COST-EFFECTIVENESS ANALYSIS USING AGENT-BASED MODELLING: A GENERAL FRAMEWORK WITH CASE STUDIES

AFFAN SHOUKAT

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN MATHEMATICS AND STATISTICS
YORK UNIVERSITY
TORONTO, ONTARIO

MAY 2019

Abstract

In recent years, agent-based modelling (ABM) has been increasingly used to elucidate complex adaptive systems. An ABM is a structural computational system that consists of a collection of abstract objects (agents) embedded in a virtual environment that interact based on a set of prescribed rules. While traditional approaches such as differential equation-based compartmental models span a vast literature, they often impose restrictive assumptions such as homogeneity and determinism that limit their application to real settings. ABM overcomes these limitations through a bottom-up approach in which macro dynamics emerge from micro level phenomena.

During the past decade, there has been a surge of interest in the use of ABM in human health and disease dynamics. While this is rapidly growing, its application to other relevant areas such as health economics is still in infancy, and frameworks that could systematically apply ABM are still lacking. In this thesis, we develop a general framework for cost-effectiveness analysis in which ABM is designed to project the system dynamics. We argue that ABM improves the empirical reliability of policy-oriented simulation models and that it presents an ideal tool to address the complexity of disease processes, project the impact of interventions and inform their optimal implementation. We use this framework in an epidemiological context to quantify the economic impact of vaccination strategies for prevention of infectious diseases.

We present two case studies for a human-to-human infection transmission (i.e., *Haemophilus influenzae*) and a vector-borne disease (i.e., Zika). In each case, we detail the construction of ABM and its utilization to conduct Bayesian cost-effectiveness analysis of potential vaccine candidates. In addition to uncovering important characteristics of these diseases in epidemic dynamics, we present their first cost-effectiveness analysis and implications for vaccination strategies in different populations settings.

To my parents

Acknowledgements

I would like to express my heartfelt thanks to my supervisor Professor Seyed Moghadas for his constant support, guidance, and positivity. His mentorship and expertise have made graduate school an enjoyable and enriching experience of my academic life. I consider myself extremely fortunate to have had him as a graduate supervisor and look forward to many exciting scientific collaborations in the future.

I also would like to express my gratitude to my supervisory and examination committee for providing guidance, suggestions, and insights throughout my PhD studies. This gratitude is extended to the faculty and staff in my home Department of Mathematics and Statistics.

Most importanly, I would like to thank my family for their guidance, encouragement, and moral support.

Table of Contents

List of Tables

List of Figures

Glossary

***in silico* Modelling** Modelling of biological processes performed via computer simulations.

Agent-Based Modelling A type of computational model for simulating the actions and interactions of autonomous agents.

Asymptomatic (or Carriage) A disease manifestation when the infected host does not show any clinical symptoms, but is infectious and can transmit the disease to others.

Attack Rate The proportion of population infected over the course of an epidemic.

Basic reproduction number Commonly denoted by $\mathcal{R}_0$, it is defined as the average number of secondary cases which a single infectious case produces in a completely susceptible population.

Bootstrap A technique used for statistical inference by repeatedly sampling a dataset with replacement.

Effective Reproductive Number Denoted by $\mathcal{R}_{\text{eff}}$, it estimates the average number of secondary cases caused by a single infectious case in a population when control measures are applied.

Generation Interval Refers to the time duration between the infection of an infected person and the infection of his or her infector.

Herd Immunity The proportion of the population that is immune against the disease.

Latent Period Defined as the time duration between disease transmission and the onset of infectiousness (i.e., when the infected host becomes infectious and can trasmit the disease). This term is sometimes interchanged with *exposed.*.

Survival Rate The percentage of people in a study or treatment group who are still alive for a given period of time after diagnosis or treatment for a disease.

Symptomatic A disease manifestation when the infected host shows clinical symptoms.

Chapter 1

Introduction

Innovations such as the exponential increase in computing power, and the evolution of information and database technologies during the preceding two decades have opened up novel vistas for the collection of complex, voluminous, and heterogeneous data that can be synthesized to address a variety of real-life challenges, such as population health. Methodologically, since mathematical and statistical models that are capable of encapsulating such data are often theoretically intractable, *computational models* have become an integral part of scientific research, and have led to dramatic changes in approaches to addressing societal issues. Such models have already had significant impacts on public policy at the global scale, and in particular for the development of intervention strategies to combat emerging disease outbreaks. However, the systematic application of computational models in understanding the underlying processes of system dynamics (e.g., network of human interactions and disease transmission mechanisms) is relatively new; the integration of these systems into the public policy and decision-making processes is even more novel. In this thesis, we aim to enhance this integration and develop a general framework for evidence-based health economic analysis, by employing a computational modelling approach that has evolved to deal with data that are more heterogeneous, less coarse (based at a community or individual level), and more complex (joint spatial, temporal and behavioural interactions). This evolution is typified by the "*Agent-Based Modelling*" (ABM) paradigm, in which the collection of autonomous decision-making entities (i.e., agents) and their interactions unveil the dynamics and emergent properties of the entire system. In the context of population health and infectious disease dynamics, the flexibility of ABM permits an effective representation of individual interaction with their own characteristics, which may impact their future decisions and outcomes. The framework developed here provides a methodology to perform cost-effectiveness analysis of public health interventions more systematically, while accounting for critical properties that are often overlooked in existing methods, yet essential for outcomes prediction and translating knowledge and evidence to action.

Traditionally, aggregate or *cohort* models such as continuous dynamical systems, decision trees, and Markov models have been used to provide the integration of disease transmission dynam-

ics and cost-effectiveness analysis. While such models are easy to implement and have been generally successful, their use is often limited due to a number of limitations such as linearity and homogeneity. To overcome these limitations, we apply ABM computational systems that can capture the system heterogeneity from micro to macro levels, while utilizing the available data and information that drive the model outcomes. In principle, ABM can theoretically encapsulate an arbitrary level of heterogeneities that can be observed experimentally, provided we understand the underlying mechanisms of the individual components of phenomena. However, in practice this flexibility comes with a complexity and thus demands stringent simulation design requirements than conventional models, especially with respect to outcome reproducibility. We address these requirements in developing our framework to ensure its sound application to cost-effectiveness analysis that remains an important component of decision-making process in public health planning and program delivery.

In this framework, we show that ABM presents an ideal tool to address the complexity of disease processes, project the impact of interventions, and inform their optimal implementation. We use this framework in an epidemiological context to quantify the economic impact for vaccination strategies for prevention of infectious diseases. We begin by providing a succinct and rigorous description of an ABM computational system in Chapter 2, including construction details and a formal mathematical structure. This is followed by an overview of cost-effectiveness analysis in Chapter 3, where we also highlight the potential flaws of current methodologies. We then apply the framework to present two case studies for a human-to-human infection transmission model and a vector-borne disease model in which ABM is used to generate system dynamics, considering potential vaccine candidates. In Chapter 4, we study the dynamics of severe community-acquired acute infections caused by *Haemophilus influenzae* serotype 'a' (Hia), with alarming incidence rates in North America, particularly among Indigenous populations. The severity and outcomes of Hia infections are reminiscent to those of invasive Hib disease in the pre-vaccine era, including pneumonia, septicaemia, and meningitis. The remarkable success of Hib conjugate vaccine suggests that the development of an Hia vaccine could be a viable prevention measure to reduce the incidence of invasive Hia disease as well as prevent the spread of disease in the general population. Recent research efforts have established the pre-clinical proof of concept for a glycoconjugate vaccine against Hia. However, quantifications of the long-term epidemiologic and economic impacts of vaccination are needed to inform decision on investment in Hia vaccine development and immunization programs.

In Chapter 5 and Chapter 6, we utilize a multi-agent ABM to uncover important characteristics

of Zika virus epidemics and transmission dynamics. Given the public health concern about its potential to cause severe outcomes and long-term sequelae, including microcephaly with brain abnormalities and neurological disorders in infants, and Guillain–Barré syndrome (GBS) in adults, the development of a preventive measure is the key to combat this vector-borne disease. Currently, a number of Zika vaccine platforms are being investigated, some of which have entered clinical trials. Understanding the impact of this vaccine and its cost-effectiveness, especially in the presence of Zika asymptomatic infection, can help inform vaccination strategies and prioritization in countries where the primary vector carrier, *Aedes aegypti*, is endemic.

The outcomes of ABM in both case studies are integrated with a Bayesian cost-effectiveness analysis to investigate the benefits of vaccination and their economic impact. We close this thesis in Chapter 7 with a discussion on the implication of our results. We believe that this research will have a significant impact on the relevant healthcare systems (including Canadian) by generating scientific, evidence-based results that inform the optimal use of health resources and improve population health, and therefore contribute to reducing health and economic burdens of preventable diseases. Although our ABM framework is useful for the exploration of uncertainties, heterogeneities, and their impact on public health decisions, we omit a number of important factors in vaccine cost-effectiveness analysis, including research and development costs of vaccine, potential adverse effects of vaccine that may incur additional costs for patient management, and emerging technologies for vaccine development. These factors may alter parameters and assumptions underlying health economics of vaccination, and should be evaluated before policy implementation. We are committed to continuously refine our models to consider these factors.

Chapter 2

Agent Based Modelling

2.1 Introduction to Agent-Based Modelling

The seminal works of the American economist Thomas Schelling (1971) [1] showed that computational and simulation approaches can be applied to understanding the universal principles of any *complex adaptive systems*[1] provided the system can be reconstructed *in silico* environment by programming the constituent components of the complex system. In Schelling's approach[2], he directly manipulated abstract computer entities representing actors and updated them iteratively. This led to a whole new field of research on socio-economic systems in which the natural unit of decomposition is the *individual* rather than the observable or equations, often termed individual-based modelling. Together with the rapid development of computational theory and the collection of vast amounts of data, this field has led to the evolution of *Agent-Based Modelling* (ABM) computational systems.

Today, methodologies studying complex adaptive systems in a qualitative sense have shifted to systematically investigate them by ABM through a disaggregation of the systems into their individual components that have their own characteristics and behaviours and by capturing the interdependencies between the individual components[3] [13, 14, 15]. An ABM computational system is a structural, dynamical system that consists of a collection of abstract objects (i.e.,

[1] While there is no single definition of a complex adaptive system, it is generally accepted that a complex adaptive system is one that has many individual parts working together in order to generate the macro dynamics of the system. Complex adaptive systems are common in both nature and society. For example, the immune system is a highly advanced biological system comprised of a complex network of individual cells and chemicals working together to produce non-linear effect, feedback loops, and other micro and macro-dynamics.

[2] In Schelling's demonstration that residential segregation can occur at a systemic level, the economist implicitly introduce the idea of a local *neighbourhood* which is a central element in the construction of agent-based models, seen later in this chapter.

[3] ABM computational systems have been applied to a range of disciplines including economics [2, 3, 4], ecology [5], healthcare [6], sociology [7], geography [8], finance [9], and even niche disciplines such as military strategies [10, 11, 12].

agents) embedded in an *in-silico* (i.e., computer) environment and interact together through a set of prescribed rules. This type of model is often implemented computationally by using deterministic input-output functions, typically coded in a structured or object-oriented programming language. The *agents* in ABM represents the individual components of the complex system under study. Each agent individually *perceives* its situation, makes *decisions*, and performs *actions* according to specific rules. These rules can be simple or complex, deterministic or stochastic, and fixed or adaptive. Thus an ABM encodes in a computer program a set of rules that describe the behaviour of agents as the system evolves in time. While often the formal set of rules are simple, large scale agent-based models can incorporate neural networks, genetic algorithms, and other machine learning techniques for realistic agent behaviour and adaptation [16]. Since each agent is modelled individually, there is no central controlling agency nor explicit language that describes the global dynamics of the system. As a consequence, ABM allows for an investigation into the universal properties of a complex system, including: *heterogeneity* since agents can be modelled individually, *adaptation* since the model is dynamic, *space and scale* since an arbitrary number of agents can be embedded in this virtual environment, and *non-linearity* since the model can track individual agents separately. This methodology has three main advantages: (i) it allows to capture global complex patterns and dynamics as a result of interactions between local, heterogeneous individual agents; (ii) it allows the construction of models in the absence of knowledge about global interdependencies of the complex system; and (iii) it provides the flexibility required to study the system's complexity in comparison to traditional equation-based aggregate-level mathematical modelling. These three points are discussed below.

Emergent global dynamics in ABM

Traditional models of complex systems are typically formulated in the general language of mathematics. Examples include dynamical systems such as the Lotka-Volterra equations, describing predator-prey interactions [17] and the *Susceptible-Exposed-Infected-Recovered* models of epidemic propagation [18]. While the utility of such models have been exemplified in a vast literature, a significant limitation of these models is the treatment of all or groups of individual components as largely homogeneous entities, i.e., a *representative entities*. For example, in differential equation models of disease transmission, all individuals in a population are equipped with the same characteristics and parameters. On the other hand, an ABM enables the generation of complex patterns and dynamics "in a bottom-up approach" in which a single unified homogeneous model is replaced with a population of individual models, each of which is an autonomous decision-maker (i.e., the

agent). This population heterogeneity includes not just the variation in individual agents, but also the interactions and network topology. Running such a model simply amounts to instantiating an agent population with initial conditions and iteratively letting the agents interact by executing the rules that define them. Of course, if the model is stochastic, then multiple realizations are necessary in order to capture the randomness. That is all that is necessary in order to *solve* an agent-based model. As a result, although the global (or macro-level) dynamics of a process are not explicitly programmed in the model, enigmatic global dynamics including fixed points, cycles, dynamic patterns, and long transients emerge from the local interactions among agents. These emergent dynamics can have properties that are decoupled from those of the individual components. The emphasis on modelling the heterogeneity of agents and the emergence of global behaviour from local interactions is an important distinguishing feature of ABM.

Absence of global interdependencies

The essential characterization of any complex system is that they are invariably qualitative in nature. That is, in many cases we do not know the full mathematical description of the complex system under study, but only the behaviour of the constituent components of the system. Agent-based models are particularly suited to complex systems in which the dynamics of its constituent components are more understood than the overall dynamics of the system. They are ideal for modelling systems in which individual component behaviour is non-linear and heterogeneous, which differential equations are unequipped to handle[4], and where individual behaviour includes learning and adaptation, including temporal and spatial correlations, and non-Markovian behaviour. In this sense, contrary to what the term *agent* suggests, the concept of an agent enables us to represent any physical object that can be programmed, provided we have a clear understanding of the object. Agents can represent particles, cells, individuals, groups, organizations, or spatial entities such as buildings and roads. Consequently ABM can, in principle, incorporate any complex behaviour that can be observed experimentally, provided we have a qualitative understanding of the underlying mechanisms and components. Of course, in practice agent-based models are limited by finite computational resources, time investment, and incomplete knowledge and therefore will inevitably require a balance between the desired complexity of a system and available resources.

[4]It is true that differential equations can be used to model arbitrary levels of heterogeneity, but the complexity of the equations and their subsequent analysis increases exponentially as the complexity of the individual behaviour increases. At some point differential equation models becomes intractable.

Flexibility of ABM in comparison to equation-based models

A common objection to agent-based models is that they are not (theoretically) as rigorous as mathematical and analytical models. That is, they do not offer a set of equations together with an algebraic solution that can be easily interpreted and analysed. Analytical or equation-based models, for example differential-equation based dynamical systems, provide a formal framework for the organization and analysis of knowledge and theoretical results. Such models are equipped with an established set of rigorous tools for analysis, for example bifurcation, sensitivity, and stability analyses. This allows an analytical model to be easily communicated because they are fully described and unambiguous by mathematical equations and formulas. On the other hand, ABM often will not make use of any explicit mathematical equations but exploits computational simulations, implemented in a programming language, to elucidate the complex system underpinning the model. However, the idea that the lack of formal mathematical tools prevents any sort of formal analysis of agent-based models is misguided. Indeed computational ABM, by virtue of being computer programs, can and do utilize a well-defined set of functions which relate inputs to outputs, in either a deterministic or stochastic fashion, and unambiguously define the global dynamics and any eventual equilibria of the system [19, 20, 21, 22, 23]. For example, in Laubenbacher et al. [22], the proposed mathematical representation of an agent-based model is a *time-discrete dynamical system* over a finite state set. In this framework, the state of the model is fully specified by a vector taking values over a finite field $\mathbb{K}$. A transition function transforms a given state into another state based on rules of the complex system underpinning the model. The model dynamics are generated by repeated iteration of this function. Another approach is taken by Veliz-Cuba, Jarrah, and Laubenbacher [23] where they derive a *polynomial dynamical system* where the input-output functions are defined by polynomials, which makes it amenable to powerful symbolic computational capabilities. That is, the computation of equilibria and analysis of the model reduces to symbolically solving a system of polynomial equations. Agent-based models are also, from a formal point of view, *Markov chains* [24, 25, 26, 27, 28], Banisch, Lima, and Araújo [25] and Gintis [27] and provide a rigorous mathematical basis of ABM by linking the micro-description of the system to the complex global behaviours in the form of a Markov chain. They establish how the corresponding global dynamics of an agent-based model are obtained by a projection construction and the model's long-term properties are given by the ergodic theorem for Markov processes. Analysing an agent-based model as a Markov chain can make apparent transient dynamics, asymptotic behaviour, and stochasticity that were otherwise not evident. More generally, ABMs can be naturally classified as hidden or latent variable models that relates a

set of observable variables to a set of latent variables [28]. In addition to formal mathematical frameworks, Grimm et al. [29] establishes a protocol for an ABM specification. They develop the so called *ODD* (Overview, Design concepts, Details) protocol which describes a standard template for the model analysis, reproducibility, and transition functions [29].

Another common objection to ABM is that a single realization of the model is just a special case, and that no formal statements on the results of the model can be stated. This is partially addressed by the works of Newell and Simon who established the *computer programs as sufficiency theorems* approach. When an ABM, call it A, produces result R, it establishes a sufficiency theorem which is the formal statement R if A [19, 30]. In other words, each run of an ABM is a logical theorem that reads: *the output of an agent-based model follows with logical necessity from applying to the input the formal rule-set that defines the model.* Nevertheless, despite the fact that each run of such a model yields a sufficiency theorem, a single run does not establish the robustness of such theorems. That is, when A yields the result R, how much change in A is necessary in order for R to no longer be derived. This problem is, however, easily treated by multiple realizations of the agent-based model, often by Monte Carlo techniques with each realization systematically varying initial conditions, parameters, and random number streams.

Despite these objections, there are several advantages of ABM approach over conventional equation-based mathematical models. Axtell [30] describes three distinct uses of ABM: (1) numerical computation of analytical models, (2) validation and robustness of analytical models, and (3) a substitute for analytical models that are intractable. Often complex equation-based models cannot be *fully solved* in order to gain insight into the system. When the solutions are not available symbolically, a set of solutions can be obtained numerically by solving the equations. ABMs can then provide a suitable validation of the numerical solutions. If the equation-based model is stochastic, the primary method of obtaining numerical solutions consists of Monte Carlo simulations. When the symbolic solution is explicitly specified, it would seem that there is no specific role for agent-based modelling. However, since the output of ABM tend to be more visual and pattern-oriented, such models can still be very effective in elucidating complex systems to individuals that have no formal mathematical knowledge.

2.2 Structure and Components

A typical agent-based model has three fundamental components:

1. A set of agents which represent the components of the system that is being modelled. Since every ABM is a computer program, agents are often implemented as virtual computational *objects*[5]. The concept of an agent as being a computational object makes it clear that this basic unit is abstract. The agent is made concrete by translating sufficient properties of the real-world component into a suitable formalization in the programming language.

2. An *in silico* environment that allows agents to change their spatial and relational associations. In most complex systems, the concept of physical space or spatial network is significant to the global dynamics. This concept is difficult to model in traditional analytic models, except in highly stylized ways. However, in ABM it is rather easy to have the agent interactions mediated by a virtual space.

3. A set of rules which define the level of connectedness and modes of interaction between agents. Each agent can be assigned a unique set of rules or a single one that can be applied to all (or a group of) agents.

To integrate these three components, an ABM requires a computational framework or a *simulator engine* (see §2.3). This simulator framework is responsible for driving the ABM by repeatedly (i.e., by iteration) executing the rules that define the agents' behaviours and interactions. This iterative process often operates over a time-step or discrete event simulation structures. In the course of these iterations, the simulator framework also calculates the aggregate results of the model which can be re-injected back into the evolving behaviour of the agents. Thus, the construction and simulation of an ABM as a whole is through a bottom-up approach from these constituent components that work together to generate the global dynamics. The structure of a typical agent-based model is shown in Figure 2.1. A mathematical description of ABM as a fully recursive system is presented in §2.4, expanding on the formalism described in [21].

Agent Structure and Properties

Since an ABM is an abstraction of a real-world phenomenon, an agent ideally represents a component of the complex system that is being modelled. When using this approach, we need to systematically recognize which components of the system can be translated into agents. Next, we need to decide on the level of abstraction and the details that are to be included for each agent.

[5] In computer science, an *object* is a data structure consisting of variables, functions, or methods, and is a value in memory which is referenced by an identifier. There is a large degree of similarity between a computational agent and the concept of an *object* (or structure) in a programming language. In fact, an ABM can be seen as a set of object classes that share the same properties and the same rules i.e., *functions*.

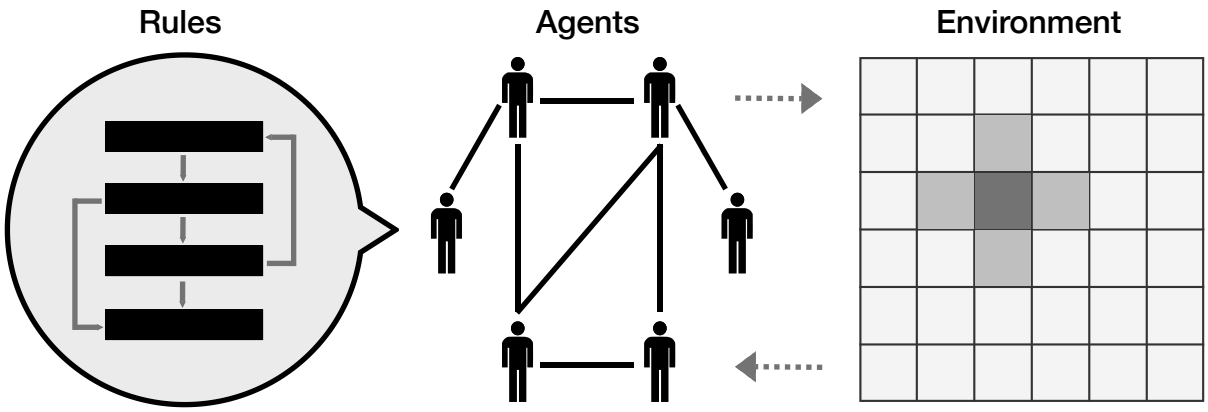


Figure 2.1: The general structure of an ABM computational system. Figure shows agents equipped with a set of properties, connected through a lattice environment system. Figure shows the interaction between a susceptible and infected individual as a result of movement across the lattice

An acceptable compromise between realism and simplicity is required. If the level of abstraction is too low, the model may fail to faithfully capture the system dynamics. On the other hand, it is neither feasible nor always desirable to model all the complexity and agents' heterogeneity. A one-to-one mapping of the real-world component to an agent is likely unnecessary, impossible, and/or computationally overwhelming. We will need to have valid hypothesis on the underlying processes or fundamental mechanisms that needs to be explained. Ideally, the level of abstraction is justified by utilizing empirical data as well as expert opinions. Doran [31] suggests two principles that should be followed for implementing a set of rules: (i) agents should be as abstract as possible subject to the requirement that any rule attributed to them must either be reliably set from empirical observation or be subject to experimental observation, and (ii) assumptions based on pre-conceptions are avoided. Several other authors [15, 19, 29, 32, 33, 34, 35] have also identified similar key characteristics:

- **Self-contained and autonomous**. Autonomy implies that there is no central authority that controls the agents' behaviour. On the contrary, an agent can be thought of a model in itself, capable of processing information and making decisions. They are free to interact with other agents and move in the virtual environment they reside in.
- **Heterogeneity**. Each agent represents a unique component of the complex system. For example, an agent representing a human in some population model can have age, sex, and location as dynamic attributes.
- **Active**. Agents can be: *goal-oriented* [34] where agents try to achieve a goal but not

necessarily maximize utility, *perception-decision-action* where they interact with each other and their environment and only perform actions when triggered to do so by some external stimulus [34, 36], or bounded rational [37] in which agents are generally assumed to be rational optimisers with complete access to information and bounded analytical ability (through heterogeneity).

- **Internal state**. Agents have an internal *state* (i.e., data represented by variables) and can communicate this state to the model by message-passing or signal protocols.

The next step is to implement a formal set of functions (the *rule-set*) which define the *perception-decision-action* cycle (PDA) for every agent [34, 36]. An agent-based model systematically and iteratively gives each agent the change to perform a PDA cycle. Formally speaking, let Λ denote an ABM and let x_a represent the represent the *internal state* of some agent a. x_a is essentially a list of quantitative variables and internal parameters associated with the agent's current situation, i.e. a vector taking values over some finite field $\mathbb{K}$ (often $\mathbb{K} = \mathbb{R}^n$) that describes the state of the agent at a given time. Let $\mathbf{X} = \{x_a\}_{a \in \Lambda}$ denote the collection of all internal states, so that $\mathbf{x} \in \mathbf{X}$ denotes the global state of the model. Then the PDA cycle of an agent a is a set of functions[6]:

- Define $\textsc{Perception}()_a : \mathbf{X} \longrightarrow \mathcal{P}$ which computes a *percept* $p \in \mathcal{P}$ using the global state of the model $\mathbf{x}$. Intuitively, the function processes the global state of the model including the environment, and returns data such as the coordinates of nearby agents or objects or possible locations for a roaming agent. Thus, elements of $\mathcal{P}$ are often an multi-dimensional vectors (or tuples) consisting of quantitative and categorical information.

- Define $\textsc{Decision}()_a : \mathcal{P} \to \mathcal{D}$ which is a core function executing the rules that defines the agent's behaviour given their $\textsc{perception}$. Decision functions are can be arbitrarily complex, encapsulating the bulk of the model's logic, ranging from simplistic *fuzzy rules* to complex behaviours modelled by neural networks, logic systems, artificial intelligence, or a hybrid/multi-layer system and are often imputed by using empirical data and expert opinion. Elements of $\mathcal{D}$ be in the form of *discrete messages*[7] or *signals* being passed between agents. In response to a message an agent may change their internal state, modify the environment, or respond back with another message, but should not be able to modify the internal state of other agents.

[6] Notation and names of the functions were adapted from Drogoul, Michel, and Ferber [36].

[7] The technical details of implementing message-passing in computational abstract objects is a technical study of computer-science. It is beyond the scope of this thesis and details are omitted.

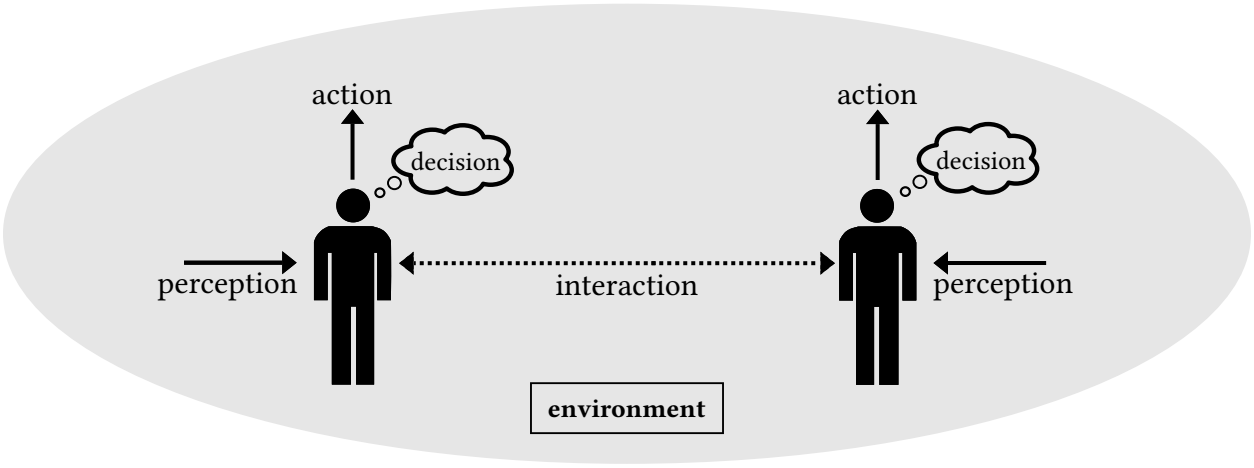


Figure 2.2: A visual representation of a *Perception-Action-Decision* cycle which formally implements the rule-set which drives an agent's behaviors and interactions.

- Finally, the action function $\textsc{Action}()_a : \mathcal{D} \times \mathbf{X} \to \mathbf{X}$ computes the new internal state $\mathbf{x}$ of the agent by acting on a decision $d \in \mathcal{D}$, updates the global state of the model, and advances the virtual time by the relevant unit.

The Virtual Environment

The second fundamental component of an agent-based model is the *in-silico* environment, which models the physical space of the real-world system. The environment defines the spatial associations of an agent and the conditions for the PDA cycle to carry out, i.e. the state of the environment is part of the input to the $\textsc{Perception}$ function of every agent. The environment may also include passive objects such as roads and buildings or resources such as wealth and healthcare. As agents move in the environment, their location can be tracked by a dynamic variable. Agents may also be spatially implicit, that is their location within the environment is irrelevant.

Modelling the environment is often via a **discrete topology** of discretized, connected, and bounded space units. The topology of the environment defines possible interactions and relationships between agents. There are two main types of a spatial environments:

1. The most common type is a discretized environment consisting of a finite grid of cells in one or more dimensions with integer coordinates, which provides a simple representation of physical space, e.g. GIS based environments. This type of environment provides an easy mechanism for agents to interact with others who share similar coordinates or reside in

neighbouring cells. The simplest form of spatial environments are described by cellular automata models where the environment consists of a square lattice, divided uniformly into 'cells'. However, in cellular automata, the cells are interpreted as agents and there is no distinction between the agents and the cells that create the lattice environment. ABMs extend this topology by decoupling the agents from their cells. Agents can then move from one cell to another and interact with different parts of the grid, resulting in a set of neighbours that constantly change as the simulation proceeds.

2. The second type is a 'relational environment' in which a link between two agents defines a network topology. Relational environments are often defined by graph- or node-based constructions and there is no distinction between agents and the nodes of the network. Both directed and undirected graphs can be supported and can be static or dynamic [15]. In static networks, links are fixed and do not change. In dynamic networks, links and nodes are determined endogenously according to the mechanisms programmed in the model.

Figure 2.3 shows examples of different environment topologies. Agents typically interact with a subset of all other agents, denoted as the agent's *neighbourhood*. In complex and large-scale agent-based models, spatial and relational environments can be intertwined to offer a more granular approach. The environment may further have its own set of properties relevant to the real-world complex systems; for example, in modelling disease-transmission at the age-group level, it might be relevant to identify sections of the environment as *school, residential* and *business*. It may respond to messages from agents (*deterministic environment*), or changes with time (*dynamic environment*), and can even generate or delete agents.

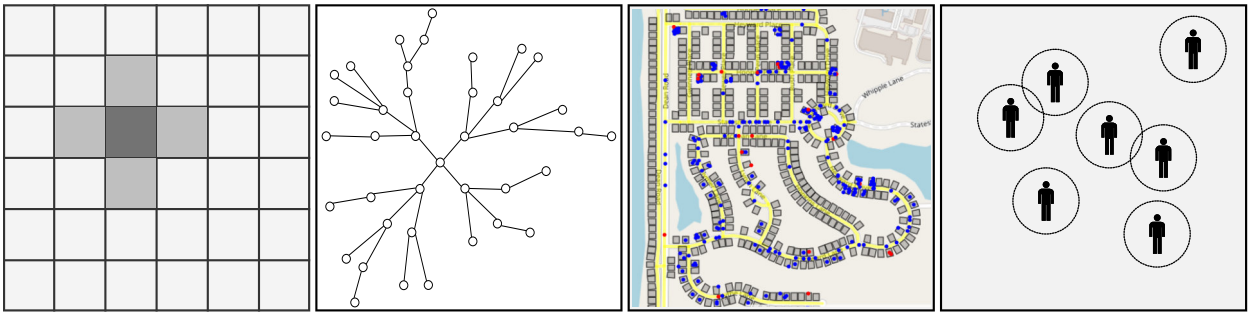


Figure 2.3: Different environment topologies for agent interactions and movement. From left to right: Cellular Automata (von Neumann neighbourhood), network relational topology, Geographic Information System (GIS), and free-roaming with neighbourhoods.

Rules of Interactions Between Agents

Recall that the DECISION and ACTION functions of every agent implement the rules that will affect their internal state and interactions with other agents in a dynamically changing environment. The rules are typically derived from published literature, expert opinion, or empirical data. For example, in modelling influenza dynamics via an ABM, the intrinsic incubation period of the disease is between three to five days, derived from observational data. At the point of infection, the latency period for each agent (if transmission occurred) is sampled from the relevant distribution. In general, rules are based upon simple *if-else* statements with agents carrying out an action once the specified condition has been met. More recently, there has been work incorporating machine-learning techniques and artificial intelligence within ABMs to better represent human behaviour. Due to these adaptive rules, ABMs are non-linear with multiple feedback-loop dynamics and can therefore have vastly different behaviours for each run.

2.3 Model Implementation

Every ABM implements a *simulator engine* that specifies the operating procedures of the agents, drives the PDA cycle, and describes the model evolution in time from an initial condition. The simulator is often implemented in a programming language and sound principles of computer science should be applied whenever possible. For example, writing independent *modules* (which represent distinct components of the complex system) makes maintenance, debugging, and code re-use easier when considering future changes to the model. Good programming principles also leads to reproducibility of the model across various platforms and operating systems.

2.3.1 Monte Carlo Simulations

ABMs are stochastic in nature and account for randomness found in real-life phenomena. This stochasticity is known as *first-order uncertainty* and relates to the natural randomness in agent behaviours, interactions, and the progression of the model. First-order uncertainty is often addressed by the use of a random number generator through a deterministic program code; however, in digital computers random numbers are not *really random*. They are generated by an algorithm that produces a sequence of numbers that is seemingly random. These numbers are referred to as *pseudo-random* and the algorithm is called a *pseudo-random number generator* (PRNG). Each sequence produced by any PRNG is uniquely identified by its *seed* s, a number which provides the initial value to the generator. The seed is usually supplied by an environmental

variable such as the computer clock, which can virtually guarantee that it is different for every simulation run. In other words, PRNGs produce numbers based on a deterministic formula which is seeded with some initial number. This allows for computer, and particularly agent-based models to simulate stochastic variables but also offer reproducible results. Several probability distributions have been established in computer-science literature[8], although most random number generators produce uniform randomness. First-order uncertainty can be reduced by running the model several times, commonly known as *Monte Carlo* simulations. The sufficient number of Monte Carlo simulations depends on the parameters, time horizon, and the structure of the model, and common ABMs can employ several hundred to several thousand independent runs.

The behavior of any ABM is further sensitive to the model parameters and on the initial conditions, often called *second-order uncertainty*. That is, while first-order uncertainty relates to stochasticity in model structure, second-order uncertainty corresponds to estimation of parameters since true values and distributions are often unknown. In order to address second-order uncertainty, underlying distributions of parameters should be utilized whenever available. Typically, however, second-order uncertainty is evaluated by formal sensitivity analysis methods such as *Probabilistic Sensitivity Analysis* using Latin Hypercube Sampling techniques.

2.3.2 Temporal Evolution

In most ABMs, modelling the temporal evolution of the environment is crucial. The model components (e.g. environment) may not only react to the agent inputs but also evolve according to endogenous factors including time. Time evolution can be modelled using three main approaches: (1) *continuous time* in which the model can compute the system state for any time input, (2) *discrete time* in which time evolves in discrete, but fixed intervals, and (3) *discrete event* in which time instantaneously jumps from one event to the next. Within each time step, specific events occur (e.g., movement of agents across the environment or transmission and progression of disease) or agents update their internal state based on interactions and endogenous factors. The decision to use continuous time or discrete time depends on the application domain. Applications of ABM in this thesis are based on the discrete time approach.

Discrete Time In discrete-time mode, the simulator advances the virtual clock by a given interval Δt. The time interval Δt will typically have a natural real-world unit associated with it,

[8] In virtually all agent-based models, the *Mersenne Twister* is an example of a pseudo-random number generator that produces a uniformly distributed random number stream, at least until the number of random draws approaches the algorithm's period of $2^{19937} - 1$.

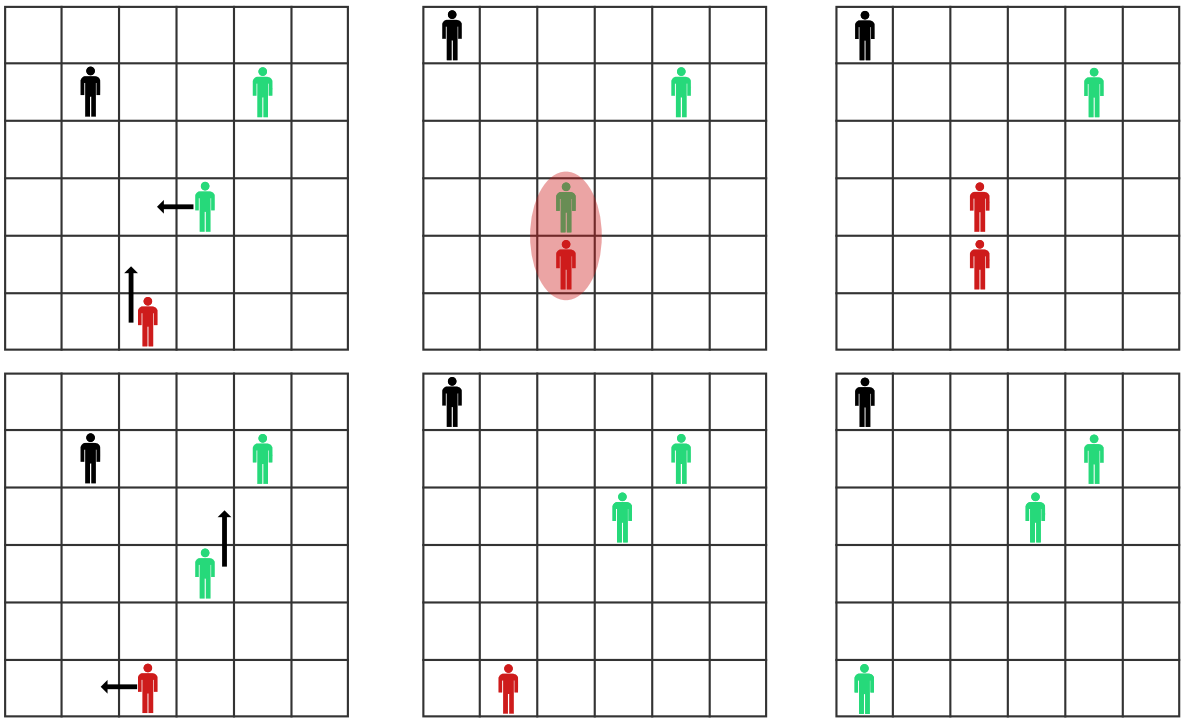


Figure 2.4: In an ABM simulation agents are updated sequentially at any time. The order in which each agent is updated may change the results obtained for the system state since the environment evolves for each action.

such as seconds, hours, days, or years depending on the complex system underpinning the model. At each Δt, the simulator engine picks agents sequentially (in some given order) and executes their PDA cycle iteratively. However, there is a significant limitation of this technique. The order in which the agents are selected (and their PDAs executed) may change the result obtained for the system, since the system state evolves with each action. It is entirely possible that a different ordering may result in very different global dynamics. Figure 2.4 illustrates this issue with three agents. Possible solutions to this limitation include selecting shuffling the order of agents at every time interval or have all agents simulate concurrently by operating on temporary variables, so that the perceived state is the same for all agents. For example, in an agent-based model that tracks the HIV patterns over long-time periods, the natural time unit is a year. A model for seasonal influenza epidemics, the natural time unit has a day resolution.

Discrete Event A discrete-event simulation considers time as discrete increments with variable magnitudes corresponding to events occurring in the model [38]. Every jump in time marks a change of state in the system. Between consecutive events, the system is idle and no change in the system is assumed to occur. A discrete-event model is programmed to maintain a list of events

(a 'queue') corresponding to the events that will occur. The program iterates over the queue, executing the event at the head of the queue. Once the event executed, it is removed from the queue, and the next event is scheduled appropriately (and sometimes dynamically). In comparison to the discrete-time approach, this is a more flexible paradigm. For instance, because discrete-event simulations do not have to simulate every time-step, they are typically run faster than the corresponding discrete-time simulations[9]. Moreover, a discrete-event can be made equivalent to a discrete-time approach via a linearly spaced sequence of events.

2.3.3 Validation and Calibration

Model validation concerns with identifying the degree of consistency between the agent-based model and the underlying system it represents. It comprises of two stages: (i) *Input validation* that refers to the realism of the assumptions used to build the model, and (ii) *output validation* which measures the plausibility of the model outcomes relative to the observations of a real-world phenomenon. Output-validation relies on a process called *calibration*, which systematically refines the model parameters so that the output data closely resemble those observed in the phenomenon.

A proposed formal framework by Marks [39] characterises a model as *useful* if it can exhibit at least some real-world observations; as *accurate* if the simulated data matches historically observed real-world data; and as *complete* if the simulated data matches all of the observed patterns of the real-world phenomenon. Based on this framework the goal of validation and calibration is to construct a model that is accurate, but also complete if possible.

