“Turning mirrors into windows”: A study of participatory dynamic simulation modelling to inform health policy decisions

Louise Freebairn
Chapter 2: Literature review

This chapter includes a detailed analysis of the literature underlying the rationale for this thesis. It is divided into seven sections describing: the challenges of evidence-informed decision making; how the knowledge mobilisation field has developed to facilitate evidence-informed policy; how systems science methods can be applied to public health issues; the limitations of traditional statistical methods to analyse complex public health problems; the core concepts and methods of dynamic simulation modelling (DSM); how participatory methods can be applied to DSM and the important gaps in knowledge to be addressed.

2.1 Challenges of evidence-informed decision making

Government and public policies have profound impact on the lives and health status of populations, and therefore it is important to ensure that policies are cost effective and mitigate the likelihood of negative outcomes [1, 2]. Ensuring that policies align with research evidence is likely to result in higher quality and effectiveness [3]. However, challenges in the use of evidence to inform policy remain [1, 3-7] and these are explored below.

Evidence-informed decision making is defined as the process of distilling and disseminating the best available evidence from research, context, and experience (political, organisational) and using that evidence to inform and improve public health practice and policy [2]. The barriers to evidence-informed policy include issues relating to the relevance
and reliability of research findings, policy makers’ skills in interpreting research evidence, costs of conducting research and poor alignment between the focus and timing of research results and the answers required for policy [7-9]. Key questions chosen by researchers may not align with the information priorities of decision makers, nor are the findings always presented in a form that is useful for or relevant to the decisions at hand [5]. A traditional investigator-driven approach to research that fails to adequately engage key stakeholders within health care systems is a known barrier to translating research into real-world settings [10]. Institutional characteristics, including the perceived importance and priority placed on research evidence, the availability and access to research, and the training of policy officers to engage with, assess and use research evidence have also been identified as important barriers to the use of evidence [2, 7, 8, 11, 12].

A frequently identified facilitator for the use of evidence in policy is the quality of relationships and collaborations between researchers and policymakers [7, 13]. Interaction and exchange between researchers and policy makers can facilitate the use of evidence in policy, however mechanisms and processes to promote these interactions are needed to extend the exchanges beyond familiar networks and existing relationships [9].

Consideration of research evidence within the context in which it will be used is also essential for effective policymaking and practice. The social and political context and the many forces at work in the policy environment provide challenges to integrating research evidence into policy and practice [9, 14]. The decision-making processes for researchers and policymakers are significantly different, both in terms of the “real-world” steps in decision making and also the factors that drive decisions [1]. Researchers rely on experimental and observational scientific studies to test specific hypotheses in a systematic
way and their influence is based on their specialised knowledge. On the other hand, policymaking is built on a complex combination of competing priorities including the political environment, history of related policies, demands from advocates and stakeholders, resource constraints and public perceptions of the value of the policy alternatives [1, 9, 14-16]. Decisions are often the result of compromise [1]. Even when based on sound scientific data, some decisions may not be considered ready for policy action due to lack of public support or competing policy issues [1].

Evidence provided to decision makers is often not in a form that is most useful for them [13]. Policymakers are looking for evidence that is timely, synthesised, contextualised for their local environment, demonstrates priority for an issue over many others, illustrates the policy implications of research findings, contrasts policy options and personalises the issue [1, 2, 11-13, 17]. Evidence dissemination preferences also vary between researchers and policymakers, with researchers ranking publication in peer review journals as their first preference for dissemination, whereas policymakers ranked seminars, webinars and workshops as their most preferred way of learning about research evidence [2].

2.2 Knowledge mobilisation to facilitate evidence-informed policy

A range of terms, including evidence-based policy/practice, knowledge translation, knowledge exchange, knowledge to action and knowledge mobilisation have been used to describe activities associated with facilitating the creation and sharing of research-informed knowledge to guide policy development [3]. Many of these terms are overlapping and are used interchangeably [18]. Some have been used as nouns to describe the process as a whole that results in the use of knowledge by decision makers whereas others are
used as verbs to represent actions or specific strategies taken to facilitate the uptake of research evidence [18]. The term knowledge mobilisation has been identified as encompassing the broadest range of activities and reflecting the non-linear, complexity of the process [3] and has been adopted for this research.

Best and Holmes have described three generations of knowledge mobilisation models: linear, relationship, and systems models [19]. The linear model involves a one-way process in which researchers produce new knowledge, which is disseminated to end users, and then incorporated into practice and policy [4]. In this model, knowledge is seen as a product that is supplied to users, can be generalised across contexts, and whose use is dependent on effective communication of results [19]. The relationship model incorporates the flow of information using principles from the linear model, however focuses on linkages and exchanges between the researchers and users of the information [3]. In the relationship model, knowledge generation is seen as a social and situational process arising from multiple sources (research, theory, policy, and practice), and not solely from the researcher [19]. In this model the use of evidence is seen to be dependent on effective relationships and processes [19].

