Dynamic Simulation Modeling Applications in Health Care Delivery Research – Emerging Good Practices Task Force Report

Task Force Chair:

Deborah Marshall, PhD, Canada Research Chair, Health Services & Systems Research; Arthur J.E. Child Chair in Rheumatology Research; Director, HTA, Alberta Bone & Joint Health Institute; Associate Professor, Department Community Health Sciences, Faculty of Sciences, Faculty of Medicine, University of Calgary, Calgary, Alberta, Canada

Task Force Leadership Group:

William Crown, PhD, Chief Scientific Officer, Optum Labs, President ISPOR, Boston, MA, USA
Mitch Higashi, PhD, Chief Economist, GE Healthcare, Barrington IL, USA
Maarten J. IJzerman, PhD, Professor of Clinical Epidemiology & Health Technology Assessment (HTA); Head, Department of Health Technology & Services Research, University of Twente, Enschede, The Netherlands
Lina Burgos Liz, MPH, MSc Ind Eng, System Dynamics Team, Universidad de los Andes, Bogotá, Colombia
Nathaniel Osgood, BS, MS, PhD, Associate Professor, University of Saskatchewan, Department of Computer Science, Associate Faculty, Department of Community, Health & Epidemiology and Bioengineering Division, Saskatoon, Saskatchewan, CA
William V. Padula, PhD, MS, Postdoctoral Fellow, Section of Hospital Medicine, University of Chicago, Chicago, IL, USA
Kalyan Pasupathy, PhD, Senior Associate Consultant, Health Care Policy & Research and Co-Scientific Director, Health Care Systems Engineering Program, Mayo Clinic Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Rochester, MN, USA
Peter Wong, PhD, MS, MBA, RPh, Vice President and Chief Performance Improvement Officer, Illinois Divisions and HSHS Medical Group, Hospital Sisters Health System (HSHS), Belleville, IL, USA
Abstract

Conventional evaluation of interventions often neglect the wider health system impacts that could be critical for achieving desired health system goals. As such, traditional health technology assessment and modeling methods are often of limited usefulness when applied to complex health systems. Health care delivery systems are inherently complex entities, consisting of multiple tiers of interdependent subsystems and processes that are adaptive to changes in the local environment and behave in a non-linear fashion.

This report presents dynamic simulation modeling methods to evaluate complex system interventions for health care delivery. It is a primer for researchers and decision makers working in health care delivery that face challenges to providing effective or efficient care delivery that can be addressed with complex system interventions. For healthcare planning, health care delivery system interventions need to incorporate the dynamics and complexities of the health care system context in which the intervention is delivered.

Three advanced simulation modeling approaches are suitable for simulating interaction between health care delivery and system interventions: agent-based modeling (ABM); discrete event simulation (DES); and system dynamics (SD) modeling. In contrast to conventional evaluations, a dynamic systems approach anticipates the upstream and downstream consequences of changes in complex healthcare delivery systems.

The specific selection of the appropriate simulation modeling method depends on a number of factors, such as whether the problem is specific to individuals or groups, the level of the problem (strategic, operational or tactical) and whether stochastic or deterministic solutions are sought. This report illustrates the application of these dynamic simulation modeling methods using specific applied examples of health care system problems and complex interventions where these methods can be useful. This report will provide needed guidance to researchers, payers, practitioners and decision makers on good practices for simulation modeling – from the selection of a simulation model method to model building and validation approaches to evaluate complex health care system interventions.
Introduction

Healthcare delivery systems are inherently complex and fragmented social systems, which are different from deterministic, animate and ecological systems. (Ackoff, 2001; Glouberman & Zimmerman, 2002; Zimmerman, Lindberg, & Plsek, 2001) Social systems are different from other systems in that people make decisions and people have agency. People interact amongst themselves but also with other parts of the system. It is hard to plan for this kind of system since decisions and choices by people may change day to day; and interactions between parts of the system and with other systems may change too. This complexity hinders the capacity decision makers have to intervene in the systems and improve them.

Differences in governmental, payer and provider structures make it difficult to offer a one-size-fits-all solution to many of the problems that healthcare delivery systems face. (W. V. Padula & Breteler, 2013) Furthermore, the needs of individual patients escalate the complexity of health care delivery systems (Berwick, 2002), requiring customizable patient-centered care that is not well defined from recent developments in evidence-based medicine. (Berwick, 2009) These issues make systems difficult to manage, and often lead to implementing complex system interventions that offer solutions for delivering patient-centered care. (Berwick, 2008)

This report presents simulation modeling methods to evaluate complex system interventions for health care delivery. It is a primer for researchers and decision makers working in health care delivery that face challenges to providing effective or efficient care delivery that can be addressed with complex system interventions. Based on the extensive literature in other fields such as engineering, operations research, and statistics, there is support to suggest that three advanced modeling approaches are suitable for simulating interaction between health care delivery and system interventions: agent-based modeling (ABM); discrete event simulation (DES); and system dynamics (SD) modeling. (Homer & Hirsch, 2006; Milstein, Homer, & Hirsch, 2010; Seila & Brailsford, 2009)

This report provides an overview of these modeling approaches and examples of healthcare system interventions where such methods could be useful. The report assists researchers and decision makers in deciding whether these simulation models are appropriate to address specific health systems problems. The report also directs readers to other resources for further education on the topic of modeling for system interventions in the emerging field of health care delivery science and implementation. Upon reviewing this report, the reader should be able to identify which simulation modeling methods work best given various health care delivery intervention characteristics. This report will also identify key steps for good practices in simulation modeling of complex interventions in health care delivery systems.
Definitions of concepts and terminology

In the field of complexity science, complexity is a property of a system, not an intervention. (MRC, 2000) A complex system is one that is adaptive to changes in its local environment, is composed of other complex systems (for example, the human body), and behaves in a non-linear fashion. (Change in outcome is not proportional to change in input.) (Plsek & Greenhalgh, 2001).

Simple systems consist of tasks that can be answered as ‘yes’ or ‘no’, whereas complicated systems consist of tasks that are based on ‘if/then’ algorithms and complex systems consist of tasks that are relationally dependent events with unpredictable outcomes. (W.V. Padula, Duffy, Yilmaz, & Mishra, 2014).

Health care consists of multiple complex systems. For instance, complex systems in health care delivery include primary care, hospitals, and long term chronic care facilities. This framework of complexity can be useful for delineating simulation modeling applications to systems depending on the level of complexity (Table 1).