Input Validation Input validation concerns the structural assumptions of the model relative to the theory it is based on. Structural assumptions include choices of the rules and behaviours that define an agent's PDA cycle, the environment, and pattern of interactions. For instance, agents can be utility maximizing or employ bounded rationality. A more complete and accurate set of assumptions correspond to a higher number of parameters and variables in the model. Parameter-rich models are often difficult to calibrate, may suffer from over-fitting, end up in *dimension hell* and might even become computationally infeasible. In order to cope with the impossibility of a complete and accurate model, it is often the case that input validation is evaluated against some *stylized facts*, i.e., focusing on a limited number of variables which are most relevant to the complex system. For example, the basic reproduction number (commonly denoted by $\mathcal{R}_0$) in

[9] If the time to some next event is very small, simulations can take a long time to run. In these cases discrete-time should be used.

epidemiology is a stylized fact summarizing succinctly the impact of a communicable disease. In this sense, one does not need to know the individual trajectories of disease transmission but can develop a functioning model based on $\mathcal{R}_0$ only.

Input-validation also often requires *program validation* which refers to the validation of the simulator engine. That is, the validation of the computer codes which implement the various components of the agent-based model. Computer codes do not automatically generate errors when a bug is encountered or if something is incorrectly implemented. On the contrary, many programs will continue to produce results, independently of how bad the code is. Careful attention is therefore needed to capture *bugs* and other artefacts and is recommended that one follows principles from computer-science, including modularity and unit-testing [34]. These tests can be performed for each functional component of the system during the model development cycle, in which simple scenarios are created to verifying that all modules of the model are working in concert when executed together.

Output Validation This process systematically determines a set of model parameters and input values which maximize the fitness of the model with the observed data, i.e., finding parameter values, assumptions and structural components that make the model *fit the data well*. A well calibrated and validated model can be used to make predictions, and can provide inference on microscopic states, which may not be inferred from standard time series or statistical methods. On the other hand, a model where the parameters have not been properly calibrated is not particularly useful for inference and may even fail to describe the dynamics of the real-world complex system.

Calibration of agent-based models can be inherently difficult due to large parameter spaces, long simulation run-times, stochasticity of the structural model, and sometimes lack of empirical data. At present, ABMs are limited to some ad-hoc, qualitative calibration of the relevant parameters. That is, comparing instances of a model with different values of parameters and choosing the ones that best fit the data. It is worth noting that calibration and validation does not necessarily imply that the model reaches to a single optimal choice for the parameters. Indeed, often confidence intervals are generated in which the true value of the parameters lie. In a Bayesian approach to calibration, parameters are equipped with a prior probability distribution reflecting the uncertainty about the parameter values based on prior knowledge. In either case, this effectively means using observed empirical data to inform parameters such that the model predicts the past and present observations well. It is clear, however, that this ad-hoc approach is limited by the quality of the

empirical data available[10]. The high degrees of freedom in an ABM further exacerbates the issue of data quality. When restrained by lack or limited amount of data, Fagiolo, Windrum, and Moneta [40] suggests the use of *stylized facts*, i.e., defining a restricted set of criteria on which the model is evaluated, to circumvent the problems of data availability and quality. With these stylized facts, the ad-hoc approach continues by calculating *statistics* of the simulated data. These are then *compared* through a suitable summary measure of the model quality, conditional on the parameters values. A common choice for such a measure is the squared difference *distance metric* $d(\text{data}_{\text{simulated}}, \text{data}_{\text{observed}}) = (\text{data}_{\text{simulated}} - \text{data}_{\text{observed}})^2$, which increasingly penalizes parameters that make simulated data more distant from observed data. The summary measure, in general, is chosen with respect to the context of the model and can be, for example, specific data points, cross-sectional averages, regression analysis, and averages over realizations. If the simulated dynamics bear resemblance to the real-world observations, then the model is a possible explanation of the underlying complex system. This naïve way of calibration and validation is often computationally demanding. Indeed, for any given point in the parameter space, a large number of Monte Carlo simulations must be run to generate a distribution for the statistics of interest. Even though we have described a systematic algorithm, calibration is still based on observed qualitative similarities between model outputs and real-world data. Nevertheless, the preceding two decades have seen an increasingly number of studies attempting calibration and estimation of agent-based models by way of optimization techniques and statistical methods [41, 42, 43]. These methods include *simulated minimum distance*, *Bayesian estimation*, *Markov Chain Monte Carlo* (MCMC), *Sequential Monte Carlo* (SMC), and *particle filters* which are all closely related to each other. However, in validating the structural components of the model, a well-established method of degeneracy tests can always be performed to evaluate the outcomes using extreme value analysis by selectively disabling portions of the model or selecting plausible values for input parameters from estimated ranges [44].

It is imperative that qualitatively derived parameters are subject to sensitivity analysis. This is because a high goodness of fit does not necessarily imply a highly predictive or explanatory power [35, Chapter 5]. In this setting, sensitivity analysis amounts to running the Monte Carlo simulations with modified parameters and initial conditions, at least within the range of some confidence interval, to determine the robustness of the simulation results. Often this analysis is implemented in three stages: (i) individual parameter sensitivity: examines model sensitivity to each parameter individually by executing simulations while varying each parameter systematically

[10]It is important to note that the rise of internet-of-things, cheap sensors and individual-level data collection tools have recently led to a much greater availability of data, almost instantly from a variety of sources and platforms.

across its full range and measuring the outcomes; (ii) parameter interactions sensitivity: identifies sensitivity of model outcomes to a number of important parameter pairs with respect to the frequency and strength of their interactions; and (iii) robustness of results: evaluates the effect of parameter uncertainty and sensitivity on the entire parameter space.

2.4 A Mathematical Formalism of Agent-Based Models

2.4.1 Partial Recursive Functions

Since an ABM is a collection of independent models (§2.1), all of which are computer programs (i.e. computable by a Turing machine), there exists a corresponding unique partial recursive function[11] [19, 21]. Therefore, in principle, one can provide a representation an ABM as an explicit set of mathematical formulas (i.e. recursive functions). A system of equations is recursive (rather than simultaneous) if the output depends on one or more of its past outputs; or in other words, the values for all state variables can be determined sequentially rather than simultaneously. If, in addition, the probability distribution of the next state only depends on the current state (and not the entire history), the system is *memoryless* and is called a *Markov Chain*. In fact, it is always possible to characterize an ABM as a Markov process [26] by redefining the *state space*, a concept we have not yet defined. Thus, all ABMs *are* Markov chains.

To begin, consider an ABM Λ with m agents which evolves over some time, which can be continuous or discrete. We are interested in the *state of the model* observed at discrete times $t_1, t_2, \ldots$, with $t_k < t_{k+1}$. Even if the underlying model runs in continuous time, the model state can be sampled at discrete observation times t_k. At any time t, an agent $i \in \Lambda$ is associated with the variable $\mathbf{x}_{(i,t)}$, which take values in some finite field $\mathbb{F}$[12]. Assuming the variables can be codified in a finite set of possibilities and are quantitative in nature (i.e., real numbers or integers), then $\mathbf{x}_{(i,t)}$ corresponds to n-dimensional vector i.e., $\mathbf{x}_{(i,t)} \in \mathbb{R}^n$. The variable $\mathbf{x}_{(i,t)}$ represents the state of an agent. For example, an agent could be described by the vector (age, sex, location), so the set of variables is a triple with a mixture of numerical and codified entries. The complete system state (i.e. the global state) at time t is the collection $\mathbf{X}_t = [\mathbf{x}_{(1,t)}\ \mathbf{x}_{(2,t)} \ldots \mathbf{x}_{(m,t)}]$ which is an $n \times m$ matrix of all individual states. The evolution of the agent's state variable through time

[11] Recursive theory is still very young, having developed only in the early twentieth century with the study of computable functions and *Turing Degrees*.

[12] For example, in the case of Boolean networks the choice of the underlying field is the Galois field $\mathbb{F} = \{0, 1\}$.

is specific by the difference equation

$$\mathbf{x}_{(i,t+1)} = f_i(\mathbf{x}_{(i,t)}, \mathbf{X}_t, \alpha_i) \tag{2.1}$$

where f_i is an agent-specific *update function* (or transition function) which implements the agent's PDA cycle, α_i is a vector of agent-specific parameters (some of which could be stochastic) and $\mathbf{X}_t$ is the global state of the model. . The set of update equations f_i, one for each agent i defines the **data-generating process** (DGP) of the model. These functions are typically complicated, possibly involving discontinuities, fuzzy logic rules, and if-else statements. In the case that each f_i is a polynomial, the resulting model is called a *polynomial dynamical system* and is amenable to the computational tools and theoretical results of computer algebra, an area that utilizes powerful symbolic computation capabilities [23]. The time evolution of the overall model is thus specified as a stochastic differential or difference equation,

$$\mathbf{X}_{t+1} = F(\mathbf{X}_t, \boldsymbol{\alpha}) + \xi_t \tag{2.2}$$

where $\boldsymbol{\alpha}$ a vector of agent-specific parameters, and $\xi_t \in \mathbb{R}^{n \times m}$ is a matrix containing all stochastic elements at time t. Since the DGP functions f_i need not be linear and stochasticity is often implemented in each f_i, the agent-based model is better represented by a more general map $\mathbf{F}$

$$\mathbf{X}_{k+1} = \mathbf{F}(\mathbf{X}_k, \boldsymbol{\alpha}, \boldsymbol{\xi}_k) \tag{2.3}$$

where $\boldsymbol{\xi}_t$ is a stochastic random matrix. The initial conditions of the system at t_0 are $(\mathbf{X}_0, \boldsymbol{\xi}_0)$. Equation (2.3) is called the *transition equation* of the system. A closer look at (2.3) reveals the Markov chain representation of agent-based models, though in practice these expressions may be extremely complex and difficult to interpret. Indeed, in most cases, the explicit set of functions f_i and $\mathbf{F}$, are not tractable. In analytical models, the transition function (2.3) often have a closed form, a simple structure, and are linear (or can be linearized), and are kept free of heterogeneity (or at a minimum). Any aggregation can be performed on variables by taking expectations over the stochastic elements. However, in ABM the specification of (2.3) have little or no restrictions. Since the state space of the model can grow large (possibly with infinite states), the transition equation often does not have an analytical representation.

Once we have specified the data generating process, we are then interested in some aggregate or macro feature of our model. Let $\mathbf{y}_t$ be a set of aggregate statistics at time t, and let h be a statistic

function (i.e., some projection function from $\mathbf{X}$ to $\mathbf{y}$) over the system state

$$\mathbf{y}_t = h(\mathbf{x}_{(1,t)},\ \mathbf{x}_{(2,t)}, \ldots, \mathbf{x}_{(m,t)}) = h(\mathbf{X}_t) \tag{2.4}$$

Regardless of the complexity and specification of each f_i, the solution (i.e., shape and form) to (2.4) for each time t can always be found by backwards iteration, which traces $\mathbf{y}_k$ back to the initial conditions. That is,

$$\begin{aligned}
\mathbf{y}_0 &= h(\mathbf{X}_0) \\
\mathbf{y}_1 &= h(\mathbf{X}_1) = h(\mathbf{F}(\mathbf{X}_0, \boldsymbol{\alpha}, \boldsymbol{\xi}_0)) \\
\mathbf{y}_2 &= h(\mathbf{X}_2) = h(\mathbf{F}(\mathbf{X}_1, \boldsymbol{\alpha}, \boldsymbol{\xi}_1)) = h(\mathbf{F}(\mathbf{F}(\mathbf{X}_0, \alpha, \boldsymbol{\xi}_0), \boldsymbol{\alpha}, \boldsymbol{\xi}_1)) \\
&\vdots \\
\mathbf{y}_k &= h(\mathbf{F}(\ldots \mathbf{F}(\mathbf{F}(\mathbf{X}_0, \boldsymbol{\alpha}, \boldsymbol{\xi}_0), \ldots)))
\end{aligned}$$

This backwards iteration uniquely relates the value of $\mathbf{y}_t$ to the initial conditions, however explicating this relationship is complicated because of the stochastic $\boldsymbol{\xi}_t$ terms. Since the DGP functions (Equation (2.3)) and the statistic function h need not be linear, these random terms cannot be averaged out by expectations. Therefore, the relationship between the initial conditions $(\mathbf{X}_0, \boldsymbol{\xi}_0)$ and the statistic $\mathbf{y}$ is only realized by Monte Carlo analysis[13]. Using Monte Carlo simulations of the agent-based model for different initial states and values of parameters, one could obtain a distribution for $\mathbf{y}$. Recall that Monte Carlo techniques for simulations relies on a pseudo-random number generator (§2.3.1) which is an inherently deterministic algorithm given the initial value of the seed s. Thus, any stochasticity implemented in the model by virtue of a PRNG has a deterministic nature, which allows to further pin down the formalism of an ABM. In particular, the stochastic term $\boldsymbol{\xi}_k$ is a deterministic function of the seed s and can be considered, conveniently, part of the initial conditions. Letting $\mathbf{Z}_0 = \{\mathbf{X_0}, s\}$, Equation (2.3) is reduced to

$$\mathbf{X}_{k+1} = \mathbf{F}(\mathbf{Z_0}, \alpha) \tag{2.5}$$

It follows that the statistic $\mathbf{y}$ is given by

$$\begin{aligned}
\mathbf{X}_t &= \mathbf{F}(\mathbf{F}(\ldots \mathbf{F}(X_0, \alpha, s))) = \mathbf{F}^t(\mathbf{Z}_0, \alpha) \\
\mathbf{y}_t &= h(\mathbf{F}^t(\mathbf{Z}_0, \alpha)) \equiv g_k(\mathbf{Z}_0, \alpha)
\end{aligned} \tag{2.6}$$

[13] Suppose the output Y of a stochastic model is completely determined by $h(X)$ where h is a deterministic function and X is a random variable, but can not be computed analytically. In a Monte Carlo simulation, many realizations of $X = x$ are made and $y = h(x)$ are computed. In this way Y is built up progressively.

Equation (2.6) is called the *input-output transformation* (IOT) function and drives the results of the ABM. When the seed s is fixed, the IOT is a deterministic mapping of inputs (initial values and parameters of the system) onto outputs[14]. While this is a convenient mathematical representation, from a practical point of view the explicit form of the IOT function is unknown and the empirical distribution of the underlying stochastic process is often obtained by Monte Carlo simulations, by selecting different random seeds together with the initial values and parameters. However, since an ABM places little restrictions on the specification of Equation (2.1), careful attention is needed to keep it simple (while controlling the complexity) and within the bounds of available computation resources. Nevertheless, we have provided a convenient formalism for agent-based modelling to bridge the gap in the alleged differences in terms of mathematical rigour between pure analytical models and computer simulations. This formalization is abstract enough to apply statistical rigour in performing quantitative analysis of the emergent properties of an agent-based model, in particular to assess stationarity and ergodicity.

2.4.2 Stationarity and Ergodicity

Stationarity and ergodicity are intuitive concepts describing the long-term properties of a process or model. Stationarity of a process, in general, implies that every observation comes from the same probability distribution and that every observation carries information about the properties of the data-generating process. A variety of well-established techniques exist for understanding stationarity of traditional models. For example, the *Dickey-Fuller* tests (unit root tests in which the presence of a stochastic trend is equivalent to check the null-hypothesis in non-stationary) and *KPSS* (checking the null-hypothesis in stationary) can be used to test the stationarity of a time series [45]. However, any conclusion derived from using parametric tests is valid only if the underlying assumptions are valid. In the framework of agent-based models where the implementation of the data-generating process may not yield an analytical form, one needs to confront with *a priori* unknown stochastic properties of the model, assumptions and applicability of parametric tests that may be too restrictive or erroneous. Therefore, non-parametric tests are in general more suited for agent-based models as they do not require any assumptions on the IOT function of the model[15].

Recall that the autonomous and heterogeneous nature of an agent can evolve the system in incon-

[14] The IOT function need not be one-to-one. Indeed, different inputs might lead to the same output.

[15] Although parametric tests are superior than non-parametric tests, their superiority stems from the assumptions about the stochastic process generating the observations. On the other hand, the limited power of non-parametric tests can be overcome by increasing the number of Monte Carlo simulations.

sistent ways and global dynamics are generated by repeatedly running the model under different initial conditions and parameters. Any equilibria obtained in agent-based models are always idiosyncratic with respect to the agents, in the sense that in distinct Monte Carlo simulations, the evolution of every agent may vary substantially. Therefore, equilibria in agent-based models can only be defined at the aggregate level and in statistical terms after the global dynamics have emerged. When the agent-based model is relatively simple so that for any values of the parameters the model is stationary and ergodic, it is generally possible to characterise its equilibria. On the other hand, non-stationarity and non-ergodicity hinders the capability of fully describing the long-term dynamics of the model. By stationarity in this section, we mean weak stationarity (also known as covariance stationarity).

Definition 1. *A stochastic process $\{w_t\}$ is weakly stationary if the first moment of w_t is independent of t, that is, $\mathrm{E}(w_t) = \mu$ and if $\mathrm{Cov}[w_t, w_{t+h}]$ exists, is finite and depends only on h and not on t.*

Recall the IOT function derived earlier $\{\mathbf{y}_k\}_{k=0}^{\infty}$ defined by $\mathbf{y}_k = g_k(\mathbf{Z}_0, \alpha)$, which relates the initial state of the system $Z_0 = (\mathbf{X}_0, s)$ to the aggregate output of the model $\mathbf{y}_k$. We call $\{\mathbf{y}_k\}_{k=0}^{\infty}$ the associated time-series of the agent-based model. Intuitively speaking, a time-series from some stochastic process is stationary if the statistical properties (mean, variance, etc) remain constant over time. In other words, the time-series has no distinguished points in time. An example of a stationary process is shown is Figure 2.5.

Definition 2. *A statistical equilibrium in an agent-based model is reached in a given time window (t^-, t^+) if the associated time-series $\{\mathbf{y}_k\}_{k=0}^{\infty}$ is (weakly) stationary. The statistical equilibrium is denoted by $\mu^* = g^*(Z_0)$ and is given by*

$$\mu^*(Z_0, \alpha) = \mathrm{E}[y_t \mid t \in (t^-, t^+)] \tag{2.7}$$

with respect to the process $\{y_t\}$, and initial conditions Z_0. An equilibrium is said to be an absorbing (or steady-state) if $\mathbf{y}_k$ is stationary in $(t^-, t^+ + \tau), \tau \to \infty$. An equilibrium is said to be a transient if $\mathbf{y}_k$ is stationary in (t^-, t^+), but no longer stationary in $(t^-, t^+ + \tau), \tau > 0$.

A model may display both transient and absorbing equilibria, but the latter shows that once the system is in this state, it can no longer move out. On the other hand, a model may oscillate between two or more transient equilibria (possibly followed by an absorbing equilibria). It follows that for any given initial conditions and parameters, there can be at most one absorbing equilibrium. It is entirely possible that a model displays no absorbing equilibrium for a given statistic of interest. If a model is stationary and converges to the same equilibria $\mu^*(Z_0, \alpha) = \mu^*(\alpha)$ irrespective of the

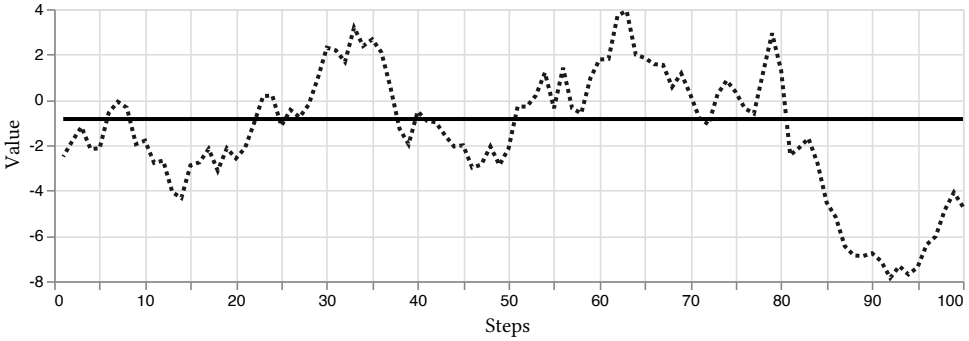


Figure 2.5: Stationarity and ergodicity are independent concepts, and one does not imply the other. A typical example of a stationary, but non-ergodic, process is drawing a number y_1 from some distribution which remains constant for the rest of the series, i.e. $y_t = y_1$ for all t (solid curve). An example of a non-stationary but ergodic process is $y_t = y_{t-1} + N(0, 1)$ (dotted curve).

initial conditions $\mathbf{Z}_0$, the process $\mathbf{y}_k$ is said to be **ergodic**. Ergodicity is sometimes defined [46] as

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathrm{Cov}(y_t, y_{t-k}) = 0 \tag{2.8}$$

which describes a property that concerns with the memory of a process. An ergodic process is characterized by weak memory (low persistence), and events far away from each other can be considered as almost independent because the effects of stochasticity fades with time. That is, a time series $\{y_k\}_{k=0}^{\infty}$ is ergodic if it exhibits the same type of qualitative behaviour. If an equilibrium is reached, it will be the same for all simulation runs, irrespective of the initial conditions, and the absorbing equilibrium will be *unique*. Different initial values (X_0, s) would not change the equilibrium value μ^*, but might change the reaching timing. Moreover, if $\mathbf{y}_k$ is ergodic, the observation of a unique time series provides sufficient information to infer the shape and form of the IOT function (2.6). That is, if the model is ergodic the properties can be analysed by using a long time-series produced by a single run of the model. If the model is non-ergodic then a set of Monte Carlo simulations (each produced by the same IOT but with different seeds) are necessary to describe, in distributional terms, the properties of the model. Non-ergodic models, on the other hand, are sensitive to their initial conditions, including the random seed. A model that is stationary but not ergodic can obtain multiple absorbing equilibria depending on the initial conditions. Furthermore, since transient equilibria are defined for a finite duration of time, these equilibria will also differ when computed for different initial conditions. Ergodicity is a powerful concept in agent-based models. If an ergodic model, one that is in a statistical equilibrium, receives a shock that moves it out of its statistical equilibrium, the system returns to the equilibrium after a finite amount of time. As a consequence, ergodic models are well suited for analytical estimation

and calibration techniques [46].

In ABMs stationarity and ergodicity tests are crucial to know whether the model reaches a statistical (unique) equilibrium state. The inherent lack of an analytical form of the data-generator process and the difficulty of dealing with unknown stochasticity requires non-parameteric statistics and tests. For agent-based models, the standard non-parametric test used to check stationarity is an application of the *Runs Test* (or Wald-Wolfowitz test) developed by Wald and Wolfowitz in 1940 [as cited in 45]. It tests the hypothesis that a given set of observations are mutually independent and randomly distributed. On the other hand, while ergodicity is crucial for understanding the long-term behaviour, literature surveying tests for ergodicity is scarce. [45] describe a modified Runs Test algorithm which considers the invariance of the moment of order k between different time-series produced by the same data-generator process, but with different random seeds. A full detailed survey of stationarity tests can be found in Phillips and Xiao [47] and Grazzini [45].

2.5 Application to Disease Dynamics

The use of ABM in the field of mathematical epidemiology has been rapidly growing, with the development of comprehensive models that incorporate various databases to address public health challenges [48, 49, 50, 51], in particular for emerging infectious diseases [52, 53, 54]. In this section, we detail an application of ABM to disease dynamics. We illustrate the construction and calibration of an ABM that describes the dynamics of disease transmission in a simple linear cascade of infection and recovery.

The model we consider here was originally developed by Kermack and McKendrick in the 1920s [18], and is referred to as the classical SIR (Susceptible-Infected-Recovered) model. In this model, the population is stratified into three different compartments (or health states) of susceptible (S), infected (I), and recovered (R). A susceptible individual leaves the S-compartment when infected and enters the I-compartment. Similarly, an infected individual leaves the I-compartment and enters the R-compartment upon recovery. An individual who recovers is assumed to have perfect immunity to the disease thereafter. When the rate of infection is proportional to the total number

of individuals in the population, the model can be represented by a set of differential equations:

$$\begin{aligned} \frac{dS}{dt} &= -\beta \frac{SI}{N} \\ \frac{dI}{dt} &= \beta \frac{SI}{N} - \gamma I \\ \frac{dR}{dt} &= \gamma I \end{aligned} \tag{2.9}$$

where β is the rate of disease transmission, γ is the recovery rate, and $N = S + I + R$. Although Equation (2.9) is written in deterministic form, it is clear that a disease transmission process involves stochasticity as contacts between individuals occur randomly, even when stochastic nature of other behavioural, host, and biological factors are omitted. Thus, the classical SIR model is built on the assumption of homogeneous mixing in the population where all individuals have equal chance to interact with others. On average, each infected individual generates βS new infected individuals per unit time.

Here, we develop an agent-based model to replicate the dynamics of the SIR model. The general framework of the model includes two main entities: (i) an *in-silico* two-dimensional lattice environment and (ii) a set of unique agents situated (fixed) in the lattice. We set the size of the lattice to 20×20 resulting in an environment with a total of 400 agents. Each agent is fully characterized by their health status of *Susceptible*, *Infected*, or *Recovered* which are programmatically codified as integer values of $0 = \text{SUS}$, $1 = \text{INF}$, and $2 = \text{REC}$, respectively. Therefore, an agent a is fully described by its associated internal state variable $x_a \in \{0, 1, 2\}$.

Running the agent-based model simply amounts to instantiating a fully susceptible agent population, introducing an infected agent as the initial condition, and iteratively letting the agents interact by executing their associated PDA cycles. The iterative process operates over a discrete time-step structure where the simulator engine advances the virtual clock by a single unit, in which the PDA cycle of each agent is carried out sequentially. The perception stage of each agent determines all possible interactions of the agent, modelled through contacts with up to 8 random agents on the lattice that are situated in neighbouring cells. The perception function returns the number of infected contacts k out of the eight random contacts of each agent at any time-step. In the decision stage, decision functions encapsulating the logic of the interactions are executed. If a susceptible agent meets an infected agent, successful disease transmission is determined using a rejection sampling-based (Bernoulli) trial where the chance of success is defined by a suitable probability distribution. Letting $x_{a,t}$ denote the internal state of an agent a at time t, the one-step

transition probability (in the Markov process) is given by:

$$\Pr[x_{a,t+1} = \text{INF} \mid x_{a,t} = \text{SUSC}] = 1 - (1 - b)^k \tag{2.10}$$

where k is the number of simultaneous contacts with infectious individuals (assuming that transmission events are independent per contact) and b is the baseline transmission probability. If the trial is successful, a susceptible individual becomes infected. In a similar way, the one-step transition from $\text{INF} \rightarrow \text{REC}$ is given by:

$$\Pr\left[x_{a,t+1} = \text{REC} \mid x_{a,t} = \text{INF}\right] = \begin{cases} 1 & \text{if } t > t_{\text{U}} \\ 0 & \text{otherwise} \end{cases} \tag{2.11}$$

where t_{U} is an empirically derived parameter representing the period of infectiousness, which in this context is sampled from a Uniform distribution between 3 to 6 time units. Finally the action stage updates the internal state of each agent as well as the global state of the model and broadcasts it to the entire lattice for the next iteration of the PDA cycle to continue. The computational implementation of the associated update function is described in Algorithm 1.

A key parameter in our model is the *unknown* transmission probability b. Typically the value of b is calibrated to a stylized fact of the underlying system such as the basic reproduction number (denoted by $\mathcal{R}_0$ as described in Section §2.3.3) or incidence rate (i.e., new infections per unit time). By running Monte Carlo simulations, the value of b could be estimated to match, for example, $\mathcal{R}_0$ obtained from simulation data with a given $\mathcal{R}_0$. It is worth noting that this process can be computationally demanding depending on the complexity of the model. One must sweep through a parameter space (which could be arbitrarily large), running Monte Carlo simulations for each value. Reduction of the parameter space to a suitable subset requires an educated initial guess.

In our example, we calibrated b to yield $\mathcal{R}_0 = 1.6$ indicating that, at the beginning of an epidemic, an infected person can infect 1.6 individuals (on average). The calibration procedure requires an initial value of b, and counting the number of secondary cases caused by the initial infected agent. If the model predicted $\mathcal{R}_0$ is not acceptable, the value of b is changed accordingly, and the process is repeated. Of course, multiple realisations are necessary for each value of b to address the first- and second-order uncertainties. Using 500 Monte-Carlo simulations, the calibration process provided an estimated value of $b = 0.047$ for which the average of realisations gives $\mathcal{R}_0 \approx 1.6$.

After calibration, we ran 500 Monte-Carlo simulation for 120 units of time to illustrate the behaviour of the system, corresponding to the spread of disease in the population. The global dynamics are

Algorithm 1: Pseudocode implementation of an agent's PDA cycle and the associated update function.

Input : agent a, time t

// when the agent is susceptible

if $x_{a,t} = SUS$ **then**
  $n \leftarrow$ Discrete Uniform $[1, 8]$ // sample a number of agents to contact
  $k = 0$ // total contacts with infectious agents
  for *1 to* n **do**
    $\hat{x} \leftarrow AgentState(\text{Discrete Uniform } [1, 400])$ // contact agent's state variable
    if $\hat{x} = INF$ **then**
      $k = k + 1$
    end
    $P = 1 - (1 - b)^k$
    if *rand()* $< P$ **then**
      $a.t_{\text{U}} \leftarrow$ Discrete Uniform $[3, 6]$ // sample length of infection t_{U}
      $x_{a,t} = \text{INF}$ // update agent state variable
    end
  end
end

// when the agent is infected

if $x_{a,t} = INF$ **then**
  if $t > a.t_U$ **then**
    // if infection duration is over, update to recovered
    $x_{a,t} = \text{REC}$
  end
end

represented by the changes in the number of individuals in different health states of the model. Figure 2.6A shows the outputs for the state variable I (i.e., the number of infections at any point in time) for each realisation. As is evident, each realisation produces a different infection curve as a result of stochasticity. It is also interesting to note that the average of realisations has a lower magnitude compared to many realisations. This is due to the fact that in many simulations, the initial infected case recovers without infecting any susceptible individuals and therefore the epidemic dies out. Figure 2.6B shows the average of realisations for all state variables.

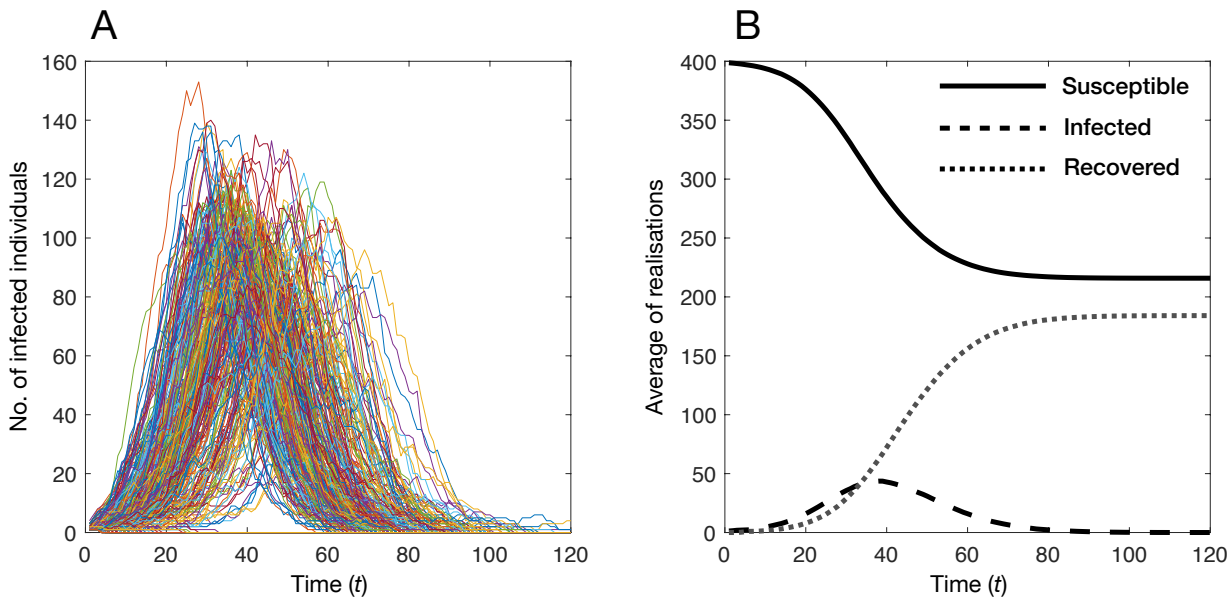


Figure 2.6: Monte-Carlo simulations of the SIR computational model: (A) independent realisations of the state variable I; (B) average of realisations for all state variables.

2.6 Concluding Remarks

The current surge of interest in ABM has gradually built up over the last twenty years [14], especially with emerging technologies in computational power and big data collection platforms that bring a higher realism to such models for simulating the real-world phenomena. The use of agent-based models has provided an additional tool for advancing quantitative science, especially in research areas (e.g., public health domain [55]) in which decision to intervene in the system dynamics may be subject to substantial heterogeneity and variability. While the capability of these models to address practical questions and inform decision-making in the face of uncertainty has been exemplified, there remain limitations to their systematic application. In particular, a more directed research is needed for expanding the theoretical aspects of ABM, by taking into account the objectives of reliability, efficiency, and adaptability which underlie the flexibility of agent-based models.

Chapter 3

Cost-Effectiveness Analysis

3.1 Introduction

Health economics (i.e., the application of economic theory to health[16]) aims to enhance and optimize the use of healthcare resources and public health policies and interventions on the basis of modern micro- and macro-economic theory. A branch of health economics is the comparison of costs and benefits of new healthcare interventions, technologies, or medications against an alternative through formal socio-economic evaluations. In the context of a health care system with limited resources and budgetary constraints, the effectiveness of a healthcare intervention is a necessary (though not always sufficient) for provision of that intervention. The costs of health care must also be considered in order to achieve maximum health gain from limited resources. There is vast health economics literature on formal, utility-maximizing evaluation techniques which offers policy-makers a means to allocate limited resources based on costs and benefits [56, 57]. Some of these techniques include cost–benefit analysis (CBA), cost-effectiveness analysis (CEA), and cost-utility analysis (CUA) [56, 57]. CBA, founded on economic welfare theory, requires all costs and benefits of interventions that are being investigated to be converted into monetary units. An intervention is considered to be economically viable if the benefits B exceeds that of the costs C (i.e., $B > C$) known as the *cost–benefit criterion*. While the basic use of the cost–benefit criterion is easy for the decision-maker, it presents challenges due to the ethical and logistical difficulties of associating monetary units with health outcomes such as mental illness-free years or additional years of life, making CBA less suitable for economic evaluations. In comparison, a CEA aims to evaluate effects of an intervention without assigning any monetary value. Instead, a CEA expresses effects using a more descriptive unit on a one-dimensional scale such as years of life saved, or increase in median survival, or survival rates. For example, a summary measure

[16] This field consists of the *economics of health* and the *economics of healthcare*, two distinct, but closely related, disciplines.

used in comparison of two interventions is the *average cost-effectiveness ratio*

$$\text{ACER} = \frac{\text{costs in units of money}}{\text{effects in natural units (e.g., life years gained)}}$$

Another more commonly used measure is the *incremental cost-effectiveness ratio* (ICER) which is the ratio of incremental costs to incremental benefits compared to the next most effective, defined as

$$\text{ICER} = \frac{\text{difference in costs}}{\text{difference in effects in natural units}} \tag{3.1}$$

Multiple interventions can now be ranked by their ICER values, which provides an easy interpretation of the results. Nevertheless, CEA comes with its own set of limitations. While CEA provides a rank order of measures, it does not decide up to which ratio an intervention should be accepted. Although this is easily addressed in the presence of a fixed budget, it is difficult under a constraint-based open budget. Another limitation is that the use of a one-dimensional measure is not suitable for comparing interventions that provide more than one effect e.g., there is no specific way to combine information about a treatment that reduces high blood pressure but also gains years of life lived, since their units are neither additive nor multiplicative. This limitation is addressed by CUA, a special case of CEA. CUA enables comparisons across different interventions and multiple effects by utilizing a common measure which encapsulates the impact of an intervention on a patient's length of life, but also the impact on their health-related quality of life. It so does this by measuring all effects of an intervention on morbidity and mortality on a multi-dimensional utility scale, through the use of appropriate weights. The best-known and most commonly used measure is called the *quality-adjusted life year* (QALY) followed by a related measure called the disability-adjusted life year (DALY), which can be used in the denominator of Equation (3.1). In addition, compared to CEA, CUA has the advantage of being applicable to different interventions (and even beyond health-care interventions) because it maps all the effects into a single utility number, and is often the preferred economic evaluation method by health care professionals and national agencies. For these reasons, CEA and CUA are sometimes not distinguished from each other in the literature and that the term *cost-effectiveness analysis* is commonly used to refer to both family of techniques. A summary of methods is presented in Table 3.1 with a formal review of CEA in §3.2.

Regardless of the economic methodology selected, any systematic evaluation requires a suitable parameterization with relevant evidence and data in order to develop policies and guidelines of new interventions and technologies. Randomized controlled trials (RCTs) provide an empirical distribution of individual patient data to allow for head-to-head comparisons of treatments in

Table 3.1: A comparison of the three types of economic evaluation frameworks. Health outcomes are a single dimensional measure such as the number of lives saved or the number of deaths averted. A utility measure is a multi-dimensional index such as the quality-adjusted life year or the disability-adjusted life year. Table adapted from Baio [57].

Evaluation type	**Costs included**		**Type of outcome**
	Direct	Indirect	
Cost-benefit	✓	✓	$
Cost-effectiveness	✓	often	Health outcome
Cost-utility	✓	rarely	Utility measure

controlled environments. While RCTs have been a crucial component in economic evaluations, there are inherent limitations in their utilization to inform policy [58]. RCT sample sizes are often too small and do not reflect standard care available to the general population. Patients are also not followed up long enough to capture the full impact of the intervention. In some scenarios RCTs can be completely ineffective in informing policy, e.g., in evaluating a potential vaccination program against aimed at reducing long-term severe sequelae caused by a disease. This inadequacy of RCTs have lead agencies such as the National Advisory Committee on Immunization (NACI) in Canada and the National Institute for Health and Clinical Excellence (NICE) in the UK to call for systematic, evidence-based methods to inform public health policies and program delivery [59, 60].

Mathematical models, as decision analytic tools, are being increasingly used within economic evaluation studies to provide an alternative approach to RCTs. The ability of mathematical models to synthesize evidence and integrate information from different sources have made them useful in situations where RCTs are not applicable or where data from RCTs is insufficient. As such they can help to inform decisions about clinical practices and health-care resource allocations. In the last ten years, with the rapid increase in the volume and heterogeneity of data, mathematical models concerning health economic evaluations have integrated several disciplines, including medical research, epidemiology, statistics, and economics. To date, the most frequently used modelling techniques have been deterministic, aggregate level models which are relatively straightforward to develop, but are limited by their homogeneity and inability to capture adaptive dynamics and possible randomness in system phenomena. For instance, in an assessment of a new vaccination programme, such models may not capture indirect benefits of an intervention such as herd immunity effects or interactions between stratified populations (e.g., different age-groups). These

models are usually complemented by dynamic transmission models (such as differential equation-based models) or stochastic aggregate models, which overcome some of the limitations, but are often intractable because of uncertain parameters and still display a high degree of homogeneity.

In Chapter 2, we argued that ABM computational systems are extremely flexible and capable of capturing (in principle) arbitrary level of heterogeneities. An ABM can incorporate the complex nature of disease transmission (e.g., co-infection by multiple pathogens) and human behaviour (e.g., sexual partnerships and contact patterns). In this chapter, we review existing methodology for CEA and present a mathematically rigorous integration of ABM within the CEA framework.

3.2 Cost-Effectiveness Analysis (CEA)

In CEA, results are characterized by the extra cost necessary to produce additional units of health benefit, i.e., *cost-per-benefit gained.* The methodology involves estimating costs, modelling intervention effects, and making an inference from the estimated costs to effects ratio. Effects of an intervention, such as illness prevention, symptoms relief, decreased medical resource utilization, and reduced loss of productivity are usually obtained from clinical trials, observational studies, academic and medical literature, and even patient interviews. However, how these effects are quantified is important. For instance, effects that can be measured in a continuum, such as life expectancy and survival time, may be estimated using survival functions [61]. We summarize the basic analysis of any CEA in four major steps:

1. The target population and the time horizon in which the analysis takes place should be clearly defined.

2. The necessary data, including quantified health outcomes, weights, and potential costs, should be collected through a systematic review of peer-reviewed literature and various other sources including government agencies and private organizations. In addition, potential data biases should be critically considered and addressed.

3. A functional disease model (e.g., Markov models or ABM computational system) that characterizes the transmission dynamics of a disease and identifies the different health states (e.g., acute or chronic, short-term or long-term sequelae) associated with disease burden should be applied with and without the intervention under study. Relevant epidemiological parameters should be sought, and whenever possible through the data collection process. This step is often the most difficult and time-consuming one, while it is also the most crucial

one.

4. Summarize the model output (i.e. relevant effects of the intervention) in a suitable, quantified measure such as the number of years gained, *quality adjusted life years*, or *disability adjusted life years*.

3.2.1 Types of Cost Variables

The cost of resources consumed by an intervention can be of two types of *direct* and *indirect*, which define the *viewpoint* or the perspective of the analysis. Direct costs include the value of all goods and services consumed in the provision of an intervention or in dealing with the immediate effects of the disease or any future consequences linked to it [62]. Common contributors to direct costs include the cost of physicians and nurses, medical testing and hospitalization, and drugs. Indirect costs are associated with societal care or impaired ability to work or engage in leisure activities [62]. Indirect costs also include loss of economic productivity due to pre-mature death caused by the disease. The extent to which indirect costs should be considered is still a matter of debate and the arguments to include them are complex [62]. Costs can further be classified as fixed or variable. Fixed costs are those that remain the same regardless of the type of disease or the intervention, e.g., the cost of buying specialised equipment for delivery of an intervention. Variable costs are those that change in the short term, e.g., increasing nursing staff because of an outbreak.