The third-generation systems model builds on the linear and relational models and acknowledges that public health issues are often best understood as complex systems. They are dynamic and constantly changing, involve interdependent systems (e.g. individual, organisation, and community), have feedback effects and intervening in one part of the system can have unexpected ripple effect on other parts of the system [19-21]. Complex adaptive systems self-organise and adapt based on experience, meaning that from the
same starting point, an intervention can potentially have several different outcomes [22, 23].

Many models, theories and frameworks have been developed for knowledge mobilisation [24]. Twenty five key frameworks were identified in a recent comprehensive review of knowledge mobilisation in the United Kingdom [3], however previous reviews identified up to 60 frameworks in use [18]. There is overlap among the frameworks and common aims, phases or domains can be identified [3, 18, 24]. These phases vary from project to project but frequently include: clarifying the issue that needs to be addressed; negotiating the purpose and goals for the knowledge mobilisation activity; identifying, reviewing and selecting the knowledge that is relevant to the problem (e.g. new research findings, knowledge synthesis or knowledge tools such as practice guidelines); considering the connections and relationships that need to be established; identifying the people and positions who should be involved; developing action plans and resources needed to operationalise the knowledge mobilisation; and considering the potentially facilitating or inhibiting effects of the local context on knowledge mobilisation efforts [3, 18].

Systems approaches are well suited to address current, complex public health issues and systems [2, 25] and can be used to analyse both complex health issues as well as the context within which they emerge. Systems approaches, such as participatory DSM, can extend and build upon key elements of knowledge mobilisation including the synthesis of diverse knowledge and evidence, investigation of dynamic and non-linear relationships within systems and exploration of adaptive, emergent behaviour of the system in response to policy interventions. The details of the synergies between systems approaches and
knowledge mobilisation are explored further in Chapters 3 and 4 and therefore not repeated here.

2.3 Systems science methods and public health

Public health issues are complex, with individual heterogeneity embedded in multilevel social and environmental contexts [26]. There are intricate networks of factors, including the physical, biological, ecological, technical, economic, social, and political, that impact on public health [22, 23]. This complexity can hinder both the generation of knowledge and the implementation of evidence-based health policies [23]. For most public health problems there are many interacting risk factors that need to be considered as important contributors to the issue [27]. However, the interactions between these risk factors and delays between exposure and outcome make it difficult to identify which risk factors are most important to target and to forecast the likely impact of policy interventions [23, 27]. Changes in behaviour also occur naturally, and continuously, as people within the system acquire new information that alters their understanding. Planned change in such a system is difficult because of these dynamic characteristics: nothing stands still while the intervention is being implemented [4].

“Systems science” is a broad term referring to a family of analytic approaches that aim to uncover the behaviour of complex systems and inform policy and program interventions [26]. Complex systems are made up of interconnected and interdependent parts. The behaviour and characteristics of a whole system cannot be anticipated, and often differs from, the behaviour and characteristics of any one element in that system when these are considered separately [22, 26, 28, 29]. Characteristics that distinguish complex systems
from more simple ones include: the presence of many interrelated components of the system, bidirectional relationships between components (also known as feedback loops), non-linear relationships among components, self-organisation or adaptation of the system in response to interventions (policy resistance), delayed effects from exposure to outcome within the system, and changes in the system behaviour over time (temporal dynamics) [22, 23, 26, 30].

In public health, such complexity is common and can be a significant challenge for the design of public health policies and interventions. The interconnected dynamics of complex systems can result in potential synergies, which may be overlooked in traditional methods of policy design [23, 31]. Tipping points, where small actions can lead to large change, are important levers that can be identified and utilised in policy decision making [32]. However, policy responses should also consider that successful interventions in one part of the system may be counteracted by negative responses elsewhere [23, 31]. Sophisticated methodological and analytic tools, such as DSM, are useful to explore policy and program scenarios, such as whether an intervention works as intended, for whom, under what conditions, at what cost, how soon, and for how long [22, 23, 26, 27, 33].

2.4 Limitations of traditional statistical methods to analyse complex public health issues

Traditional epidemiological statistical techniques have contributed substantial knowledge in health research, however their inherent assumptions and characteristics result in important limitations when applied to complex systems, like those in public health policy. These limitations include reductionism, assumptions of independent associations between
attributes and effects, reliance on methods that assume linear relationships and an inability to accommodate time related dynamics of the system. These limitations and their impact on the policy decision making discourse are discussed in detail below.

Complexities within disease prevention science have commonly been dealt with by employing reductionist analytic approaches that focus on reliably estimating each component of a system [26] and reducing the system to a series of isolated and independent associational effects from which causal processes are inferred [22, 34]. Invaluable knowledge has been gained with these reductionist empirical approaches, including the discovery of the link between smoking tobacco and lung cancer and asbestos exposure and mesothelioma [22]. However, sole reliance on them may result in failure to achieve adequate understanding of broader system behaviour shaping some of the most pressing public health and disease prevention problems [22, 26, 33].