Table 1: Concepts & Terminology Definitions

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex System</td>
<td>One that is adaptive to changes in its local environment, is composed of other complex systems, and behaves in a non-linear fashion. (MRC, 2000) Tasks that are relationally dependent events with unpredictable outcomes. (Glouberman &amp; Zimmerman, 2002; Zimmerman et al., 2001)</td>
</tr>
<tr>
<td>Health Care Delivery Systems</td>
<td>Health care delivery systems are inherently complex entities, consisting of multiple tiers of interdependent subsystems and processes, as well as varying degrees of private and public elements throughout different regions. (W.V. Padula et al., 2014)</td>
</tr>
<tr>
<td>Complex Intervention</td>
<td>A complex intervention is built up from a number of components, which may act both independently and inter-dependently. (MRC, 2000)</td>
</tr>
</tbody>
</table>

Health care delivery systems are inherently complex entities, consisting of multiple tiers of interdependent subsystems and processes, as well as varying degrees of private and public elements throughout different regions.
regions. (W.V. Padula et al., 2014) More recently, with the focus on patient-centered care, the complexity has increased in these systems. The Institute of Medicine’s seminal report, “Crossing the Quality Chasm: A New Health System for the 21st Century,” highlights patient-centered care as a necessary component of good practice, emphasizing the importance of continuous health relationships, knowledge sharing, and free information flow across segments of the health system. (Berwick, 2002)

According to Berwick’s Triple Aim, health care system improvement requires the simultaneous pursuit of three aims in iterative cycles: improving both the experience of care and the health of populations, and reducing per capita costs of health care. (Berwick, Nolan, & Whittington, 2008) However, providers are challenged with implementing evidence-based practices, such as checklists, that lack the facility to incorporate patient preferences. (Gawande, 2009; Pronovost & Vohr, 2010) Thus, the focus on patient-centered health care delivery also underscores the need for simulation modeling methods of complex systems that can capture patient preferences and simulate patient and provider behavior.

The adoption of the evidence-based practices for patient care is hampered by the current fragmentation of existing health care delivery systems. System redesign is an essential step to achieving consistent implementation of evidence-based practice interventions. As an administrator at the Centers for Medicare and Medicaid Services, Berwick piloted a new systems-based design founded on the Triple Aim framework – the Accountable Care Organization (ACO) (McClellan, Mckethan, Lewis, Roski, & Fisher, 2010). If ACOs remain sustainable in the U.S., then system levels would require additional definitions or redefinition of current systems to include the influence of an ACO as a stakeholder in patient-centered care. (Fisher et al., 2009; Fisher, McClellan, & Safran, 2011)

In the Medical Research Council’s framework for the evaluation of complex interventions, a complex intervention is “built up from a number of components, which may act both independently and inter-dependently.” Consequently it can be challenging to determine what aspect(s) of the intervention effect change in the system. (MRC, 2000).

When evaluating complex interventions in these settings, the MRC report emphasizes that it is necessary to consider the wider ramifications of intervening and to be aware of the interaction that occurs between components of the intervention as well as between the intervention and the context in which it is implemented. This includes the operations, structures, and relationships that exist in each setting and the implications that contextual effects have for designing and evaluating interventions.
Why is Simulation Modeling Relevant to Health Care Delivery Research in the Context of Complex Systems?

As with the evaluation of policy interventions, simulation methods can be informative in health care delivery research. Health care systems are characterized by nonlinearities, feedback loops, and large numbers of variables and that evolve dynamically over time. The possible interactions among components of systems can be extremely complex. Simulation models can identify the critical functional and relational aspects of a system. With improvements in the availability of clinical and health care utilization data, the ability to assemble separate estimates of components of treatment into systems of care is improving rapidly. Thus, simulation modeling can help us understand why a system behaves the way it does as a function of its organization and the relationships amongst components of the system.

What is simulation modeling used for?

**Health Care Delivery Research in Complex Systems**
- Model building process and simulation are learning processes themselves
- Identify critical functional and relational aspects in complex systems.
- Understanding system behavior as a function of its organization (structure).
- Shift paradigms and mental models

**Design and Evaluation of Health Care Delivery System Interventions**
- Evaluate intended and unintended consequences of an intervention using “what if...?” scenarios
- Tool for designers (e.g. policy design, system design and re-design) that is more prescriptive in nature by informing decision making.

What/How can simulation modeling contribute to better design and evaluation of healthcare delivery system interventions?

Significant advances in disease diagnosis and treatment and care have been generated by technical and scientific progress. Many of these advances are being implemented in healthcare, without addressing the unintended or unanticipated consequences of these interventions at the system level. In fact, policies that are implemented to address difficult challenges in healthcare sometimes fail to solve persistent problems or
create new problems. This phenomenon is known as *policy resistance*: the tendency for interventions to be defeated by the system’s response to the intervention itself (J. D. Sterman, 2006). Framed as a complex system, policy resistance, and a general resistance towards new innovations in healthcare arises from the mismatch between the complexity of these systems and our capacity to understand them (J. D. Sterman, 2006).

It is widely accepted in health care that generating reliable scientific evidence requires conducting experiments, comparing and differentiating hypotheses and obtaining results that are replicable (Brownson, Fielding, & Maylahn, 2009; J. D. Sterman, 2006). However, generating reliable scientific evidence becomes more difficult as complexity increases and is not always feasible due to ethical, physical, or technical reasons. Simulation models are virtual worlds that offer decision makers the capability of conducting experiments and evaluating system interventions. Simulation models provide low risk and low cost laboratories to learn and gain understanding about health care systems and the effects that interventions may have on them.

Simulation modeling methods using ‘what if’ scenarios can then be used to estimate the downstream outcomes associated with systems of care that are too complex to anticipate based upon piecemeal analyses of the system components. In the virtual world of the simulation model, decision makers can push the system to extreme conditions, extend time of observation, strengthen and relax assumptions, which is often impossible or infeasible in the real world. Simulation models provide immediate feedback to decision makers, allowing them to gain years of simulated experience and knowledge about the system and interventions by revealing dynamics and mechanisms that are otherwise not obvious. (Richmond, 1993; J.D. Sterman, 2000)

Through simulation modeling, decision makers can observe effects that interventions can have on different parts of the system concurrently; it engages decision makers into systemic thinking and focus on interdependencies - broadening their perspective and enhancing their understanding of interventions and overall system (Ackoff, 2001; Richmond, 1993). This understanding comes in the form of shift in paradigms and mental models.

Traditional approaches and statistics provided *descriptive* ways of measuring and testing individual relationships from the past. As massive amounts of data are collected and warehoused, the *descriptive* analyses are replaced by *predictive* models, which strive to foretell the future. These models are widely being developed in areas of data mining and predictive analytics. Simulation modeling takes it one step further to become *prescriptive* in nature, such that the models prescribe what actions/interventions to take, based on scenarios tested through experiments (Brown, Patrick, & Pasupathy, 2013; Pasupathy, 2010).
Moreover, when decision makers are deeply involved in the process of model building, the process and its products, i.e., preliminary models, become tools for discussion that help better define the scope of the problem at hand and elicit ideas, solutions and interventions from decision makers and other stakeholders (Vennix, 1999). Being engaged in this iterative process, decision makers find their mental models and preconceived ideas about the system challenges and are obliged to think broadly about the problem at hand and reflect on the system in which it is embedded (Brown et al., 2013; Pasupathy, 2010). Hence, decision makers are forced to engage in operational thinking and develop intuition about the system thinking about the nuts and bolts of the system and how it really works; hence informing the design of the system and interventions realistically and more accurately (Richmond, 1993). System redesign is an essential step to achieving sustainable implementation of complex interventions.

For example, arthritis is associated with a significant societal burden, both economically and in terms of health status. Due to reduced mobility, people living with arthritis lose independence and are more likely to suffer falls and fractures. Early recognition and intervention for both osteoarthritis (OA) and rheumatoid arthritis (RA) prevents or minimizes permanent, irreparable joint damage which results in functional impairment. Ensuring timely access to appropriate and effective care is the first step in preventing the deleterious, progressive effects of these diseases. In theory, the pooling of patients should help decrease the variability in the system and improve access to arthritis care.