3.2.2 Quality-Adjusted and Disability-Adjusted Life Years

Quality-adjusted life year The quality-adjusted life year (QALY) is a commonly used summary measure (or an outcome measure) to quantify the effects of an intervention, incorporating the impact on both the quantity and quality of life through the use of generic utility weights. QALYs provide a high degree of standardization for the comparison of interventions, insomuch that the use of QALYs is now required by the National Institute for Health and Clinical Excellence (NICE) in the UK for health intervention assessment [60]. Today, QALY calculation methodologies have advanced from the qualitative analysis as originally introduced by Klarman, Francis, and Rosenthal [63], Torrance, Thomas, and Sackett [64], and Fanshel and Bush [65] to utilizing statistical models deeply rooted in expected utility analysis. For instance, Miyamoto [66] formulates six classes of QALY utility models and axiomatizes these models under expected utility (EU) and rank-dependent utility (RDU) assumptions. In general, QALY utility models are now widely used in the EU analysis

of health interventions because they capture quantitatively concrete outcomes such as death, but also specific improvements in the state of health, such as reduced pain or improved ability to walk. Specific states of health are often associated with *utility weights*, where a more desirable health state will receive greater utility and will be favoured in the analysis. Derivations of utility weights are discussed later in this chapter.

By considering utility weights as a single index of morbidity and mortality, the basic calculation of QALY is quite simple. Consider an individual who is burdened with some health state H that does not change up to time t (in years). Let the utility of spending t years in health state H be denoted by $u(H,t)$, which, under an expected utility framework [56], must have the from $u(H,t) = tv(H)$ where $v(H)$ is the utility weight assigned to state H. The function $v(\cdot)$ is cardinal and is unique up to positive affine transformations and can be chosen such that $0 \leq v(\cdot) \leq 1$. This leads to a quite intuitive interpretation of QALY: A year of life lived in perfect health is worth 1 QALY while death is assigned 0 QALYs. A year of life live in all other health states is worth less than 1 QALY. States worse than death can exist and they would have a negative value and subtract from the number of QALYs [67]. QALYs are calculated by simply multiplying the duration of time spent in a health state by the associated utility of that state, For instance, if an individual is in a health state H^* for 10 years with $v(H^*) = 0.6$, this would generate six non-discounted QALYs (i.e. 0.6 multiplied by 10 years). In most cost-effectiveness analysis, future QALYs are discounted[17] to present values, incorporating the idea of positive time preference, i.e., that individuals prefer to receive health benefits now rather than in the future. The standard time discounting expression is $e^{-r(x-a)}$ where r is the discount rate. The *quality-adjusted life expectancy* (QALE) of an individual in health state H at age a is then given by

$$\text{QALE} = \int_{x=a}^{a+L} v(H)e^{-r(x-a)} = v(H)\frac{1-e^{-rL}}{r} \tag{3.2}$$

where L is the life expectancy (relative to age a) and a the year in which the weight will be applied. However, the direct use of Equation 3.2 to calculate QALYs is rare. The main use of QALYs in a cost-effectiveness analysis is to assess the improvement in QALE obtained through an intervention relative to a situation in which either no intervention or a standard alternative intervention is provided. Suppose an intervention improves an individual's quality of life from health state H_1 to

[17] Discounting is a mathematical procedure for adjusting future costs and outcomes to *present value*.

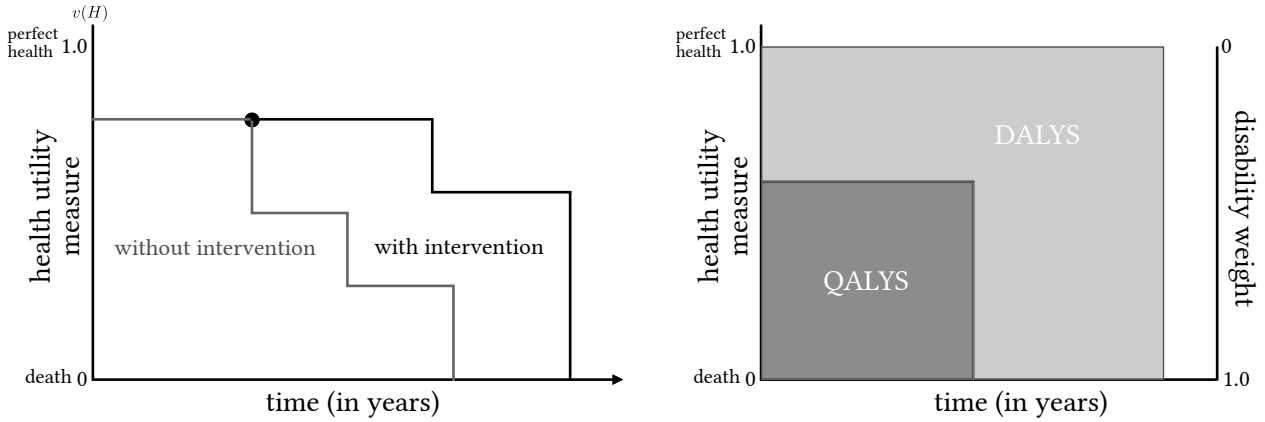


Figure 3.1: Visual interpretation of quality-adjusted life years (QALYs) and disability adjusted life years (DALYs). In the left figure, the total number of QALYs with and without intervention are the areas under the relevant polygons. The right figure shows the relationship between QALYs and DALYs.

H_2. The outcome of interest is, then, the number of *QALYs gained*, determined by

$$\begin{aligned} \text{QALYs gained} &= \int_{x=a}^{a+L_2} v(H_2)e^{-r(x-a)} - \int_{x=a}^{a+L_1} v(H_1)e^{-r(x-a)} \\ &= v(H_2)\frac{1-e^{-rL_2}}{r} - v(H_1)\frac{1-e^{-rL_1}}{r} \end{aligned} \tag{3.3}$$

where L_2 is the period over which an intervention affects an individual's quality of life and $v(H_2)$ is the associated utility weight of quality of life with intervention, while L_1 and $v(H_1)$ are the corresponding parameters without intervention. However Equation 3.3 is based on the unrealistic assumption that quality of life remains constant throughout an individual's life. A more general formula, given by [68], is

$$\text{QALE} = \sum_{m=1}^{N} Q_m \frac{e^{-r(t_m-a)} - e^{-r(t_{m-1}-a)}}{r} \tag{3.4}$$

In this formula, an individual's life expectancy is divided into N time periods t_m (with $1 \leq m \leq$, $t_0 = a$, $t_N = a + L$) with possible different durations, each affected by quality of life Q_m. The number of QALYs gained follows immediately,

$$\text{QALYs gained} = \sum_{p=1}^{p} Q_p^i \frac{e^{-r\left(t'_p-a\right)} - e^{-r\left(t^i_{p-1}-a\right)}}{r} - \sum_{m=1}^{N} Q_m \frac{e^{-r(t_m-a)} - e^{-r(t_{m-1}-a)}}{r} \tag{3.5}$$

Although we have presented a simplistic framework here, a decision-theoretic analysis based on expected utility theory including risk aversion and uncertainties can be found in [56].

In the context of the above framework, measurements of utility weights, across different studies, must ensure the same units for all possible health states. Several methods have been established to accomplish this including direct methods such as *Rating Scale*, *Time Trade-off*, and *Standard Gamble* and indirect methods also known as generic preference-based measures (see Figure 3.2) [56, 57, 69]. The rating scale method is the simplest direct approach and consists of a line with clearly defined end points describing the best and worst health states. Respondents are asked to evaluate a certain disease or health state by indicating where on the scale they consider the health state to be. The corresponding QALY weight of the health state is then read, after normalization, from the $[0, 1]$ scale. This method is not very popular due to the inherent *end-of-scale* and *spacing-out* bias. The Time Trade-Off procedure presents respondents with two alternative scenarios and ask which they prefer. The choice is between choosing to live x number of years in a impaired health state or giving up years of life to live for a shorter y period in full health[18]. The time spent in full health y is varied until the respondent is indifferent between the alternatives. The standard gamble procedure is similar. Here, the choice is between the certainty of remaining in a particular health state or taking a gamble (say with probability p) on a treatment/intervention that may award perfect health or lead to immediate death. The probability p is varied until the individual is indifferent between the certainty and the gamble. Generally, these methods are time consuming and, in some cases, unethical. As a result, indirect methods for measuring health outcomes have also been developed, often "off-the-shelf" questionnaires such as the *Short Form 36* (SF-36), the *Nottingham Health Profile*, and the *Sickness Impact Profile* [70].

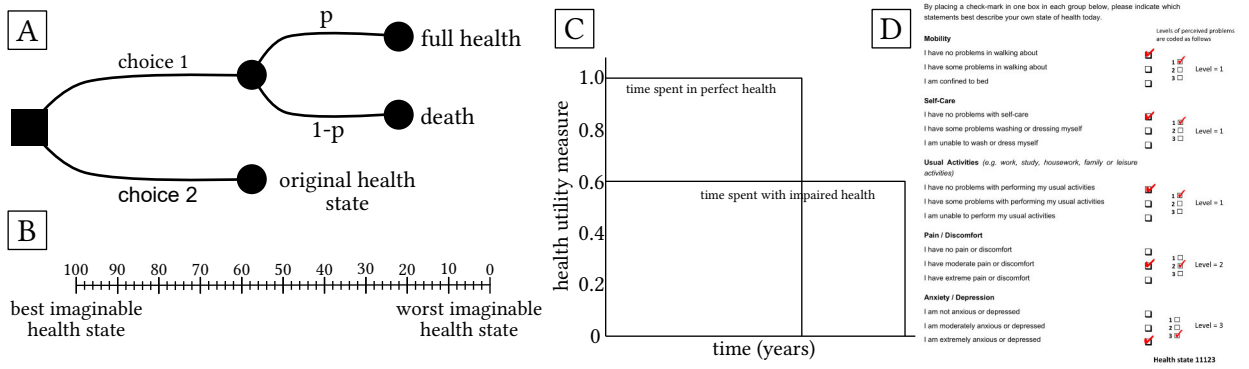


Figure 3.2: Methods of measuring health utility weights including The Standard Gamble (A), Rating Scale (B), Time Trade-Off (C), and the SQ-5D form (D).

Disability-adjusted life year The *disability-adjusted life year* (DALY) measure is an alternative to the QALY framework developed in the early 90s, as a means of estimating the global burden of

[18]It is assumed that the only possible treatment is free and would cure perfectly.

disease [71, 72]. DALYs have been the key measure in four Global Burden of Disease (GBD) studies, each assessing the worldwide impact of disease and injury [73, 74]. DALYs are a two-dimensional, time-based measure that combines years of lifetime lost through premature mortality and the number of years lived with any mental or physical disability caused by a disease or injury. One DALY has the intuitive interpretation of losing one year of perfect health. The burden of disease, defined by *disability weights*, can be thought of a measurement of the gap between an impaired health state and the ideal situation where everyone lives into old age, free of disease and disability. Disability weights (and their correct elicitation) are a crucial component of DALY calculations as they enable comparison of morbidity and mortality using a common unit. Although DALYs can be considered as a variant of QALYs, the disability weights used in DALY calculations differ from the health-related utility weights used in QALY calculations that often rely on preference-based utility measures usually elicited from surveys and questionnaires. DALY weights, on the other hand, are based on a universal set of standard weights based on expert valuations and judgements. Another significant difference is that, although measured on similar scales, disability weights represent levels of loss of functioning caused by a disease or injury whereas QALYs represent the levels of quality of life in particular health states. That is, QALY weights are normally measured on a scale in which 1 represents full health and 0 represents death; while DALY weights are measured on a scale in which 0 represents no disability. DALYs are therefore a measure of something *lost* rather than *gained*, and unlike QALYS, they are not desired themselves, but what is sought is their reduction. An advantage of DALY models is that they incorporate age-weighting function, as opposed to QALY models which assume one QALY has always the same interpretation, regardless of the age. Another advantage is that DALY calculations are essentially used in an aggregated context, where the aggregation is carried out by summing each incidence of the disease or injury.

DALY calculations are composed of the morbidity component, i.e., *years lived with disability* (YLD) and the mortality component *years lost due to premature death* (YLL). The morbidity component for a single individual is calculated by

$$\text{YLD} = \text{duration till perfect health (or death)} \times \text{disability weight}$$

The mortality component for a single individual is the difference of life expectancy and age at time of death. A DALY is simply the sum of YLDs and YLLs, i.e., $\text{DALY} = \text{YLD} + \text{YLL}$. The expressions for YLD and YLL can be extended by applying social weighting expression such as *age weighting* $Cx^{-\beta x}$ and *time discounting* $e^{-r(x-a)}$ where r is the discount rate. Thus, the general formula for

the DALY measure due to sequelae j at time t for the person i is:

$$\Delta_{i,j}^{t} = \int_{a_i^t}^{a_i^t + L(a_i)} K\delta_j C x \mathrm{e}^{-\beta x} \mathrm{e}^{-r\left(x - a_i^t\right)} \mathrm{d}x \tag{3.6}$$

where x is the time, a_i is the age of disease onset, δ_j is the disability weight of sequelae j, β and K are the age-weighting parameter and the age-weighting modulation factor, respectively, and $L(a)$ is the duration of condition in the case of disabilities or average life expectancy in the case of death. The aggregate DALY at time t for a given sample of the population with size N_t due to total disease burden is then

$$\Delta_t = \sum_{i=0}^{N_t} \Delta_i^t \tag{3.7}$$

Summing over the relevant time period T (with time discounting), the total number of DALYs attributable to a disease or injury is

$$\Delta = \int_0^T \Delta_t \mathrm{e}^{-rt} \mathrm{d}t \tag{3.8}$$

The effectiveness of some intervention is then assessed by the reduction in DALYs in the presence and absence of the intervention, i.e.,

$$\text{effectiveness} = \Delta^{\text{Without intervention}} - \Delta^{\text{With intervention}}$$

Although QALYs and DALYs share the same conceptual framework, they are not interchangeable as they are based on different assumptions and methodologies, for example, elicitation of utility weights for quality of life as compared to the expert valuation of disability weights. Moreover, while considered to be the cornerstone of economic evaluations, exemplified throughout literature, both methodologies have been under debate in recent years. Concerns relating to QALYs range from the theoretical foundations of the framework [56, 75] to problems in the multiplicative model which underlies the generation of QALY values [76, 77]. Similarly, the idea of DALY as expressing burden of disease in a single index is tempting; however, several studies have questioned both the validity of the results as well as the underlying value-judgements [78, 79, 80]. For instance, Anand and Hanson [80] exposes the inherent inequities: discounting future health gains and losses is disadvantageous for future generations, age-weighting disfavours children and seniors, and the chosen estimates for life expectancy tend to disfavour women.

3.2.3 Incremental Cost-Effectiveness Ratio

The results of a cost-effectiveness analysis are usually summarized by several available measures including the cost-effectiveness ratio, cost-effectiveness acceptability curves, and net benefit ratio. The most common, *incremental cost-effectiveness ratio* (ICER), is defined by Equation (3.1), where the difference in costs appear in the numerator and difference in effects of interventions (usually measured by QALYs or DALYs) are in the denominator. ICER values are more commonly used when the interventions are mutually exclusive, e.g., in a scenario with two incompatible medications of different costs for the same health condition, one needs to consider the rate at which higher expenses brings additional benefits. In other words, the ICER of an intervention is defined as the ratio of incremental costs and incremental benefits relative to the next best available alternative or "nothing". Most common ICER calculation frameworks take a "population" level perspective as opposed to individual level.

Let $i_1 = (e_1, c_1)$ denote the effectiveness and cost of a new intervention t_1 that is compared with the alternative $i_0 = (e_0, c_0)$. For example, (e_i, c_i) could represent sample statistics of the effect measure such as average number of QALYs. This data is usually obtained from randomized control trials or observational studies. Then we have the following scenarios:

1. $e_1 > e_0$ and $c_1 < c_0$. The new intervention is more effective and costs less, in which case it is said to be *dominant*.
2. $e_1 > e_0$ and $c_1 > c_0$. The new intervention is more effective and costs more.
3. $e_1 < e_0$ and $c_1 < c_0$. The new intervention is less effective and costs less.
4. $e_1 < e_0$ and $c_1 > c_0$. The new intervention is less effective but costs more, in which case it can discarded.

If $e_1 = e_0$ or $c_1 = c_0$, we accept the intervention that *minimizes costs* or *maximizes benefits*, respectively. It is clear that in scenarios 1 and 4, the choice between the two interventions is simple. For scenarios 2 and 3, an intervention has higher cost but also yields a greater benefit or a lower effectiveness but is also cheaper to implement. In this case, decisions can be made in light of the ICER value, i.e., the additional cost for each unit of benefit gained by the new intervention over its alternative. The ICER value in the comparison of t_0, t_1 is defined as

$$\text{ICER} = \frac{\Delta_c}{\Delta_e} \tag{3.9}$$

where $\Delta_c = c_1 - c_0$ and $\Delta_e = e_1 - e_0$, provided that the denominator is not zero. An intervention is considered acceptable by comparing the ICER value against some threshold, often based on *willingness-to-pay* (WTP) criteria. In a societal perspective that accounts for the total costs to all payers for all subjects, and the World Health Organization suggests using the per-capita gross domestic product (GDP) as a WTP threshold [81]. ICER values up to the per-capita GDP are considered *very cost-effective*, and for a WTP up to 3 times the per-capita GDP as *cost-effective*. For a WTP greater than 3 times the per-capita GDP, the intervention is considered to be not cost-effective [81][19]. Other thresholds include individual WTP (2 times of salary) [83] and value of a statistical life [84].

While the measures of interest constructed so far (i.e. QALYs, DALYs, and ICER) are seemingly straightforward, estimates of the cost-effectiveness of healthcare interventions are subject to uncertainty, which should be taken into account during the decision-making process [85]. Essentially, such analysis relies on statistical models and assumptions on the underlying distributions of costs and effects [61, 85, 86, 87, 88]. The choice of distributions used in practice is often determined by convenience, for example, on the basis of familiarity or ease of computation. Usually, the choice of normality is commonly assumed for describing cost and benefit data [86, 87, 88], or at least a large enough sample size for the sample means to be normally distributed; in addition, most approaches to estimation of cost-effectiveness adopt essentially a frequentist approach. Such assumptions are rarely realistic. For instance, data obtained from individual-level datasets (such as those collected in RCTs) trials are unlikely to be normally distributed; clinical outcomes are often measured on a binary scale, such as the eradication of a symptom, or on a ordinal scale, such as questionnaires. Similarly, cost data will often have a large presence of *structural zeros* [89], and are typically positively skewed (or even bimodal) [90, 91]. In such cases even the use of Lognormal or Gamma models becomes impractical, since these distributions are defined for strictly positive parameters. Assumptions on normality have also lead to several troubling problems in the interpretation and estimation of the ICER, in addition to the statistical difficulties with estimation of a ratio parameter [87, 88].

In order to address these concerns, several authors have now established a general framework

[19] Although there is widespread acceptance of using threshold values to assess cost-effectiveness, it is argued that thresholds based on per capita GDP have major shortcomings as guides for policy-makers [82]. An alternative approach which avoids the limitations and focuses instead on getting the largest health impact for the budget

for CEA that can incorporate arbitrary distributions and employ a Bayesian approach[20] which treats model parameters as random quantities while accounting for patient level data available (as produced by RCTs or observational studies) [57, 85, 92, 93, 94, 95]. In the next section, within the purview of this thesis, we extend this Bayesian statistical framework of [57] and [96] to integrate an alternative source of data, mainly the data-generation process of an ABM computational system.

3.3 Integration of CEA in the ABM Framework

Consider an ABM computational system of which the data-generating process (i.e. Equation 2.1) produces the set of observables $\mathcal{D}^i = \{\mathbf{x}_{ij} \mid j = 1, 2, \ldots, n_i\}$ where each $\mathbf{x}_{ij}$ is, possibly multivariate, an observation of subject j (of the *in-silico* population) receiving intervention $i \in \mathcal{I}$, with $\mathcal{I} = (0, 1, 2, \ldots)$ a set of interventions to be evaluated and n_i is the number of individuals given intervention i. Typically $\mathbf{x}_{ij}$ will be represented by two numbers: the effectiveness of a suitable clinical outcome e (e.g., e measured in terms of QALYs or DALYs), and the measure of the individual specific costs c, including the cost of intervention. We can then formally write $\mathbf{x}_{ij}$ as a vector of two elements, $\mathbf{x}_{ij} = (e_{ij}, c_{ij})$. Without loss of generality, we consider the comparison of two interventions $\mathcal{I} = (0, 1)$ where $i = 0$ represents the status quo, standard programme that is already available, and intervention $i = 1$ is suggested to replace it or implemented simultaneously, either to the entire population or to a specific sub-group of individuals. The entire observable dataset is then referred to as $\mathcal{D} = \mathcal{D}^0 \cup \mathcal{D}^1$. Denote by $\mathbf{e}_i, \mathbf{c}_i$ the sample means of effectiveness and costs, respectively, of intervention i. Define the *increment in mean effectiveness and costs* as

$$\begin{aligned} \Delta_{\mathbf{e}} &= \mathbf{e}_1 - \mathbf{e}_0 \\ \Delta_{\mathbf{c}} &= \mathbf{c}_1 - \mathbf{c}_0 \end{aligned} \tag{3.10}$$

then the ICER statistic $\hat{R}$ is

$$\hat{R} = \frac{\Delta_{\mathbf{c}}}{\Delta_{\mathbf{e}}} \tag{3.11}$$

If $\hat{R} < \lambda$ and $\Delta_{\mathbf{e}} > 0$, or if $\hat{R} > \lambda$ and $\Delta_{\mathbf{e}} < 0$, where λ represents some willingness-to-pay parameter for an additional health benefit, then the intervention is said to be cost-effective. That is, the criteria of acceptability not only depends on the cost-effectiveness ratio being less than λ but also depends on the sign of $\Delta_{\mathbf{e}}$. This point is illustrated by the *cost-effectiveness plane* (Figure 3.3)

[20] O'Hagan, Stevens, and Montmartin [92] points out that the construction of the proposed *cost-effectiveness acceptability curves* by van Hout et al. [86] which plots the probability of *net benefit* against the threshold willingness-to-pay parameters is essentially Bayesian.

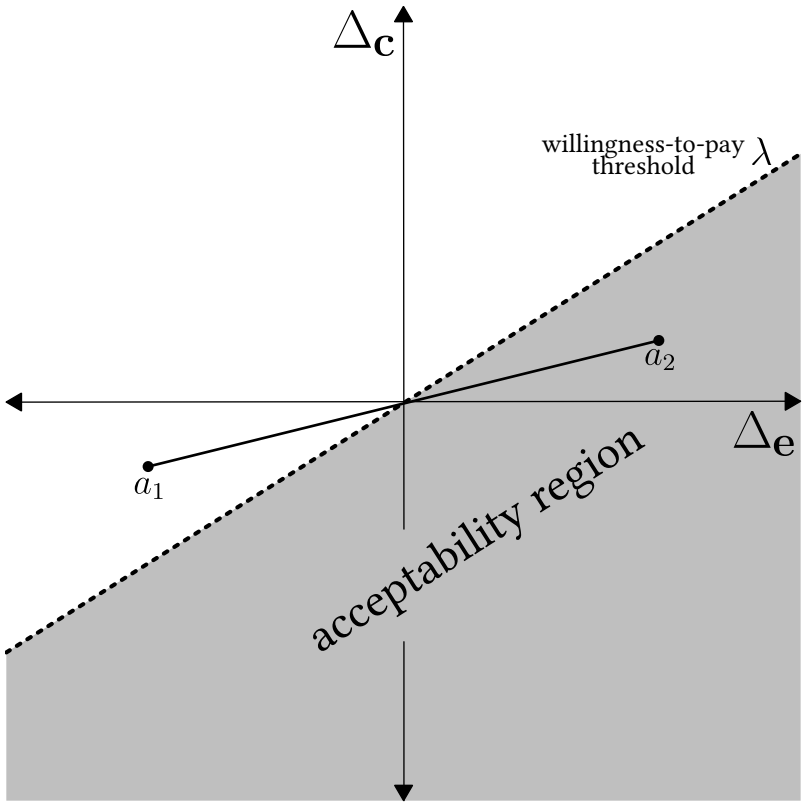


Figure 3.3: The cost-effectiveness plane. The dotted line with slope λ divides the plane into two regions, cost-effective (grey region, lower right) and not cost-effective (white region, upper left) regions. Points a_1 and a_2 have ICER values less than λ, but a_1 falls in the rejection region and a_2 in the acceptance region.

which plots possible $(\Delta_{\mathbf{c}}, \Delta_{\mathbf{e}})$ pairs with a *threshold* line of slope λ passing through the origin. The shaded area below this threshold line indicates the region of acceptability. Inference for the $\hat{R}$ is limited to constructing confidence intervals, however generating confidence intervals of such a ratio statistic is not straightforward since the variance of a ratio can not be obtained in closed form and the sampling distribution of a ratio is often unknown. For instance, Wakker and Klaassen [87] note that, under the assumption of normality for $\Delta_{\mathbf{e}}$ and $\Delta_{\mathbf{c}}$, the sampling distribution of $\hat{R}$ is Cauchy distributed and thus standard statistical techniques of inference are not applicable[21]. These concerns can be addressed by an alternative, but equivalent, formulation. The region of acceptability of an intervention in the cost-effectiveness plane can be expressed as the region in which

$$\beta(\lambda) := \lambda\Delta_{\mathbf{e}} - \Delta_{\mathbf{c}} > 0 \tag{3.12}$$

[21]Several studies have established methods for constructing statistically rigorous confidence intervals including parametric methods (normal theory methods), nonparametric methods (e.g. standard bootstrap, bootstrap percentile), and Bayesian methods [88, 97].

where $\beta(\lambda)$ is referred to as the *Monetary Net Benefit*, which expresses the cost-effectiveness criterion on a single, monetary scale by converting the $\Delta_{\mathbf{e}}$ units of effectiveness into $\lambda\Delta_{\mathbf{e}}$ units of money. In practice however, the explicit value of λ is unknown and subject to uncertainty. As such, decision makers often infer the relative cost-effectiveness by means of a *Cost-effective Acceptability Curve*, introduced by van Hout et al. [86], which plots the probability of $\beta(\lambda) > 0$ over a range of suitable λ values, based on available evidence. Although the CEAC was originally formulated under a frequentist approach, unknown parameters such as λ are often better understood in a Bayesian framework where appropriate probability distributions can be used to address the uncertainty. That is, the probability of $\beta(\lambda) > 0$ is only meaningful in a Bayesian framework.

The Bayesian extension to the above framework is relatively simple. We suppose that a general observation $\mathbf{x} = (e, c)$ follows a distribution from a family $\mathfrak{F} = \{f(\cdot \mid \boldsymbol{\theta})\}$, indexed by the population parameter $\boldsymbol{\theta} = \{\theta_0, \theta_1\}$ where the true parameter value for intervention i is θ_i. The likelihood of observing the data is then,

$$p(\mathcal{D} \mid \boldsymbol{\theta}) = \prod_{i \in \mathcal{I}} \prod_{j=1}^{n_i} f(\mathbf{x}_{ij} \mid \theta_i) \tag{3.13}$$

The uncertainty about $\boldsymbol{\theta}$ can be formally described by probability distribution by starting from a suitable prior distribution $\pi(\boldsymbol{\theta})$ that represents beliefs on $\boldsymbol{\theta}$ prior to observing any data. The posterior joint density is then given by Bayes' theorem,

$$p(\boldsymbol{\theta} \mid \mathcal{D}) \propto p(\mathcal{D} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta}) \tag{3.14}$$

from which it is possible to obtain the marginal distributions $p(\theta_i \mid \mathcal{D})$. With a functional form for the posterior distribution, future (i.e. yet unobserved) health responses be evaluated by drawing independent $\mathbf{x}$ from $p(\boldsymbol{\theta} \mid \mathcal{D})$, thereby taking into account prior information and all individual and population variability.

In order to assess the relative cost-effectiveness of interventions $i \in (0, 1)$, we consider the means of effectiveness and costs for each intervention. Denote by $\boldsymbol{\alpha}(\boldsymbol{\theta}) = (\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\gamma}(\boldsymbol{\theta}))$ the mean of the distribution $f(\cdot \mid \boldsymbol{\theta})$, where $\boldsymbol{\mu}(\boldsymbol{\theta}) = (\mu(\theta_0), \mu(\theta_1))$ is the population mean effectiveness, and $\boldsymbol{\gamma}(\boldsymbol{\theta}) = (\gamma(\theta_0), \gamma(\theta_1))$ is the population mean costs. Accordingly, an intervention i (say $i = 1$) is cost-effective relative to intervention $i = 0$ if it is more effective and cost less, i.e., if $\mu(\theta_1) > \mu(\theta_1)$ and cheaper, i.e., if $\gamma(\theta) < \gamma(\theta_0)$. We may also consider, similar to the Equation 3.10, the *increment*

in mean effectiveness and costs,

$$\Delta_\mu = \mu_1 - \mu_0 \tag{3.15}$$

$$\Delta_\gamma = \gamma_1 - \gamma_0 \tag{3.16}$$

For a given WTP threshold λ, the net benefit is given by $\beta(\lambda) := \lambda\Delta_\mu - \Delta_\gamma$, by which intervention $i = 1$ is more cost-effective than intervention $i = 0$ if $\beta(\lambda) > 0$. However since $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ are inherently functions of the random variable $\boldsymbol{\theta}$ with associated posterior distribution $p(\boldsymbol{\theta} \mid \mathcal{D})$ (Equation 3.14), β is also a function of $\boldsymbol{\theta}$, and its posterior distribution is derived from the posterior distribution of $\boldsymbol{\theta}$. We are interested in the probability of positive net benefit, i.e.

$$Q(\lambda) = \mathcal{P}(\beta(\lambda) > 0 \mid \mathcal{D}) \tag{3.17}$$

which is evaluated with the associated posterior distribution. A plot of $Q(\lambda)$ against λ is called the *cost-effectiveness acceptability curve*, which provides a visual representation of the uncertainty of the cost-effective analysis. The CEAC is particularly useful in presenting the results of a cost-effectiveness analysis as it provides a summary of the probability of cost-effectiveness by varying WTP values, as decision-makers are often not ready to commit to a single value of λ. The ultimate aim of the Bayesian extension is, then, to provide a framework for computing $Q(\lambda)$ for various assumptions and scenarios, utilizing prior information to inform on the parameters $\boldsymbol{\theta}$ [22]. However, while this approach provides a well-justified interpretation for a CEAC, it is not without inherent limitations. For instance, since the calculation of a Bayesian CEAC requires the specification of the prior distribution, there exists many CEAC plots, one for each unique prior distribution chosen, with no "correct" one. Using a non-informative prior also creates further potential areas of question in the analysis.

Another criterion to investigate cost-effectiveness of interventions is the *expected net benefit*,

$$\text{ENB} = \mathrm{E}[\beta(\lambda) \mid \mathcal{D}] = \mathrm{E}[\lambda\Delta_\mu - \Delta_\gamma] = \lambda\mathrm{E}[\Delta_\mu] - \mathrm{E}[\Delta_\gamma] \tag{3.18}$$

where the expectations are now over the distribution of $\boldsymbol{\theta}$. It is clear to see that if $\text{ENB} > 0$ then

[22] The R package BCEA developed by Baio, Berardi, and Heath [98] describes in detail and provides the relevant algorithms on how to perform health economic evaluations from the perspective of a Bayesian statistical approach. The package can be used present the results of a Bayesian cost-effectiveness model, producing standardised and highly customisable outputs. I have actively started to port this code over to Julia programming language, a fresh new approach to numerical computing.

$$\lambda > \frac{\mathrm{E}[\Delta_\mu]}{\mathrm{E}[\Delta_\gamma]} \tag{3.19}$$

which is similar to the approach to cost-effectiveness analysis based on the incremental cost-effectiveness ratio, except inference about ENB is more straightforward to evaluate since it is a ratio of expectations as opposed to an expectation of a ratio.

Although the above Bayesian approach accounts for individual variations and uncertainty of the parameters by utilizing prior information, traditional economic evaluations have resorted to additional sensitivity analysis methods to test the robustness of the results, particularly given the irreversibility of decisions and the large financial commitments of health care interventions. Various different methods for SA have been recognized in health-economic literature. In *Scenario Analysis*, likely values are selected for the parameters and the model is evaluated under all these different scenarios. Although this leads to a spectrum of results, fails to consider the possible correlation between the parameters of interest or the underlying uncertainty and thus no probabilistic meaning can be placed on the results. More in line with the Bayesian approach, and an alternative to Scenario Analysis is *Probabilistic Sensitivity Analysis* in which all input parameters are considered as random variables and are therefore associated with relevant probability distributions. Probabilistic Sensitivity Analysis is often conducted using a simulation approach, such that for each simulation $s = 1 \dots S$, a value $\boldsymbol{\theta}_{(s)}$ is simulated from the distribution $p(\boldsymbol{\theta} \mid \mathcal{D})$ and used in the cost-effectiveness analysis. An alternative approach to PSA is based on the *value of information* analysis, in which the overall value of the decision process is compared to that obtained in the actual evaluation [57].

Sensitivity analysis methods are often required for when the data-generating process is inadequate for a robust analysis, e.g., using randomized clinical trial data in which censored data is present such as time to death. In our approach of using an ABM to model epidemiological relevant scenarios, the computational model provides a data-generating process which naturally accounts for first- and second- order uncertainties, both on the individual and population level, through Monte Carlo simulations. As such, it addresses some of the limitations of traditional data-generating processes including censored data and small sample sizes, by incorporating more realistic prior information and hence reaching stronger conclusions.

3.4 Discussion

The integration of ABM computational systems with Bayesian cost-effectiveness analysis, along with rapid developments in computational power and simulation models, provides a robust framework for approaching more intuitive and complex problems in health care systems. In such a framework, independent individuals ("agents") are assigned context-specific attributes such stage and severity of disease, and move through the model experiencing events at discrete times, induced by independent decisions and localized interactions. The effects of implementing a health intervention, e.g. vaccination, can then affect the probabilities of experiencing these events or even generate new events, significantly changing the outcomes of the model. Cost-inducing events can be categorized and averaged over the relevant time horizon. Similarly, effectiveness of the intervention can be quantified in terms of individual QALYs or DALYs. By modelling at the individual level, ABM provides more flexibility and offers greater realism over traditional methods. The data-generation process of ABM enhances the utility of existing methods such as randomized clinical trials in which data may be inadequate for decision-making and capturing important heterogeneities.

Chapter 4

Case Study 1: Cost-Effectiveness of a Vaccine for *Haemophilus Influenzae* Serotype 'a'

In this case study, we highlight a human-to-human infection agent-based model of *Haemophilus influenzae* serotype 'a' (Hia) to simulate epidemic dynamics taking into account the relevant clinical and epidemiological outcomes of Hia disease. This approach takes into consideration the age-dependent individual characteristics and population heterogeneities, as well as the herd immunity generated by naturally acquired or vaccine-induced protection. The model is then utilized to conduct cost-effectiveness analysis of a potential vaccine candidate to inform government decision-making and program delivery. This cost-effectiveness analysis was conducted in the context of Nunavut, Canada where pre-dominantly the aboriginal population is affected by Hia.

4.1 Background

Haemophilus influenzae (*H. influenzae*) is a Gram-negative pathogenic bacterium that normally resides in the upper respiratory tract and is responsible for a wide range of invasive infections [99, 100, 101]. *H. influenzae* is divided into typeable and nontypeable strains based on the presence or absence of a polysaccharide capsule. Typeable strains are further classified into six serotypes ('a' to 'f') based on their ability to react with antisera against recognized polysaccharide capsules [99, 100]. Typeable strains tend to cause invasive diseases such as meningitis, bacteremic pneumonia and septic arthritis, while nontypeable strains generally cause non-invasive infections.

Among encapsulated serotypes, the serotype *b* (Hib) was one of the leading causes of invasive disease with severe long-term sequelae in paediatric population and immunocompromised adults worldwide prior to the introduction of universal infant immunization in the late 1980s [99, 102, 103].

Since the introduction of Hib conjugate vaccines, the incidence of Hib has dramatically decreased [104, 105], although it has not been eliminated and instances of resurgence have occurred [106, 107, 108]. However, surveillance programs, and clinical and epidemiological studies indicate that serotype *a* (Hia) has now emerged as a significant cause of invasive disease in some populations and geographic regions, especially among indigenous communities of the North American Arctic, including Alaska and northern Canada [109, 110, 111, 112, 113, 114, 115, 116]. A recent study shows that in Northwestern Ontario, Canada, with a relatively high (82%) indigenous population, the incidence of invasive Hia disease exceeds that of Hib in the pre-Hib vaccine era [117]. The severity and outcomes of Hia infections are reminiscent to those of invasive Hib disease [118, 119]. The reasons for increased susceptibility of these specific populations to *Haemophilus influenzae* infections are still unknown [118, 120].

The global success of Hib immunization programs over the past 20 years suggests that a protein-polysaccharide conjugated vaccine may be a solution to prevent Hia disease before it can open the niche to spread in the general population [120, 121, 122]. In 2016, the National Research Council, Public Health Agency of Canada, and the Canadian Institutes of Health Research organized a workshop to examine the current state of Hia disease epidemiology, summarize immunology and vaccine research, and identify potential vaccine solutions [121, 122]. The meeting included representatives from academia, government public health agencies, hospital laboratories, and federal departments involved in Aboriginal health. It concluded with a list of recommendations and identified the key components to focus on in the development of an Hia vaccine including completing pre-clinical studies (i.e. choice of protein carrier, obtaining regulatory approvals) and policy to demonstrate value of a Hia vaccine [121]. As such, recent research efforts have now established the pre-clinical proof of concept for a glycoconjugate vaccine against Hia; in a first study by Cox et al. [123], they show that antibodies to encapsulated Hia can be generated via a conjugation strategy and that these antibodies can facilitate bactericidal killing of Hia strains. However, the cost-effectiveness and economic impact of a potential vaccine candidate is a major factor in decisions regarding vaccine production and implementation of immunization programs. To address this knowledge gap, we developed an ABM to conduct a cost-effectiveness analysis from a government perspective. In this chapter, we detail the modelling process and its analysis, and present the results.

4.2 Model Details and Parameterization

The general structure of the model is a single-agent, discrete-time ABM in which an agent represents an *individual human*, characterized by a time-dependent vector of variables, including their demographic information, health status, and immunity levels. The key characteristics of each agent is illustrated in Algorithm 2.

Algorithm 2: The agent structure for human in HIA.

```
agent structure {
    basic variables {
        id ;                                                    // ID of the human
        health ;                                              // current health status
        age ;                                        // age in days - 365 days per year
        expectancy ;                                             // life expectancy
        expectancyreduced ;      // expectancy can be reduced due to invasive disease
        agegroup
        gender
    };
    model specific variables {
        invtype ;                                                // invasive sequelae
        invdeath ;                                  // death due to invasive disease
        plvl ;                    // immunity level following vaccination or recovery
        meetcnt ;                                      // total no.  of interactions
        pvaccine ;                            // if primary vaccine series is received
        bvaccine ;                                  // if booster vaccine is received
        dosesgiven ;                                          // no.  of doses given
        vaccineexpirytime ;   // duration of vaccine and naturally acquired protection
    };
    associated functions {
        func initialize()
        func interact()
        func update()
    };
} end;
```

4.2.1 Disease Model, States, and Outcomes

The disease model is based on the natural history of *Haemophilus influenzae* infection, which includes the states of latent (infected but not yet infectious), carriage (infectious without symptoms),

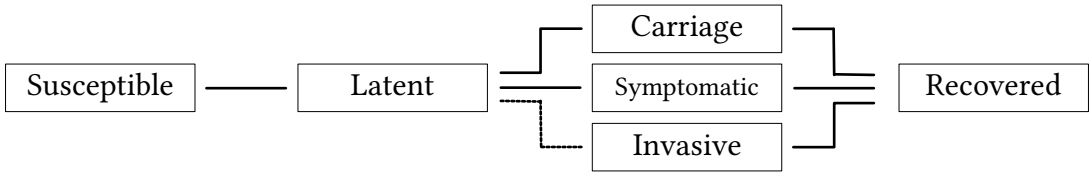


Figure 4.1: Schematic diagram for transitions and natural history between epidemiological states in the disease component of the ABM framework. The model does not include recurrent episodes of invasive disease. We structured the health status of the individuals in the model based on the epidemiological and clinical characteristics of Hia disease reported in previous studies [107, 124, 125].

symptomatic non-invasive disease, and symptomatic invasive disease. This infection stage of each agent is stored as a dynamic attribute. Recovery from infection provides a high level of protection, although there is a possibility of reinfection. Similar to vaccine-induced immunity, naturally acquired immune protection is assumed to wane over time. The model does not include the recurrence of invasive disease after the first episode or during partial protection following recovery from infection or vaccination. A schematic diagram for the model for infection dynamics is illustrated in Figure 4.1. Clinical presentations of invasive Hia disease are analogous to those caused by Hib, including meningitis, bacteraemic pneumonia, septic arthritis, and osteomyelitis [118]. Bacterial meningitis is often associated with long-term sequelae even after full recovery from the disease. Survivors of bacterial meningitis are at risk of developing life-long neurological and behavioural deficits leading to an impaired quality of life. These factors have been considered in the cost-effectiveness analysis of an Hia vaccine.

All parameters governing the dynamics of the model are drawn from previously published literature. Their descriptions are provided below and their relevant sources are provided in Table 4.3 and Table 4.4.

4.2.2 Population and Demographics

An in-silico population of 100 000 agents ("individuals") was generated, with a demographic distribution identical to Nunavut, Canada, considering that Hia was the predominant serotype causing invasive disease in the region during 2000 – 2012. Nunavut spans over 1,750,000 km^2 of the Canadian Arctic Archipelago, with a population of approximately 36,000 primarily inhabited with Indigenous Inuit. We used the sero-epidemiological data reported for Nunavut to parameterize and calibrate the model to the incidence of invasive Hia disease in different age groups [111, 114, 116]. During a 13-year period from 2000 to 2012, a total of 89 cases were serotyped, of which 43 were Hia with an overall rate of 13.7 per 100,000 population for the incidence of Hia disease [111].

Age specific annual incidence rates were 274.8 per 100 000 for < 1 year old, and 61.2 per 100 000 for 1–4 years of age [111]. Taking this into account, we stratified our population into five age groups of <1, 1–5, 6–10, 11–59, and 60+ years of age to match the age-specific incidence rates and other population characteristics used for model parameterization. In the simulation model, all natural or disease-induced deaths were replaced by newborns to maintain a constant population size.