Traditional research methods such as randomised control trials may be viewed as the most rigorous scientific design for evaluating intervention effectiveness, however they also have limitations for real-world, policy-relevant research, when the exposure (i.e. policy issue) cannot be randomised, may be subject to time delays or may emerge over time as the system adapts and changes [1, 26]. For example, obesity is an important public health issue which is impacted by multiple interdependent systems eg. biological, behavioural, social and environmental [33]. Exposure to risk factors for obesity, such as family history or built environment, cannot be randomised and may occur many years before the onset of the condition e.g. dietary habits established in infancy and early childhood may result in the onset of obesity in adolescence or early adulthood. Reliance on reductionist methods, that attempt to isolate causal relationships between risk exposure and development of disease,
may hinder insights about these complex systems that could be important for effective intervention design or management of systems within which interventions are delivered [26].

Multilevel statistical analyses are often used to summarise data and estimate “independent” associations with individual-level outcomes to test hypotheses [35]. These techniques apply statistical controls for individual attributes, e.g. age or education status, believed to be simultaneously related to a health outcome of interest, e.g. development of diabetes mellitus, to isolate and investigate the impact of an independent variable e.g. obesity. However, these regression-based approaches necessarily simplify complex interrelations [36]. The focus is on decomposition of variability and estimating “independent” effects and this necessarily isolates elements from each other and ignores feedback loops e.g. the reinforcing feedback loop that results in increases in bodyweight associated with increasing age [35, 37].

Regression approaches are, therefore, not equipped to investigate the processes embedded in complex systems characterised by dynamic interactions between heterogeneous individuals and between individuals and their environment with multiple feedback loops and adaptation [27, 35]. By attempting to isolate the effect of changing a single factor while holding all the other features of the system constant, the context of dynamic interactions and feedback loops is excluded. This results in findings that may not be generalisable to other contexts, i.e. the effects of changing a single factor may be contingent on, or influenced by, dynamic relationships within the context of the system [35].
Another frequently used statistical approach is the attributable fraction, that estimates the comparative burden each risk factor contributes in a given population and the proportion of that condition that could be averted by targeting specific prioritised risk factors [38]. Attributable fractions are often used as a static measure that considers a fixed scenario, for a specific point of time, and assumes a risk distribution that remains unchanged over time [39]. The assumptions underpinning the attributable fraction are that exposure variables are independent, and relationships between exposures and outcomes are unidirectional, linear and constant through time [36, 38]. This can result in overestimation of the potential effect of an intervention, for example Page et. al. compared an attributable fraction approach with system dynamics modelling to assess the impact of suicide prevention interventions [36]. The authors demonstrated that, by artificially assuming that the population prevalence and incidence of suicidal ideation remained constant over time, the use of attributable fractions inflated the estimated effectiveness of the suicide prevention programs. In contrast, the system dynamics modelling approach allowed for dynamic movement of people in and out of states of suicidal ideation over time and this impacted on the assessment of intervention effectiveness [36].

Effective policy decision making requires approaches that combine mechanisms and explore their interactions, as no one relationship or mechanism is independently able to completely explain all important aspects of the issue [31, 33]. Disease systems involve complex relationships between causes and outcomes [23, 34]. Adaptivity means that individual and population behaviour can evolve based on past history and feedback loops and causal effects can be magnified (i.e., positive feedback) or dampened (i.e., negative feedback) as disease processes progress or social systems adapt [31, 40]. Contextual effects
mean that health outcomes are shaped by specific social, economic, and political contexts and have a high degree of sensitivity to initial conditions of the system [34].

These traditional analytic methods yield few insights into both the dynamic processes of systems, particularly when they involve feedback loops and adaptation, and the strength of associations between risk factors [22, 27]. Temporal dynamics are also important but not captured well using traditional epidemiological approaches [22, 26]. For example, time from an exposure to disease, and time from a given intervention to its impact on disease, are not considered in standard statistical techniques, such as attributable fraction estimates [36]. Understanding these complex processes and temporal dynamics is important for predicting the effects of the intervention under other scenarios and for identifying alternate interventions that may achieve the desired effect [22, 33, 35, 36, 41].

The limitations of these statistical techniques are often acknowledged in epidemiological papers, however may not be made explicit in the communication of results, and in managing expectations of intervention effectiveness, among stakeholders and policy planners [36]. This is important from a policy planning and resourcing perspective, as policy makers need to have confidence in the statistics intended to inform policy decisions [1, 27]. They also need guidance regarding the length of time that an intervention will take to have an effect, how long the effect might last, the impact of behavioural aspects of uptake and participation in the intervention, potential intervention implementation issues and impact of scaling up interventions - and this information is not commonly provided by traditional statistical techniques [36, 42].

These methodological challenges also limit the ability to explore complex causal factors and evaluate “up-stream” policies that target social determinants of health [22]. This can result
in the development of multi-sectoral, comprehensive strategies to tackle complex public health problems in the hope that if more risk factors are targeted in strategies for prevention, they are more likely to be effective [27]. However, comprehensive strategies may not represent the most efficient or effective approach to reducing disease burden at the population level. Rather, they may spread finite resources less intensively over a greater number of programs and initiatives, resulting in a reduced “dose” effect and diluting the potential impact of the investment [27]. For example, “Get Healthy Philly” was introduced in Philadelphia as a multisectoral initiative targeting healthy food access and affordability, tobacco control, built environment facilitators of physical activity, in multiple settings including workplaces, schools and other childcare settings [33]. It involved a range of approaches including partnerships with business, public health messaging and increasing walkability and rideability [33]. Using traditional methods of evaluation, it was not possible to determine which interventions included in the initiative were producing significant effects and which were having minimal or no effect [33]. Yet this would be important information to guide future intervention planning.