A centralized intake system involves pooling of wait lists to create a single first-come, first-serve, but severity prioritized queue, from which patients are directed to an appropriate service provider. Suppose policy makers are considering the possibility of implementing a new centralized system for the intake of patients with OA and RA due to present challenges in access to timely and appropriate arthritis care. However, a central intake system can be structured in a variety of ways, and the impact of different structures on patient outcomes and costs are not obvious. Simulation modeling allows policy makers to evaluate these different structures and alternative scenarios. This ability to evaluate system redesign and alternative interventions arising from policy and clinical decision making is a critical, but largely missing tool in health services research.

**What are the Differences between Health Economic Models in Health Technology Assessment and Simulation Models in Health Care Delivery Systems?**

Health technology assessment (HTA) is defined as "the systematic evaluation of the properties and effects of a health technology, addressing the direct and intended effects of this technology, as well as its indirect and unintended consequences, and aimed mainly at informing decision making regarding health
Traditionally, health economic models used in HTA are based on clinical evidence and analyse economic consequences of that specific technology compared to usual care. Most HTA reports have a limited scope with regard to the consequences to the healthcare delivery system. For healthcare planning, health care delivery system interventions need to incorporate the dynamics and complexities of the health care system context in which the intervention is delivered. Conventional evaluation of interventions in health care are often limited because they neglect the wider health system impacts that could be critical for achieving desired health goals.

In contrast, as described previously, a systems approach anticipates the downstream consequences of changes in the healthcare system. This enables health service planners to identify upstream points of leverage through experimentation with various ‘what if’ scenarios without actually having to implement the policy first. Simulation is used to model interventions before the cost-intensive design and development and implementation phases. Thus, effects on patient care, the health care system, as well as health economics aspects can be estimated and anticipated (Gantner-Bar, Djanatliev, Prokosch, & Sedlmayr, 2012).

For instance, a health economic model comparing tissue engineered and biodegradable gels for repair of small cartilage defects might evaluate the clinical effectiveness and economic consequences compared to physiotherapy, surgery or watchful waiting in terms of cost per quality adjusted life years. But an HTA does not typically address the consequences to the healthcare system beyond a budget impact analysis. For health care planning and delivery, other questions may arise, such as the required healthcare facilities to deliver this minimally invasive therapy, and the change in hospital service due, for instance, a delay of whole joint replacement in case of severe osteoarthritis. In addition, the health system is likely to adopt minimally invasive interventions, but the diffusion may differ widely between regional health facilities, depending on interaction between healthcare payers, providers and physicians. The impacts of such interactions are not typically accounted for in HTA.

### Issues Relevant to the Health Care Delivery System that Simulation Modeling Methods can Address

Simulation modeling can be used for a large number of policy decisions regarding health care delivery planning. To supplement the HTA types of reports and analyses, there are multiple specific situations where simulation modeling may be useful for policy makers, for example:

**a) Simulation modeling can estimate consequences of healthcare delivery system interventions**

Many interventions in healthcare have an impact on the healthcare delivery system, which is typically not considered in health economic models. Simulation modeling is proposed to better estimate the downstream consequences once a health policy or delivery intervention is implemented, accounting for feedback loops...
to support the adaptive nature of the healthcare delivery system. These models can also be used to estimate the consequences of demographic change, or, for instance, ageing of the population (Ansah et al., 2014).

b) Simulation modeling allows the incorporation of behavioral aspects and personalized health care decisions

One of the limitations of most health economic models is that they are built upon “health states” rather than “events”. This prevents a proper representation of the unique pathways of individual patients through the health care system. Yet, these individual pathways are important to consider as individuals make decisions about when they will see a doctor, if they comply with their medication regimen, or if they are willing to co-pay for expensive treatment. Simulation models and agent-based modeling in particular, allow more flexibility to incorporate the dynamics of people making decisions affecting population health outcomes and thus efficient planning of healthcare interventions.

c) Simulation models are much more flexible to consider consequences of co-morbidities and healthcare utilization

Most health economic models assume an underlying disease for which a treatment is evaluated. However, many people with chronic disease suffer from multiple morbidities and experience multiple episodes of interactions with the healthcare system. Potentially, networks of related diseases can be defined, similar to networks of underlying genetic mutations and networks of social activities (Barabasi, 2007). If such networks can be identified and modeled, the consequences of healthcare delivery interventions on the health system can be evaluated.

d) Simulation models can consider the spatial consequences of a healthcare delivery intervention

Many healthcare interventions also have a spatial component, such as infectious disease policies (Davis, Sevdalis, & Drumright, 2014) or remote health services like tele-monitoring. If health services are delivered at home, or if general hospitals specialize into healthcare centers, this has a large impact on the amount of patients travelling to healthcare facilities. At least, it will impact the case-mix of patients in the hospital, and simulation modeling can be applied to estimate the consequences on hospital admissions and support further capacity planning (Hulshof, Boucherie, Hans, & Hurink, 2013).

e) Simulation modeling addresses system problems that are too complicated to enable an analytic solution.

Health care consists of multiple complex systems. The inherent feedback loops that reflect interactions amongst the operations, structures, and relationships in the health care system evolve dynamically over
time and cannot always be captured in an analytic solution. But simulation methods can be applied to model such relationships.

**Different kinds of healthcare delivery problems where simulation modeling approaches are valuable**

The feasibility and relevance of simulation modeling methods to inform health system planning and decision making for improving system efficiency have been demonstrated (Brailsford, Harper, Patel, & Pitt, 2009). The specific selection of the appropriate simulation modeling method depends on a number of factors, such as whether the problem is specific to individuals or groups, the level of the problem (strategic, tactical, or operational) and whether stochastic or deterministic solutions are sought (Gunal, 2012).

**Table 2. Examples of Problems addressed with simulation modeling approaches to evaluate complex Health Care Delivery Interventions**

<table>
<thead>
<tr>
<th>System Level</th>
<th>Types of Problems</th>
<th>Examples of Potential Simulation Modeling approaches</th>
<th>Intervention Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategic level</td>
<td>Policy</td>
<td>System Dynamics, Agent Based Modeling</td>
<td>Informing regional or national policy regarding implementation of a centralized intake system for referral to an appropriate provider for assessment and specialist consultation if needed, for patients with musculoskeletal pain.</td>
</tr>
<tr>
<td>Tactical level</td>
<td>Management</td>
<td>Agent Based Modeling, Discrete Event Simulation</td>
<td>Wait time management for referral for a specific service e.g., consultation with orthopaedic surgeon or rheumatologist</td>
</tr>
<tr>
<td>Operational level</td>
<td>Logistics</td>
<td>Agent Based Modeling, Discrete Event Simulation</td>
<td>Scheduling surgical dates for joint replacement in the operating room Evaluating the introduction of a new service using tissue engineered and biodegradable gels for repair of small cartilage defects with respect to the change in required healthcare facilities and hospital services.</td>
</tr>
</tbody>
</table>
Evaluating the change in hospital services due to a delay of total joint replacement in cases of severe osteoarthritis.