4.2.3 Transmission and Infection Dynamics

Disease transmission (infection) occurs through contacts between susceptible and infectious individuals in the state of carriage or symptomatic (but non-invasive) disease. Since invasive disease often requires hospitalization, we assume that this state of disease is not a significant contributing factor to disease transmission.

In order to determine the contact structure between individuals, the population was stratified to four age groups (0 to <2 years of age; 2 to <5 years of age; 5 to <10 years of age; and 10+ years of age). The population contact structure was then derived by converting a relevant *Who Acquires Infection From Whom* (WAIFW) matrix [107] for Alaska Native populations, whose elements correspond to the product of the annual rate at which persons of age group i encounter persons of age group j and the probability of transmission between a susceptible contact in age group i and infectious contact in age group j, into a probability distribution of individuals in age group i encountering individuals in age group j. See Table 4.1. Disease transmission occurred as a result of

Table 4.1: (a) *Who Aquired Infection From Who* matrix adapted from [107]. The elements correspond to the product of the yearly rate at which individuals of age group i encounter persons of age group j and the probability of transmission given contact between susceptible in age group i and infectious in age group j. (b) Cumulative probability distribution for individuals in age group i encountering individuals in age group j.

		age group of infectious person			
		0-2	**2-5**	**5-10**	**>10**
age group of susceptible person	**0-2**	2.11	0.15	0.53	0.03
	2-5	0.55	0.40	0.50	0.12
	5-10	0.56	3.68	3.61	0.13
	>10	0.55	0.55	0.81	1.43

		age group of person j			
		0-2	**2-5**	**5-10**	**>10**
age group of person i	**0-2**	0.75	0.80	0.98	1.0
	2-5	0.35	0.60	0.92	1.0
	5-10	0.07	0.53	0.98	1.0
	>10	0.17	0.33	0.57	1.0

rejection sampling-based (Bernoulli) trials where the chance of success is defined by a probability

distribution. Given a contact between a susceptible (or recovered person) i and infectious person j, the probability transmission of disease between this susceptible-infectious pair was calculated by

$$\mathcal{P}_{\text{transmission}}(j \to i) = C\beta_{\text{ag}}(1 - \rho_i) \tag{4.1}$$

where β_{ag} is the age-specific transmission rate of person j calibrated by fitting the model to reported incidence rates (see Table 4.3), C is the reduction in transmission for individuals in carriage state that assumed to be 50% less infectious compared to those in symptomatic state [126], and ρ_i is the effect of naturally-acquired or vaccine-induced protection acting as a reduction factor in the baseline transmission depending on the level of immune protection at the time of contact (Table 4.3).

Upon successful transmission of the bacteria, infected individuals move to the latent state. The latent period for an infected individual (following exposure and colonization) was sampled from a truncated log-normal distribution with shape and scale parameters of 0.588 and 0.458, with the mean of 2 days [127]. After the latent period has elapsed, individuals become infectious and experience one of the clinical states of carriage, symptomatic, or invasive disease. The transition to carriage or symptomatic (and invasive) states is determined through a probability distribution defined by decision tree analysis, taking into account their infection stage, age, natural immunity levels, and vaccination status. The decision tree paths are illustrated in Figure 4.2. For example, a susceptible individual who has previously experienced infection has a 60%–90% chance of transitioning to a carriage state. The duration of carriage varies in reported estimates, but it can last from several days to several months [107, 128]. We sampled the carriage period from a uniform distribution in the range of 14–70 days. The period for symptomatic infection was sampled from a Poisson distribution with a mean of 2 days post symptoms onset. The symptomatic (non-invasive) infection was considered non-communicable 2 days after the start of effective antibiotic treatment [129]. Individuals who further progressed to invasive disease (manifested as meningitis, pneumonia, or non-meningitis-non-pneumonia) were assumed to receive critical care (i.e., hospitalization).

The length of hospital stay, obtained from the Canadian Institute of Health Information databases [130], varies by age and depends on the type of invasive disease (Table 4.2). About 25% of deaths caused by invasive disease due to bacterial meningitis occur within 2 days of hospitalization [131]. To corroborate this estimate for individuals with fatal outcomes, the time spent in the hospital before death was sampled from a truncated Poisson distribution with an average of 4 days and maximum of 10 days. The case fatality ratio was set to 9.1% for invasive disease [111].

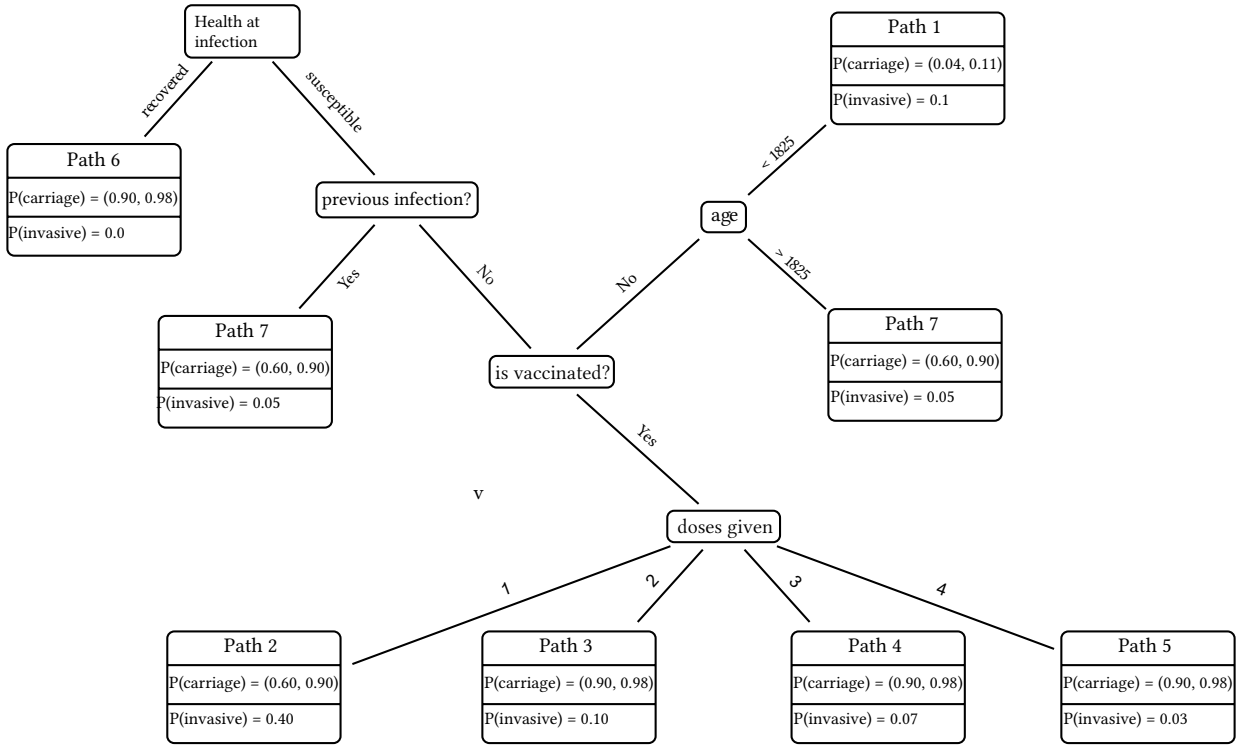


Figure 4.2: Progression through the Hia disease model by an infected agent.

Table 4.2: The length of hospital stay estimates for meningitis, pneumonia, and non-meningitis-non-pneumonia (NMNP) outcomes of invasive disease due to Hia.

Parameter Description	**Baseline value (range)**		**Source**
Length of hospital stay (days) for invasive disease			
Age Group	Meningitis	Pneumonia	NMNP
<1 years	11.8 (10-14)	5 (3-7)	9.1 (7-11)
1-7 years	9.3 (7-11)	4 (2-6)	5.6 (5-7)
8-17 years	4.9 (3-7)	5.6 (4-7)	6.7 (5-9)
18-59 years	6.9 (5-9)	8.1 (6-10)	9.5 (8-11)
60-70 years	11.2 (9-13)	8.8 (7-11)	11.7 (10-14)
80+ years	19 (17-21)	8.4 (6-10)	12.2 (10-14)

[130]

Clearance of infection upon recovery was assumed to confer a transient immune protection of 95% that wanes over time between 2 and 5 years, thereby increasing the level of susceptibility to re-colonization [107, 126, 127]. This protection level was considered adequate to prevent invasive disease if infection occurred.

4.2.4 Vaccination Schedules and Dynamics

Prevention of infection and invasive disease following vaccination was implemented based on the number of vaccine doses and induced protection levels. In the context of Hib, clinical studies have shown that a high level of protection requires at least two doses of conjugate vaccines [132, 133], followed by a booster. In the absence of Hia vaccine data, we assumed that an Hia vaccine will be rolled out in a similar schedule to the routine infant immunization programs against Hib in Canada[23] [134]. We therefore implemented vaccination in primary series with 3 doses scheduled at 2, 4, and 6 months, followed by a booster dose at 18 months of age. Furthermore, considering the similarity between Hia and Hib immune dynamics, vaccine efficacy was estimated based on Hib conjugate vaccines against colonization and invasive disease. Vaccine specific parameters are summarized in Table 4.3.

The durations of high protection following primary series (1–3 years) and booster vaccination (6–10 years) were also sampled from estimates reported for vaccination against Hib [107]. For the coverage of infant immunization, we used estimates of Hib vaccine coverage in Canada for primary series (77%) and booster vaccination (93.5% of primary vaccinated infants) [135], though recognizing that this coverage is even lower in remote regions like Nunavut. Since bacteria with identical or similar polysaccharides to *Haemophilus influenzae* can induce cross-protective antibodies [136, 137], the accumulated exposure to such bacteria may raise some level of pre-existing immunity. We therefore assumed a 50% protection against colonization for individuals older than 5 years of age.

4.2.5 Model Calibration

The model was calibrated by fitting baseline age-specific transmission probabilities β_1, β_2, β_3, β_4 to reported incidence rates (Table 4.3), by running simulations over a 30-year period in a no-vaccine scenario corresponding to the period of 1991–2020. In particular, we ran simulations for the first 10 years as a warm-up period to reach a stationary state in the model. In the next 13 years, transmission probabilities β_1, β_2, β_3, β_4 were systematically adjusted over a 4-dimensional parameter space to generated results that match incidence rates reported for different age-group, with an overall rate of 13.2 per 100,000 population. All calculations were based on the average of 500 Monte-Carlo independent realizations. Each simulation was seeded independently with an

[23]The schedule for a routine Hib vaccination varies in different countries depending on the type of vaccine and the region's public health recommendations. For example, vaccination schedules for individual EU countries and specific age groups can be found in https://vaccine-schedule.ecdc.europa.eu/.

Table 4.3: Description of *Haemophilus Influenzae* serotype 'a' model parameters and associated ranges. Model parameters were largely derived from published literature.

Parameter description	**Baseline value (range)**		**Source**
Transmission rates for infection β			
<1 year	β_1 = 0.0793		calibrated to age-specific incidence rates
1 – 5 years	β_2 = 0.0545		
6 – 10 years	β_3 = 0.0491		
10 – 60 years	β_4 = 0.0799		
60+ years	β_5 = 0.0491		
Relative transmission of carriage	0.5 (0.3 - 0.7)		[127]
Infection Parameters			
Latent period following colonization	Mean: 2 days (Lognormal)		[127]
Period of communicability following the start of treatment for (non-invasive) symptomatic	2 days		[129]
Infectious period for carriage	14 – 70 days		[107, 128]
Disease Outcomes			
Case fatality ratio	9.1%		[111]
Probability of carriage	Depends on age and immunity		[132, 138]
Probability of invasive disease	Depends on age and immunity		[138]
Length of hospital stay	Depends on age and immunity		[130]
Time spent in hospital before death	Mean: 4 days (Poisson)		[131]
Immune Protection levels	Against colonization	Against Invasive	
After 1st dose	50%	60%	[132, 133]
After 2nd dose	80%	90%	
After 3rd dose	85%	93%	
After booster dose	(85%-95%)	97%	

Table 4.3: Description of *Haemophilus Influenzae* serotype 'a' model parameters and associated ranges. Model parameters were largely derived from published literature.

Parameter description	Baseline value (range)	Source
After recovery from infection	95% 100%	
Duration of immune protection against colonization		
Naturally acquired	2 - 5 years	[107, 126]
After completing primary series	1 - 3 years	
After receiving a booster dose	6 - 10 years	
Vaccine Coverage		
Primary Series <1 year	77%	[135]
Booster Dose <2 years	93% of primary vaccinated individuals	

individual in the latent state of the infection, and the events and outcomes were recorded over time.

4.2.6 Cost-Effectiveness Analysis

For the cost-effectiveness analysis, we considered sequelae for individuals who develop invasive disease through short and long-term financial burden together with associated disabilities, and possible reduction of life expectancy [139].

Direct costs borne by government were considered for the cost-effectiveness analysis, including physician visits, immediate hospitalization for invasive disease, and long-term care for patients with neurological sequelae. In addition, infants with invasive disease can't be treated in their home community due to the general lack of medical resources, expertise, equipment, and facilities in the Canadian North and have to be medically evacuated to hospitals by the Medivac program [140]. As these transportation costs can be significant (upwards of $55,000 CAD), our analysis takes considers the associated costs of the Medivac program. For minor sequelae, we considered costs associated with special programs in pre-school (0-5), school years (6-18), and adult training programs up to 22 years of age [141]. Bacterial meningitis is often associated with major sequelae even after full recovery from the disease. Survivors of bacterial meningitis are at risk of developing life-long neurological and behavioural deficits leading to an impaired quality of life, and often

require lifetime care [141] which carries a significant financial burden.

Costs of vaccine doses and administration were estimated based on a monovalent Hib conjugate vaccine in Canada. Indirect costs associated with lost productivity and those incurred by households were not included in the cost-effectiveness analysis. All costs were converted to year 2017 Canadian dollars using the health and personal care component of the Canadian Consumer Price Index [142] and future costs and health outcomes were discounted at 3% [143]. All cost parameters and values are presented in Table 4.4.

Table 4.4: Cost parameters of *Haemophilus Influenzae* treatments and interventions used in the cost-effectiveness analysis.

Parameter description		**Estimated Costs**		**Source**
Medivac per evacuation		$55,000		[121]
Physician visit		$ 60		[130]
Hospitalization per night	Meningitis	Pneumonia	NPNM	
<1 years	$11,076	$8,739	$10,237	[130]
1 – 7 years	$8,856	$7,554	$7,088	
8 – 17 years	$6,833	$9,649	$7,508	
18 – 59 years	$9,994	$13,278	$11,696	
60 – 79 years	$16,088	$13,093	$13,645	
80+ years	$24,479	$9,983	$12,866	
Long-term sequelae				
Major (lifetime)		$109,664/year		[141]
Minor (up to age 22)				
Pre-school		$21,434/year		
School years		$26,917/year		
Adult training		$13,957/year		
Vaccination				
Vaccine dose		$20		[144]
Administration		$8		[144]
Wastage		3%		Assumed

Our model measures effectiveness in disability-adjusted life years (DALYs) averted, which was estimated using the method recommended in the 1996 global burden of disease study [72]. Although there is no data comparing the long-term disability rates of Hia and Hib, we considered existing

weights (as measures of impairment of quality of life) for major and minor sequelae used for Hib, corresponding to standard global burden of disease categories. These include cognitive deficit, hearing loss, motor deficit, seizures, visual impairment, and multiple impairments [145, 146, 147]. For the scenarios with years of life lost due to bacterial meningitis, we assumed a reduction of lifetime in the range 2–10 years [139]. The disability weights for minor and major sequelae due to invasive Hia disease are presented in Table 4.5, along with their sources. To determine cost-effective scenarios for vaccination, we calculated ICER values and confidence intervals over a 10-year period following the start of vaccination. Since the ICER is neither a sufficient nor an unbiased statistic, the uncertainty around the point estimates was assessed by applying a non-parametric bootstrap method [87, 97]. A cost-effectiveness plane was then utilized to offer a visual representation of the joint distribution and uncertainty along with 95% confidence intervals.

Table 4.5: Disability weights for long-term major and minor sequelae of bacterial meningitis due to invasive *Haemophilus Influenzae* disease.

Parameter Description	**Parameter value (range)**			**Source**
Long-term sequelae	Disability weights	Major Sequelae	Minor Sequelae	
Cognitive difficulties	0.469	0.01	0.024	
Seizure disorder	0.099	0.015	0	
Hearing loss	0.223	0.032	0.006	
Motor deficit	0.388	0.012	0.013	[145, 146]
Visual disturbance	0.223	0.01	0.001	
Clinical impairments	0.359	0.07	0.008	
Multiple impairments	0.627	0.019	0.008	

4.3 Results

We implemented two vaccination strategies for infants during the warm-up period in the model simulations, with primary vaccination coverages of 77% and 90%. For booster doses, we considered coverages of 90% and 93%, and this coverage applied to those who have completed the primary series. We also considered scenarios in which the expected individual lifetime was reduced due to invasive disease outcomes [139]. Individuals who suffer from invasive disease with major sequelae can die prematurely before reaching expected life expectancy. This was modelled by a uniform

Table 4.6: Summary of Hia simulation scenarios and their associated figures.

Scenario	Primary Coverage	Booster Coverage	Life-time Reduction	Associated Figures
I	77%	90%	No	4.3, 4.4, 4.5, 4.6
II	90%	93%	No	4.7, 4.8, 4.9, 4.10
III	77%	90%	Yes	4.11, 4.12, 4.13, 4.14
IV	90%	93%	Yes	4.15, 4.16, 4.17, 4.18

distribution with minimum of 2 years and a maximum of 10 years of life lost. These scenarios are summarized in Table 4.6.

No lifetime reduction The results of the scenario with 77% coverage of primary series and 90% coverage of booster dose and without any lifetime reduction due to invasive disease are illustrated in Figure 4.3. In this figure, the overall and age-specific incidence rates are shown over a 40-year time horizon. Years 10 to 22 correspond to a calibration period fitting the model to Nunavut incidence rates between 2000 and 2012. At year 30, the model diverged to two alternative settings. The first setting was in the absence of vaccination. The second setting included a routine infant vaccination schedule. The results show that the vaccination program reduced the overall incidence of invasive disease by 63.8% on average after 10 years of vaccination from 9.97 to 3.61 cases per 100 000 population. While there is a steady decline in the incidence of invasive disease in infants following the start of vaccination, we observed an initial increase in the incidence of disease in other age groups. This is explained by the effect of partial protection conferred by primary series in infants, who are more likely to experience carriage if infected. Because the carriage period is significantly longer than symptomatic period, it can lead to more opportunities for infection transmission in the population. With continuous vaccination and the rise of herd immunity, the initial increase in the incidence of other age groups is followed by a sharp decline several years after the onset of the vaccination program. The cost-effectiveness analysis was performed over the course of 10 years post vaccination. Our results show that a routine vaccination program with 77% primary coverage reduces the overall costs of disease management by 53.4% on average by the tenth year, from CDN $1.863 million (95% CI: $1.229–$2.519) to CDN $0.868 million (95% CI: $0.627–$1.120) (Figure 4.4). The cost categories included in the analysis were hospitalization, MediVac, physician visits, and long-term disability care caused by major and minor sequelae (Figure 4.5). The largest costs are associated with hospitalization and long-term care of major sequelae. The distribution of costs in the presence and absence of vaccine were significantly

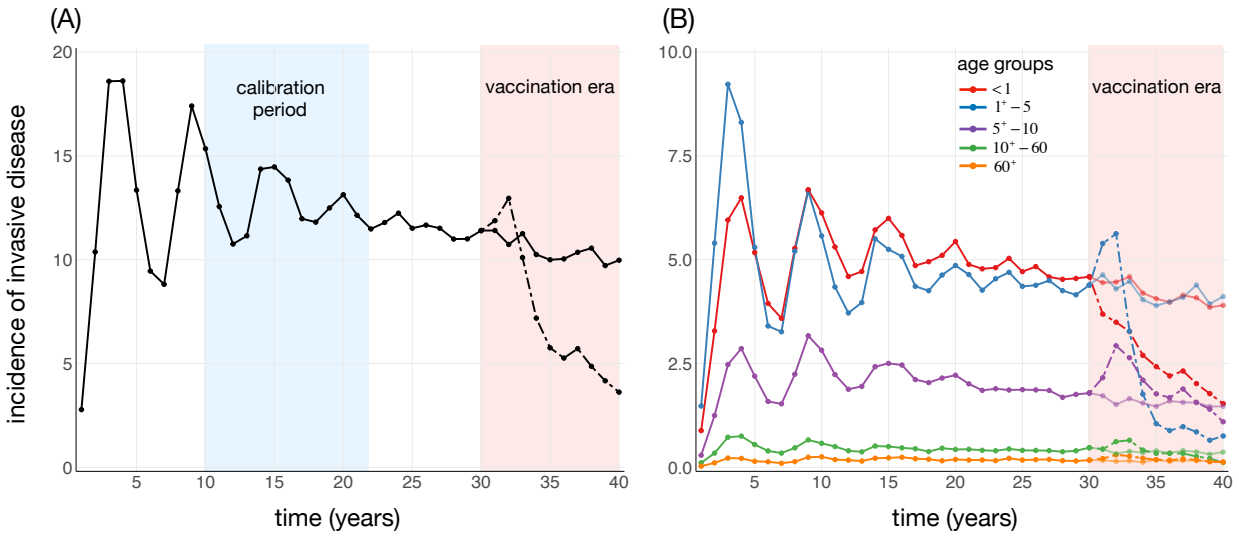


Figure 4.3: **Scenario I.** Overall (A) and age-specific (B) incidence rates over a 40-year simulation time period. Years 10 to 22 correspond to a calibration period fitting the model to Nunavut incidence rates between 2000 and 2012. At year 30, two alternative scenarios were run in the absence of vaccination (solid curves), and with routine infant vaccination schedules (dashed curves).

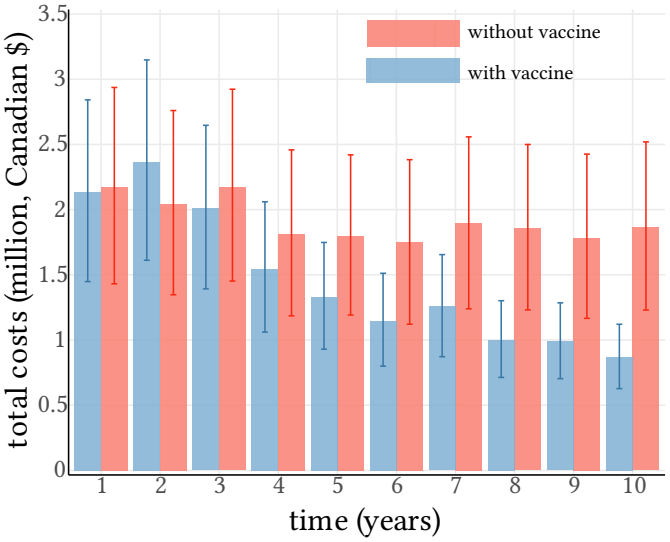


Figure 4.4: **Scenario I.** Overall average annual costs without vaccination (red) and with vaccination (blue) over a 10-year period. Average costs and associated 95% confidence intervals were computed by performing nonparametric bootstrapping method over 500 independent simulations. Direct costs included physician visits, hospitalization, MediVac, major and minor disability, and vaccination costs of doses per individual, administration, and wastage. All costs are in 2017 Canadian dollars.

different (Mann-Whitney U test, p-value < 0.001).

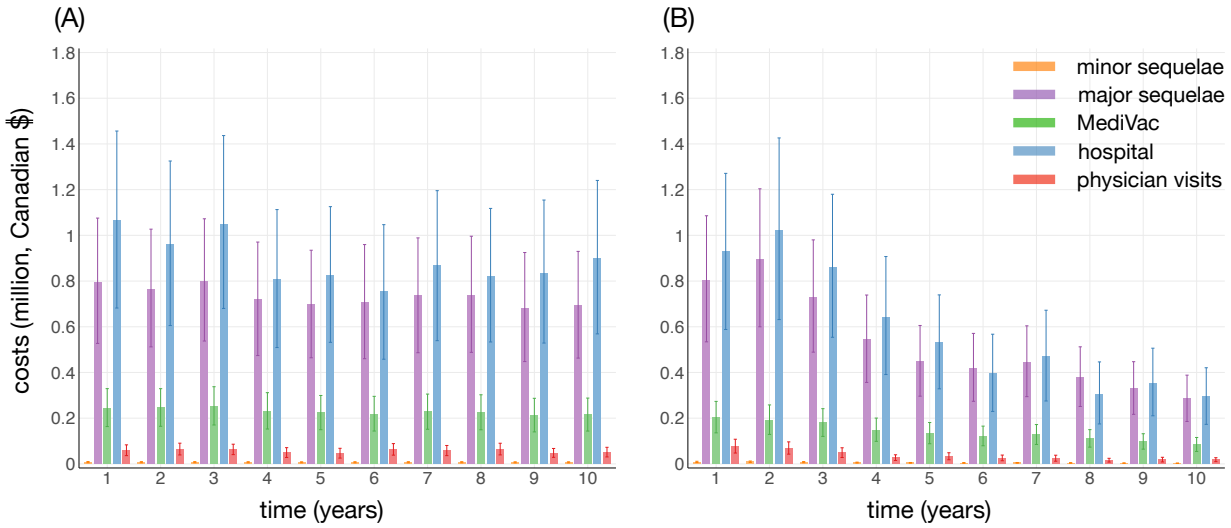


Figure 4.5: **Scenario I.** Distribution of different cost categories with their 95% confidence intervals, obtained using nonparametric bootstrap over 500 independent simulations without vaccination (A), and with routine infant vaccination schedule (B). All costs are in 2017 Canadian dollars.

The associated cost-effectiveness plane, constructed using 5000 bootstrap replicates, of the net costs and net effects (measured by DALYs) is illustrated in Figure 4.6. This figure, based on a $30 vaccine cost per individual (including vaccine dose, administration and wastage) derived from current Hib vaccine prices, shows the cost-effectiveness results with the inclusion and exclusion of the MediVac program. All ICER values are clustered in the dominant region of the cost-effectiveness plane, suggesting that a routine infant immunization program is expected to be very cost-effective in both scenarios. Further interpretation of these results suggests that in populations with similar incidence rates, vaccination would still be very cost-effective even when critical care resources for management of invasive disease are available and exorbitant costs of MediVac are averted.

Since an Hia vaccine has not yet been developed (and licensed), vaccination costs per individual are undetermined. We therefore performed the same analysis for a plausible range of vaccine costs per individual from $10 to $50 per dose. The results of ICER values, illustrated in Figure 4.6, show that our conclusions of vaccine cost-effectiveness for a routine infant immunization program remain intact.

Increasing the coverage of primary series to 90% did not alter the conclusion obtained for 77% coverage of primary vaccine. Figure 4.7 shows the rapid decline in incidence by the tenth year of vaccination and Figures 4.8 and 4.9 illustrate the cost savings associated with the vaccination

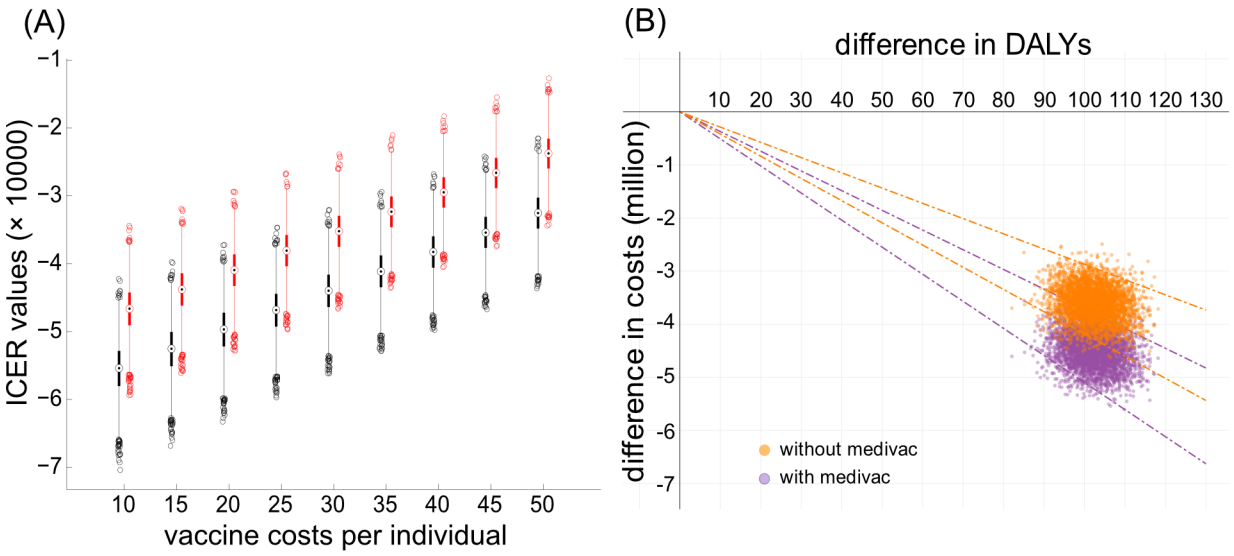


Figure 4.6: **Scenario I. (Left)** Boxplots for ICER values with the inclusion (black) and exclusion (red) of Medivac costs, as functions of vaccine costs per individual. Each box contains 50% of data points for ICER values between the first and third quartiles of the bootstrap sampling distribution. Whiskers represent the remaining 50% data points. **(Right)** Cost-effectiveness plane using average costs (y-axis) and average DALYs averted (x-axis) calculated on daily bases for two scenarios in the presence and absence of MediVac program. The dashed lines indicate 95% confidence intervals for the ICER values. The ICER values, clustered in the southeast quadrant of the cost-effectiveness plane, indicate that the vaccination program is very cost-effective.

programs. In Figure 4.10, ICER values are clustered in the dominant region of the cost-effectiveness plane, suggesting that a routine infant immunization program is expected to be very cost-effective with or without MediVac programs. Figure 4.10 describes the results of cost-effectiveness analysis for a plausible range of vaccine costs.

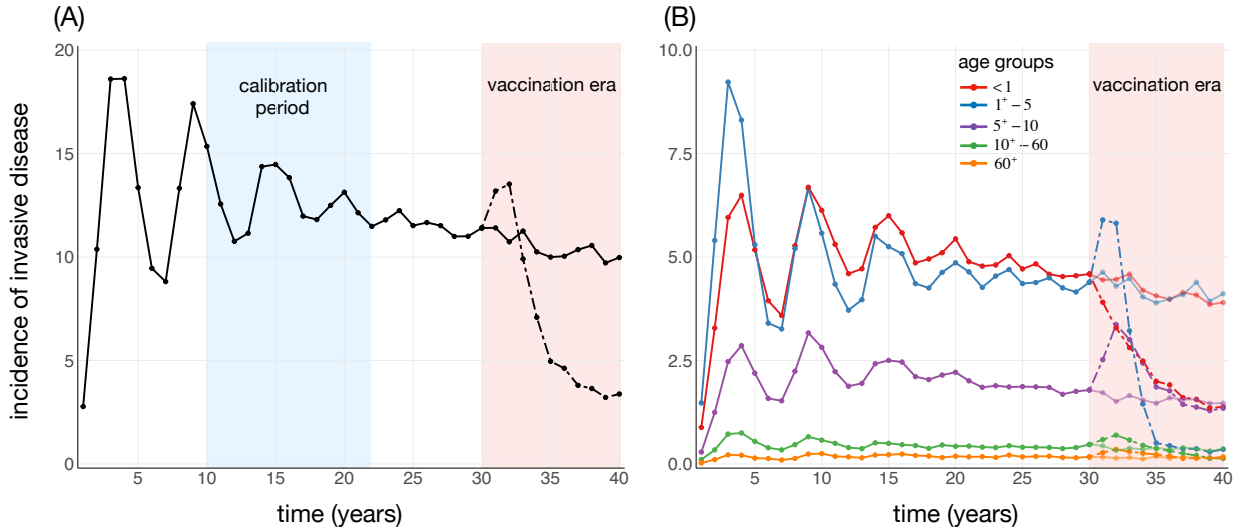


Figure 4.7: **Scenario II.** Overall (A) and age-specific (B) incidence rates over a 40-year simulation time period. Years 10 to 22 correspond to a calibration period fitting the model to Nunavut incidence rates between 2000 and 2012. At year 30, two alternative scenarios were run in the absence of vaccination (solid curves), and with routine infant vaccination schedules (dashed curves).

With lifetime reduction Similar results for incidence and vaccine cost-effectiveness were obtained by considering a reduction in lifetime expectancy due to major sequelae. The reduction

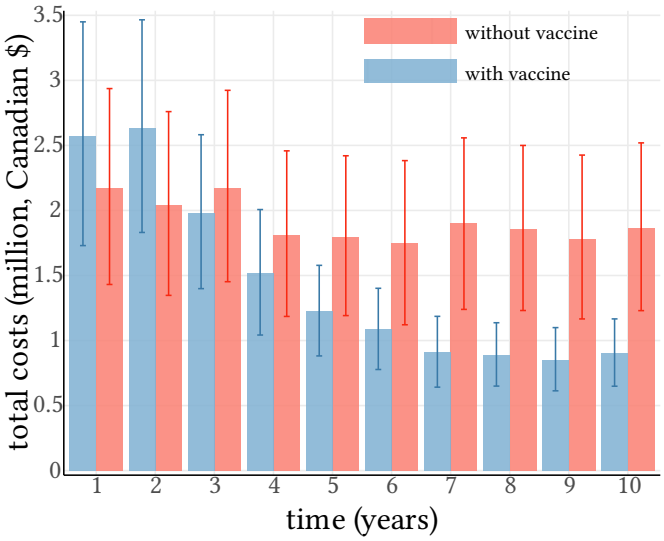


Figure 4.8: **Scenario II.** Overall average annual costs without vaccination (red) and with vaccination (blue) over a 10-year period. Average costs and associated 95% confidence intervals were computed by performing nonparametric bootstrapping method over 500 independent simulations. Direct costs included physician visits, hospitalization, MediVac, major and minor disability, and vaccination costs of doses, administration, and wastage. All costs are in 2017 Canadian dollars.

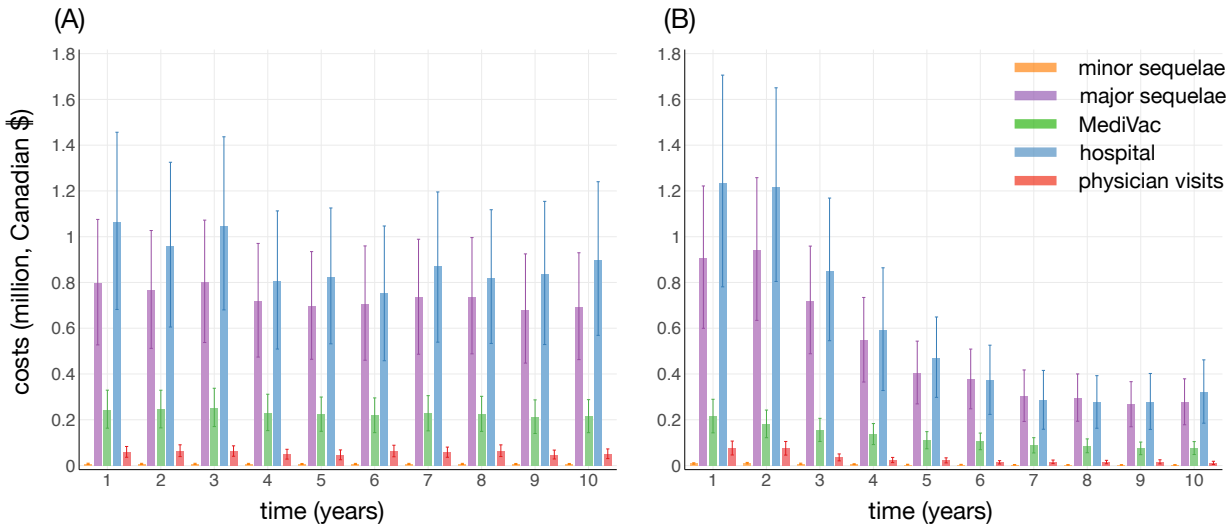


Figure 4.9: **Scenario II.** Distribution of different cost categories with their 95% confidence intervals, obtained using nonparametric bootstrap over 500 independent simulations without vaccination (A), and with routine infant vaccination schedule (B). All costs are in 2017 Canadian dollars.

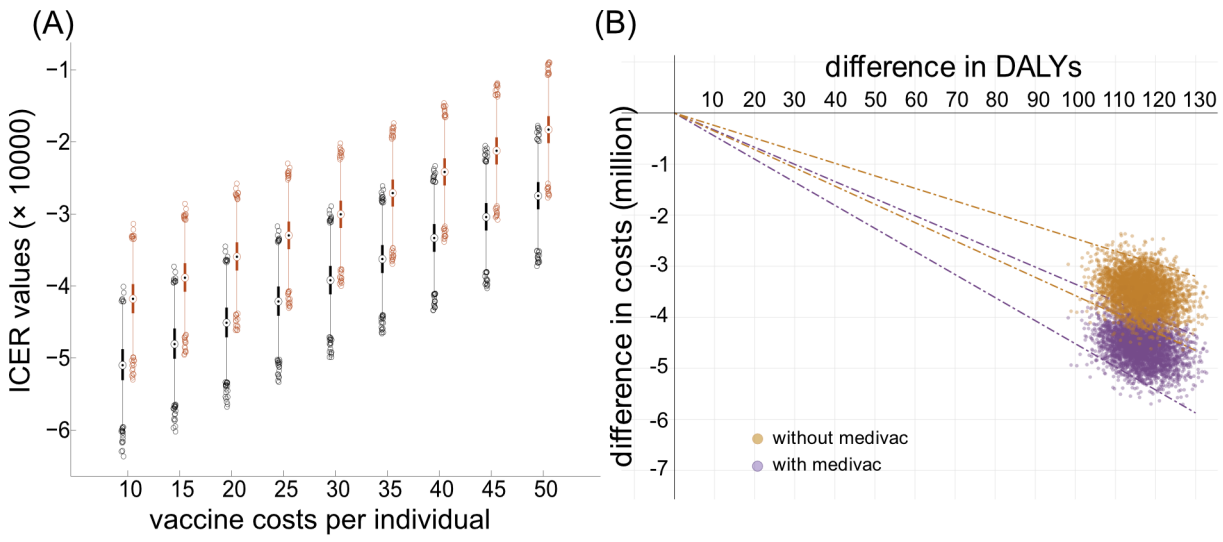


Figure 4.10: **Scenario II. (A)** Boxplots for ICER values with the inclusion (black) and exclusion (red) of Medivac costs, as functions of vaccine costs per individual. Each box contains 50% of data points for ICER values between the first and third quartiles of the bootstrap sampling distribution. Whiskers represent the remaining 50% data points. **(B)** Cost-effectiveness plane using average costs (y-axis) and average DALYs averted (x-axis) calculated on daily bases for two scenarios in the presence and absence of MediVac program. The dashed lines indicate 95% confidence intervals for the ICER values. The ICER values, clustered in the southeast quadrant of the cost-effectiveness plane, indicate that the vaccination program is very cost-effective.

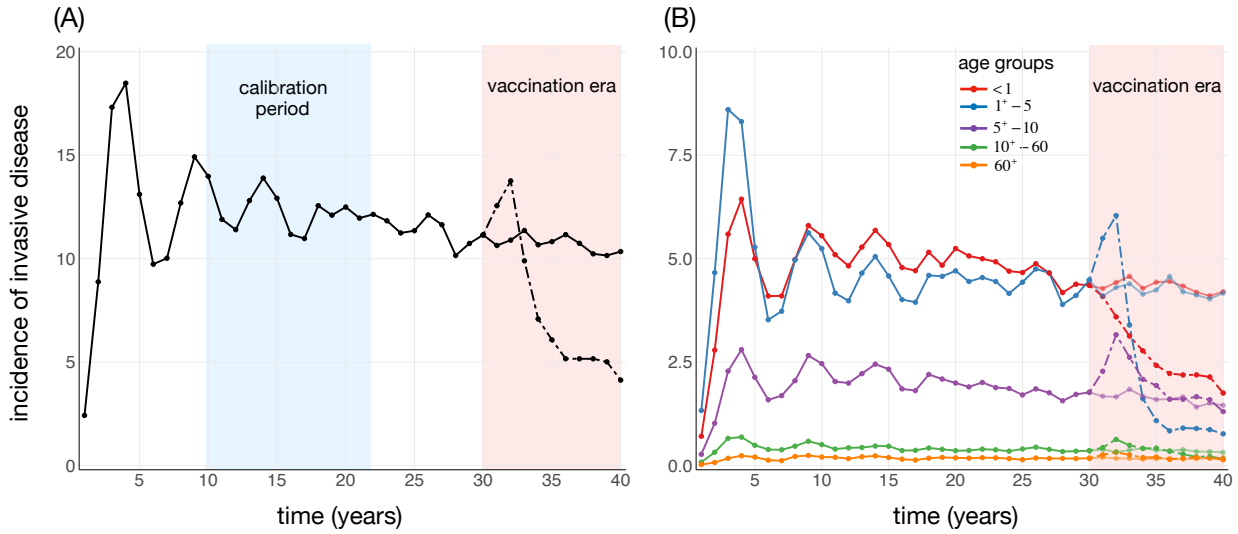


Figure 4.11: **Scenario III.** Overall (A) and age-specific (B) incidence rates over a 40-year simulation time period. Years 10 to 22 correspond to a calibration period fitting the model to Nunavut incidence rates between 2000 and 2012. At year 30, two alternative scenarios were run in the absence of vaccination (solid curves), and with routine infant vaccination schedules (dashed curves).

in lifetime was modelled by a uniform distribution with minimum of 2 years and a maximum of 10 years. The results for vaccination coverages of 77% plus a 93% booster coverage are shown in

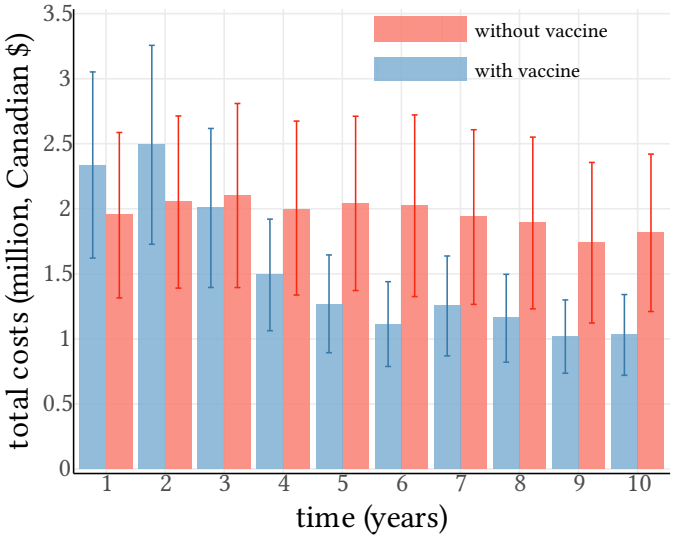


Figure 4.12: **Scenario III.** Overall average annual costs without vaccination (red) and with vaccination (blue) over a 10-year period. Average costs and associated 95% confidence intervals were computed by performing nonparametric bootstrapping method over 500 independent simulations. Direct costs included physician visits, hospitalization, MediVac, major and minor disability, and vaccination costs of doses, administration, and wastage. All costs are in 2017 Canadian dollars.