Failure to recognise the dynamics and feedbacks of the system, and the way the system adapts and responds, can lead to or exacerbate policy resistance as policy makers persistently react to the problem situation, intervening at low leverage points and triggering delayed and distant, but powerful feedbacks [32, 43]. As a problem intensifies, pulling the same policy levers can trigger a vicious cycle moving the system response further from rather than closer to our goals [43].

Systems science methods seek to “put the pieces back together” so as to understand characteristic system behaviour, not only at the level of the smallest components, but to
also provide insight into the system as a whole [26], helping decision makers understand how multiple variables, factors and interventions interact [44]. Dynamic simulation modelling methods can be used to explore the dynamic complexity that characterises many public health issues and provide guidance on when and how to intervene and likely unanticipated consequences of policy decisions [22, 44, 45].

The ability to test the potential impact of programs and policies in the “safety” of a virtual environment before they are implemented, saves time, effort, costs and resources [44, 46]. Systems science methods can capture “emergent behaviours” of the system, that is, system-wide behaviour that is observed but cannot be attributed to the behaviour of any individual component [22, 26].

2.5 Dynamic simulation modelling

Dynamic simulation modelling is a systems science method that can be used to explore and understand problems that appear in the real world using computer simulations [23, 30, 46, 47]. Dynamic simulation modelling can provide a mechanism to represent a complex system in a simpler form that is more accessible for direct study and experimentation [48]. Complex systems are often counterintuitive, with causes and effects separated in both time and space, and modelling allows experiments to be conducted to see how a system behaves under different conditions and scenarios [23, 47]. DSMs can account for temporal dynamics in estimating the likely population-level impacts of interventions over time [22, 36].
Once the model is built or even during the process of building, it can be used to explore and test our understanding of the behaviour of the system [23, 26, 32, 43, 47, 49, 50]. In many situations we are unable to use real-world experimentation to compare alternative solutions to problems because it would be unethical, too expensive, dangerous or impossible, and in these situations, models of the real system can be used to facilitate understanding [23, 30, 46, 47].

Models should not be viewed as “crystal balls” that can precisely predict the future but as tools that can enhance learning about complex issues and forecast likely outcomes for defined scenarios [20, 47, 50]. Dynamic models can help us more quickly identify inconsistencies between our understanding of the issue and the empirical evidence [50]. Models can also guide future data collection, raise new questions and hypotheses, facilitate the identification of important leverage points in a system and bring scientific rigour to thinking about an issue [23, 32, 43, 47, 51].

Methods in dynamic simulation modelling

The models that were developed in the participatory simulation modelling case studies examined in this thesis applied three commonly used methods in DSM: System Dynamics (SD), Agent-Based Modelling (ABM) and Discrete Event Simulation (DES). Each method was chosen as the most appropriate tool to capture the mechanism being modelled. With advances in modelling software the methods can be used in combination within a single model and this opportunity was leveraged in this research with “hybrid” models utilising multiple modelling methods being developed. These methods, their history and application are outlined below.
System Dynamics

System dynamics is a method for understanding how systems change. It models the relationships between elements in a system and how these relationships influence the behaviour of the system over time [23, 29, 30, 51]. Important elements of system dynamic models include feedback loops, stocks and flows. Feedback loops represent the circular causality in a system i.e. how elements in the system have positive or negative reinforcing effects on other elements [23, 30, 52]. Causal loop diagrams are used to conceptualise the system, identify key variables and important feedback loops [52].

System dynamics is considered a strategic modelling method where the system is modelled at an aggregate level [30, 52]. In this modelling method individuals, for example, people or products, do not appear in the model as individuals, they are represented in “stocks” or accumulations. Individual events, such as decisions or recovery from a disease, are similarly not considered, they are aggregated in “flows” [23, 30, 52].

Jay Forrester created the system dynamics method in the 1950’s. His first dynamic model explained the large fluctuations in production, inventories, headcount and profit in the appliance division of General Electric. Towards the end of the 1960s, his work increasingly turned to public policy issues and the more general term “System Dynamics” replaced “Industrial Dynamics” [53].

The combination of the field’s three defining elements, namely feedback, computer simulation, and engagement with mental models, facilitated the adoption of system dynamics across a wide range of applications [53]. For the first element, feedback, Forrester placed prime importance on patterns of behaviour of feedback systems and the
policies that produced them. He identified that decisions in different parts of an organisation created repercussions elsewhere and that those repercussions eventually fed back to impact on the originator [53].

The second element, computer simulation, brought Forrester's theory to practical realisation [43, 53]. The “what if ” analysis capability of computer simulation brought scientific rigour to policy makers and managers considering the effects of decisions and facilitated the identification of leverage points [32, 43]. As mentioned previously, leverage points are those places in the system where a small shift in one element can produce large changes in other parts of the system [32]. System dynamics practitioners have argued that many policies debated in corporations or government are low leverage and unlikely to result in significant impact or, where leverage points have been intuitively identified, without a systems approach, then actions taken may in fact impact on the leverage point in the opposite direction to that required [32, 43, 53].