The following examples illustrate how these simulation modeling methods have been applied to health care delivery interventions:

- The Mayo Clinic’s Center for the Science of Health Care Delivery has applied health care delivery systems thinking to redesign its practices, particularly in the domains of cardiac surgery (Marmor, Rohleder, Cook, Huschka, & Thompson, 2013) and outpatient practice (Kuchera & Rohleder, 2011). The Center used system dynamics modeling for high level planning of primary care staffing that incorporated new care delivery modes. The model allowed for “what-if” scenarios to be evaluated showing projected access performance for measures such as time to appointment and corresponding staffing requirements.

- Another example is the ReThink Health model. This system dynamics model simulates the behavior of a health system - tracking changes in health status, utilization, and costs. It has been used to analyze various health policy strategies (expanding health insurance coverage, delivering better preventive and chronic care, and improving environmental conditions) to reduce deaths and improve the cost-effectiveness of interventions (Milstein, Homer, Briss, Burton, & Pechacek, 2011; Milstein et al., 2010). (Milstein’s 2011 paper was awarded Public Health Systems Research Article of the Year by Academy Health.)

- In Ontario, the median waiting time for total hip and knee joint replacements in Ontario was greater than 6 months, longer than clinically appropriate. To inform decisions to reduce waiting times and improve waiting list management, the team developed a discrete event simulation model of the Ontario total joint replacement system to evaluate the effects of 4 management strategies on waiting times: (1) reductions in surgical demand; (2) formal clinical prioritization; (3) waiting time guarantees; and (4) common waiting list management. Using the DES model, they concluded that increases in the number of surgeries provided greater than those observed historically or reductions in demand were needed to reduce waiting times to clinically acceptable levels within 10 years. (Cipriano, 2008)

- GE Healthcare has applied agent-based modeling, combining demographic, economics and epidemiological data, to support resource allocation decisions about the optimal delivery of care for heart disease(GEHealthcare, 2013).
The Alberta Health Services Bone and Joint Strategic Clinical Network was seeking a sustainable solution to balancing access, effectiveness and efficiency in delivering health services to patients with osteoarthritis across the continuum of care. Simulation modeling methods with system dynamics can also be useful to inform system-wide planning for a single disease, such as osteoarthritis (Marshall, 2010).

**Comparison and contrast of key simulation modeling methods (SD, DES, ABM) with other modeling methods (e.g. optimization, Markov models) – differences and complementarities**

Simulation models in the sense that we are discussing here (dynamic simulation rather than statistical simulation) are distinguished by their explicit representation of a system state and the mechanics of its evolution over time. Such state might, for example, include the health status and risk behaviors of population members, cumulative societal cost and quality adjusted life years. In contrast to the situation for analytic models – where the trajectories associated with system evolution are specified as an explicit function of time – for simulation models, this evolution implicit by specifying the rules governing that system evolution.

In contrast to statistical models involving observables, such rules aspire to characterize the posited “physics” of the system, causal drivers hypothesized to characterize “how the system works”. The International Society of Pharmacoeconomics and Outcomes Research (ISPOR) and the Society for Medical Decision Making (SMDM) published the ISPOR-SMDM Modeling Good Research Practices Task Force Reports providing guidance on state-transition models, such as Markov models and discrete event simulation modeling. (Caro, 2012; Siebert, 2012; Pitman, 2012)

The resulting simulation models – like the systems that they characterize – are often non-linear in character, a feature with several implications. First, the non-linearity of the models and systems characterized implies that understanding the behavior of the system to a portfolio of interventions requires simulating those interventions together, so as to capture situations where such interventions work synergistically and at cross-purposes with one another. In contrast to Markov models, which are commonly used to characterize evolution of isolated cohorts – simulation model non-linearity generally implies that individuals or cohorts cannot be simulated as solitudes, but must instead be simulated in a population context.

Second, the non-linearity leads such models to exhibit emergent behavior, behavior where the behavior of the whole can be very distinct from – and cannot be reduced to – that of its parts. Such emergent behavior is often surprising, counter-intuitive, and often differs strikingly at different temporal and spatial scales.

Third, while Markov models and linear systems models can be solved to provide a description of the
system's evolution a priori using known statistical interactions, to understand simulation models requires
that they be executed over time in a mechanistic fashion accounting for the dynamics in the system.

There is a large variety of simulation models, some of which share similar capabilities. We focus on three
primary types in this report (system dynamics modeling, discrete event simulation modeling and agent
based modeling), selected based on their suitability to address problems in health care delivery systems
and ability to simulate dynamically the interactions between operations, structures, and relationships in the
health care system.

Individual-based simulation modeling is associated with two major traditions: 1.) micro-simulation -
originating in economics and emphasizing evolution based on empirically grounded, statistical relationships,
and 2.) agent-based modeling - originating in computer science and traditionally depending on algorithmic
and rule-based formulations in richer, dynamic, environments. While their origin, emphasis, and preferred
patterns of practice differ, the methods overlap in content and underlying concepts, and we consider them
here together.

There are many other types of related simulation modeling methods. Reflecting the important role networks
have come to play in many agent-based models, we further consider aspects of dynamic social network
analysis as specializations of agent-based models. Similarly, we consider diverse compartmental modeling
techniques – such as those prevalent in mathematical epidemiology since its inception in the 1920s – under
the rubric of system dynamics. Furthermore, simulation models can be used to evaluate and optimize a
healthcare intervention given constrained resources. For example, optimization models can consider the
demand for imaging in the context of the limited availability of imaging capacity and scanning time to
optimize the use of imaging services. Likewise, these modeling studies can be applied to improve
scheduling and hence, to minimize waiting time for patients (Vanberkel et al., 2011)

Simulation Modeling Methods

System Dynamics (SD)

What it is: a simulation modeling method used for representing the structure of complex systems and
understanding the behavior over time

When was it developed: In the 1950s by Jay Forrester at the Massachusetts Institute of Technology
(MIT). Forrester, electrical engineer, developed this methodology as he was faced with the challenges
of managerial problems when involved at the managerial level at General Electric (GE). Managers at GE
did not understand why the employment at their appliance plants in Kentucky exhibited a significant
three year cycle.
From hand simulations of the stock-flow-feedback structure of GE plants, including the existing corporate decision-making structure for hiring and layoffs, Forrester was able to show that the instability of GE employment was due to the internal structure of the firm and not to an external force such as the business cycle. He reasoned that social systems are much harder to understand and control than physical systems.

Later on, in collaboration with Richard Bennett the first computer modeling language (SIMPLE) for system dynamics was created and later refined by Phyllis Fox and Alexander Pugh (DYNAMO). In 1961, Forrester published the first book in the field, Industrial Dynamics (Forrester, 2007).

**What are the main/core concepts, assumptions and characteristics:** The core elements of SD are feedback, accumulation (stocks), rates (flows) and time delays. Stocks are accumulations or aggregations of something e.g. people, beds, oxygen. Flows are rates; these feed in and out of stocks and have the same units of stocks per time unit e.g. people per hour, beds per year, oxygen per minute. Social systems contain feedback processes both reinforcing and balancing.