Figure 4.11, Figure 4.12, Figure 4.13. Cost-effectiveness results of these scenarios are shown in Figure 4.14. The results for vaccination coverages of 90% plus a 93% booster coverage are shown in Figure 4.15, Figure 4.16, Figure 4.17. Cost-effectiveness results of these scenarios are shown in Figure 4.18.

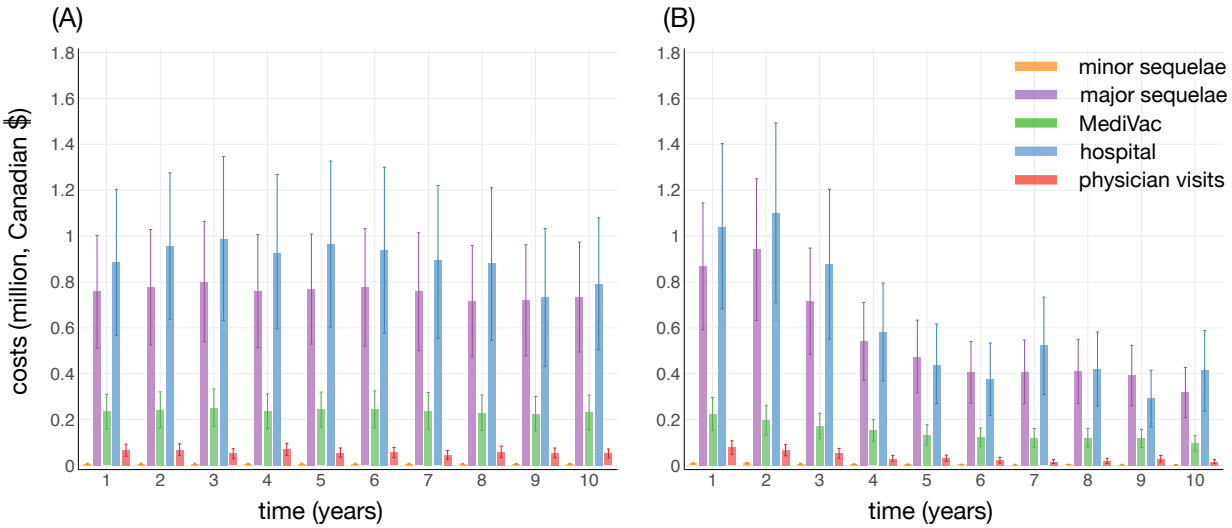


Figure 4.13: **Scenario III.** Distribution of different cost categories with their 95% confidence intervals, obtained using nonparametric bootstrap over 500 independent simulations without vaccination (A), and with routine infant vaccination schedule (B). All costs are in 2017 Canadian dollars.

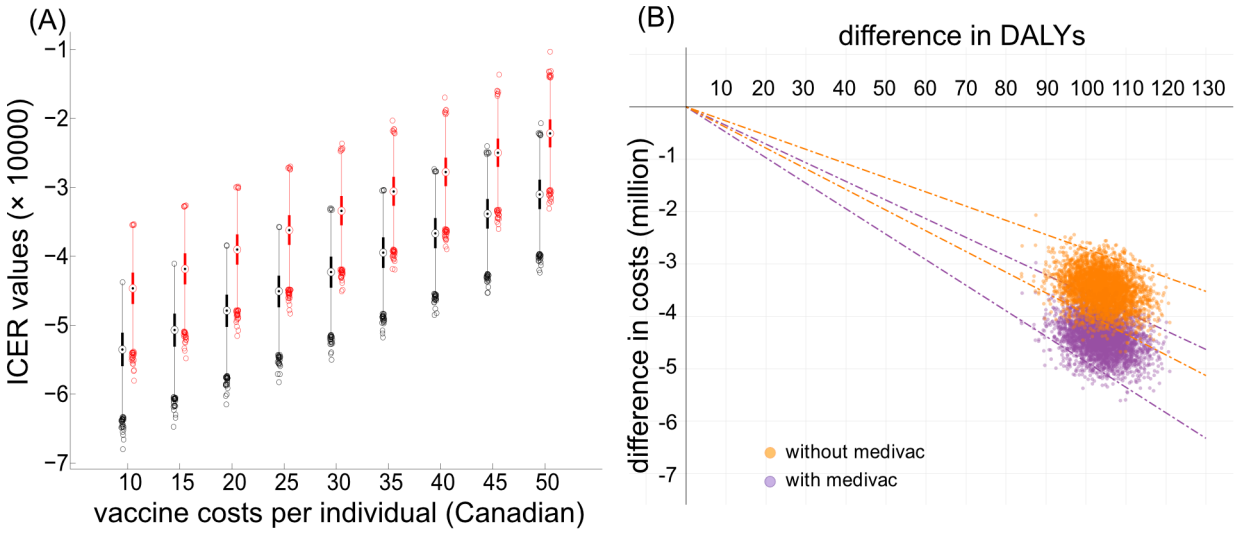


Figure 4.14: **Scenario III.(A)** Boxplots for ICER values with the inclusion (black) and exclusion (red) of Medivac costs, as functions of vaccine costs per individual. Each box contains 50% of data points for ICER values between the first and third quartiles of the bootstrap sampling distribution. Whiskers represent the remaining 50% data points. **(B)** Cost-effectiveness plane using average costs (y-axis) and average DALYs averted (x-axis) calculated on daily bases for two scenarios in the presence and absence of MediVac program. The dashed lines indicate 95% confidence intervals for the ICER values. The ICER values, clustered in the southeast quadrant of the cost-effectiveness plane, indicate that the vaccination program is very cost-effective.

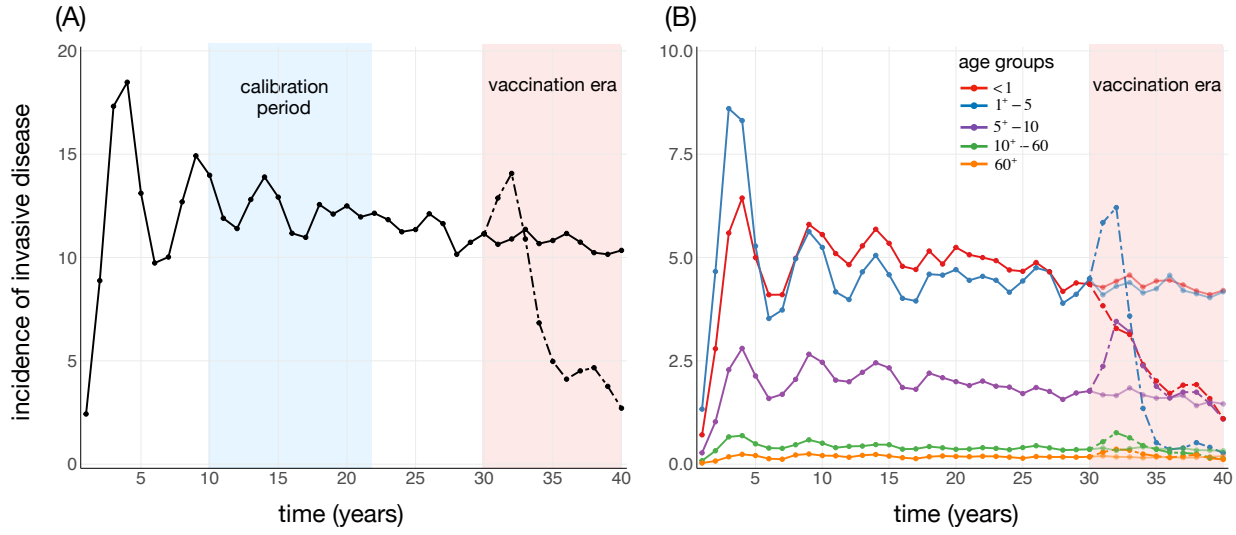


Figure 4.15: **Scenario IV.** Overall (A) and age-specific (B) incidence rates over a 40-year simulation time period. Years 10 to 22 correspond to a calibration period fitting the model to Nunavut incidence rates between 2000 and 2012. At year 30, two alternative scenarios were run in the absence of vaccination (solid curves), and with routine infant vaccination schedules (dashed curves).

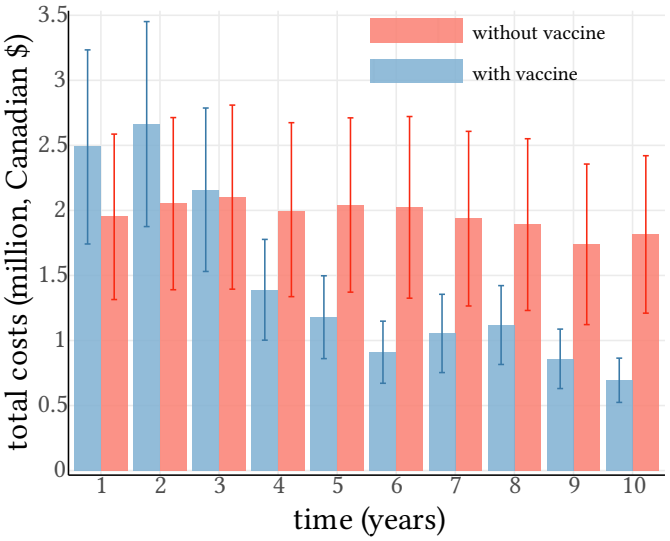


Figure 4.16: **Scenario IV.** Overall average annual costs without vaccination (red) and with vaccination (blue) over a 10-year period. Average costs and associated 95% confidence intervals were computed by performing nonparametric bootstrapping method over 500 independent simulations. Direct costs included physician visits, hospitalization, MediVac, major and minor disability, and vaccination costs of doses, administration, and wastage. All costs are in 2017 Canadian dollars.

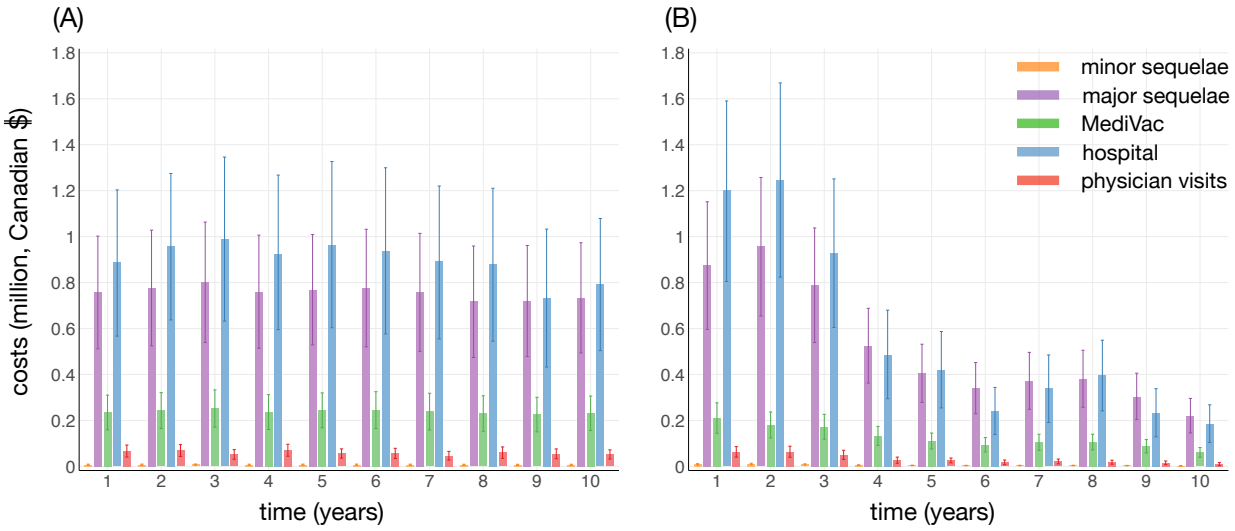


Figure 4.17: **Scenario IV.** Distribution of different cost categories with their 95% confidence intervals, obtained using nonparametric bootstrap over 500 independent simulations without vaccination (A), and with routine infant vaccination schedule (B). All costs are in 2017 Canadian dollars.

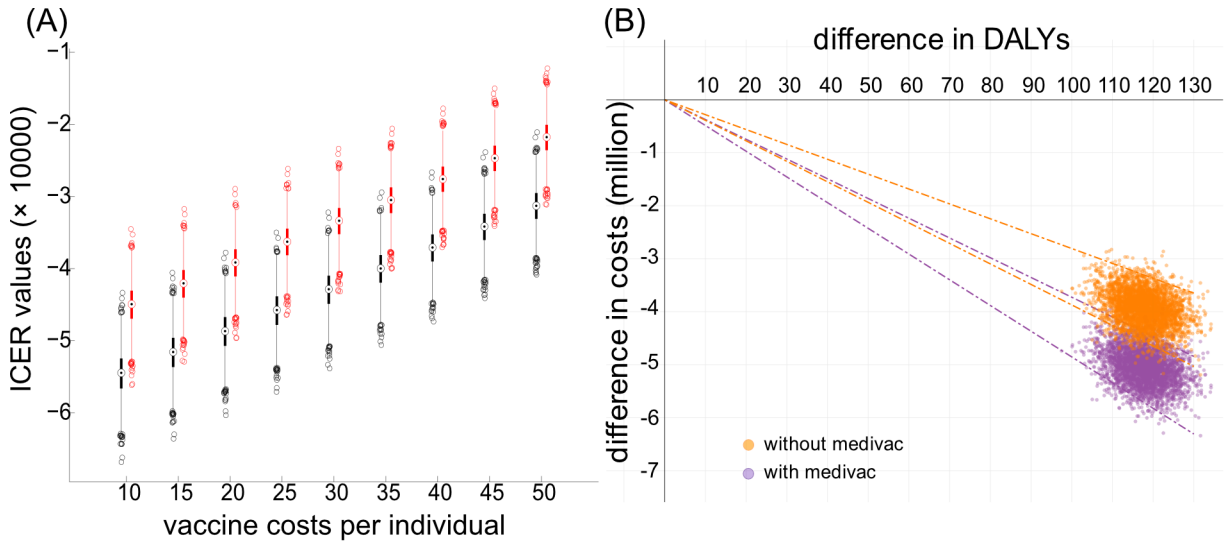


Figure 4.18: **Scenario IV. (A)** Boxplots for ICER values with the inclusion (black) and exclusion (red) of Medivac costs, as functions of vaccine costs per individual. Each box contains 50% of data points for ICER values between the first and third quartiles of the bootstrap sampling distribution. Whiskers represent the remaining 50% data points. **(B)** Cost-effectiveness plane using average costs (y-axis) and average DALYs averted (x-axis) calculated on daily bases for two scenarios in the presence and absence of MediVac program. The dashed lines indicate 95% confidence intervals for the ICER values. The ICER values, clustered in the southeast quadrant of the cost-effectiveness plane, indicate that the vaccination program is very cost-effective.

4.4 Discussion

In this chapter, we sought to investigate the cost-effectiveness analysis of a potential Hia vaccine candidate. Our motivation stems from a number of studies which have reported the emergence and increasing rates of invasive Hia in the northern regions of Canada and Alaska [109, 110, 111, 112, 113, 114, 115]. It is now established that Hia is a significant health burden in northern Canada, particularly among aboriginal communities in Nunavut, with over 90% of cases being less than two years of age. Surveillance of IMPACT (Canada's Immunization Monitoring Program ACTive) study between 2007 and 2015 reports a number of invasive Hia cases in major urban centres [121], indicating that the disease is not just a risk to indigenous or northern populations. Immunization with Hib conjugate vaccine confers no cross-protection against Hia, and therefore timely treatment is essential to alleviate disease outcomes [99, 123]. The incidence rates of invasive Hia disease now underscore the urgent need to prevent primary infection, especially among pediatric population. The experience with Hib vaccines demonstrate that even highly vulnerable populations can be successfully protected using immunization with protein-polysaccharide conjugated vaccines.

To this end, we developed an ABM system, which implements a population-based probabilistic approach, taking into account the herd immunity effects, in addition to encapsulating heterogeneities in mixing patterns and transmission rates. We used the simulation results at the individual level to calculate disability adjusted life years and calculate ICER values to quantify the benefits of a vaccination program. We employed a nonparametric bootstrap method to infer statistical properties of ICER measuring the sensitivity of the model outcomes to parameter variations, which indicates the robustness of our results.

Using the cost estimates associated with treatment, vaccine administration, and long-term sequelae of Hia, our results show that the introduction of a 3-dose primary series plus a booster dose vaccination program is (dominantly) cost-effective. The total costs of such immunization program is significantly less than the costs required to provide life-time care of severely debilitated survivors of invasive Hia. Our analysis suggests an overall 53.4% reduction in costs by the tenth year of the vaccination program, with significant decreases across all cost categories, including immediate hospitalization and long-term disability. We assumed incremental increase in vaccine efficacy with 50% (first dose), 80% (second dose), and 85% (third dose) protection against infection in primary series, and 85%-95% protection following a booster dose. These protection levels are conservative compared to estimates reported in previous studies, indicating that vaccine efficacy against invasive Hib disease after one, two or three doses of vaccine was 59%, 92% and 93%, respectively [133].

The immunization coverage for Hia vaccine in our model is based on the estimates of Hib vaccine coverage of 77% in Canada [135]. This coverage is even lower in northern populations with an estimate of 68% in Nunavut [135]. These relatively low rates may be explained by the rising prevalence of vaccine hesitancy in Canada [148]. The perceived (but untrue) lack of safety in the Diphtheria, Tetanus, Pertussis, Hepatitis B, Polio, and *Haemophilus influenzae* type b (DTaP-HB-IPV-Hib) combination vaccine might be a contributing factor to reduced vaccination coverage [149] There is also evidence of reduced immunogenicity of Hib in the DTaP-HB-IPV-Hib combination vaccine compared to the monovalent Hib vaccine [150, 151, 152]. Thus, a possible solution to increasing coverage and conserve immunogenicity would be to decouple the Hib vaccine from other antigens and offer a conjugate bivalent Hib/Hia vaccine with an appropriate composition of carrier proteins [120, 153]. In our study, increasing Hia vaccine coverage to 90% did not alter our conclusion, and the vaccine remains dominantly cost-effective.

Despite the strengths of the ABM approach for cost-effectiveness analysis, this particular study

has several limitations which the model did not account for, but warrant further investigation. Currently, there is a clear lack of literature detailing the risks and financial burdens of long-term sequelae caused by Hia invasive disease. Furthermore, epidemiological and clinical databases are mainly established for invasive cases, and therefore rates of non-invasive symptomatic or carriage remain undetermined. For missing information on Hia, we parameterized our model with available estimates pertinent to Hib infection. There is also evidence that young children, especially infants may experience repeated invasive episodes after the initial one [154]. Our model does not include recurrence of invasive disease, but considers the possibility of recurrent symptomatic infection or carriage. We expect that repeated invasive disease would argue in favor of an Hia vaccine being cost-effective. Our analysis was carried out in the context of the Canadian healthcare system with publicly funded immunization programs. We conducted this analysis from a governmental perspective, and not a societal perspective; yet we understand that invasive disease and its outcomes can lead to significant socioeconomic burden (e.g., loss of productivity). We also did not consider the costs associated with research and development of a vaccine candidate prior to its availability and use in immunization programs. Moreover, our model did not include possible costs associated with potential adverse side-effects of vaccination. However, based on low adverse rates of Hib and other conjugate vaccines [155], we do not expect the inclusion of these costs in the model to change the conclusion of our cost-effectiveness analysis. Despite these limitations, our results highlight the importance of vaccination against Hia, and indicate that a routine infant immunization program will be highly cost-effective. While informing decision-making on vaccination policies, this study provides a modelling framework for future efforts in vaccine cost-effectiveness analysis. Given our results [50], we believe research and development of an Hia vaccine candidate is an important public health investment.

Chapter 5

Case Study 2 (Part I): Dynamics of Zika Infection

In this chapter, we detail the construction of a multi-agent model in which two distinct types of agents interact to generate the overall dynamics of the system. We present a case study for Zika virus (ZIKV) infection (vector-borne disease), which employs populations of humans and mosquitoes as agents in the chain of disease transmission. The model is utilized to uncover important characteristics of ZIKV epidemics and transmission dynamics.

5.1 Background

ZIKV, an arbovirus from *Flaviviridae* family, is a mosquito-borne virus that is phylogenetically similar to other important mosquito-borne flaviviruses such as West Nile, dengue, and yellow fever viruses [156, 157, 158]. In most cases, the infection presents no symptoms or only mild symptoms, including mild fever, rash, arthralgia, arthritis, myalgia, headache, conjunctivitis, and edema [156, 157]. However, prenatal ZIKV infection has been linked to adverse pregnancy and birth outcomes, most notably microcephaly and other serious neurological disorders [159, 160, 161, 162]. ZIKV is transmitted to humans primarily through the bites of infectious mosquitoes in the subgenus *Stegomyia*, particularly *Aedes aegypti* [163]. However, a number of cases have been reported as a result of sexual contacts [164, 165, 166, 167] and blood transfusion [168] which highlights the potential significance of human-to-human transmission, especially when clinical symptoms of ZIKV infection are absent.

In 2013-2014, the largest documented outbreak of ZIKV occurred in French Polynesia with an approximated 11% of the population seeking medical attention for ZIKV related complications [158]. Following this initial outbreak, ZIKV spread to 69 countries and territories worldwide [158, 169], causing the World Health Organization (WHO) to declare a *public health emergency*

of international concern [170]. Ideal climate conditions and the lack of countermeasures such as vaccination or treatment intensified the ZIKV outbreaks, especially in regions such as Latin America where the primary transmitting vector (i.e., *Aedes aegypti* mosquito) is endemic [171]. The outbreaks spread to more northern latitudes, including several southern parts of the United States [172, 173] and Canada [174]. Although ZIKV outbreaks have diminished in affected countries and the WHO has ended its declaration, the risk of future outbreaks cannot be discounted. Sporadic cases of ZIKV infection have occurred [175] and the threat of large outbreaks continues to exist in the absence of countermeasures such as vaccination or prophylactic drugs. Although vector-control programs can mitigate the impact of disease, ZIKV still remains an important public health concern due to its potential to cause severe outcomes and long-term sequelae, especially in the absence of estimates for the levels of herd immunity generated during the 2015-2016 outbreaks and unknown transmissibility of asymptomatic compared to symptomatic ZIKV infection.

A significant portion (up to 80%) of ZIKV infection is estimated to be asymptomatic without presenting any clinical symptoms of illness [162, 176]. However, the extent to which asymptomatic infection contributes to the overall disease incidence and its impact on the size of outbreaks has not been quantified, which introduces substantial uncertainty into modeling studies of ZIKV transmission dynamics and control interventions. For instance, this quantification is required in understanding the levels of herd immunity in the population, which can prevent large-scale outbreaks if it is sufficiently high.

To understand the effect of asymptomatic transmission to the overall ZIKV dynamics, a comprehensive multi-agent ABM was constructed which encapsulates age-dependent individual attributes, population heterogeneities, and the effect of herd immunity to simulates disease spread in humans through vector (i.e., mosquitoes) and sexual encounters. In particular, using a scaled-down population with demographic characteristics resembling those of Colombia, one of the most Zika-affected countries in South America, we generated simulations of the daily incidence of ZIKV infection over a 2-year period. We also investigated the likelihood of observing a second wave of ZIKV infection, estimated the cumulative attack rates for difference levels of the relative transmissibility of asymptomatic infection (compared to symptomatic infection), and calculated the effective reproduction number of ZIKV infection at the end of first wave. We demonstrate that the occurrence of a second wave of ZIKV depends heavily on the relative transmissibility of asymptomatic infection.

5.2 Model Structure and Parameterization

The general structure of our model is a multi-ABM consisting of an *in-silico* population of *human* and *mosquito* populations, characterized by a time-dependent vector of relevant variables and parameters including their infection stage, demographic information, and disease outcomes. The functional structure of the agents as implemented programatically are illustrated in Algorithm 3. Each of the populations resides on a one-dimensional *in-silico* lattice. Due to the lack of individual movement data, our model does not include mobility patterns (which may influence the level of exposure to the vector), but we did consider individual interactions only for the implementation of sexual transmission (to be described later). The number of human and mosquito agents in each population was determined based on a range of estimated basic reproduction numbers $\mathcal{R}_0$ of ZIKV reported in previous studies [177, 178]. Particularly, using daily counts of confirmed ZIKV cases obtained from the Secretary of Health of Antioquia, Colombia during January-April 2016, [177] estimates the basic reproduction number in the range 1.9 to 2.8 with a mean of 2.216. Thus, the abundance of mosquito, considered as the ratio of mosquito population to human population, was varied between 2, 5 and 10 corresponding to scenarios of $\mathcal{R}_0 = 1.9$, $\mathcal{R}_0 = 2.2$ and $\mathcal{R}_0 = 2.8$. Given the short simulation time horizon, we ignored the individual births and deaths in the populations, and therefore the population size remained constant.

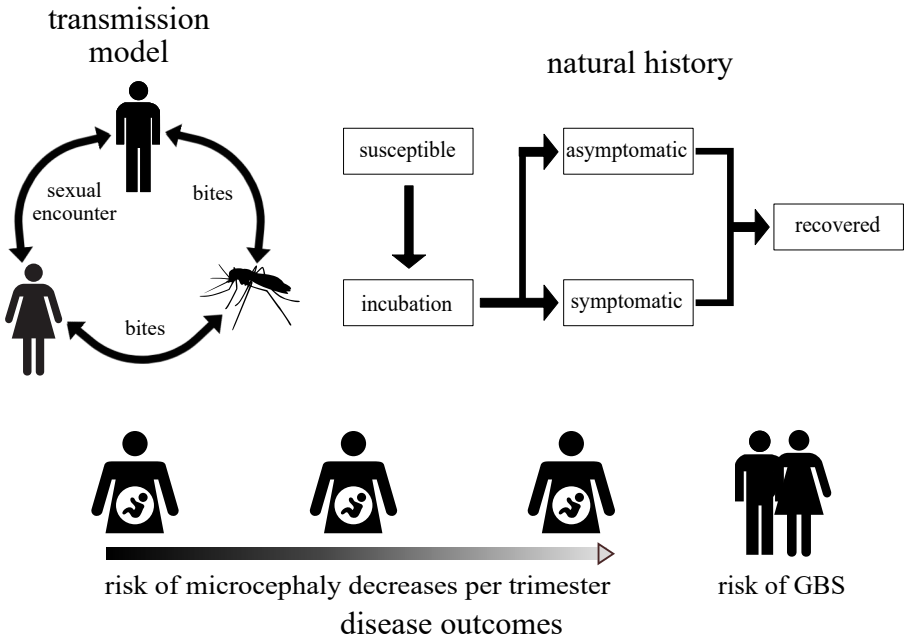


Figure 5.1: Schematic model diagram and natural history ZIKV infection.

Algorithm 3: The dynamic variables and functions that characterize both types of agents (human and mosquito) in the model.

```
1  human agent structure {
2      basic variables {
3          ID ;                                          // ID of the human
4          health ;                                  // infection stage of human
5          gender
6      };
7      model specific variables {
8          partner ;                              // monogamous partner of human
9          sexfrequency ;                            // total frequency of sex
10         sexprobability ;                       // probability of weekly sex
11         cumulativesex
12     };
13     associated functions {
14         func initialize()
15         func interact()
16         func update()
17     };
18 };
19 mosquito agent structure {
20     basic variables {
21         health ;                               // infection stage of mosquito
22         age
23         ageofdeath ;                                          // lifespan
24     };
25     model specific variables {
26         numofbites ;                        // number of bites over lifespan
27         bitedistribution ;                  // how the bites are distributed
28     };
29     associated functions {
30         func initialize()
31         func interact()
32         func update()
33     };
34 } end;
```

The model is initialized by the corresponding functions which set up the one-dimensional lattice environments, apply relevant demographics and sexual interaction rules, and mosquito-human bite interactions. The evolution of time was in discrete time steps, representing a single day of

the real-world system. This pattern was repeated over a 2-year time horizon, with 2000 Monte Carlo iterations for each scenario. All simulations started at day "0" in a high-temperature season. In the next few sections, we describe the general components of our model. All baseline values of the model parameters are based on published estimates and are summarized in Table 5.1. A schematic diagram describing the multi-agent model is illustrated in Figure 5.1.

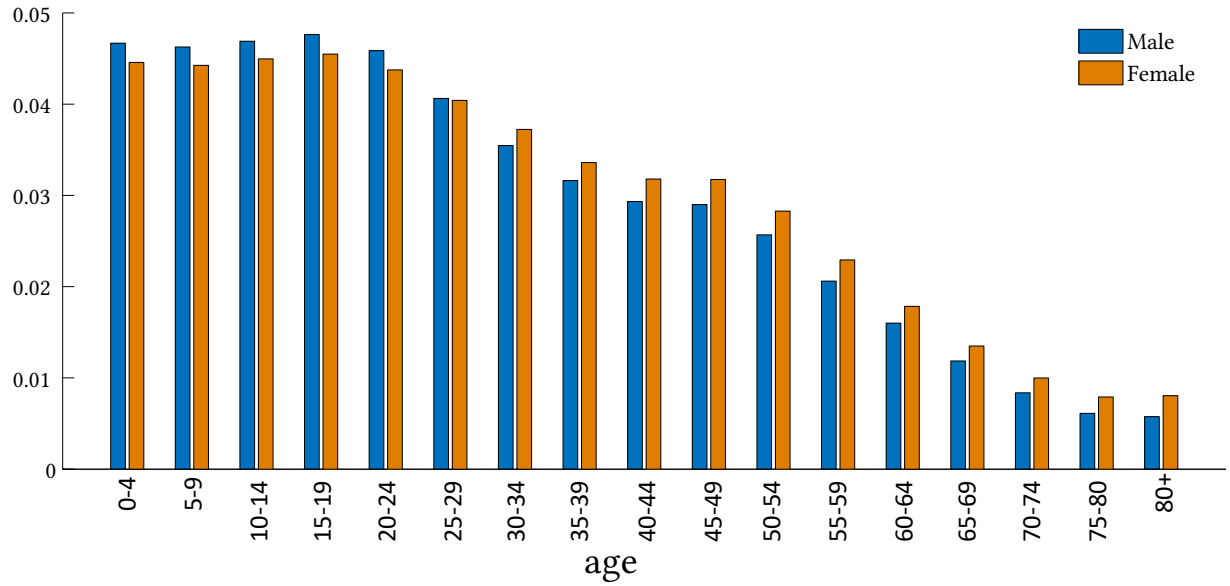


Figure 5.2: Population age and sex distributions of Colombia derived from census data [179].

Human Demographics For the human population, individual age and sex attributes were sampled from relevant distributions of population demographics of Colombia (Figure 5.2). Natural death for the human population was not implemented given the relatively short time-span of the epidemic dynamics. An epidemiological report from the US Center for Disease Control and Prevention indicates that about 1% of Zika cases resulted from sexual contact with travellers to affected areas [180]. Previous studies [52, 181] have omitted this route of transmission for ZIKV infection dynamics due to its low risk [182]. Here, we include the possibility of ZIKV transmission through sexual contact in the model, and consider the range of 1–5% for the risk of transmission to account for its variability. To implement ZIKV sexual transmission dynamics in the model, we considered individuals above age of 15, and created partners in a monogamous context. The frequency of sexual encounters per week for partnered individuals was sampled from their associated distributions corresponding to sex and age of the individuals. Figure 5.3

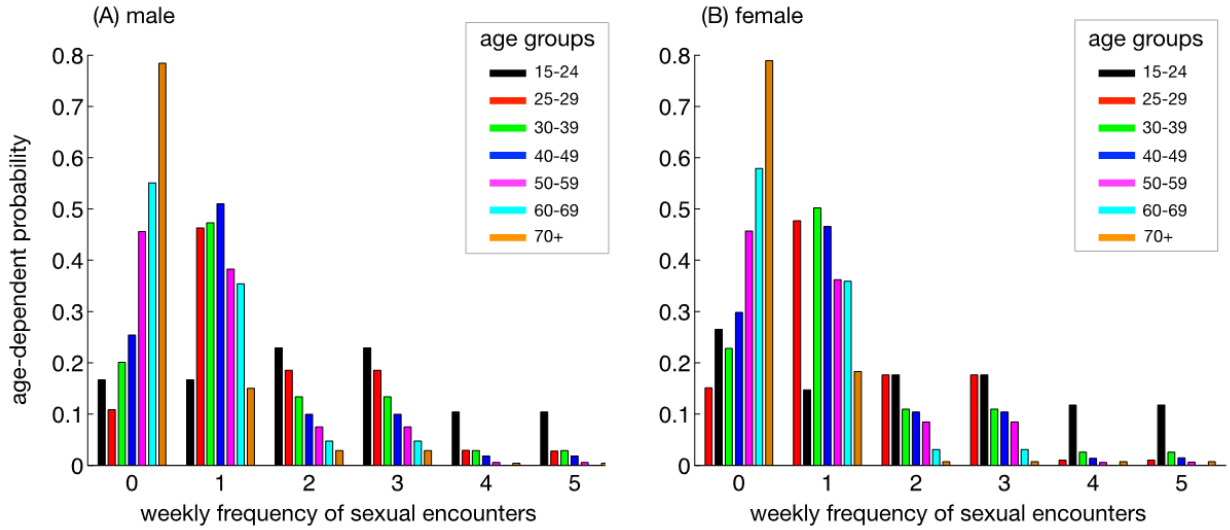


Figure 5.3: Age-dependent probability distributions of weekly frequency of sexual encounters among adult men and women [183, 184].

represents weekly frequency of sexual encounters derived from a national probability sample among adult men and women in the United States [183, 184]. For an individual of age a_i years, the partner was selected with an age in the range $a_i \pm 5$. We assumed the same risk of ZIKV sexual transmission for infectious individuals to their susceptible partners. Demographic related variables were static and did not change during the simulation.

Mosquito Lifespans and Biting Process Due to similarities between ZIKV and dengue infections, being primarily transmitted through the bites of infectious *Aedes aegypti* mosquitoes, we relied on parameter estimates reported in the literature for dengue infection. We assumed that mosquitoes have a lifespan [185] determined by a hazard function given by [185]

$$H(t) = \frac{ae^{bt}}{1 + \frac{as}{b}(e^{bt} - 1)} \tag{5.1}$$

Using Equation (5.1), we generated discretized distributions for sampling lifetime of mosquitoes shown in Figure 5.4. For the season with a high temperature, the lifetime of mosquitoes was sampled with $a = 0.0018$, $b = 0.3228$, and $s = 2.146$, having the mean of 19.6 days [185]. The longevity of mosquitoes for the season with a low temperature was sampled from the distribution generated using $a = 0.0018$, $b = 0.8496$, and $s = 4.2920$, with the mean of 11.2 days. After 180 days of simulations, mosquito lifespans were sampled from the distribution corresponding to the low temperature season. All deaths in the mosquito population were replaced, thus maintaining a

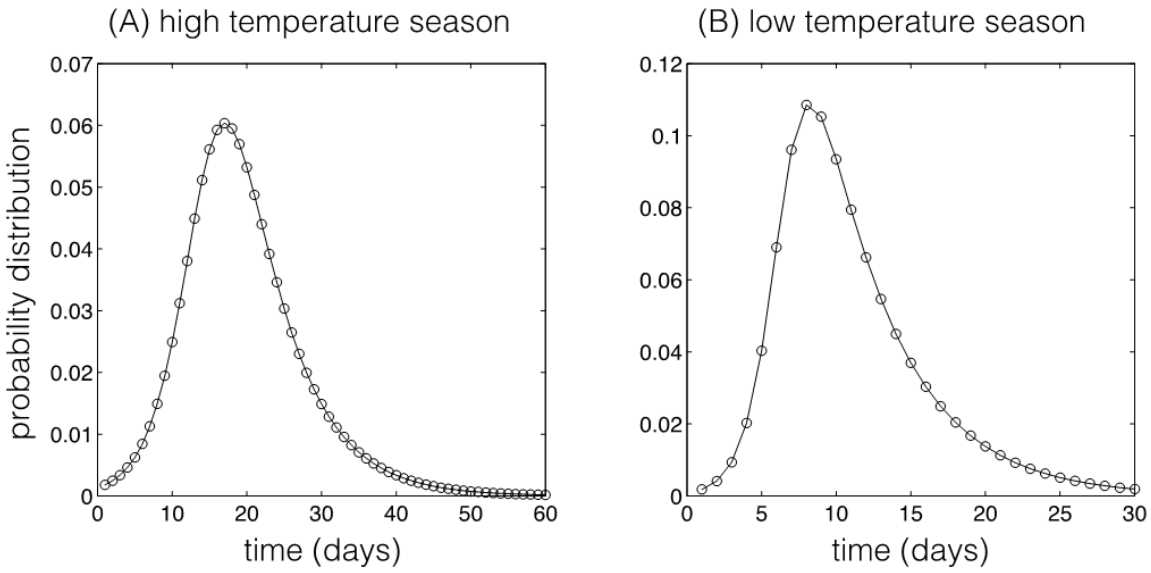


Figure 5.4: Distributions of mosquitos lifespan during seasons with: (A) high temperature (a=0.0018, b=0.3228, s=2.1460) and (B) low temperature (a=0.0018, b=0.8496, s=4.2920).

constant population size.

The mosquito bites were implemented as a Poisson process, with a biting rate of 0.5 per day within the reported range 0.33–1 in previous studies [186, 187, 188]. This corresponds to an average of 1 bite every 2 days. We considered the half-life of a single mosquito as the mean of a Poisson distribution, from which the number of bites was sampled. Bites for each mosquito were randomly distributed over the mosquito lifetime, with a maximum of 1 bite per day. We assumed that a single bite corresponds to a full blood meal for a mosquito and has the potential for disease transmission.

Infection Stages The human population was constructed to encapsulate several epidemiological statuses of susceptible, exposed and incubating, infectious (i.e., symptomatic and asymptomatic), and recovered. The infection stages of vector population included compartments of susceptible, infected and incubating, and infectious. We assumed that recovered individuals are fully protected against ZIKV re-infection (at least for 3 years).

Transmission Dynamics ZIKV transmission from humans to mosquitoes occurred as a result of Bernoulli trials where the chance of success is defined by a transmission probability distribution. This probability was calculated at the time of bite from a susceptible mosquito to an infectious

human (or vice versa) by

$$\mathcal{P}[\mathrm{M} \to \mathrm{H}] = \mathcal{P}[\mathrm{H} \to \mathrm{M}] = 1 - (1 - \beta_{\mathrm{vec}})^N$$

where β_{vec} is the baseline probability of transmission for symptomatic cases and N is the number of bites of a single mosquito to an infectious individual. This baseline probability of transmission was determined by calibration of the model to the given basic reproduction number in the range 1.9 to 2.8 as estimated in [177]. Recent studies suggest that symptomatic cases reached levels of molecular viral load that were significantly higher than asymptomatic cases [189]. Although this level depends on the time of sampling, it may be an indication of lower transmissibility of asymptomatic infection. We therefore considered scenarios in which the reduction factor in the transmissibility of asymptomatic infection compared with symptomatic infection was in the range 0.1 to 0.9. This was implemented as a reduction factor in transmission parameter β_{vec}. Similarly, any potential interventions to blunt transmission such as reducing the number of mosquitoes or use of condoms was also included in the model as a reduction factor in β_{vec}. Transmission values β_{vec} ranged from 0.2851 to 0.3947 (obtained from calibration, §5.2.1) depending on the assumed relative transmissibility of asymptomatic infection compared to symptomatic infection from 0.9 to 0.1.

For sexual transmission of ZIKV, we considered the probability

$$\mathcal{P}[\mathrm{H} \to \mathrm{H}] = 1 - (1 - \beta_{\mathrm{sex}})$$

where β_{sex} is the risk of sexual transmission per encounter. If an infectious individual was at least 15 years old and had a sexual partner, we used Bernoulli trials for each sexual contact where the weekly frequency of contacts with the susceptible partner was sampled from age and sex-dependent distributions with a maximum of one encounter per day. We assumed the risk of sexual transmission was included only during the infectious period. Although this risk may continue for several days or weeks following recovery [190, 191], our assumption is justified due to uncertainty in the duration of sexual transmission at the individual level. The associated pseudocode describing the transmission dynamics is illustrated in Algorithm 4.