Forrester also valued the importance of engaging with the mental models of managers and decision-makers. Mental models are our cognitive understanding of a system [52]; people rely on mental models to understand and manage situations ranging from simple every day decisions, like what to cook for dinner, to developing a high level strategy for a complex policy issue [20, 52]. The purpose of computer simulation is not to provide “the answer” but to create a process through which stakeholders interact with a model to learn about the complex dynamics of the systems in which they were embedded, improve their intuition and create a new mental model which can then become the shared basis for action [20, 52, 53].
The system dynamics approach has two main advantages relative to other policy informatics approaches. First, by emphasising feedback, system dynamics can identify and highlight potential areas of policy resistance [43]. Models illustrate how policy actions can trigger reactions, which can be delayed and unanticipated, that feed back to undermine original policy objectives and even exacerbate original problems [23, 32, 43]. An understanding of the sources of policy resistance is essential for the design of improved public policies.

Second, the feedback approach enables system dynamics models to capture complex dynamics with minimum detail. In contrast to other modelling techniques that generate complexity from detailed depictions of individual agents, the system dynamics approach allows modelers to isolate those dynamics generated by the broader feedback structure of systems [23, 32]. This approach can produce models that are small enough to easily communicate core insights to policy makers, yet sophisticated enough to replicate counterintuitive behaviours [54].

*Discrete Event Simulation*

Discrete Event Simulation (or Discrete Event Modelling) models systems as processes and can be used to explore the impact of policy and program decisions for constrained and non-constrained resource systems [55]. The core concepts of discrete event simulation (DES) are entities, attributes, events, resources, queues, and time.

Entities are objects that have attributes, experience events, consume resources, and enter queues over time as they move through the model [30, 55]. In health care applications, entities are often people with a disease or patients in a service.
Attributes are features specific to each entity that allow it to carry information and, in a health context, could include age, sex, ethnicity, health status, past events, and accumulated costs [55]. These values may be used to determine how an entity responds to a given set of circumstances, for example, the timing and type of past events may influence the likelihood and timing of subsequent events [55]. Attribute values may be modified at any time during the simulation, aggregated with those of other entities, or analysed further outside the simulation (e.g., to estimate mean cost and effect) [30, 55].

Events are generally defined as things that can happen to an entity or the environment. An event can be the occurrence of clinical conditions such as onset of a condition, a diagnostic test result, or progression of a disease to a new stage; resource use (e.g., outpatient clinic visit or admission to hospital); clinical decision (e.g., change in dose); or even experiences outside of health care (e.g., failure to show up at work) [55]. Events can occur, and recur, in any logical sequence.

A resource is an object that provides a service to an entity. DES represents resource availability at relevant points in time (e.g., an emergency department with resourcing for six beds can treat people more quickly than one with resourcing for two beds). In representing resources, DES can capture spatial factors, such as the number of available consulting rooms or distance between a ward and an operating theatre [30]. Queues are an important concept in DES and occur when several entities compete for a specific resource for which there is a constraint [41]. When a resource is “occupied” it cannot be accessed by an entity and the entity must wait, forming a queue [30, 55]. The simulation can be used to identify the utilisation of resources, the time spent in the system or part of the system,
waiting times, queue lengths, system throughputs, bottlenecks and costs of entity processing [30, 55].

Time is also an important component of DES. An explicit simulation clock (initiated at the start of the model run) keeps track of time making it possible to track periods between events (e.g., hospital length of stay, time spent with symptoms, survival) [55]. In health care, events occurring to an individual and how that individual interacts with others, the health care system, and the general environment can be modelled in DES simultaneously [55]. The term “discrete” refers to the fact that DES moves forward in time at discrete intervals (i.e., the model jumps from the time of one event to the time of the next) and that the events are discrete (mutually exclusive) [55]. DES operations can include delays, services provided by different resources, and choices between process branches [30, 41].

The level of abstraction for DES is much lower than for system dynamics. The process diagrams reflect the physical steps that happen in the real-world system. Entities and resources are passive in DES, that is, they have no behaviour of their own, they just carry data. Anything that happens to them is defined by the process flowchart [30, 55].

DES was first used in the 1950’s in manufacturing companies to assist in the improvement of production processes [56]. In 1961, IBM engineer Geoffrey Gordon developed the GPSS (General Purpose Simulation System) which is considered the first software implementation of discrete event simulation [30]. A time-consuming feature of early DES was designing, writing and de-bugging the model’s code [56]. From the earliest days of simulation, there has been interest in creating the means to make this more rapid and more reliable and with advances in technology this has been realised [30, 56].
DES is widely used in business, logistics and manufacturing. It is one of the more common modelling methods used in healthcare to model health services. A recent umbrella review of systematic reviews found 586 papers published for health care applications of DES compared with 103 for system dynamics, 47 for agent-based modelling and 1 for hybrid modelling [57]. DES has been used to model biologic processes [58], emergency department flows [40, 59], health system performance [60], and economic impact of changes to population health screening methods [61].