These feedback processes describe the reactions of agents and the environment to decisions that affect them and their goals. Therefore, our decisions and actions today affect our actions and decisions tomorrow as the environment and other conditions change. An important concept in SD is nonlinearity. This concept is tied to the existence of feedback processes and it means that an effect is seldom proportional to the cause.

One of the core assumptions in SD is that the behaviour of the system is due to its structure and not to external forces or factors. The structure of the system can be understood as the feedback loop structure (processes) and the structure of accumulations and rates which are the ones that generate the behaviours. An example of this is the employment instability at General Electric that Forrester studied. At a more technical level, characteristics of SD are: a) although mathematically SD uses time steps, it is thought of as continuous time; b) higher level of aggregation compared to other methodologies.

**Type of problems it can contribute to solving:** SD can be used for policy analysis and design for problems in complex social, managerial, economic and ecological systems. Any dynamic system is characterized by interdependence, mutual interaction and feedback. Some applications can be categorized as: a) recognition and identification of behavioural patterns in a system, e.g., in an organization; b) gain insight into the processes of a system and the consequences of decisions; c) identification of leverage points and/or structures in the system to generate change; d) reproduction of a given behaviour (reference mode) (J.D. Sterman, 2000).

**Model outputs and level of insight:** Are very varied and dependent on the purpose of the model and the type of problem. In general terms, SD can produce patterns and trends, as well as mean values. SD
allows for the elicitation of mental models from stakeholders involved in the discussions and also from those involved in the model building process. This methodology generates high level of insight about the problem and the system under study at the strategic and policy levels.

**Interpretation of outputs:** Interpretation depends also on the type of problem and the purpose for which the model is designed. The model building process in itself can offer knowledge about the problem, the system and its processes. The patterns and trends result of simulation experimentation with different policies or strategies (“what if” questions) can be analysed by modellers and stakeholders to inform decision making.

The model will not give a unique answer or optimal answer to a problem. Instead, the model will offer the possibility of experimentation with parameters and variables to test alternative strategies for system intervention and observing their potential outcomes.

**Discrete-event Simulation (DES)**

**What it is:** a simulation method used to characterize and analyse specific processes and use of resources originally in business and industrial settings. However, healthcare has become a common setting for its application.

**When was it developed:** DES first emerged in the late 1950s and was developed by K.D. Tocher when working at the Operational Research group at United Steel Companies in the UK. Tocher was faced with constructing a simulation model of one of the steel plants. Even though the company had several steel plants, i.e. different technologies, Tocher envisioned a standard model that would permit to represent any of the plants through changes in parameters.

The concept of General Simulation Program (GSP) was conceived in 1957 and was the first identifiable specialist package. However, it was limited due the machine-level language used to write the codes which made it costly and non-portable. During the late 60s and the 70s several improved versions were created and high-level languages were used. Development of software continues today (Hollocks, 2006).

**What are the main/core concepts, assumptions and characteristics:** The core concepts in DES are events, entities, attributes and resources. An event is a something that happens at a certain time point in the environment and that can affect resources and/or entities. Entities are objects that have attributes and consume resources while experiencing events. Attributes are features or characteristics unique to an entity. They can change over time or not. Resources are objects that provide a service to an entity. Queues are another important concept in DES. Queues occur when several entities compete for a specific resource for which there is a constraint. At a more technical level, time is discrete and change happens when an ‘event’ occurs.
**Type of problems it can contribute to solving:** Most problems or questions that DES can help analyze are those regarding resource utilization and queues i.e. wait times. In addition, in healthcare specifically, DES can be useful to also analyze effects on health related outcomes. DES is also useful for problems where it is particularly relevant to be able to capture the changing attributes of entities e.g. patients; and where the processes to be characterized can be described by events. (Karnon et al., 2012)

**Model outputs and level of insight:** The outputs of DES are generally mean values and distributions of values. This methodology generates high level of insight about the problem and system under study at logistic/operational level.

**Interpretation of outputs:** outputs of DES can be interpreted or used for system performance indicators such as resource utilization, wait times, number of entities in queues, and throughput of services or products. Also, scenarios with different strategies and policies (“what if” questions) can be tested. The mean values or distributions can be thought of as accurate; however, not necessarily optimal.

---

**Agent Based Modeling (ABM)**

**What it is:** As described by Gunal MM, ABM is a simulation method for modeling dynamic, adaptive, and autonomous systems. It is employed to discover systems by using ‘deductive’ and ‘inductive’ reasoning. At the core of an ABM model, there are ‘autonomous’ and ‘interacting’ objects called agents. Agents are social and interact with others and they live in an environment and their next actions are based on the current state of the environment. In addition, an agent senses its environment and behaves accordingly based on simple rules defined. Agents may have explicit goals to maximise or minimise, may learn and adapt themselves based on experience. The definition of agent behaviours ranges from simple ‘if - then’ statements to complex models, for example cognitive science or artificial intelligence.

**When was it developed:** In 1971, Thomas Schelling used ABM to propose a theory to explain the persistence of racial segregation even though the legal and cultural environment was one of growing tolerance. Using a basic ABM model with simple if-then statements (e.g. individuals will tolerate racial diversity, but will not tolerate being in a minority in their locality) he was able to show via colored squares on a matrix, that segregation will still be the equilibrium situation.

**What are the main/core concepts, assumptions and characteristics:** The three core concepts that form the basis for ABM are: agency, dynamics, and structure (Borshchev & Filippov, 2004). Agency means that agents (e.g. patients) have goals, beliefs and can act. These agents can move through space and time, interact with each other, learn, and disseminate new learnings to other agents in their social network. Dynamics means that both the agents and their environment can change, develop, or
evolve over time. Structure is emergent from agent interaction. For example, how human populations will tend to aggregate in certain locations based on pre-defined behaviors that have been coded into the agents. All of the above factors can be directly modelled.

**Type of problems it can contribute to solving:** ABM has been applied in a variety of modeling scenarios: market forecast, human migration and movement patterns, resource management (e.g. water), and political mobilization. The widest use of ABM related to population health has been to model large-scale anthropogenic or natural disasters, such as a chemical spill, infectious disease outbreak, fire, hurricane, or flooding. The response of the affected population is driven by available information about the event, behaviors (e.g. evacuation) and containment strategies (e.g. vaccination or quarantine) (IPLER, 2008).

As an emerging best practice for health outcomes researchers, ABM is potentially well suited to informing public health planning and policy, as well as investment decisions in healthcare infrastructure. The attainment of specific population health goals can be simulated at the population level, and the investments needed to achieve these goals can be estimated. Primary goals can be defined by disease outcomes, efficiency measures, ROI, or costs (Kruzikas et al., 2014).

**Model outputs and level of insight:** As applied to healthcare systems, ABM model outputs can include health outcomes (e.g. QALYs, mortality), disease patterns and trends (e.g. viral transmission, diabetes), costs, resource utilization, and labor productivity (e.g. patients treated per day, bed occupancy). ABM is well suited to generate insights into the health of large populations over time.