Upon successful transfer of ZIKV, individuals experience an intrinsic incubation period (IIP) before becoming infectious [192, 193]. We sampled the IIP for an infected individual from a log-normal distribution with the shape and scale parameters of 1.72 and 0.21, and mean of 5.7 days (95% CI: 4–8), as estimated for dengue infection [192]. In the infectious stage, individuals are classified as

Algorithm 4: Pseudocode implementation of ZIKV transmission dynamics.

```
   input  : β_vec = calibrated transmission probability
   input  : H = humans with an assigned partner
 1 for each mosquito m do
 2     if m will bite then
 3         h = random(non-isolated humans) ;
 4         if h.health = SUSC and m.health = INF then
 5             r = rand()
 6             r ≤ β ⟹ h.health = INF
 7         end
 8         if h.health = SYMP and m.health = SUSC then
 9             r = rand()
10             r ≤ β ⟹ h.health = INF
11         end
12         if h.health = ASYMP and m.health = SUSC then
13             r = rand()
14             α = reduction factor
15             r ≤ α · β ⟹ h.health = INF
16         end
17     end
18 end
19 for each human h ∈ H do
20     p = partner ;
21     if h.health = INF and p.health = SUSC then
22         β_sex = calculate probability(h)
23         if P = true then
24             β_eff = h.sexprobability · (1 − p.condom reduction)
25         end
26     end
27 end
```

either symptomatic or asymptomatic, with (40–80%, sampled uniformly) of infected individuals experiencing asymptomatic infection without developing clinical symptoms [162, 176, 193]. The infectious period was sampled from a log-normal distribution with the shape and scale parameters of 1.54 and 0.12, and mean of 4.7 days (95% CI: 3.8–5.7) [194]. For a bite through which a mosquito was infected, the extrinsic incubation period (EIP) was sampled from a log-normal distribution with the shape and scale parameters of 2.28 and 0.21, and mean of 10 days (95% CI: 7–14) [195]. Once this period elapsed, the mosquito became infectious for its remaining lifespan.

Table 5.1: Zika model parameters values and their associated ranges. Transmissibility values ranged depending on the assumed relative transmissibility of asymptomatic infection compared to symptomatic infection from 0.1 to 0.9. Risk of infection through sexual encounter was assumed to be low [182].

Parameter Description	**Baseline value (range)**	**Source**
Transmission rates for infection		
Human to mosquito Mosquito to human	0.2851 to 0.3947 based on relative transmission	calibration process
Relative transmissibility of asymptomatic infection	0.1 to 0.9	assumed
Human infection parameters		
Intrinsic incubation period	Lognormal($\mu = 1.72, \sigma = 0.2$)	[178, 193]
Infectious period	Lognormal($\mu = 1.54, \sigma = 0.12$)	[196]
Risk of infection through sexual encounter	1% - 5%	assumed
Fraction of infected cases experiencing asymptomatic infection	40% - 80%	[162, 176]
Mosquito lifespan and infection parameters		
Lifespan (high temperature)	mean: 19.6 days	[185]
Lifespan (low temperature)	mean: 11.2 days	[185]
Extrinsic incubation period	Lognormal($\mu = 2.28, \sigma = 0.21$)	[195]
Number of mosquito bites	Poisson($\lambda =$ mosquito half-life)	[187, 188]

5.2.1 Calibration

The baseline transmission probability β_{vec} was determined by calibrating the model to estimated basic reproduction numbers of Antioquia, Colombia. These estimates were in the range 1.9–2.8 with the mean of 2.2 [177, 178], determined using daily counts of confirmed ZIKV cases. Accordingly, we calibrated for three scenarios of basic reproduction numbers $\mathcal{R}_0 = 1.9$, $\mathcal{R}_0 = 2.2$ and $\mathcal{R}_0 = 2.8$. For each $\mathcal{R}_0$ scenario, we considered a reduction factor in the transmissibility of asymptomatic infection compared with symptomatic infection in the range 0.1 to 0.9.

Calibration simulations were started by seeding the simulation with 1 infected human in the IIP stage and a fully susceptible mosquito population. To obtain the average $\mathcal{R}_0$, each simulation

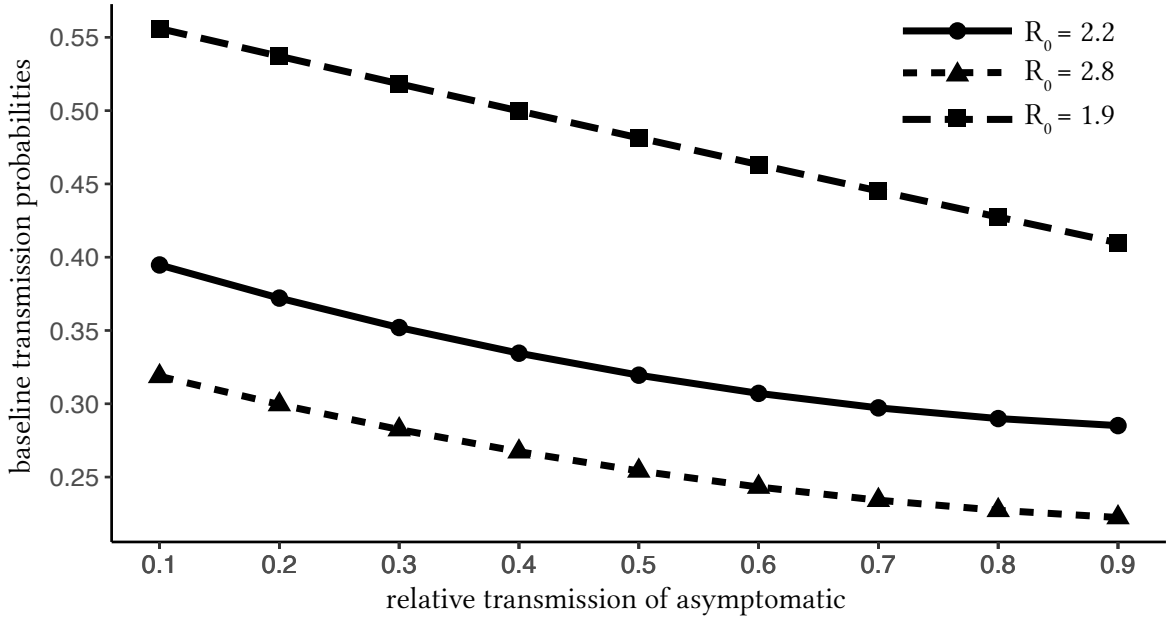


Figure 5.5: Calibrated probabilities of ZIKV transmission based on the reproduction number and ratio of mosquito to human populations.

was run until this initial infected person recovered. The basic reproduction number $\mathcal{R}_0$ was then calculated by counting the average number of secondary infections over 1000 Monte Carlo simulations, generated by the initial case either through mosquito bites or sexual transmission. This was repeated while sweeping over a range of transmission values and relative transmissions of asymptomatic infection. A linear regression was applied to the resulting data in order to determine β_{vec} values, illustrated in Figure 5.5.

5.3 Results: The Role of Asymptomatic Infection

Disease outcomes and cumulative daily incidence throughout each simulation were recorded and averaged for each scenario to estimate attack rates during an outbreak. We defined an outbreak if the cumulative incidence of infection during the third disease generation was greater than $\mathcal{R}_0$, which implies a sustained transmission within the first two disease generation intervals with $\mathcal{R}_0$ exceeding 1. Simulations in which there was no more transmission of ZIKV after the third generation interval were excluded in estimating the attack rate. We assumed a gamma distribution for the generation interval with the mean of 14 days and standard deviation of 2 days [197]. The

effective reproduction number R_{eff} at the end of an outbreak was estimated using the formula

$$\mathcal{R}_{\text{eff}} = \left(\frac{\text{number of susceptibles at the end of outbreak}}{\text{number of susceptibles at the start of outbreak}} \right) \mathcal{R}_0$$

In addition, the probability of a second wave was calculated by considering the fraction of simulations that resulted in an outbreak in the second year following the outbreak in the first year.

Simulations were run to obtain the daily case incidence of ZIKV infection for calibrated scenarios, corresponding to basic reproduction numbers $\mathcal{R}_0$ of $1.9, 2.2$ and 2.8. For each scenario, we further considered disease spread when the contribution of symptomatic ZIKV infection to disease transmission through mosquitoes was reduced by 10%, 30% and 50%. This reduction was implemented probabilistically for each mosquito bite when the infectious case was symptomatic.

For $\mathcal{R}_0 = 2.2$, Figure 5.6 shows the daily incidence of symptomatic infection over a two year period for various levels of the relative transmissibility of asymptomatic infection and different reduction levels of Zika transmission from symptomatic cases to mosquitoes. As the contribution of symptomatic infection to disease spread reduced (e.g., due to vector control interventions), the occurrence of a second wave of outbreak required a higher level of the relative transmissibility of asymptomatic infection. The probability of a second wave of infection occurring as a function of the relative transmissibility of asymptomatic infection is illustrated in Figure 5.7.

With 10% reduction of transmission from symptomatic cases, the probability of a second wave occurring increased from 0.19 to 0.48 when the relative transmissibility of asymptomatic infection increased from 10% to 90%. We observed the same increasing trend (with lower probabilities) for higher levels of transmission reduction.

We also estimated the effective reproduction numbers $\mathcal{R}_{\text{eff}}$ and the cumulative attack rates at the end of the first wave of ZIKV outbreak. Figure 5.9 shows boxplots for the range of $\mathcal{R}_{\text{eff}}$ estimates as a function of the relative transmissibility of asymptomatic infection when $\mathcal{R}_0 = 2.2$ at the onset of the outbreak. With a 10% reduction of ZIKV transmission from symptomatic cases, the median R_{eff} was estimated at 2.04 (95% CI: 1.64–2.18) for 10% relative transmissibility of asymptomatic infection. The mean attack rate for this scenario (Figure 5.10) was estimated at 9% (95% CI: 8.4%–9.4%). When the relative transmissibility increases to 90%, the median $\mathcal{R}_{\text{eff}}$ was 1.99 (95% CI: 1.50, 2.18) with an average attack rate of 11% (95% CI: 10.5%–11.3%). As the level

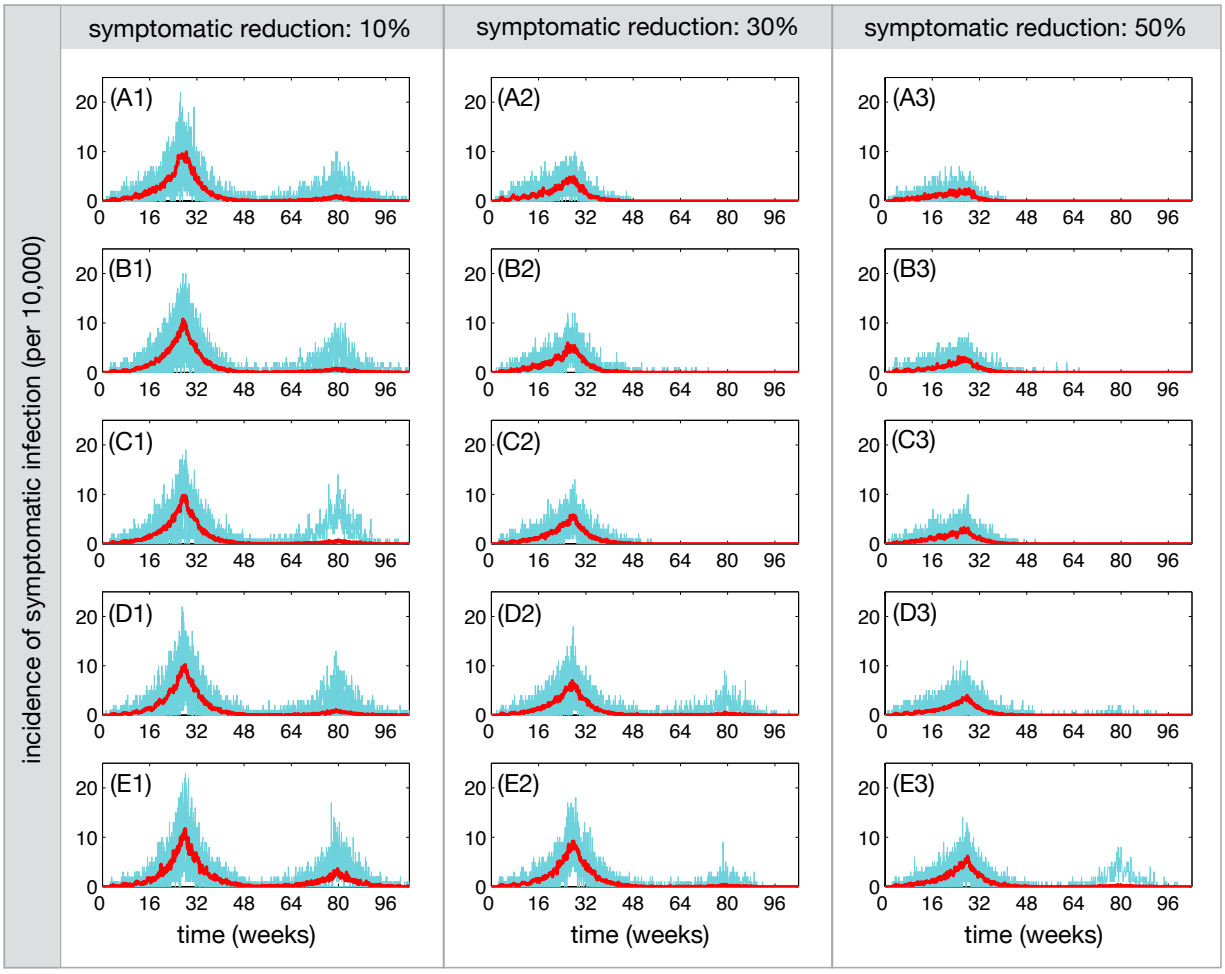


Figure 5.6: Incidence of symptomatic ZIKV infection over a two year period for the first and second waves, when the contribution of symptomatic ZIKV infection to disease transmission through mosquitoes was reduced by 10% (A1–E1), 30% (A2–E2), and 50% (A3–E3). The relative transmissibility of asymptomatic infection is 10% (A1,A2,A3), 30% (B1,B2,B3), 50% (C1,C2,C3), 70% (D1,D2,D3) and 90% (E1,E2,E3). The red curve represents the average of sample realizations for incidence curves. Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 2.2$.

of ZIKV transmission from symptomatic cases decreased, the median $\mathcal{R}_{\text{eff}}$ declined towards $\mathcal{R}_0$ (Figure 5.9), giving lower attack rates for the first wave (Figure 5.10). Overall, we estimated attack rates to range from 2.2% to 11% for the scenarios simulated here with $\mathcal{R}_0 = 2.2$. These estimates are consistent with those reported for Colombia during outbreaks through February 2017 [52].

We also estimated the cumulative number of Zika infection resulted from the virus transmission through sexual encounter. Figure 5.8 shows the range of these estimates for different relative transmissibility of asymptomatic infection in the absence of any control measure. For a low

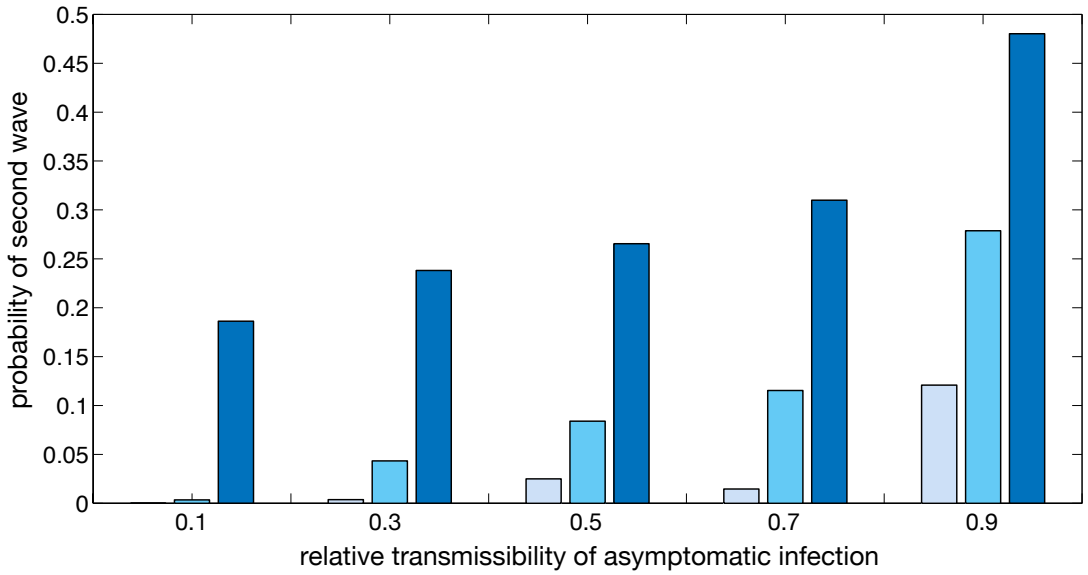


Figure 5.7: The probability of a second wave of ZIKV outbreak occurring as a function of the relative transmissibility of asymptomatic infection. Color bars correspond to scenarios in which infection transmission from symptomatic cases to mosquitoes was reduced by 10% (dark blue), 30% (light blue), and 50% (grey). Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 2.2$.

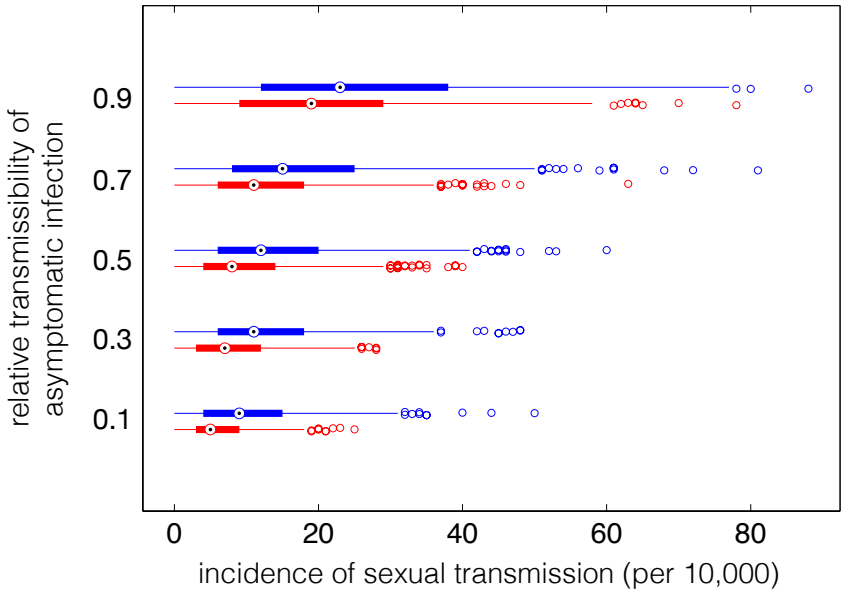


Figure 5.8: Estimated range of cumulative incidence of sexual transmission during the first wave of ZIKV infection as a function of the relative transmissibility of asymptomatic infection, in the absence of condom use (blue) and 50% condom use during symptomatic infection (red). Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 2.2$.

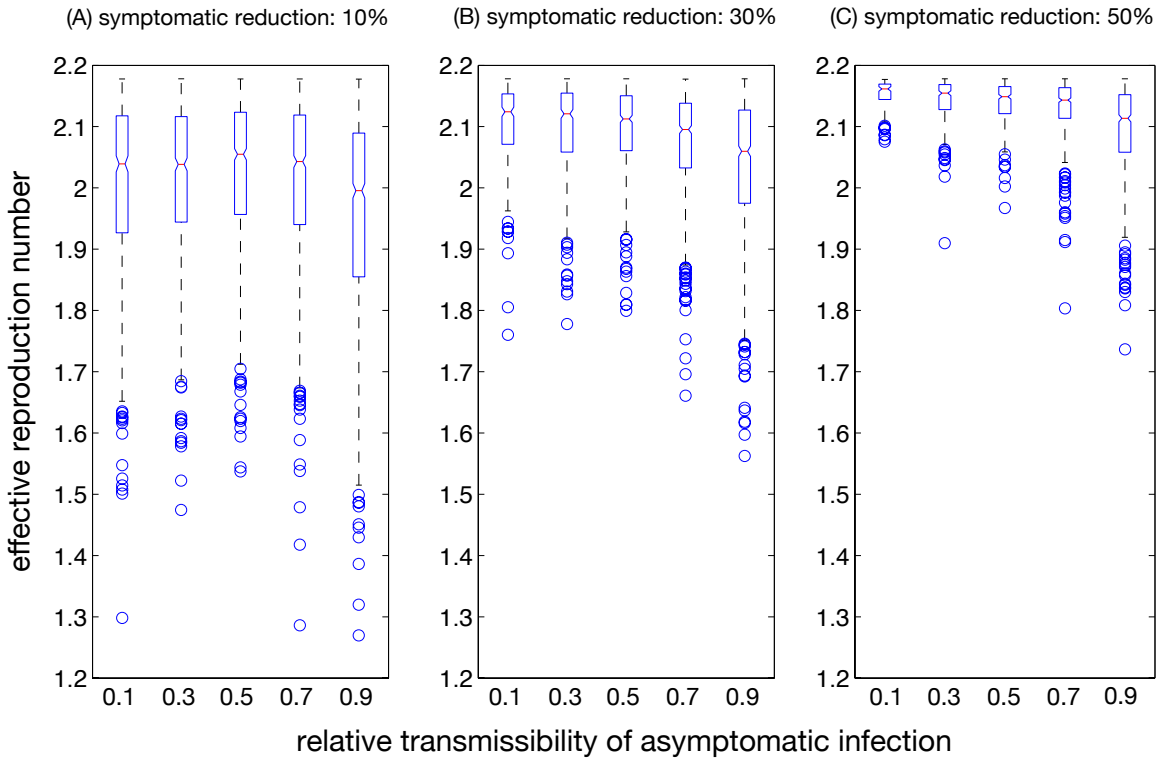


Figure 5.9: Effective reproduction number $\mathcal{R}_{\text{eff}}$ at the end of the first wave as a function of the relative transmissibility of asymptomatic infection. The contribution of symptomatic ZIKV infection to disease transmission through mosquitoes was reduced by 10%, 30%, and 50%. Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 2.2$.

relative transmissibility (10%), the median number of sexual transmission is 10.6 (95% CI: 0, 31.5), which accounts for 1.16% of the cumulative incidence (95% CI: 0, 2.29%). When the relative transmissibility increased to 90%, the median number of sexual transmission increased to 23 (95% CI: 0, 77). This corresponds to 2.4% of the cumulative incidence (95% CI: 1.02, 3.88%). These results suggest that the previous work in a deterministic context [198] may have overestimated the upper bound of the fraction of cases due to sexual transmission.

We observed similar results for scenarios of $\mathcal{R}_0 = 1.9$ and $\mathcal{R}_0 = 2.8$. Incidence of symptomatic ZIKV infection over a two year period are shown in Figure 5.11 and Figure 5.12. The probability of a second wave of ZIKV outbreak and effective reproduction numbers at the end of the first wave are shown in Figure 5.13 and Figure 5.14. Attack rates (cumulative incidence per 10,000) of ZIKV infection over a 2-year period are shown in Figure 5.15 and Figure 5.16. Estimated range of cumulative incidence of sexual transmission during the first wave is shown in Figure 5.17. These results suggest that the relative transmissibility of asymptomatic infection is a key parameter in

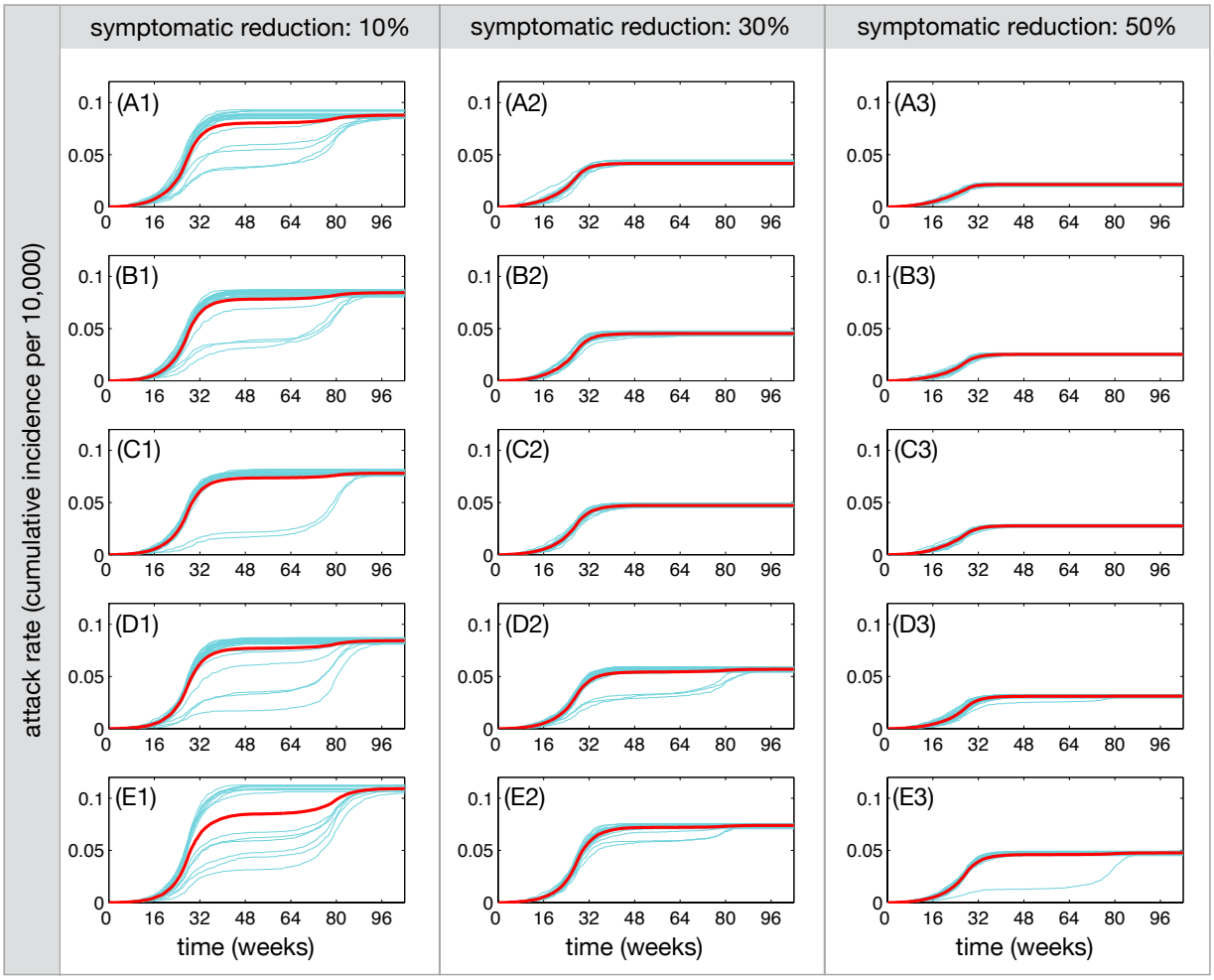


Figure 5.10: Attack rates (cumulative incidence per 10,000) of ZIKV infection over a 2-year period for the first and second waves, when the contribution of symptomatic ZIKV infection to disease transmission through mosquitoes was reduced by 10% (A1–E1), 30% (A2–E2), and 50% (A3–E3). The relative transmissibility of asymptomatic infection is 10% (A1,A2,A3), 30% (B1,B2,B3), 50% (C1,C2,C3), 70% (D1,D2,D3) and 90% (E1,E2,E3). The red curve represents the mean attack rate in each scenario within its 95% confidence interval. Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 2.2$.

estimating the burden of disease through different modes of transmission (i.e., mosquito bites and sexual interactions) and evaluating the probability of a second wave of ZIKV infection.

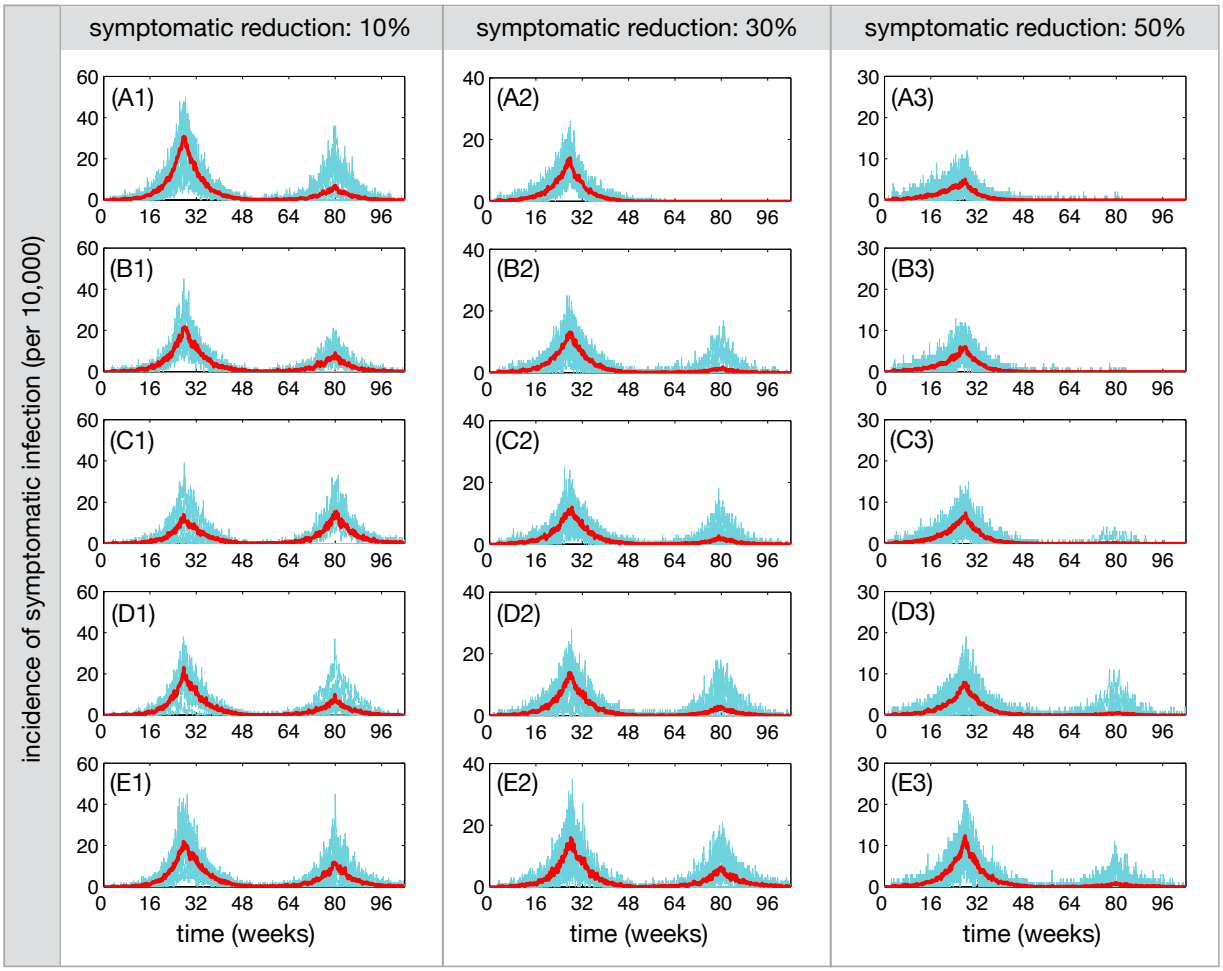


Figure 5.11: Incidence of symptomatic ZIKV infection over a two year period for the first and second waves, when the contribution of symptomatic ZIKV infection to disease transmission through mosquitoes was reduced by 10% (A1–E1), 30% (A2–E2), and 50% (A3–E3). The relative transmissibility of asymptomatic infection is 10% (A1,A2,A3), 30% (B1,B2,B3), 50% (C1,C2,C3), 70% (D1,D2,D3) and 90% (E1,E2,E3). The red curve represents the average of sample realizations for incidence curves. Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 2.8$.

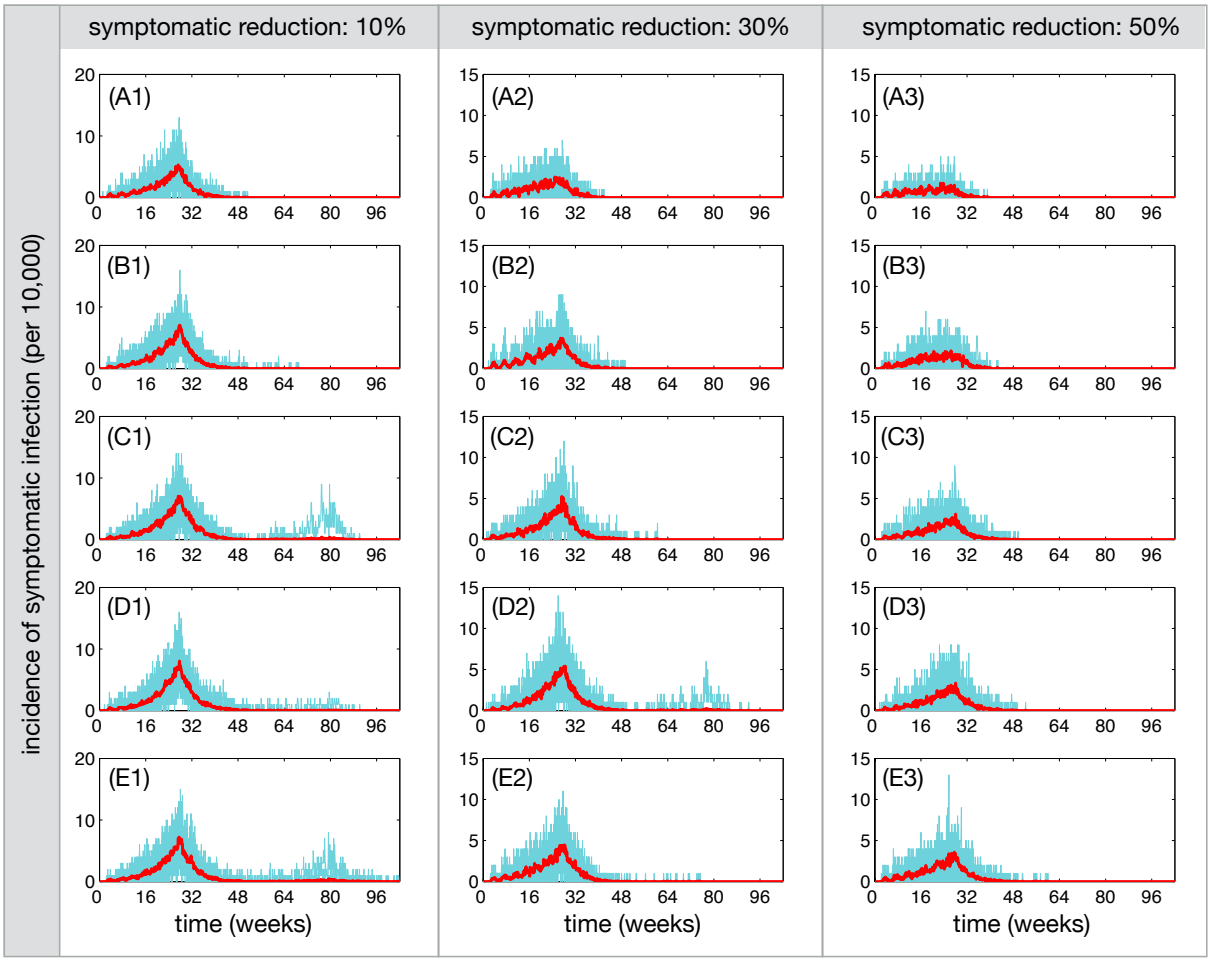


Figure 5.12: Incidence of symptomatic ZIKV infection over a two year period for the first and second waves, when the contribution of symptomatic ZIKV infection to disease transmission through mosquitoes was reduced by 10% (A1–E1), 30% (A2–E2), and 50% (A3–E3). The relative transmissibility of asymptomatic infection is 10% (A1,A2,A3), 30% (B1,B2,B3), 50% (C1,C2,C3), 70% (D1,D2,D3) and 90% (E1,E2,E3). The red curve represents the average of sample realizations for incidence curves. Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 1.9$.

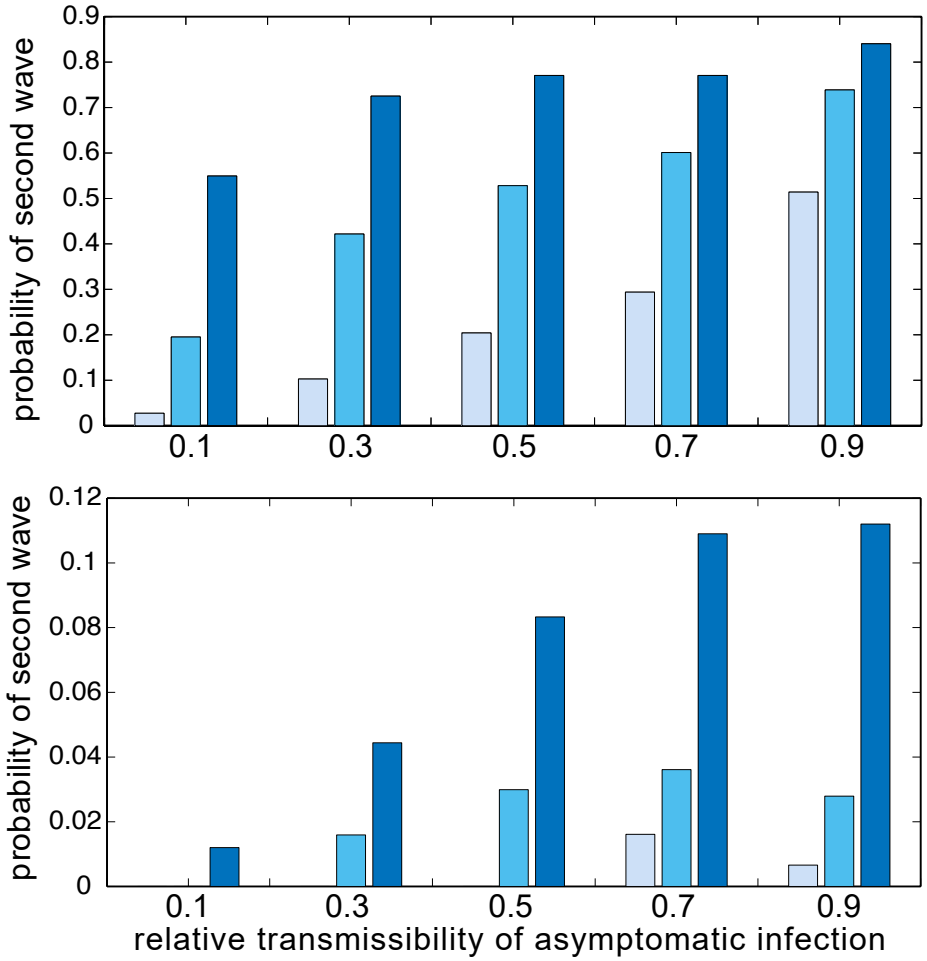


Figure 5.13: The probability of a second wave of ZIKV outbreak occurring as a function of the relative transmissibility of asymptomatic infection. Color bars correspond to scenarios in which infection transmission from symptomatic cases to mosquitoes was reduced by 10% (dark blue), 30% (light blue), and 50% (grey). Top and bottom figures correspond to model scenario calibrated to basic reproduction numbers $\mathcal{R}_0 = 2.8$ and $\mathcal{R}_0 = 1.9$, respectively.

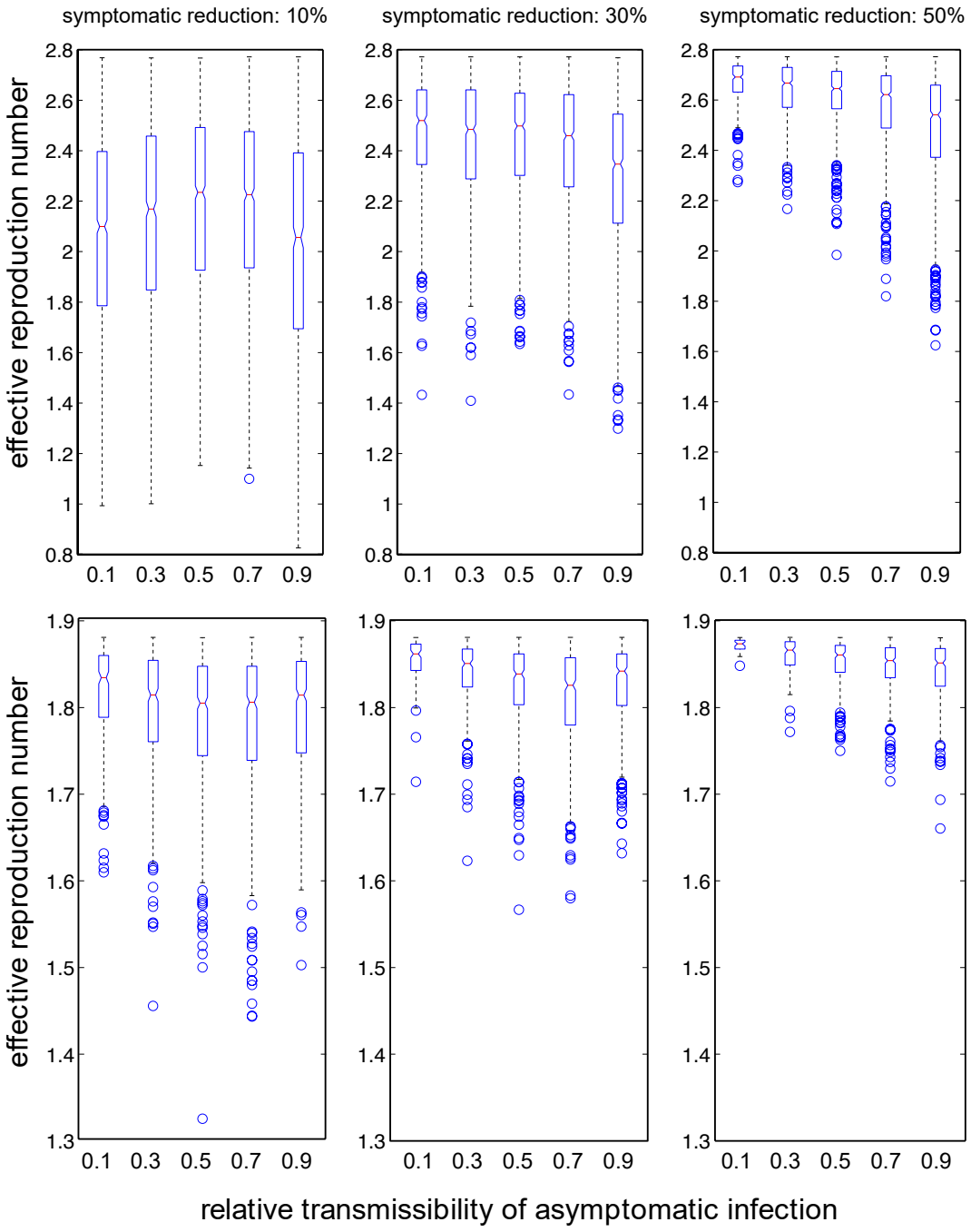


Figure 5.14: Effective reproduction number at the end of the first wave as a function of the relative transmissibility of asymptomatic infection. The contribution of symptomatic ZIKV infection to disease transmission through mosquitoes was reduced by 10%, 30% , and 50%. Top and bottom figures correspond to model scenario calibrated to basic reproduction numbers $\mathcal{R}_0 = 2.8$ and $\mathcal{R}_0 = 1.9$, respectively.