Agent-based modelling

Agent-based modelling (ABM) is the most recently developed of the three modelling methods. In ABM, system-level phenomena are observed through explicit modelling of an individual, their behaviours and their interactions with each other and with the environment [62]. The models can be used to uncover complex causal effects, identify underlying mechanisms behind complex systems, and make sense of large amounts of existing evidence and data [63]. The adoption of ABM by simulation practitioners increased from the early 2000’s, triggered by computer science led advances in modelling technology, rapid growth in computing power and memory and a motivation to gain deeper insights into systems that could not be well captured by other modelling methods [30].

In an ABM, actors in a system are represented as autonomous individuals. They are given a starting configuration and rules that govern their behaviour, including adaptation, and interaction with each other and with their environment through time [62, 64]. The ABM then simulates both individual trajectories and population-level patterns or outcomes, which are generated from the bottom up by the actions and interactions of the agents. This
modelling method provides mechanistic mapping from individual-level assumptions to evolving population-level dynamics [62]. Assumptions can be informed by data or theory, and outcomes at both the individual and population levels can be compared statistically. ABM allows enormous flexibility in assumptions, and agents can be modelled at any level (or multiple levels) of scale. [64]

Agents can represent virtually anything in agent-based modelling [30]. For example, agents can be people, vehicles, equipment, health services, projects, investments or products. Agents may or may not interact in a social context and they may be active or passive within the model. Agents may be positioned in a spatial context (or not), may or may not interact with each other and may be very many or very few [30, 63]. The characteristics of agents may be defined in the model or they may emerge from the model dynamics and stochastics [62-64].

By modelling populations of individuals, ABM can also capture the interaction of actors with each other and with their evolving environments. This type of interaction and feedback between individual and social levels of scale is important for the study of such phenomena as interacting social influence and social selection processes, strategic social marketing, and the bidirectional influence of social norms and individual behaviour [30, 62, 64].

The level of abstraction in an ABM is also flexible and is determined by the level of abstraction of the “agents”. If agents are individuals, then the ABM will be more detailed, however if the agents are developed at a high level of abstraction e.g. Projects, companies or concepts, then the model will also be at a high level of abstraction [30]. ABM allows for multiple levels of aggregation to coexist within a given model or to be modified easily if
necessary. ABM can be useful when the appropriate level of complexity or abstraction is not known ahead of time and requires exploration during model development [62].

An important benefit of ABM is that it can be used to explore emergent phenomena which results from the interactions of individual entities [62]. Stochasticity can be applied to the agents’ behaviour with sources of randomness incorporated where appropriate, as opposed to a noise term added more or less arbitrarily to an aggregate equation in other modelling methods [62].

ABM is of use when describing the system from the bottom up, that is, from the perspective of the individual agents’ activities. It is useful when the behaviour of individuals is complex and cannot be clearly defined in an aggregated form and when activities rather than processes are a natural way of describing the system [62]. Experts can easily ‘connect’ to the model facilitating the crucial process of validating and calibrating the model through expert judgement [62]. ABMs strive to represent detailed reality by including individual behaviour, social networks and interactions, geographies, environmental variations, and evolution. Thus, the computational model underlying a “realistic” ABM might contain thousands of rules and model parameters [65].

ABMs allow us to generate hypotheses that articulate complex causal pathways which may include latent variables. Therefore, the ways variables are instantiated (represented) differ from those of other statistical approaches [65]. Good modelling practices suggest instantiating the model parameters with empirical data whenever possible. When such data are not available, models may be parameterised based on values derived from subject matter experts. As a last resort, the parameters for which there are no data or even expert opinion may be derived through calibration [65]. The relative importance of quantifying
unspecified parameters can be assessed using sensitivity analysis of simulation data, which demonstrates the magnitude of the impact the missing information is likely to have on the model outcomes of interest [62, 65]. Moreover, ABM and other systems science methodologies are capable of handling bidirectional relationships and feedback loops; non-linear, networked relationships; and heterogeneity, which are difficult to handle using statistical methods [65].

An early example of ABM was the Schelling model of segregation [66]. This was an abstract study of the interactive dynamics of discriminatory individual choices. Schelling demonstrated, using models of racial dynamics in neighbourhoods, that individual members of two recognisable groups distributing themselves in accordance with a preference that one’s neighbours be members of the same group as themselves or even a preference for a mixture, but “only to some limit”, led to complete segregation. Once the minority share (or number of people in the opposite group) in a neighbourhood reached the “tipping point”, then the existing residents moved away, and more minority group members moved into the neighbourhood. This model demonstrated that the behaviour at an individual level (micro-behaviour) was different to the population level (macro-behaviour). For example, even when individuals had only a 60% preference to have neighbours who were like them the resulting segregation at the population level was 100%.

Other early applications of ABM included the study of evacuation, traffic and customer flows; market behaviour including stock markets, internet service provider business models, and shopbots (online price comparison software); operational risk management and organisational design; and how individuals are influenced by their social context [62]. More recent examples of health applications in ABM include models of obesity [31],
tobacco use [64], diabetes [63], cardiovascular disease [61] and for the reduction of alcohol related harms [67, 68].