**Interpretation of outputs:** The strength of interpretation lies in the conduct of sensitivity analyses. By varying the interventions applied to the healthcare system, (e.g., introducing a new diabetes prevention program vs. lowering the co-pay for diabetes medications), ABMs can be a powerful tool to test assumptions, assist planning, and forecast the effects of different health system scenarios on population health.
Characterization by resource requirements for each of the three simulation modeling methods –
time, data, money, knowledge about problem.
Depending on the scope of the model the requirements might change. According to Jun et al. 2011 the
resource requirements are:

<table>
<thead>
<tr>
<th></th>
<th>Time</th>
<th>Data (quantitative and qualitative)</th>
<th>Cost</th>
<th>Knowledge about the problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD</td>
<td>from days to years</td>
<td>from none to extensive</td>
<td>From tens to thousands</td>
<td>from moderate to complete</td>
</tr>
<tr>
<td>DES</td>
<td>from weeks to months</td>
<td>Good quality statistics</td>
<td>From hundreds to hundred thousands</td>
<td>from moderate to complete</td>
</tr>
<tr>
<td>ABM</td>
<td>From weeks to months</td>
<td>from none to extensive</td>
<td>From hundreds to hundred thousands</td>
<td>Moderate to complete</td>
</tr>
</tbody>
</table>

DES requires good quality data in the form of distributions and probabilities. SD and ABM can handle many
types of data. Quantitative attributes such as patient age, compliance with drug therapy, and size of
hospital (e.g. number of beds) can be modelled. Qualitative attributes such as ordinal or categorical data
(e.g. ethnicity), or relational (e.g. I am linked to him and her), or data requirements that are more vague (e.g.
A sends B a message about one time in three). As a result, ABM is able to model a variety of outcomes,
such as epidemiological disease burden, population socio-demographics, health status, system utilization,
patient preferences, healthcare provider preferences, behaviors and costs (Kruzikas et al., 2014; Sobolev,
Harel, Vasilakis, & Levy, 2008).

Why the need for Good Practices for Simulation Modeling?
The translation of evidence into policy and clinical care through implementation in the healthcare system are
core issues facing healthcare delivery system transformation around the world. Implementing evidence
based practices can be achieved through the aid of operations research methods to reengineer healthcare
delivery systems and improve patient outcomes and health system performance (Berwick, 2009). The
hierarchical relationship between the health system, providers and the patient demand a level of complexity
that can be captured using simulation modeling methods.

Although operations research methods are widely used in industrial and business operations optimize
processes and improve effectiveness and efficiency, they are still relatively new in health applications (J.D.
Sterman, 2000). There is a lack of clear and accessible guidance for using simulation modeling methods to
evaluation interventions in healthcare delivery systems. This guidance on good practices in simulation
modeling methods is important to the scientific field because traditional health technology assessment and modeling methods are often of limited usefulness when applied to health systems.

Recently, there has been noticeable growth of studies applying simulation modeling methods in health research and health system management. The feasibility and relevance of these methods to inform healthcare delivery system planning and decision making for improving system efficiency have been demonstrated (Brailsford et al., 2009). Developing good practice guidelines will advance the application of simulation modeling methods in health.

### Simulation Modeling Good Practices

**Criteria for selecting simulation modeling method**

Selection of simulation modeling method for the entirety or portions of a model will generally consider a variety of considerations. Most fundamentally, these include a focus on model purpose, the problem being investigated. This focus on model purpose reflects two facts. First, all models – like maps – are abstractions that are “wrong” in the sense that they are unable to address all details. Second, it reflects the fact that the modeling methodologies discussed here differ even more fundamentally in terms of their aims and the questions that they prioritize than in the details of the formalisms. For example, while SD modeling emphasizes representations and processes that help shift stakeholders’ mental models, agent-based modeling emphasizes agent-agent and agent-environment interaction and multi-scale insights, and discrete event modeling emphasizes insights into the impact of resources on process efficiency and throughput.

Also of key importance in method selection are the degree to which one is seeking to capture inter-individual interactions, the availability of requisite skill sets (e.g., recourse to software engineering expertise for ABM), available level of process-related knowledge and empirical data, the level of performance viewed as acceptable, flexibility sought in model scope (e.g., types of heterogeneity incorporated – more flexible for individual-based models than for SD/compartmental models), the nature of the interventions or counterfactual situations to be represented, the character of the outputs of interest, and the importance of differences between individuals according to characteristics, history, and spatial, network context, the importance of insights at multiple scales, the need to support scaling to large or highly heterogeneous populations, and the degree to which one is seeking the simulation to reproduce statistical variability, and whether one seeks to use tools to assist in model analysis or to reason about the possible behavior of the simulation over broad ranges of parameters, rather than concerning discrete particular scenarios.

Modeling approaches individuating system actors are notably attractive for capturing information regarding or implementing interventions or governing processes that depend on agent history (e.g., on an individual's past care pathways), and learning and memory effects. The capacity to maintain such longitudinal
information raises particular opportunities for calibration and validation against data individual-level sources. Such individual-based models also shine in representing large amounts of heterogeneity, in contrast to the coarse-grained disaggregation that occurs in aggregate model, which scales poorly with the rise in the number of distinctions by which we wish to capture heterogeneity.

As a result, individual based models confer substantial advantages in capturing not only diverse continuous and discrete attributes (e.g., sex, income, BMI, birth weight, and preferences), but especially evolving conditions such as co-morbidities, and capturing (possibly dynamic) situated context (spatial or network position and connections, with associated exposures, localized perception, resource availability, choice sets, influencing local factors). ABMs support not only great scalability but also great flexibility in representing both discrete and continuous heterogeneity. Adding – or removing – a new dimension of heterogeneity to an ABM is a lightweight, modular operation.

**Figure 1. Criteria for selecting simulation modeling method**

1. **PURPOSE**
   - What is the purpose of the model? What is the problem being investigated? Why are we building this model?

2. **OBJECT**
   - What is the scope of the model (boundary)? What are we modeling? Is it feasible?*

3. **METHOD**
   - How to model the object to achieve the purpose?

*Thickness of lines mean stronger relationship
This contrasts with the situation in aggregate models, where a similar change is a heavy-weight operation affecting the structure extending across much of the model. As a result, ABMs support more nimble experimentation with the degree to which heterogeneity is considered. Of particular note in ABMs is the capacity to capture empirically grounded, rich models of individual decision making (for example, using elements of discrete choice theory, or cognitively science inspired models of decision making), which can aid in endogenously capturing behavioral responses of the population to interventions.

The ability to capture such heterogeneity can aid not only in capturing behavioral variability in underlying processes, but also in evaluation of targeted interventions in examining transfer effects of interventions. Moreover, the actor-centric character of agent-based modeling further supports the straightforward and transparent construction of multi-level models, whose structure mirrors that of the external world. Such models can then be used to characterize emergent behavior at multiple distinct levels of a system – opening opportunities not only for understanding intervention impact on and across multiple levels of intervention, and enhancing calibration and validation mechanisms.

The stochastic character typical of individual-based models imposes represents both an asset and a liability. On the positive side, the presence of such stochastics supports substantial analysis insights, such as explaining empirical variability, testing of interventions and scenarios under expected uncertainties. However, the presence of the stochastics also imposes a substantial performance burden atop the already heavy computational demands of agent-based models. In a similar fashion, the flexibility of agent-based models is a double-edged sword, permitting a tremendously wide repertoire of possible model designs, but in a way that current software technologies require recourse to software engineering expertise. For larger models, this can require individuals with programming (preferably, software engineering) background to be involved in model construction, maintenance, debugging, and quality assurance.