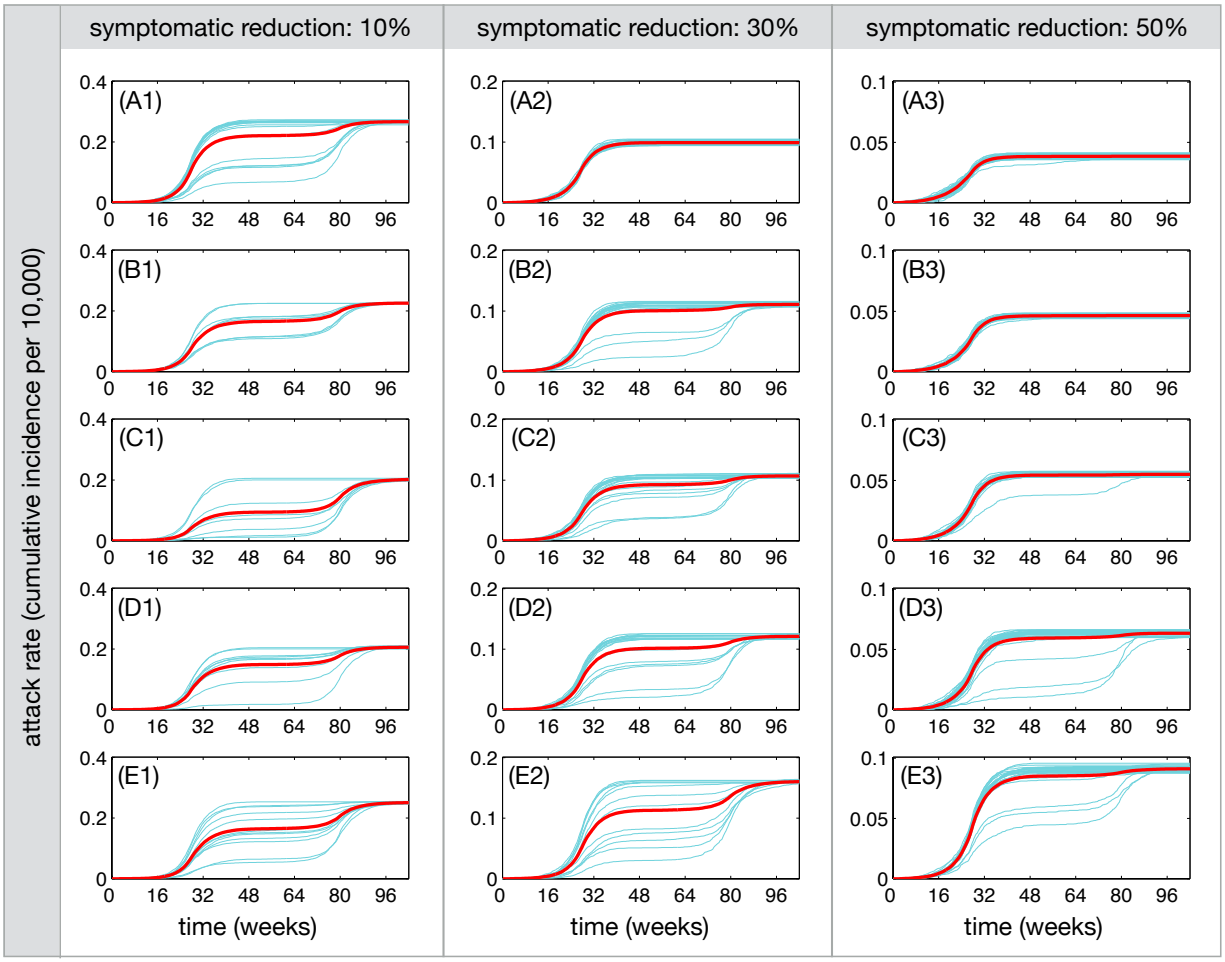


Figure 5.15: Attack rates (cumulative incidence per 10,000) of ZIKV infection over a 2-year period for the first and second waves, when the contribution of symptomatic ZIKV infection to disease transmission through mosquitoes was reduced by 10% (A1–E1), 30% (A2–E2), and 50% (A3–E3). The relative transmissibility of asymptomatic infection is 10% (A1,A2,A3), 30% (B1,B2,B3), 50% (C1,C2,C3), 70% (D1,D2,D3) and 90% (E1,E2,E3). The red curve represents the mean attack rate in each scenario within its 95% confidence interval. Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 2.8$.

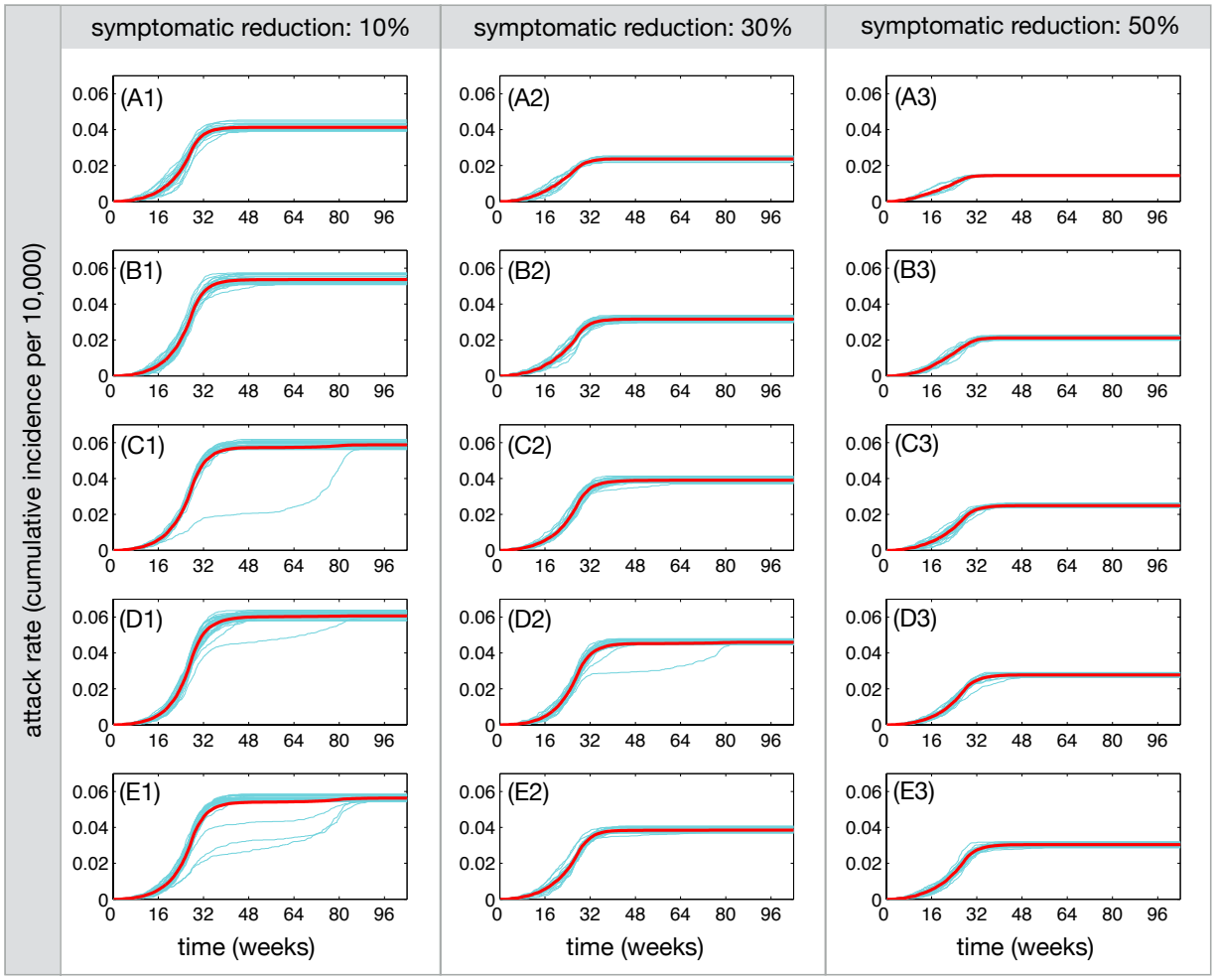


Figure 5.16: Attack rates (cumulative incidence per 10,000) of ZIKV infection over a 2-year period for the first and second waves, when the contribution of symptomatic ZIKV infection to disease transmission through mosquitoes was reduced by 10% (A1–E1), 30% (A2–E2), and 50% (A3–E3). The relative transmissibility of asymptomatic infection is 10% (A1,A2,A3), 30% (B1,B2,B3), 50% (C1,C2,C3), 70% (D1,D2,D3) and 90% (E1,E2,E3). The red curve represents the mean attack rate in each scenario within its 95% confidence interval. Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 1.9$.

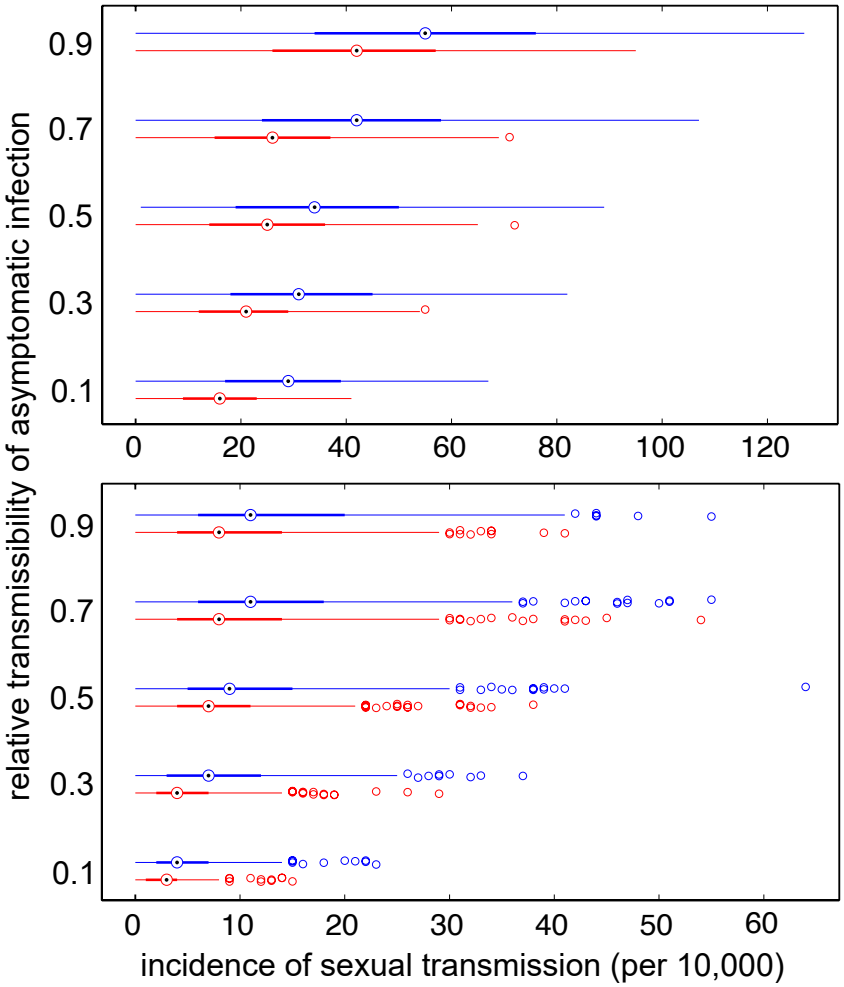


Figure 5.17: Estimated range of cumulative incidence of sexual transmission during the first wave of ZIKV infection as a function of the relative transmissibility of asymptomatic infection, in the absence of condom use (blue) and 50% condom use during symptomatic infection (red). Top and bottom figures correspond to model scenario calibrated to basic reproduction numbers $\mathcal{R}_0 = 2.8$ and $\mathcal{R}_0 = 1.9$, respectively.

5.4 Discussion

The results above show that the relative transmissibility of asymptomatic infection is a key epidemiological parameter that can significantly influence disease dynamics, especially in the context of intervention strategies. We considered scenarios in which the contribution of ZIKV transmission from symptomatic cases is reduced as a result of isolation via hospitalization, self-isolation, or through disease interventions. For instance, interventions to reduce exposure to infectious bites may include mosquito avoidance through full clothing, mosquito repellents, and spraying and larviciding. Similarly, condom use was considered as an effective intervention to

prevent sexual transmission of the Zika virus. In an exploratory analysis, we found that the use of condoms could significantly reduce the risk of sexual transmission (Figure 5.8, red bars); however, this reduction depends on the level of condom use. We observed that when interventions are absent or their effectiveness is very low in blunting the contribution of symptomatic cases to ZIKV transmission, a second wave is more likely to occur as the relative transmissibility of asymptomatic infection increases (Figures 5.6, 5.12, 5.11). Furthermore, the occurrence of a second wave of ZIKV infection requires higher values of the relative transmissibility as the effectiveness of interventions increases.

The relative transmissibility of asymptomatic infection has also important implications for the use of the effective basic reproduction number $\mathcal{R}_{\text{eff}}$ in determining the potential for a second wave. For example, the probability of a second wave occurring is over 26% for a relative transmissibility of 0.5 when the transmission of ZIKV from symptomatic cases is reduced by 10% on average (Figure 5.6, E1). In this case, the estimated $\mathcal{R}_{\text{eff}}$ has the median of 2.05 (95% CI: 1.71, 2.18), suggesting that the herd immunity is relatively low to prevent a second wave (Figure 5.6, C1). In fact, the median attack rate is estimated at 6.6% (95% CI: 6.2%, 6.9%). However, for the same relative transmissibility, the corresponding probability for the scenario in which the transmission of ZIKV from symptomatic cases is reduced by 30% remains below 9%. In this case, while $\mathcal{R}_{\text{eff}}$ is above 1 (median: 2.11) due to the effectiveness of interventions and low attack rates of the first wave, the second wave is unlikely to occur (Figure 5.6, C2). Increasing the relative transmissibility leads to higher attack rates of the first wave (Figure 5.6), but may also increase the probability of a second wave unfolding (Figure 5.7). These results indicate that the level of herd immunity in the population cannot be accurately measured without quantitative estimates of the contribution of asymptomatic infection.

In the context of the 2015–2016 ZIKV outbreaks in the Americas, previous work suggests that Zika spread may have contributed to the generation of herd immunity, which prevented the occurrence of a second wave of widespread ZIKV infection in the presence of sustained control efforts [199]. A recent stochastic model of ZIKV spread through the Americas estimates reporting and detection rates of 1–2% [52]. Without considering the effect of interventions or behavioural changes due to increased awareness, the model in [52] projects a significant variation amongst attack rates in different countries, and illustrates the importance of seasonal factors in the introduction and occurrence of multiple waves of ZIKV infections. As expected [52, 200] and shown in our simulations, these epidemic waves coincide with the seasonal pattern of mosquito lifetime. However, our results also indicate that the occurrence of a second wave depends on other factors, such as transmission reduction measures that largely influence the contribution of symptomatic

infection to disease spread, and more importantly, the silent transmission of the Zika virus from asymptomatic infection. Quantifying asymptomatic transmission requires specific data on the magnitude and duration of infectiousness in infected individuals, combined with measures of exposure to biting mosquitoes during the course of infection [201]. While we do not address the contribution of asymptomatic infection to herd immunity, our study highlights its importance in understanding the disease dynamics and the epidemiological trends observed in countries affected by the Zika virus. In a recent study [202], the authors highlight the potential for large errors that can arise in quantifying the contribution of asymptomatic infection to the overall cumulative incidence in an infectious disease outbreak These considerations call for further biological, clinical, and epidemiological studies to provide estimates of the relative transmissibility of asymptomatic infection, given its central role in determining the levels of herd immunity.

The importance of asymptomatic transmission has also been recognized in other vector-borne diseases including dengue and malaria [203, 204]. While infectiousness and severity of the disease are strongly correlated with viremia, outbreaks of dengue associated with low viremia have been reported [205, 206, 207]. It has been shown that, at a given level of dengue viremia, infected individuals with no symptoms or prior to the onset of symptoms are more infectious to mosquitoes than those with symptoms [203]. In the case of Zika, asymptomatic cases with low viremia may also play a role in silent transmission of infection through sexual contacts and blood transfusion. Within the context of previous studies [52, 181, 199, 200], our results underscore the need to characterize and quantify the transmission potential for asymptomatic ZIKV infection. Quantitative modelling can be used to predict the risk of infection more accurately, identify effective public health measures, and suggest strategies to counter vector-borne diseases with similar characteristics.

Chapter 6

Case Study 2 (Part II): Cost-Effectiveness of a Zika Vaccine

In this chapter, we extend the previously developed ZIKV transmission model in Chapter 5 to include vaccination dynamics and Zika-associated sequelae during pregnancy. We utilize the extended model to inform on the socio-economic impact of a vaccination program in 18 countries in the Americas affected by the 2015-2017 ZIKV outbreaks, with a particular focus on Colombia which was one of the most ZIKV affected country. We perform this analysis using country-specific parameter estimates extracted from published studies with a plausible range of costs of vaccine administration.

6.1 Background

Several studies have shown the potential for ZIKV to cause severe outcomes and long-term sequelae, including microcephaly with brain abnormalities and neurological disorders in infants, and Guillain-Barré syndrome (GBS) in adults [160, 161, 162]. Furthermore, congenital ZIKV syndrome has been reported to occur in the same proportion of women with asymptomatic infection as symptomatic ZIKV infection during pregnancy [208]. These outcomes have instigated global efforts for the development of a safe and effective Zika vaccine. A prophylactic vaccine has the potential to reduce disease incidence and eliminate birth defects of prenatal ZIKV infection in future outbreaks.

At its first consultation in March 2016, the World Health Organization and United Nations International Children's Emergency Fund set out a framework to facilitate the development of a safe and effective ZIKV vaccine, and its strategic implementation [169, 209]. They proposed a *Target Product Profile* (TPP) for vaccination as a response measure for future ZIKV outbreaks [210]. The TPP recommends vaccinating women of reproductive age and childbearing women to minimize

the incidence of microcephaly and related neurological disorders in infants. In addition to women of reproductive age, vaccination of males is also recommended, provided vaccine supply and resources are available. The TPP describes two scenarios:

- *Outbreak Response*: A mass vaccination campaign of the target group to prevent ZIKV infection and associated complications for an ongoing epidemic or an imminent outbreak.
- *Routine/Endemic use*: A routine immunization program for the general population.

The TPP is more oriented towards addressing the first scenario.

Since the development of the TPP, there are a number of vaccine candidates being investigated using a variety of vaccine platforms [211], including purified inactivated, live attenuated, viral-vectored, virus-like particles, recombinant subunit, DNA, self-replicating RNA, and mRNA [211, 212]. Experience with the development of other flavivirus vaccines suggests that generating a preventive Zika vaccine should be feasible [213, 214]. Indeed, several vaccine candidates have now advanced to clinical trials and shown to be safe and well-tolerated in generating humoral immune responses [215, 216]. However, the economic impact of a vaccine candidate will be a major factor in decisions regarding implementation and strategic use of vaccines in immunization programs. We sought to investigate the cost-effectiveness of a potential Zika vaccine candidate in 18 affected countries in the Americas, where the estimated attack rates (i.e., the proportion of the population infected) during the 2015-2017 outbreaks exceeded 2% [52, 217]. We considered the World Health Organization's recommendation of vaccine prioritization of women of reproductive age including pregnant women, and assessed the vaccination's impact on prenatal ZIKV infection and microcephaly as well as other severe brain anomalies [210].

6.2 Model Details and Parameterization

The basic *agent structure* from Chapter 5 was expanded to account for vaccination dynamics, pregnancy, and microcephaly disease outcomes (Algorithm 5). We briefly describe the model again. For infection dynamics, human population remained stratified into susceptible, exposed and incubating, infectious (i.e., symptomatic and asymptomatic), and recovered compartments. Similarly, mosquito population remained divided into susceptible, exposed and incubating, and infectious groups. The model simulated disease spread via two main modes of transmission, including vector bites and sexual interactions. Human to mosquito transmission (or vice versa) occurred as a result of rejection sampling-based (Bernoulli) trials, where the chance of successful transmission is given

Algorithm 5: The extended *human agent* from algorithm 3 with fields relevant to vaccination and pregnancy dynamics.

```
human agent structure {
    basic variables {
        ID ;                                          // ID of the human
        health ;                                      // infection stage of human
        gender
    };
    model specific variables {
        partner ;                                     // monogamous partner of human
        sexfrequency ;                                // total frequency of sex
        sexprobability ;                              // probability of weekly sex
        cumulativesex
        ispregnant ;                                  // is agent pregnant?
        timeinpregnancy ;                             // total time in pregnancy
        isvaccinated ;                                // is agent vaccinated?
        protectionlvl ;                               // efficacy of vaccine
    };
};
```

by $\mathcal{P}[\mathrm{M} \to \mathrm{H}] = \mathcal{P}[\mathrm{H} \to \mathrm{M}] = 1 - (1 - \beta_{\mathrm{vec}})^N$ where β_{vec} is the calibrated baseline probability of transmission for symptomatic cases and N is the number of bites of a single mosquito to an infectious individual. The number of bites for each mosquito was individually sampled from a Poisson distribution with the half-life of the mosquito as the mean of the distribution. The bites over the lifespan of a mosquito were also implemented as a Poisson process with an average of one bite every two days, and a maximum of 1 bite per day. Sexual transmission of ZIKV was implemented for individuals older than 15 years of age and in a monogamous context. The frequency of sexual encounters for partnered individuals was sampled from age-dependent distributions. Upon successful ZIKV transmission, susceptible individuals entered an intrinsic incubation period (IIP), sampled for each individual from the associated distribution. After the IIP elapsed, a fraction (sampled between 40% to 80%) of infected individuals entered asymptomatic infection without developing clinical symptoms. A schematic diagram of the model for transmission dynamics, natural history of ZIKV infection, and disease outcomes are provided in Figure 5.1

The model was parameterized with country-specific demographics (i.e., age and sex distributions and fertility rates). Each simulation was seeded with a single case of Zika in the incubating stage and run for a time horizon of 1 year (360 days), beginning with a high temperature season. We ran

2000 Monte Carlo simulations of ZIKV infection dynamics with a scaled-down population of 10 000 individuals and 50 000 mosquitos. Given the short (one-year) simulation timelines, we ignored the individual births and deaths in the populations, and therefore the population size remained constant. However, births following pregnancy were considered in the model implicitly for the effect of microcephaly if Zika infection occurred. Disease and vaccination outcomes were recorded throughout each simulation and used to calculate ICER values and cost-effectiveness acceptability probabilities. Only epidemic curves that had at least one secondary case by the end simulations were considered in cost-effectiveness analysis. Infection and disease transmission parameters were as described previously in Table 5.1. Vaccination, disease outcomes, and cost-effectiveness parameters are summarized in Table 6.1 and Table 6.4.

Calibration and Transmissibility

Based on the calibration process in the previous chapter, we initially present cost-effectiveness results of a ZIKV vaccine in Colombia. This calibration was based on the estimates of the reproduction number of $\mathcal{R}_0 = 2.2$ and $\mathcal{R}_0 = 2.8$ for Antioquia, Colombia with the mean attack rate of 8% (95% CI: 4% and 26%) [177, 178]. Given that relative transmissibility of asymptomatic infection is a key epidemiological parameter, we therefore considered two scenarios with reduction factors of 0.1 (low) and 0.9 (high) to quantify this relative transmissibility. We also considered two scenarios with reduction factors of 0.1 and 0.5 for symptomatic transmission to account for decreased mobility and lower exposure to mosquito bites through full clothing, mosquito repellents, or possible isolation during symptomatic infection.

Later studies [52, 217] provided ZIKV burden in countries affected by the disease, and estimated attack rates for these countries. We therefore expanded our cost-effectiveness analysis for 18 countries in the Americas, where the estimated attack rates (i.e., the proportion of the population infected) during the 2015-2017 outbreaks exceeded 2%, by calibrating the model to these estimated attack rates (considering both symptomatic and asymptomatic infections). Here we assumed the same transmissibility for both asymptomatic and symptomatic infection, with any transmission reduction in asymptomatic infection accounted for in the calibration process. In the main simulations, these attack rates were considered as the proportion of the population immune (i.e., representing the level of herd immunity) at the start of simulations for each country in the evaluation of vaccination scenarios.

Calibration was performed to determine the transmissibility of the disease corresponding to country-specific attack rates in the absence of any control measures. We seeded the calibration

Table 6.1: Zika model parameters values and their associated ranges.

Parameter Description	**Baseline value (range)**	**Source**
Risk of microcephaly		
First trimester (97 days)	5% - 14%	[159, 208, 218]
Second and third trimester	3% - 5%	[159, 208, 218]
Risk of Guillain-Barre Syndrome	0.025% - 0.06%	[219]
Life expectancy		
Without microcephaly	70 years	[220]
With microcephaly	35 years	[220]
Probability of survival past first year of life for infants with microcephaly	0.798	[221]
Pre-existing level of herd immunity		
From previous outbreaks	8% (2.2% - 11%)	[52]
Vaccination coverage and efficacy		
Non-pregnant women of 15 to 49 years of age	60%	
Pregnant women	80%	assumed
Other individuals from 9 to 60 years of age	10%	
Preventing infection	60% - 90%	
Cost-effectiveness rates		
Disability weight for microcephaly	0.16	[147]
Annual discount rate	3%	assumed

for an appropriate initial value of the transmissibility, and ran 2000 Monte-Carlo simulations. The simulations were averaged after a 2-year time horizon to estimate the attack rate. This process was iterated over a range of transmission values so that estimated attack rates were covered for each country. We then performed curve fitting to the simulated data to determine the transmission values that correspond to the country-specific attack rates for simulation scenarios. A general formula of transmission values as a function of attack rates was found using Matlab's curve fitting toolbox. These attack rates and calibrated transmission values are shown in Table 6.2.

Table 6.2: Estimated attack rates for the 2015-2017 ZIKV outbreaks [52, 217], and corresponding transmissibility values via model calibration.

Country	Attack Rate	Estimated Transmissibility
Belize	21%	0.2884
Bolivia	10%	0.2761
Brazil	18%	0.2859
Colombia	12%	0.2792
Costa Rica	2%	0.2476
Ecuador	8%	0.2723
El Salvador	16%	0.2839
French Guiana	18%	0.2859
Guatemala	14%	0.2817
Guyana	15%	0.2829
Honduras	14%	0.2817
Mexico	5%	0.2641
Nicaragua	17%	0.2849
Panama	15%	0.2829
Paraguay	17%	0.2849
Peru	4%	0.2602
Suriname	22%	0.2891
Venezuela	19%	0.2868

6.2.1 Disease Outcomes and Microcephaly

There is evidence that associates the risk of microcephaly in infants to ZIKV infection in all trimesters of pregnancy, although the risk is significantly higher in the first and second trimesters [208, 218]. Previous studies, considering possible over-reporting, have quantified the risk of developing microcephaly in both symptomatic and asymptomatic pregnant women [159, 222]. We considered the associated risks in a probabilistic approach to determine the microcephaly outcome in pregnant women at the time of infection. The risk of microcephaly was highest in the range 5% to 14% during the first trimester (which ends at 97 days of pregnancy), and reduced to 3% to 5% during the second and third trimesters [159, 208, 218]. We set a probability of 0.798 for survival past first year of life for infants with microcephaly [221]. Infants with microcephaly who survived their first year of life were assumed to have significantly lower life expectancy [220, 221],

Table 6.3: Age-specific fertility rates per 10,000 women of reproductive age [224].

Country	**Age Groups**						
	15-19	*20-24*	*25-29*	*30-34*	*35-39*	*40-44*	*45-49*
Belize	69.7	150.9	142	98.5	49.3	16.2	1.3
Bolivia	72.6	146.9	148.6	115	80.9	36.5	8
Brazil	68.4	107.6	90.6	55.8	29.2	10.2	1.9
Colombia	57.7	112.3	96.8	65	37.7	14.7	2.7
Costa Rica	59.1	101.1	87.2	70.3	40.5	10.7	1.3
Ecuador	77.3	139.3	124.6	90.9	55.1	24.4	5.7
El Salvador	66.8	108.1	97.6	70.5	37.2	12.6	1.7
French Guiana	82.6	156.2	182.5	151.3	88.7	33	2.6
Guatemala	84	173.2	159.4	124.2	80.1	32.8	6.4
Guyana	90.1	156.3	118.7	87.2	49.7	13.1	4.7
Honduras	68.4	134.8	113.7	87.3	56.2	28.2	5.3
Mexico	66	126.4	127.5	83	44	9.2	1.8
Nicaragua	92.8	122.5	108.7	76.1	42.8	16.1	4.6
Panama	78.5	149.1	132.2	87.9	38	9.1	0.9
Paraguay	60.2	129.8	130.3	102.9	65.3	26.1	5.1
Peru	68	110	113	104	73	25	3
Suriname	48.1	117	128.6	101.1	59	24.3	1.9

and reduced by 50% from 70 years to 35 years on average [220]. In addition to microcephaly, we considered the risk of developing GBS in ZIKV-infected individuals [219]. The risk of GBS in adults was sampled in the range 0.025%–0.06% [219]. We also considered the effect of neurological and behavioural deficits due to microcephaly, leading to an impaired quality of life, quantified by disability weights provided in the Global Burden of Disease study [223]. The total number of pregnant women was calculated based on the country-specific fertility rate of population in each simulation (Table 6.3). Ignoring fatal complications, the number of pregnant women at any point in time for each simulation was calculated by

$$\text{number of pregnant women} = \left(\frac{\text{nWRA}}{1000}\right)(\text{fertility rate} \times 0.75 + \text{abortion rate} \times 0.167) \quad (6.1)$$

where nWRA is the number of women of reproductive age. We used an estimated abortion rate of 12% for WRA [225]. The probability of birth for 9 months was assumed 0.75, and the probability

of abortion for 2 months was assumed 0.167.

Given the flexibility of an ABM, we considered trimesters to implement risk of microcephaly in the model. At the onset of each simulation, the trimester for each pregnant woman was randomly selected according to the respective distributions [226]. The beginning of the first trimester was set for newly pregnant women during simulations. Microcephaly occurred according to the risk associated with infection in each trimester, which was implemented at the time of infection.

6.2.2 Vaccination Dynamics

Based on the recommendations outlined in the WHO/UNICEF ZIKV Vaccine Target Product Profile (TPP) for vaccine prioritization [210], we implemented vaccination in the model for women between 15 and 49 years of age. The vaccination coverage was set to 60% for this group at the onset of simulations. For pregnant women in the same age group, the vaccination coverage was set to 80% throughout the simulations. We also considered a vaccination coverage of 10% for other individuals in the population between 9 and 60 years of age, in order to reduce the risk of disease transmission to pregnant women. This implementation strategy corresponds to the *outbreak response strategy* in the TPP document [210], prioritizing women of reproductive age (WRA).

While some ZIKV vaccine candidates have entered clinical trials, there is currently no data available to indicate the level of vaccine-induced protection and number of vaccine doses required. We therefore assumed that a single vaccine dose provides a protection efficacy in the range 60% – 90% against infection, which was sampled for each vaccinated individual and implemented as a reduction factor in disease transmission. We assumed that ZIKV infection following vaccination (if occurred) was asymptomatic without clinical manifestation and that vaccination has no effect on the risk of microcephaly in pregnant women if infection occurred. Naturally acquired immunity was assumed to provide full protection for sufficiently long period of time, so that the risk of re-infection within the same epidemic season was eliminated.

6.3 Cost-Effectiveness

We conducted the cost-effectiveness analysis from a government (or a public-payer) perspective and considered both short- and long-term medical costs specific to each country, summarized in Table 6.4. Short-term costs included physician visits and diagnostic tests for symptomatic ZIKV

infection in pregnant women. For microcephaly in infants and GBS in adults, we considered lifetime direct medical costs related to hospitalization, treatment, and other associated outcomes. Based on the estimates for other flavivirus vaccines, we considered a range of US$2 to US$100 for vaccination costs per individual (VCPI) [228], including vaccine dose, administration, and 3% wastage.

The health impact of microcephaly and GBS on an individual's quality of life was captured by disability-adjusted life years (DALYs) with disability weights extracted from the Global Burden of Disease [223]. We used disability weights associated with *i.e., severe intellectual disability*, which was considered a good proxy for microcephaly, also used in previous work [220]. We understand that the disability weights may be subject to uncertainty, such uncertainty is not quantified, and

Table 6.4: Direct medical costs, and GDP per capita for affected countries in the Americas. Physician visits only applied to symptomatic individuals. Data obtained from [227].

Country	**Cost Categories**			
	Microcephaly	*GBS*	*Physician Visit*	*GDP per capita*
Belize	$103,586	$32,709	$61	$4,955
Bolivia	$80,974	$25,569	$57	$3,097
Brazil	$100,068	$31,599	$57	$8,694
Colombia	$78,990	$24,943	$68	$5,900
Costa Rica	$124,203	$39,220	$63	$11,563
Ecuador	$98,759	$31,185	$60	$6,084
El Salvador	$124,203	$39,220	$63	$3,719
French Guiana	$91,925	$29,027	$65	$18,036
Guatemala	$91,173	$28,790	$59	$4,032
Guyana	$98,974	$31,253	$57	$4,325
Honduras	$88,351	$27,899	$57	$2,358
Mexico	$93,867	$29,640	$67	$8,867
Nicaragua	$72,383	$22,856	$56	$2,109
Panama	$107,620	$33,983	$63	$14,009
Paraguay	$81,542	$25,749	$58	$4,094
Peru	$88,850	$28,056	$61	$6,042
Suriname	$95,294	$30,091	$63	$7,298
Venezuela	$120,582	$38,076	$69	$7,766

we therefore relied on established values used in previous work. For given vaccination costs per individual (VCPI), we calculated incremental cost-effectiveness ratios (ICER) averaged over simulations. Negative ICER values are always considered to be cost-saving since the intervention provides additional benefit over the alternative while costs less than the alternative. For positive ICER values, we considered the World Health Organization standards of using the per-capita gross domestic product (GDP) as a threshold of Willingness to Pay (WTP) [81]. The vaccination program was considered *very cost-effective* and *cost-effective* for ICER values up to the per-capita GDP and 3 times the per capita GDP, respectively. We also considered a range of WTP values to inform decisions on vaccine cost-effectiveness in settings where the per-capita GDP threshold may not be applicable. We calculated the average incremental cost-effectiveness ratio (ICER) and the associated 95% confidence interval using a non-parametric bootstrap method of 2000 replicates, and constructed the cost-effectiveness plane and acceptability probabilities to offer a visual representation of the joint distribution of costs and benefits. A discount rate of 3% was applied to both the costs and DALY calculations to consider preference for present value.

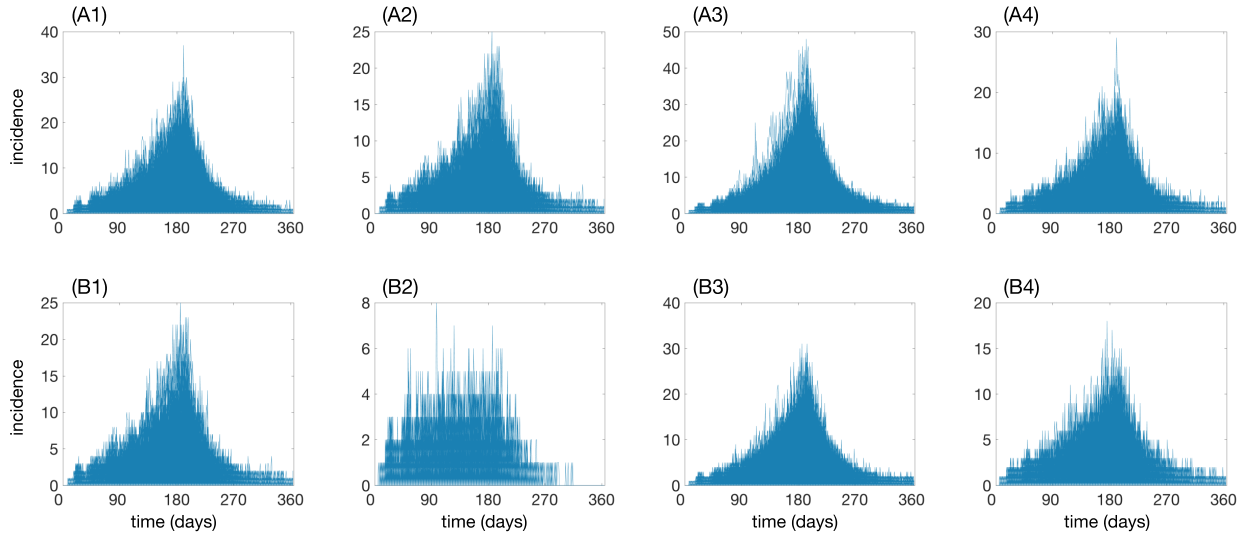
6.4 Results

We considered a plausible range of $\$2$ to $\$100$ for VCPI to account for vaccine dose, administration, and wastage based on the estimates for other flavivirus vaccines [228]. Our results show that for a sufficiently low VCPI, a single-dose vaccination program is cost-saving for all countries studied here. We describe our results below.

6.4.1 Vaccination in Colombia

We first present the results for Colombia since it is one of the most Zika-affected countries in the northwest of South America. We considered the scenarios of $\mathcal{R}_0 = 2.2$ and $\mathcal{R}_0 = 2.8$, taking into consideration pre-existing herd immunity in the population. Figure 6.1 and Figure 6.2 show disease incidence of 2000 independent realizations for a time horizon of one year. Each figure shows simulation-level incidence (blue curves) with and without vaccination, and in the absence and presence of herd immunity in the population. Subfigures correspond to different relative transmissibility of asymptomatic infection. In the next few sections, we describe the results of the cost-effectiveness analysis however one may find a succinct summary of the analysis in Table 6.5 which summarizes simulation outcomes for VCPI in each scenario.

Disease incidence without vaccination



Disease incidence with vaccination

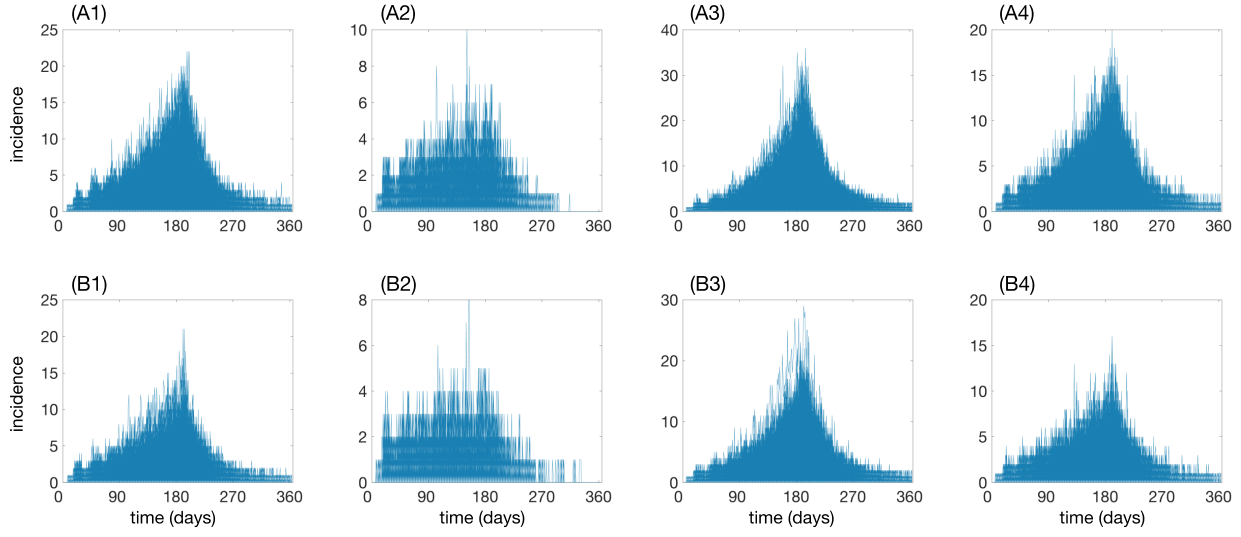
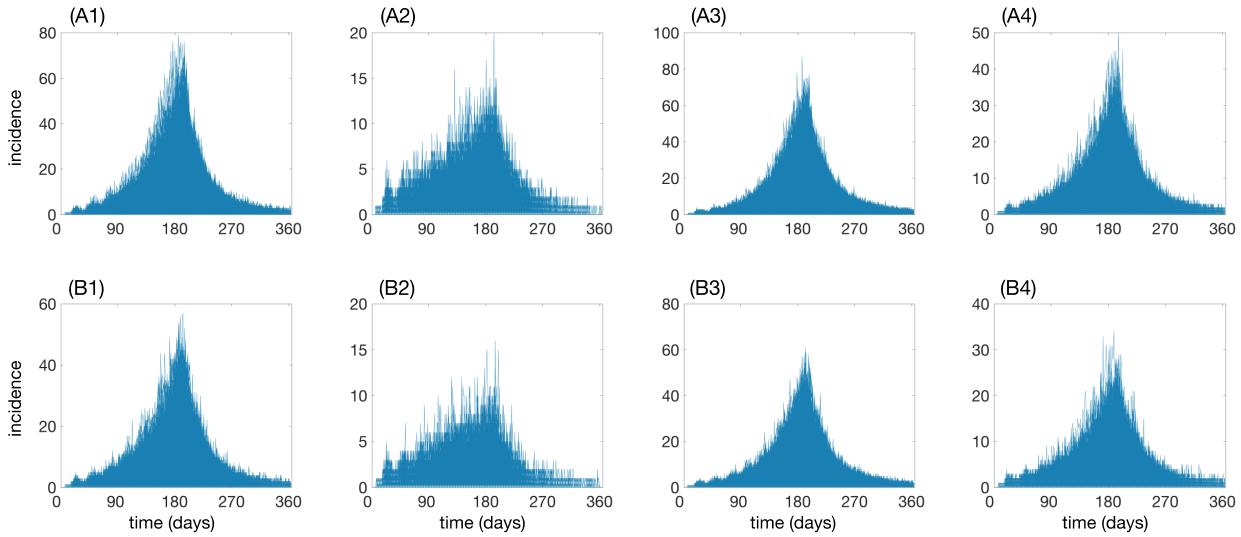


Figure 6.1: Incidence of infection for 5000 independent realizations without (top) and with (bottom) vaccination in the absence of herd immunity (A1-A4) and in the presence of 8% herd immunity (B1-B4) in the population. Simulations were run considering the relative transmissibility of asymptomatic infection and the reduction of transmission by symptomatic infection to be respectively: 0.1 and 0.1 (A1,B1); 0.1 and 0.5 (A2,B2); 0.9 and 0.1 (A3,B3); 0.9 and 0.5 (A4,B4). Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 2.2$.