Multimethod modelling

Advances in simulation modelling software now allow the modelling methods described above to be combined into multi-scale hybrid models. This means that different components of the system can be modelled at the appropriate level of abstraction [30]. For example, in the model developed for diabetes in pregnancy in this thesis, agents have individual differences, including age, weight status and ethnicity. They can interact with health services represented by process (DES) components and system dynamics occur within an agent to represent the physiological dynamics underlying the development of impaired glucose regulation.

This flexibility is advantageous for several reasons: it allows each of the methods to be chosen and implemented when they are best suited in terms of level of abstraction [30]; it facilitates stakeholder understanding and learning [50, 69] and it can maximise computational efficiency [50].

The involvement of stakeholders in the model development process can facilitate the mobilisation of model-based learning into policy and program decision making. Policy modelling is likely to be of limited value if done without strong and iterative engagement with the users of the model outputs, i.e. decision makers [46, 69, 70]. Modellers must engage with users in a meaningful, ethically informed and iterative way. This is introduced briefly below and explored in detail in Chapters 3 to 6.
2.6 Participatory simulation modelling

The process of participatory simulation modelling involves engaging multidisciplinary stakeholders in a group model-building process and can be used in conjunction with multiple modelling methods including system dynamics, discrete event simulation and agent-based modelling [49, 52, 71, 72]. Various terms have been used to describe these activities including: participatory modelling, group model building, companion modelling (ComMod), and participatory simulation [72]. The tools and methods used in these different approaches may differ, however the underlying principles are in essence very similar, and subscribe to the same basic aim, to engage end-users and other stakeholders actively in model development to increase the robustness, validity, utility of and trust in the models and facilitate their use to support decision-making processes [50, 52, 69, 71-73]. The term participatory modelling has been adopted in this research. Participatory modelling has been an important method in system dynamics modelling almost since its inception [71] and has been widely adopted in environmental modelling projects [52, 73-78].

Osgood (2017) describes the history of stakeholder engagement in modelling projects as being divided into two eras, with a third era just starting to emerge [50]. In the first era “Bring us your problem, and we’ll tell you what decision is best”, modelling projects were mostly conducted inside academic or specialist organisations, with the primary outcome delivered to stakeholders being findings from model explorations conducted in isolation by such specialists. This was replaced by a second era in which models were often delivered to the stakeholder team for use, but generally as a “black box” [50]. Stakeholders were able to interact with the model (for example, through a web-based or desktop interface), but only
for pre-defined scenarios and with constrained outputs. Within this second era of modelling, the internals of the model, including the assumptions made, typically remained invisible to end users, and were unable to be modified by them [50]. Even in those cases in which the modeller was embedded within the end user team, requests to evolve the model, even for modest changes to the assumptions, or to add certain outputs or new types of scenarios, had to be referred to the modeller for action [50].

Advances in modelling technologies allow more transparency and a third era of participatory modelling is developing which is increasingly being undertaken within interdisciplinary teams [49, 50]. Although modelling experts are still required, modelling is no longer restricted to computer science experts, and models are being designed to be broadly accessible across team members [63]. Team members are able to proactively inspect and critique the assumptions of a model, locally modify those assumptions, run the model, and increasingly to supplement previously defined model outputs with those that they develop themselves [50]. Such broader access to models can support faster model evolution and learning, particularly in identifying discrepancies between model results and empirical observations or knowledge concerning the world, and in helping to refine mental models across the team [20, 23, 43, 52].

It is difficult to understand and forecast in advance the impact of policy decisions on system behaviour as a whole [23, 32]; however, an unambiguous model specified on a computer can play out the logical implications of the assumptions captured within that model [50]. From this perspective, the discovery of an inconsistency between what the model suggests in simulation results and empirical knowledge is not a failure of the model, but a success of the modelling process to facilitate learning, in that the process helps refine that
understanding, making it more robust [49, 50, 77]. Within this third era of modelling, embedded transparent models within teams helps harness the knowledge across the breadth of the team and can enhance their ability to identify areas where their knowledge falls short, and contribute to making it more robust [50].

Principles and functions of participatory modelling

Many frameworks, guidelines and principles for participatory modelling have been developed within the environmental sciences field where it has been widely acknowledged that sustainability issues involve social processes and stakeholder engagement is necessary to support effective action [69, 70, 72, 79, 80]. Frameworks and guidelines for participatory modelling have ranged from highly prescriptive scripts used for Group Model Building associated with system dynamics modelling [52, 71, 73, 81-83] to more general guidelines and considerations [20, 49, 69, 72, 84].

Participatory modelling projects are diverse and flexible principles guiding the conduct of participatory processes that are also easily modifiable and applicable across sectors have been proposed as a practical approach to inform existing and future practice [69, 72]. The following principles have been emphasised:

**Planning stakeholder engagement** - There should be careful consideration of the selection of stakeholders to include as participants, what their level of involvement and function will be in the participatory model development and at what phases of the project they will be involved [69, 84]. The specific skills, knowledge and domains of influence that each stakeholder or group of stakeholders bring to the process also need to be taken into
account, including the skills and knowledge of the modelling team, to ensure the right mix is available to guide model development [85].