While system dynamics models can be applied at a wide variety of levels of aggregation – with some classic models involving individual-level use of such models, they are most commonly used as compartmental models at the aggregate level. In contrast to agent-based models, smaller system dynamics models are often far faster to construct, maintain, and – often – to understand. Both because they involve fewer “moving parts” and are not stochastic, such models are also typically far more rapid to execute than are ABMs and DES models. Because of its reliance on creative use of a small modeling vocabulary – most centrally, stocks and flows – less computationally specialized skillsets are required for working with system dynamics models.

Due to the lower software engineering burden, often projects can spend considerably more time learning from the model, rather than maintaining it, and that the turnaround in learning from model changes is considerably faster. The presence of a spectrum of mechanisms – from qualitative techniques such as causal loop diagrams to fully specified and parameterized simulation models – often means that system
dynamics modeling projects deliver early and ongoing insights. The design of system dynamics formalisms
to be readily understandable not only eases model building, but also distinguishes the technique in terms of
its support for participatory processes.

By virtue of its use of both qualitative and quantitative mechanisms that can be viewed and understood by
those from diverse walks of life, system dynamics supports refined and time-tested processes for use in
group-model building, thereby securing diverse benefits, including helping to elicit stakeholder and
community mental models, break down barriers to effective communication amongst those groups, and can
aid greatly in securing buy-in and energizing a group. This low execution burden not only strengthens the
learning curve, but also supports participatory settings that leverage rapid interaction with scenarios
formulated by stakeholders, community members, or other non-technical participants. While many system
dynamics practitioners do not seek to conduct formal mathematical analyses to their models, system
dynamics is distinguished by its capacity for supporting closed-form analysis of models.

This ability not only allows for general reasoning about simpler model – for example, supporting an
understanding of all possible behaviors of that model, over wide ranges of assumptions, or in clearly
revealing the long-term behavior of that model under similar condition – but also supports a wide variety of
analytic tools in software that leverage such analysis to provide powerful insights to the modeller – e.g.,
Kalman Filtering to recurrently reground a model in as new empirical data becomes available, analyses to
identify model feedbacks governing model behavior at the current time or to identify particularly fundamental
assumptions, etc.

Discrete event simulation offers particularly great strength in the context of defined (although potentially
contingent) workflows, associated with multiple stages of processing for some class of discretely defined
entities (patients, vials of vaccine, etc.), and where this processing is contingent upon the availability of
resources of one or more types. Such resources may be fixed in space (an MRI machine, an examination
room), portable (a blood pressure cuff, wheelchair, or an IV drip), or be mobile with agency (a clinician, a
nurse’s aide, etc.). DES can be an exquisitely effective and concise, and straightforward tool for capturing
situation with rather more passive agents, particularly capturing queuing behavior, the impact of resource
availability, arrival time distributions, etc. on waiting times, throughput, queue length, resource utilization,
quality of care conferred and other outcome measures of interest in health services research.

By mapping agents to space, DES can further represent the impact of the physical environment – for
example, facility layout and resource placement – on such outcome measures as an emergent
phenomenon. While DES – like agent-based modeling – is characterized at an individual-based level, it
would require far more time, effort, and software engineering skills to capture such interactions between
resources, agents, spatial layout, queueing, etc. within an agent-based based model. By contrast, while
DES shines in representing comparatively more passive entities that are “operated upon” by processes, it
offers comparatively less flexibility for representing situations when entities need to interact in a flexible way with each other, with the environment, or otherwise exert a high degree of agency.

**Model design and assumptions**

The most fundamental decision involving design of a given model version is that related to model scope, particularly with regards to the reflective delineation of factors falling into each of 3 categories: a.) endogenous factors (calculated as part model operation), b.) exogenous factors (represented in the model, but according to pre-specified assumptions, such as given by using constant values or time series), and c.) factors that are consciously ignored.

For making decisions involving model scope, it is of great importance to maintain an understanding of model purpose, and highly advisable to adhere to an incremental model development methodology (see below). Such discipline is particularly important for the ABM approach, whose flexibility raises the risk of scope creep and overly casual inclusion of additional factors. Other fundamental decisions to be decided up front include the scope of the model population, temporal and spatial scales, including time horizon, spatial extent and topology, and any discretization imposed.

**Documentation: pre, during, post**

Maintaining model documentation is essential not only for communicating and sharing models, but is further important for avoiding model defects, enhancing transparency of the model, reducing work associated with model changes, supporting new members of the team, and facilitating model evolution. Documentation of model scope has been discussed above; documentation of the model population and scales are also key importance.

In addition to the clear delineation of both data sources and parameter assumptions, it is best to clearly document the formulations used to derive of such model parameter values – for examples, steps of and sources for “backing out” calculations, aggregation and transformations applied. It can further be extremely valuable to identify the units of measure associated with model quantities. This can be valuable not only for identifying model defects (automated by some software packages), but also for combating the “curse of dimensionality” associated with sensitivity analysis and model calibration by reducing the count of parameters required for such processes, for judicious choice of model scenarios, and for creating scale models.

During model construction, a clear record of model changes (as maintained manually or by version control systems) can be instrumental in resolving model defects. Given the important role that reproducibility of scientific results has traditionally played in scientific research, it is desirable to publish a sufficient degree of detail on a model’s formulation as to permit recreation of that model; this includes not only the fundamental specification (“source code”) of the model, but also the framework in which it was built, model parameters, any external sources of data used, and initial conditions. Beyond this basic criterion, it is important to put in
place abstract specifications of a model in terms of building blocks whose semantics is precisely understood, and which declaratively specify the model – characterizing it in terms of what is represented, rather than all the details as to how that is captured.

Most importantly, for the growing number of models whose implementation relies more heavily on algorithms and computer code, this entails specification beyond the associated code, preferably in terms of mathematical formalisms (e.g., finite state machines or probabilistic automata, flow-charts, various forms of UML diagrams, etc.), or – at the very least – in terms of pseudo-code. Throughout a model's construction and refinement, reproducibility concerns make it essential to maintain successive versions of the model under software-supported or manual version control.

During model use in scenario examination, it important to routinely report metadata indicating time of model start, stop and elapsed CPU and clock time, model version and parameter assumptions as part of or cross-linked to scenario results. It is further important to report hardware and software platform used to run the scenarios (including modeling software and operation system versions), model software settings (e.g., specifying prioritization the handling of simultaneous equal-prioritized events, time steps and numerical integration routines employed), random number seeds used, as well as the objective of model scenarios. For model scenarios which are defined as part of larger structures (e.g., sensitivity analyses/parameter sweeps, Monte Carlo ensembles, MCMC methods, extreme value tests, systematic investigation of intervention options or exogenous shocks, etc.), the structure and parameters of that system should additionally be defined in terms of metadata to that for particular discrete scenarios (intention behind the run, time start and finish, model version, etc.) both input and output. Input concerns include the factors such as the sampling mechanism used (e.g., MCMC algorithm and parameters), definition of the sensitivity grid involved and count of iterations, realization count per iteration, distributions from which parameter values are drawn, etc.

