev Public Health **2018**; 39:77–94.
7. Willem L, Verelst F, Bilcke J, Hens N, Beutels P. Lessons from a decade of individual-based models for infectious disease transmission: a systematic review (2006–2015). BMC Infect Dis **2017**; 17:612.
8. Canfell K, Kim JJ, Kulasingam S, et al. HPV-FRAME: a consensus statement and quality framework for modelled evaluations of HPV-related cancer control. Papillomavirus Res **2019**; 8:100184.
9. van Kleef E, Robotham JV, Jit M, Deeny SR, Edmunds WJ. Modelling the transmission of healthcare associated infections: a systematic review. BMC Infect Dis **2013**; 13:294.

10. Grimm V, Berger U, Bastiansen F, et al. A standard protocol for describing individual-based and agent-based models. Ecol Modell **2006**; 198:115–26.

11. Grimm V, Berger U, DeAngelis D, Polhill JG, Giske J, Railsback S. The ODD protocol: a review and first update. Ecol Modell **2010**; 221:2760–8.

12. Lanzas C, Dubberke ER. Effectiveness of screening hospital admissions to detect asymptomatic carriers of Clostridium difficile: a modeling evaluation. Infect Control Hosp Epidemiol **2014**; 35:1043–50.

13. Bintz J, Lenhart S, Lanzas C. Antimicrobial stewardship and environmental decontamination for the control of clostridium difficile transmission in healthcare settings. Bull Math Biol **2017**; 79:36–62.

14. Ferrer J, Salmon M, Temime L. Nosolink: an agent-based approach to link patient flows and staff organization with the circulation of nosocomial pathogens in an intensive care unit. Procedia Comput Sci **2013**; 18:1485–94.

15. Dawson DE, Farthing TS, Sanderson MW, Lanzas C. Transmission on empirical dynamic contact networks is influenced by data processing decisions. Epidemics **2019**; 26:32–42.

16. Grimm V, Railsback SF. Individual-based modeling and ecology. Princeton. NJ: Princeton University Press, **2005**.

17. Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals. Princeton, NJ: Princeton University Press, **2008**.

18. Wilensky U. NetLogo. Evanston, IL: Northwestern University, **1999**.

19. Bigbee T, Cioffi-Revilla C, Luke S. Replicating the classic Sugarscape in MASON. Complex behavior in economics: modeling, computing and mastering complexity. Aix-en-Provence, France: Third Aix en Provence Complexity Workshop (COMPLEXITY2006), **2006**. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.105.2460

20. Caron-Lormier G, Humphry RW, Bohan DA, Hawes C, Thorbek P. Asynchronous and synchronous updating in individual-based models. Ecol Modell **2008**; 212:522–7.

21. Eisinger D, Wiegand K. SERGE: a spatially explicit generator of local rainfall in southern Africa. South African J Sci **2008**; 104:37–42.

22. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD; ISPOR-SMDM Modeling Good Research Practices Task Force. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM modeling good Research Practices Task Force Working Group-6. Med Decis Making **2012**; 32:722–32.

23. Turner L, Shamseer L, Altman DG, Schulz KF, Moher D. Does use of the CONSORT statement impact the completeness of reporting of randomised controlled trials published in medical journals? Cochrane Syst Rev **2012**; 1:60.

24. den Boon S, Jit M, Brisson M, et al. Guidelines for multi-model comparisons of the impact of infectious disease interventions. BMC Med **2019**; 17:163.