**Being aware of social and group dynamics, special interests, power and hierarchies** - A participatory modelling process should always consider the reasons and intentions of stakeholders in becoming involved as well as the reasons and intentions of modellers (and other professionals) in proposing the involvement of stakeholders [49]. The social dynamics within the participant group need to be considered, including, for example, how powerful stakeholders might permit, facilitate, or encourage other actors to participate, or alternatively, how they might prevent them from participating [49].

**Flexibility** - Participants are involved in the process from the very beginning, having a say in the goals of the study, and also in the choice of methods, models and scenarios, and the scope of the study [49, 72]. In many cases the participation in the study becomes the most important and productive part of the project. Unexpected changes in goals and priorities (particularly those that arise from learning from the model) should be expected and accommodated within the process [49, 72]. Stakeholder motivation is important for the success of projects, and stakeholders may be demotivated if they are forced into a predefined protocol or procedure [72].

**Openness and transparency** – being open both scientifically and socially. Learning to work with stakeholders throughout the whole project, and providing tools that they require, they choose, and they are willing to use is necessary to encourage stakeholders to use the models to inform decision making [72]. It may not be possible to identify which tools will be needed at the start of a project, the decision-making process itself needs to be collaborative, and that in this process a range of models and modelling methods may be
needed [72]. Existing models need to be tested, documented, and archived in such a way that would make them available if stakeholders require them, and the models should be kept open so that they can be easily modified if such modifications are needed [72].

**Iterating and refining** - participatory modelling needs to be collaborative, iterative and agile [46, 49, 72]. This approach facilitates a sense of ownership of the model and encourages commitment from users about what they may come to see as ‘their’ model, rather than some black box that someone else is imposing on them [46]. Participant input also helps to prevent modellers making naive assumptions about the focus topic for the model, which is easy to do if one is not a domain expert [46].

**Accepting uncertainty and encouraging learning** - In participatory modelling, the model is always evolving, and uncertainty is an important consideration and discussion point [49, 72]. Through collaboration, the modellers are educated about the complexities of the system they are trying to represent, but equally, the users are educated about the capabilities, limitations and uncertainty in the model that they are helping to develop [46]. Active engagement of stakeholders can help parameterise and check the logical consistency of models, even where ‘hard’ data is sparse [46]. Lack of data should not be used to justify a decision not to model, but the approach needs moderation [46]. An iterative, participative approach to modelling allows data needs to be identified and ways of addressing these developed [46, 49, 72]. Rather than being viewed as “crystal balls” that are assessed as either being accurate and successful or flawed and a failure, models have significant potential to assist learning through the participatory process by bringing together best evidence, data and knowledge and consolidating and testing a shared hypothesis [20, 50].
A recent review of participatory modelling projects identified a number of functions enabled by the engagement of stakeholders in model development processes [69]. The most frequently reported function was gaining access to specific domain knowledge, followed by facilitating group processes and social learning, and yielding socially robust solutions (i.e., those that are accepted by decision makers and the general public) [69]. Joint problem framing so that real-world problems are addressed, developing scenarios and indicators to capture the relevant concepts from both science and practice perspectives, gaining access to and informing decision makers about state-of-the-art science; presenting results and facilitating use of the model results to inform decisions were also identified as important functions of participatory model development processes [69].

2.7 Important gaps in knowledge

The advances in technology described in the preceding sections are leading to increased adoption of tools and methods capable of integrating diverse evidence sources and exploring the dynamics of complex systems to inform policy decision making [83, 86]. However, most participatory modelling projects do not explicitly reflect on the participatory process component of the project [69]. Therefore, many of the challenges in aligning these technological advances with real-life policy making had not been examined in detail and were unresolved [69]. Questions also remained regarding how to facilitate participatory processes effectively and encourage the acceptance of participatory modelling processes [46, 72]. Exploration of different participatory modelling methods have been needed to produce more detailed understanding of what motivates stakeholder participation, in both the short and long term, with particular emphasis applied to understanding the value (or lack thereof) participants obtain from participation, and how
participatory model building and model-based reasoning can result in improvements to
decision making [87]. No standard template for participatory modelling processes has
emerged [69] and the literature is mostly theoretical [88]. It has been argued that lessons
to improve participatory modelling approaches will likely come from “craft knowledge”,
gained from experience [46].

Many participatory simulation modelling projects have been conducted in the
environmental science field. Therefore rigorous evaluation of the acceptability, perceived
value and utility of these methods and tools in the health sector has been required if their
adoption to support evidence-informed policy and planning is to be achieved [41, 89]. To
date, while DSM has been applied to health sector issues, the potential of participatory
DSM in the health sector has not been adequately explored. In particular, stakeholder
engagement and involvement of end-users in health-related simulation model
development has been lacking [86, 90] or, when engagement has occurred, the
participatory process has not been analysed and reported [91].

These gaps in knowledge and resulting research questions are described in more detail in
the research protocol presented in Chapter 3.
References