In terms of output aspects of such structures, documentation should include discussion of aggregation, summarization, statistical methods applied (preferably with associated references) and their parameters. Where possible, both input factors (e.g., preprocessing code, databases, text files, spreadsheets) and output factors (e.g., code, syntax files, scripts, spreadsheets, and database queries) used for reporting results should be cross-linked to outputs, copied or otherwise made available in an immutable fashion, and placed under version control.

Many models go through calibration processes, whereby simulation model parameters are adjusted such that the emergent behavior of a simulation model compares most closely with empirical data. For such processes, information should at least be provided regarding the calibrated values of model parameters. However, it is further desirable to thoroughly document the calibration process, to the point of reproducibility. This includes not only the parameters adjusted, their bounds and other constraints, and
empirical data, but also the optimization algorithm(s) used and associated parameters, objective functions employed, the weighting used for discrepancies involving different data points, etc.

**Model building strategies**

Given the counter-intuitive behaviors frequently associated with simulation models and their complexity, modellers are advised to proceed incrementally and adaptively, cycling through steps of adding small pieces to a model, running the model for insight (sometimes with the new piece first disabled, cross-checked for invariant behavior, and then enabled). This approach helps to maximize learning by fostering understanding the specific contributions of newly added model pieces to model behavior. It further aids stakeholder morale and involvement. In addition, incremental development enhances model quality by helping to ensure that latent defects are spotted as quickly as possible (both due to visible behavior and when running suites of formal automated tests).

Finally – and perhaps most importantly – the flexibility associated with the incremental development process helps contribute to a more responsive model and greater prospects for learning, by allowing the models' evolution to be shaped by learning throughout the process. The model that emerges from such a process of continual learning is likely to be considerably more responsive to project needs than would have been anticipated up front.

**Data Requirements**

Simulation models use empirical data in two primary capacities. The first use is for model parameterization (with the data being incorporated directly or indirectly – for instance, via backing out – into model formulation). The second use lies in model calibration, where the data is used as evidence to match against emergent behavior of a model. For the first of these uses, appropriate documentation is particularly important, and probabilistic sensitivity analysis can be of high value in examining response of model outputs – and particularly cross-intervention trade-offs – in light of uncertainties concerning param

**Model validation**

Both validation and verification processes are key for models; as we use the terms here, validation focuses on the correspondence between a model and the real-world phenomena under investigation or to be addressed, while verification seeks to understand to which the model actually implemented is true to what was intended to be built (i.e., to its design). There is rarely an opportunity to genuinely validate a model; rather, the model building process is often associated with a confidence-building process as the model faces successive opportunities to be falsified and cross-checked by additional observations.

Traditional model validation approaches vary widely among the different modeling methodologies discussed here, with discrete event modeling having a particularly extensive body of practice on the subject; we note here just a few. Of particularly great interest is the use of cross-validation, which tests the predictive power
of a model by reserving a subset of empirical data for testing the model’s predictive abilities, having used the balance of that data for constructing that model. The importance of unit checking has been alluded to above, and retain particular importance because of the high risk of combining values with incompatible units or (much worse) of incompatible dimensions – yielding numerical values that are precise but lack meaning. Extreme values tests help demonstrate reasonable behavior of the model under a wide variety of assumptions.

The model calibration process – seeking to match emergent model output against empirical data – often provides much additional confidence into model suitability and fitness for purpose. Both calibration and cross-validation processes have particular texture for individual-based models, as they can leverage additional types of data (e.g., longitudinal data, spatial and topological patterns, and patterns at multiple scales), but the data is almost never sufficient to unambiguously estimate model state. Finally, primary data collection to check model predictions – particularly following interventions – can be particularly valuable for raising confidence in a model.

Model verification plays a key additional role in helping to enhance confidence in a model, and reduce risk that a well-designed model will be implemented incorrectly. Implementation defects can run the gamut from misunderstanding of operator precedence (and misplaced or absent parentheses) to use of a division operator where a multiplication operator was intended, forgetting or neglecting to include a new variable in a relevant total, to misunderstanding the meaning of or order of the arguments associated with function to draw a uniformly distributed random number. While technically an aspect of model design, we will also consider here defects associated with neglect of numerical analysis concerns (e.g., numerical instability caused by an overly coarse timestep).

The theory of model verification draws heavily on principles and practices of software quality assurance. Some of the most important factors promoting model quality have to do with process commitments on the part of the modeling team. Examples include adherence to regular peer review in both informal (e.g., pair modeling, peer desk check of models) and formal (formal model inspection). While such peer review efforts are the rule in software projects, model inspection is worthy of particular note for modeling, as it is distinguished as a recognized best practice in the software development, and has been shown to be both more efficient and more effective than other powerful quality assurance practices, such as software testing.

Where a model is being worked on by more than one party, such best practices further include practices such as buddy testing (where one party tests the model components contributed by another party), continuous integration efforts and their associated smoke tests, automated testing, stylistic analysis, etc. While little applied in modeling, formal strategies for estimating the occurrence of latent model implementation defects can also confer considerable value for larger model development efforts.
Important technical best practices for verification include use of assertions to check model assumptions (e.g., that stocks of physical quantities are non-negative, that one model variable is strictly less than another), automated and unit and model-wide test suites. Where possible, use of techniques such as mocking (to allow for more rigorous and precise testing). Models which include considerable levels of algorithmic specification will also benefit from strict adherence to quality coding standards (widespread use of commenting, short methods/functions and widespread abstraction mechanisms for common patterns of code and conditions, elimination of manifest constants), architectural principles (separation of concerns – for example, model observer processes from visualization specifics from governing processes from Database layer operations, and possible use of aspect-oriented mechanisms to capture cross-cutting concerns). Such models will also benefit greatly from periodic refactoring of code – improving the clarity, modularity, transparency, flexibility, and generality of the implementation code without changing its current behavior.

**Analysis of outputs**

A key use of models lies in analysis of outputs associated with model scenarios. Sensitivity analysis – structure and parameters (including one way, multi-way parameter sweeps and probabilistic analysis) – is highly important in understanding the variability of a model under different assumptions. The details associated with model analysis differ considerably between modeling types, with stochastic models typically requiring ensemble-based analyses. Dimensional analysis can be particularly of particular value to allow for the most judicious scenario selection, selecting the variables to be varied so as to maximize learning. As noted above, recourse to unambiguous information on model version and parameter assumptions is key for reliable learning from model outputs.

**Reporting**

Reporting of model results should place a premium on reproducibility. The basic mechanisms that foster this were emphasized in the sections above, and maintaining such information in the modeling project fosters its sharing when reporting model results. Also of key concern in many contexts is communicating model limitations. For all models, it is important to articulate model scope, purpose, limitations and assumptions to stakeholders. Such communication is particularly key for deterministic models, such that those consuming reports of model output recognize that there are broad bands of uncertainty associated with model results.
Results and conclusions - Identify areas where continued methodological development or standard definitions are needed. [This section will be completed after review comments are received and discussed.]

Recommendations – Uptake of this research by healthcare practitioners – how to apply in practice and confer value. [This section will be completed after review comments are received and discussed.]
References