Disease incidence without vaccination



Disease incidence with vaccination

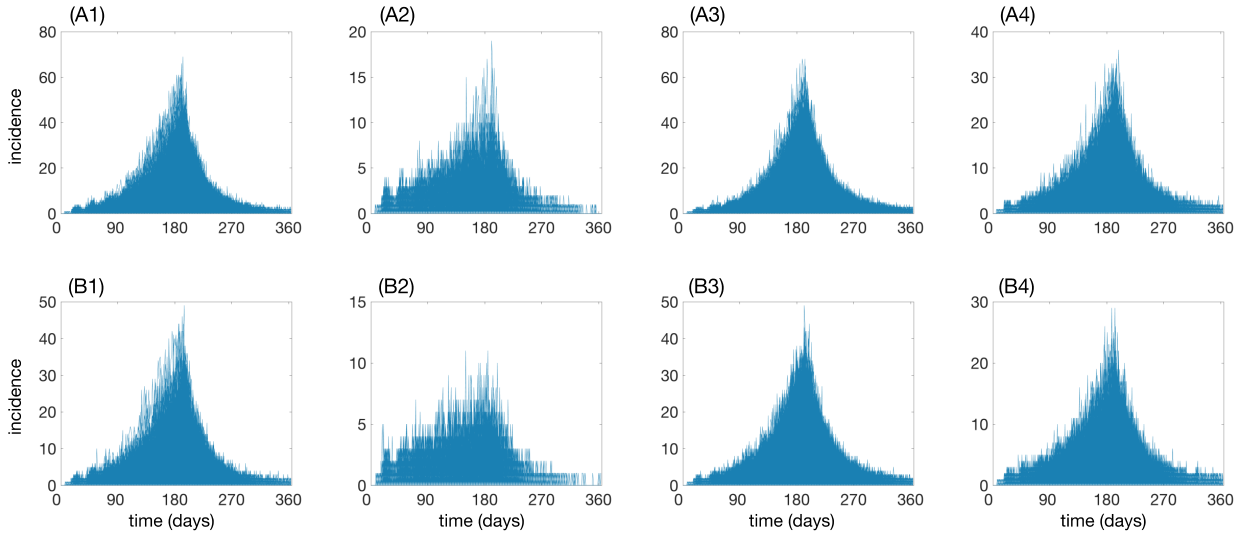


Figure 6.2: Incidence of infection for 5000 independent realizations without (top) and with (bottom) vaccination in the absence of herd immunity (A1-A4) and in the presence of 8% herd immunity (B1-B4) in the population. Simulations were run considering the relative transmissibility of asymptomatic infection and the reduction of transmission by symptomatic infection to be respectively: 0.1 and 0.1 (A1,B1); 0.1 and 0.5 (A2,B2); 0.9 and 0.1 (A3,B3); 0.9 and 0.5 (A4,B4). Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 2.8$.

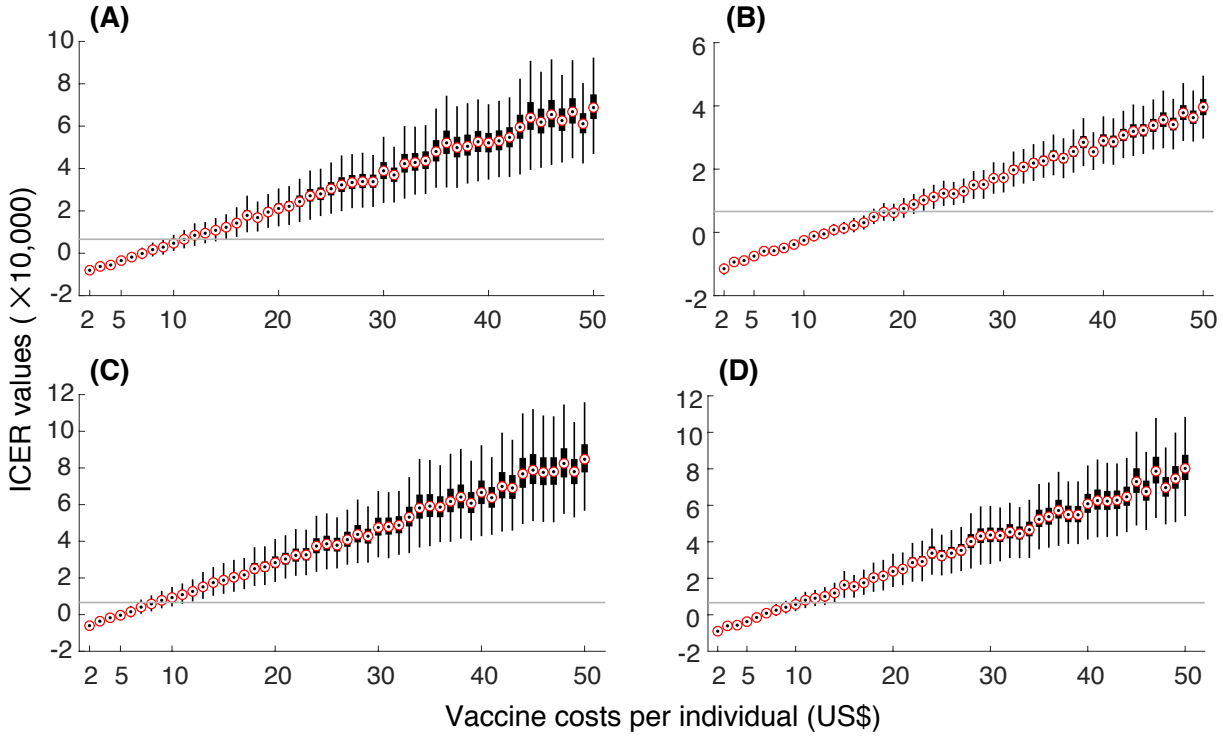


Figure 6.3: Boxplots for ICER values obtained using bootstrap method for a range of VCPI. Subplots correspond to the scenarios without pre-existing immunity (A,B), and with an average of 8% pre-existing immunity (C,D) in the population. The relative transmissibility of asymptomatic infection was set to 10% (A,C) and 90% (B,D). Solid (grey) line represents the willingness-to-pay threshold corresponding to the average of per capita GDP of Colombia between 2013 and 2017. Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 2.2$.

Vaccine cost-effectiveness Estimated ICER values (corresponding to $\mathcal{R}_0 = 2.2$) are shown in Figure 6.3. In this figure, boxplots of calculated ICER values and their associated confidence intervals obtained using the non-parametric bootstrap method are shown for a range of VCPI. The grey line corresponds to Colombia's GDP of $\$6610$ per DALY averted, corresponding to the average per-capita GDP of Colombia between 2013 and 2017, which was considered as the threshold value for cost-effectiveness. Subplots correspond to the scenarios of no pre-existing immunity and with an average of 8% pre-existing immunity in the population. In a fully susceptible population, with a 10% relative transmissibility of asymptomatic infection (i.e. Figure 6.3A), the ICER values and their associated ranges remained negative for 100% of simulation results when VCPI is $\$6$ or less, thus suggesting that the vaccine is cost-saving regardless of the thresholds of WTP. For VCPI with positive ICER values, the vaccine is very cost-effective with a probability of at least 90% if VCPI is $\$10$ or less. Increasing the threshold to $\$19832$ (three times the average GDP) [228],

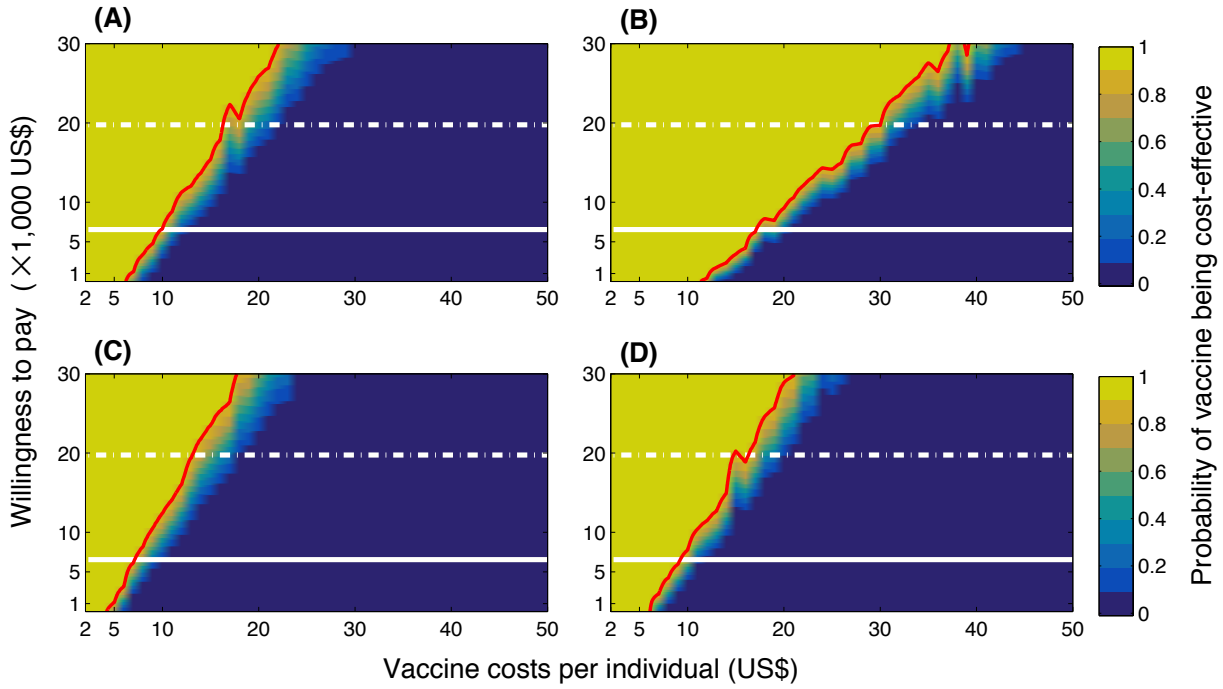


Figure 6.4: Probabilities of vaccine being cost-effective for a range of VCPI and willingness-to-pay. Subplots correspond to the scenarios without pre-existing immunity (A,B), and with an average of 8% pre-existing immunity (C,D) in the population. The relative transmissibility of asymptomatic infection was set to 10% (A,C) and 90% (B,D). Solid line represents the willingness-to-pay threshold corresponding to the average of per capita GDP of Colombia between 2013 and 2017. Dashed line represents three times the average of per capita GDP of Colombia. The red curve represents the 90% probability of vaccine being cost-effective for a given VCPI. Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 2.2$.

our results suggest that vaccination is still cost-effective for VCPI up to $\$16$. The probability of cost-effectiveness is sensitive to VCPI and decreases sharply from 90% to below 10% with marginal increase in the VCPI (Figure 6.4). When the transmissibility of asymptomatic infection is relatively high, i.e., 90% (Figure 6.3B), then vaccination is cost-saving for VCPI up to $\$12$, as suggested by negative ICER values. For positive ICER values, vaccination is very cost-effective if VCPI is $\$16$ or less. For three times the GDP threshold of willingness-to-pay, vaccination is still cost-effective for VCPI up to $\$29$. We also investigated the presence of pre-existing herd immunity as a result of previous outbreaks. We used estimates of 8% (95% CI: 4% – 26%) attack rate to account for herd immunity in the population [52]. When the relative transmissibility of asymptomatic infection was low (10%), the ICER values and their associated ranges are negative for VCPI up to $\$4$ (Figure 6.3C). In the presence of herd immunity, for positive ICER values, vaccination remains very cost-effective (with a probability of at least 90%) at the $\$6610$ threshold of WTP per DALY

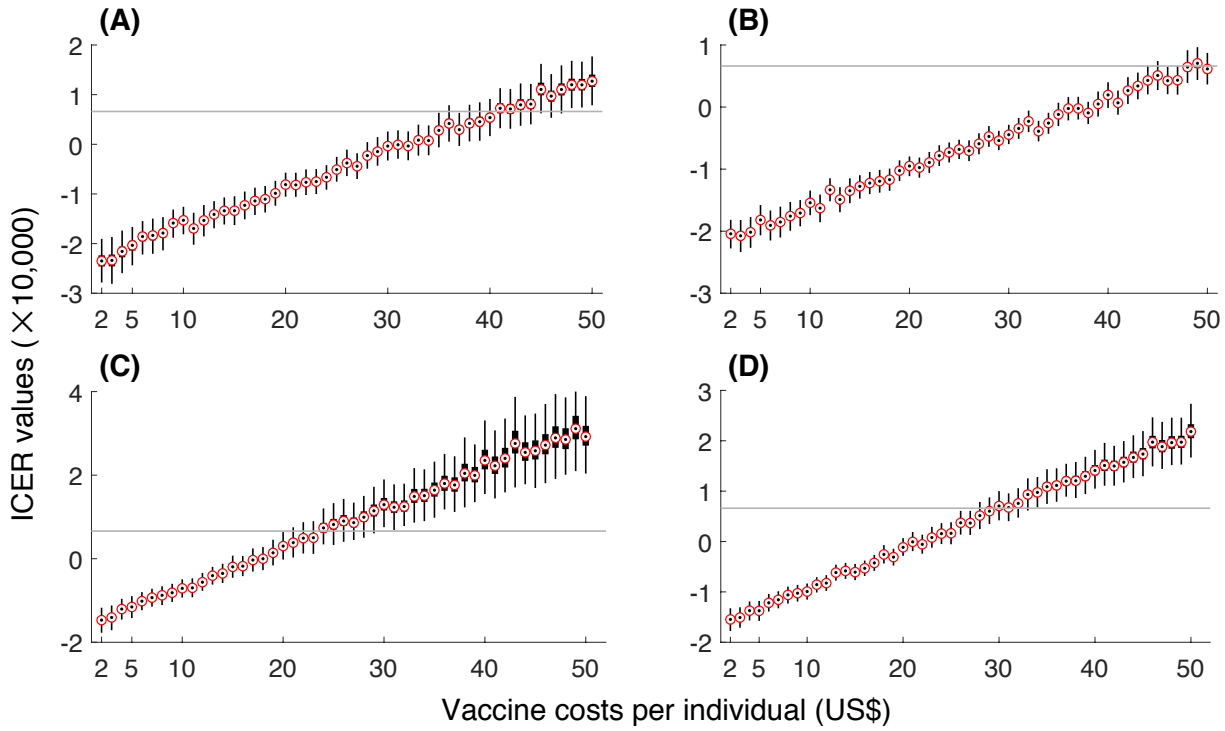


Figure 6.5: Boxplots for ICER values obtained using bootstrap method for a range of VCPI. Subplots correspond to the scenarios without pre-existing immunity (A,B), and with an average of 8% pre-existing immunity (C,D) in the population. The relative transmissibility of asymptomatic infection was set to 10% (A,C) and 90% (B,D). Solid (grey) line represents the willingness-to-pay threshold corresponding to the average of per capita GDP of Colombia between 2013 and 2017. Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 2.8$.

averted if VCPI does not exceed $\$7$. At the threshold of three times the average GDP, vaccination is still cost-effective for VCPI up to $\$13$. With the same level of herd immunity (i.e., 8%), but a higher relative transmissibility of asymptomatic infection (90%) (Figure 6.3D), vaccination is cost-saving (with negative ICER values) for VCPI up to $\$6$. When ICER values are positive, vaccination is very cost-effective if VCPI is $\$8$ or less, and cost-effective if VCPI is $\$14$. In Figure 6.4, the associated probabilities of vaccine being cost-effective for a range of VCPI and WTP values for $\mathcal{R}_0 = 2.2$ are presented for all scenarios. In this figure, the red curve represents the 90% probability of vaccine being very cost-effective for a given VCPI. Solid line represents the WTP threshold corresponding to the average of per capita GDP of Colombia. Dashed line represents three times the average of per capita GDP of Colombia. Subplots correspond to the scenarios with and without pre-existing immunity and with reductions of 10% and 90% relative transmissibility of asymptomatic infection. We also evaluated the vaccine cost-effectiveness scenarios in a population setting with a higher

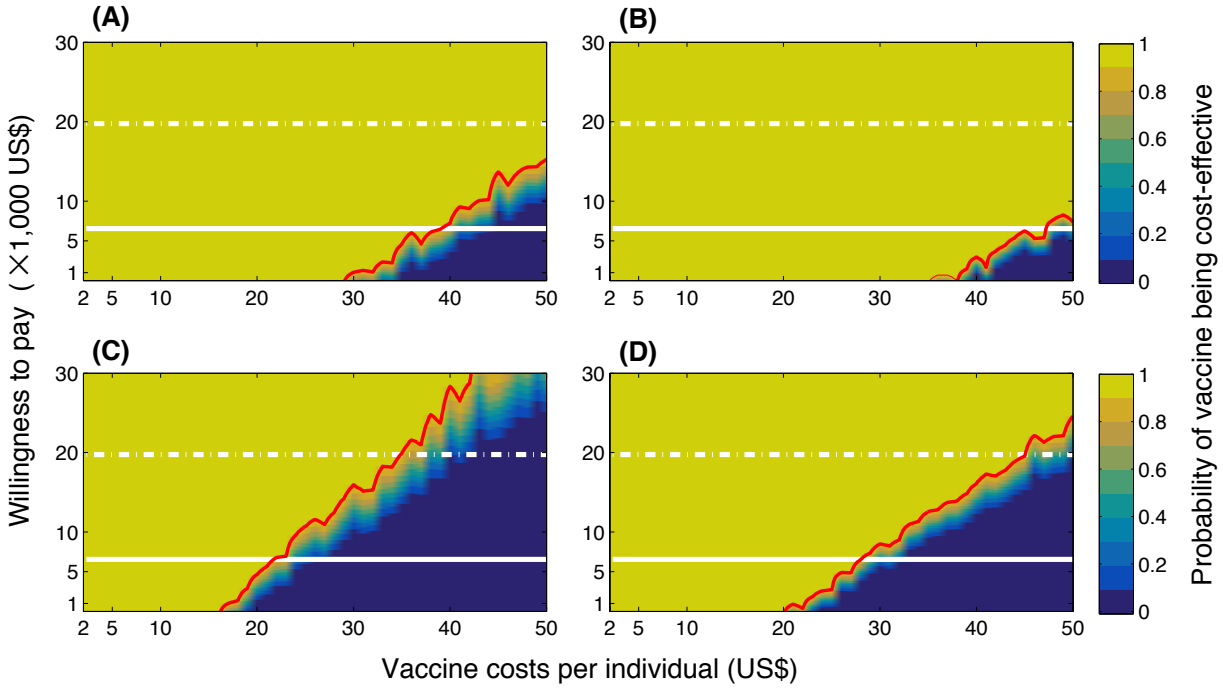


Figure 6.6: Probabilities of vaccine being cost-effective for a range of VCPI and willingness-to-pay, with $\mathcal{R}_0 = 2.8$. Subplots correspond to the scenarios without pre-existing immunity (A,B), and with an average of 8% pre-existing immunity (C,D) in the population. The relative transmissibility of asymptomatic infection was set to 10% (A,C) and 90% (B,D). Solid line represents the willingness-to-pay threshold corresponding to the average of per capita GDP of Colombia between 2013 and 2017. Dashed line represents three times the average of per capita GDP of Colombia. The red curve represents the 90% probability of vaccine being cost-effective for a given VCPI. Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 2.8$.

transmissibility of $\mathcal{R}_0 = 2.8$. The estimated ICER values and associated confidence intervals shown in Figure 6.5. Compared to $\mathcal{R}_0 = 2.2$, Figure 6.5 and Figure 6.6 indicate that vaccination is very cost-effective for a larger range of VCPI, in particular, when the reduction of transmission from ZIKV symptomatic infection is relatively low (10% on average). Figure 6.5A and Figure 6.5B show that in the absence of pre-existing herd immunity the vaccine is cost-saving (as suggested by negative ICER values) up to a VCPI of $\$29$ when the relative transmissibility of asymptomatic infection is 10%, and up to $\$35$ when the relative transmissibility of asymptomatic infection is 90%. Similarly, Figure 6.5C and Figure 6.5D show that in the presence of herd immunity the vaccine is cost-saving when VCPI is $\$16$ and $\$20$ for low and high relative transmissibility, respectively. For positive ICER values, vaccination remains very cost-effective (with a probability of at least 90%) up to VCPI of $\$38$ and $\$45$ in the absence of herd immunity and up to VCPI of $\$22$ and $\$27$ in the presence of herd immunity, for 10% and 90% relative transmissibility of asymptomatic infection.

In Figure 6.6, the associated probabilities of vaccine being cost-effective for a range of VCPI and willingness-to-pay for $\mathcal{R}_0 = 2.8$ are presented.

Effect of vaccination on microcephaly We calculated the reduction of fetal microcephaly during pregnancy by comparing the simulation scenarios in the presence and absence of vaccination. We used cumulative number of fetal microcephaly cases following ZIKV infection during pregnancy at the end of each simulation and calculated percentage reduction of microcephaly due to vaccination using non-parametric bootstrap sampling. The results are presented in Figure 6.7 and Figure 6.8, in which the distribution of the percentage reduction is shown. In all scenarios investigated for vaccine cost-effectiveness, the median percentage reduction of microcephaly exceeded 64%, suggesting that a vaccine with protection efficacy as low as 60% could significantly reduce the incidence of microcephaly.

6.4.2 Vaccination in the Americas

In addition to Colombia, we applied the model to assess the cost-effectiveness of a potential ZIKV vaccine in 18 other countries in the Americas. All estimates are based on the attack rates reported in Table 6.2 as the level of pre-existing herd immunity in the population for each country. Incidence and attack rates (which serve to validate the model) for different countries in the absence of vaccination are illustrated in Figure 6.9 and Figure 6.10.

Our results show that for a sufficiently low VCPI, a single-dose vaccination program is cost-saving for all countries studied here (Figure 6.11, green). In this figure, green curves correspond to the upper range of VCPI for which the vaccine is cost-saving, i.e. ICER < 0. Similarly, red and black curves correspond to the upper range of VCPI for which the vaccine is very cost-effective and cost-effective, respectively. The lowest VCPI was estimated for Costa Rica, where

Table 6.5: Upper range of VCPI (US dollar) for a Zika vaccine candidate to be cost-saving (ICER<0), very cost-effective (WTP of per capita GDP) or cost-effective (WTP of three times per capita GDP).

	0% herd immunity						8% herd immunity					
RTA	10%			90%			10%			90%		
$\mathcal{R}_0$	<$0	$6,610	$19,832	<$0	$6,610	$19,832	<$0	$6,610	$19,832	<$0	$6,610	$19,832
2.2	$6	$10	$16	$12	$16	$29	$4	$7	$13	$6	$8	$14
2.8	$29	$38	$53	$35	$45	$66	$16	$22	$35	$20	$27	$45

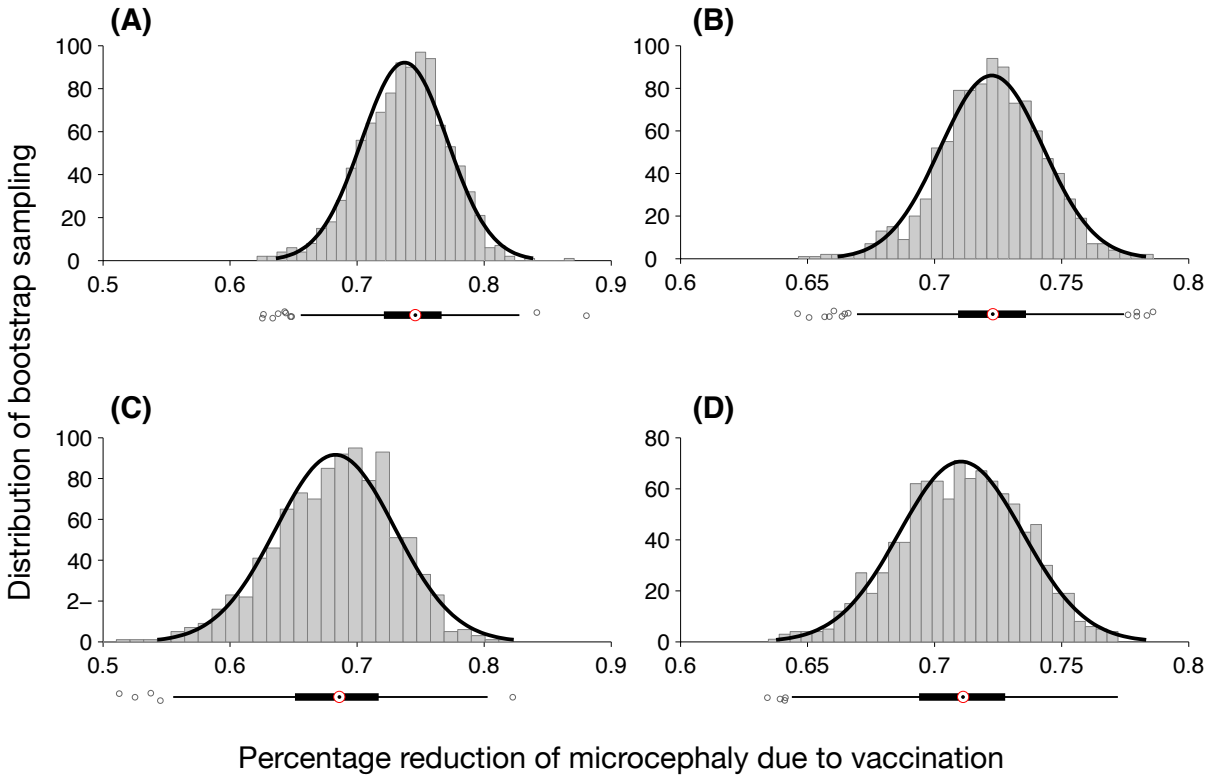


Figure 6.7: Distribution of percentage reduction of microcephaly obtained using bootstrap method. Subplots correspond to the scenarios without pre-existing immunity (A,B), and with an average of 8% pre-existing immunity (C,D) in the population. The relative transmissibility of asymptomatic infection was set to 10% (A,C) and 90% (B,D). The median percentage reduction is (A) 0.739 (IQR: 0.715 – 0.759); (B) 0.723 (IQR: 0.709 – 0.736); (C) 0.687 (IQR: 0.652 – 0.717); (D) 0.711 (IQR: 0.694 – 0.728). Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 2.2$.

the vaccine is cost-saving with a probability of at least 90% for VCPI up to $10, derived from the cost-effectiveness acceptability curves (Figure 6.12). With the same probability, the highest VCPI was estimated at $25 for Guatemala and Panama under which the vaccine is cost-saving. The upper range of VCPI for a cost-saving scenario in other countries is estimated between $14 and $24. For positive ICER values, the vaccine is very cost-effective with a probability of at least 90% at VCPI of $16 or less in Costa Rica (mean incremental cost of $7352/DALY averted; 95% CI: $1280–$9234) and $47 or less in French Guiana (mean incremental cost of $14475/DALY averted; 95% CI: $10016–$16653), with other countries having an upper value of VCPI in this range (Figure 6.11, red). For the threshold of three times the per capita GDP, the vaccine is still cost-effective (with a probability of at least 90%) with VCPI up to $24 (mean incremental cost of $4829/DALY averted; 95% CI: $2395–$6068) in Nicaragua and $96 (mean incremental cost of

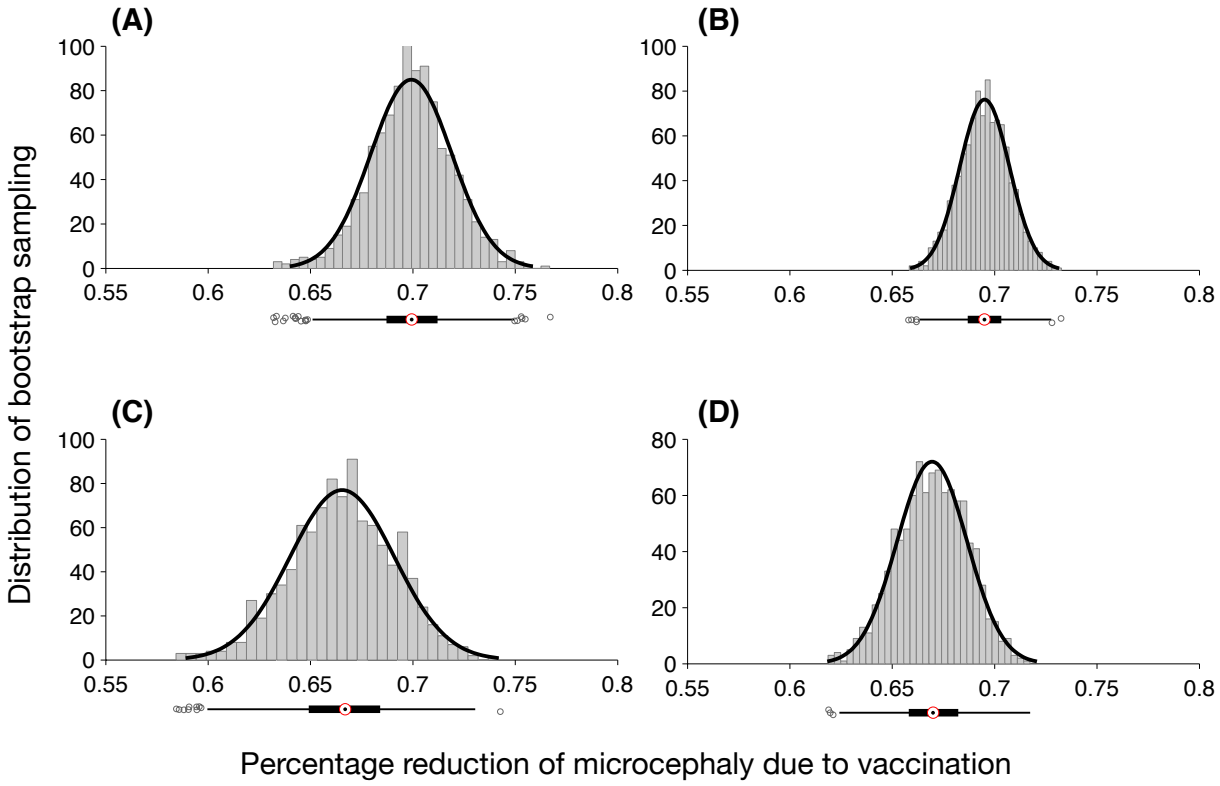


Figure 6.8: Distribution of percentage reduction of microcephaly obtained using bootstrap method, with $\mathcal{R}_0 = 2.8$. Subplots correspond to the scenarios without pre-existing immunity (A,B), and with an average of 8% pre-existing immunity (C,D) in the population. The relative transmissibility of asymptomatic infection was set to 10% (A,C) and 90% (B,D). The median percentage reduction is (A) 0.699 (IQR: 0.687 – 0.712); (B) 0.695 (IQR: 0.687 – 0.704); (C) 0.666 (IQR: 0.649 – 0.683); (D) 0.670 (IQR: 0.658 – 0.682). Figure corresponds to model scenario calibrated to basic reproduction number $\mathcal{R}_0 = 2.8$.

$49934/DALY averted; 95% CI: $36523–$53661) in French Guiana, with other countries having an upper value of VCPI in this range (Figure 6.11, black). The VCPI for scenarios of cost-saving, very cost-effective, and cost-effective for each country are provided in Table 6.6. The corresponding incremental cost per DALY averted with 95% confidence intervals are reported in Table 6.7. The associated cost-effectiveness acceptability curves are presented in Figure 6.12. In this figure, the red curve represents the 90% probability of vaccine being cost-effective for a given VCPI. Solid line represents the willingness-to-pay threshold corresponding to the average of per capita GDP of each country between 2015 and 2017. Dashed line represents three times this average of per capita GDP.

We also calculated the reduction of fetal microcephaly during pregnancy by comparing the

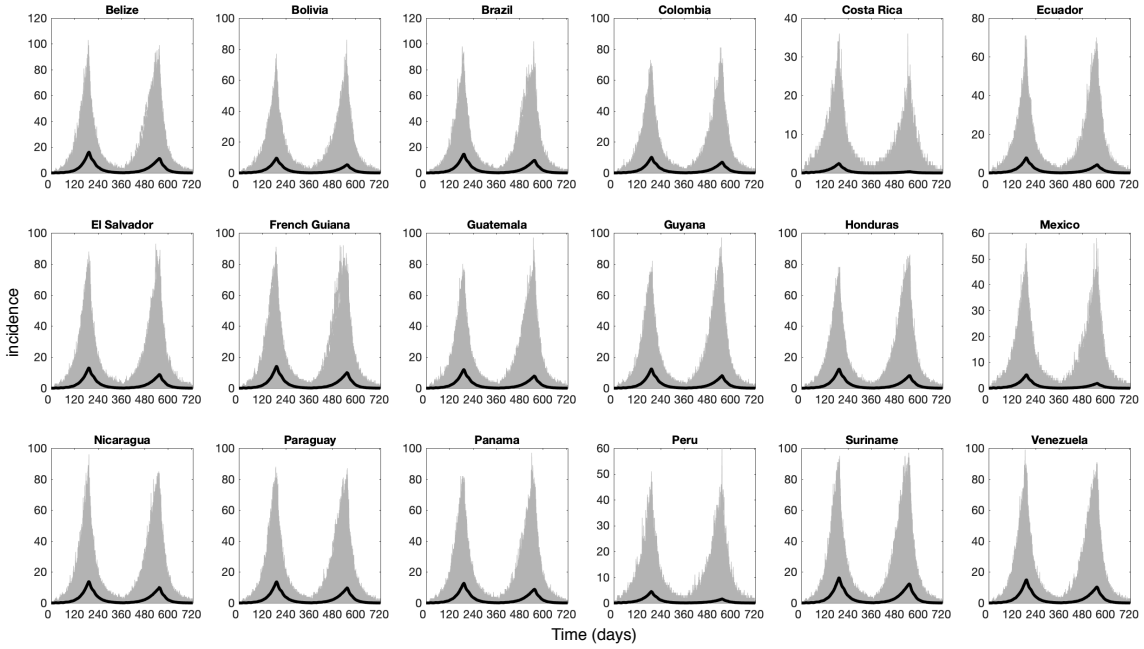


Figure 6.9: Incidence of ZIKV infection for each country with estimated attack rates for two years in the absence of vaccination (corresponding to the main scenario). The black curve shows the average of 2000 realizations.

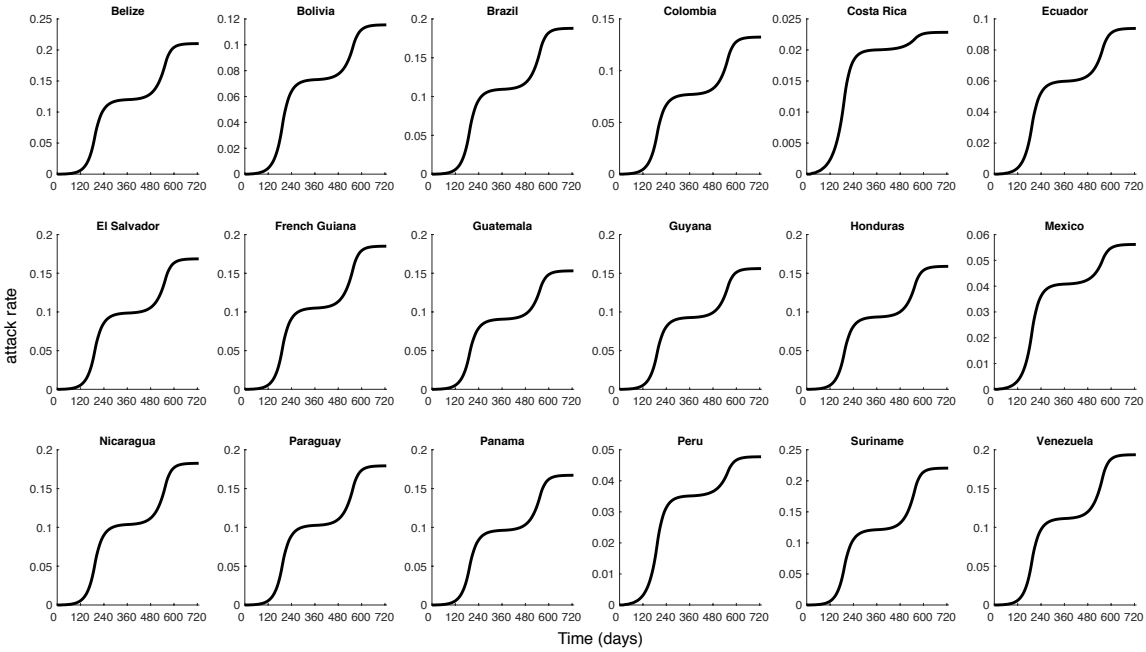


Figure 6.10: Attack rates (average of 2000 realizations) of ZIKV outbreaks for two years in the absence of vaccination (corresponding to the main scenario).

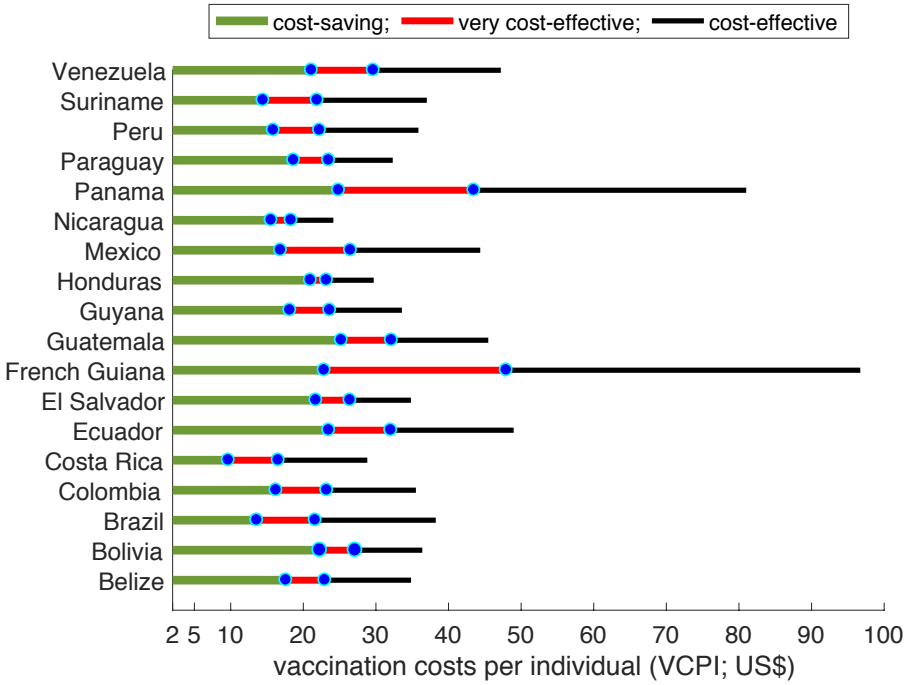


Figure 6.11: Upper range of VCPI (US dollar) for the scenarios of cost-saving (green), very cost-effective (red), and cost-effective (black). All estimates are based on the level of pre-existing herd immunity in the population for each country.

simulation scenarios in the presence and absence of vaccination. We found a marked reduction in cases of microcephaly within the range of 74%–92% due to vaccination, with the median percentage reduction exceeding 80% in all countries (Figure 6.13). This suggests that a ZIKV vaccine with a prophylactic efficacy as low as 60% could significantly reduce the incidence of microcephaly.

Given that the attack rates in future outbreaks may be different from those estimated for the 2015–2017 outbreaks, we further conducted cost-effectiveness analysis for two additional scenarios (Table 6.8). As such, we conducted cost-effectiveness analysis for two additional scenarios. In the first scenario, we calibrated the model to an increase of 4% to the estimated attack rate for each country. In the second scenario, the model was calibrated to a 4% decrease in the estimated attack rates, with a lower bound of 1%, for each country. The levels of pre-existing herd immunity at the onset of simulations remained the same as those in Table 6.2. In the scenario with increased attack rates, we found that vaccination is very cost-effective with a probability of at least 90% at VCPI of $\$20$ or less in Nicaragua (mean incremental cost of $\$1067$/DALY averted) and $\$50$ or less in French Guiana (mean incremental cost of $\$14914$/DALY averted). The upper VCPI for other countries

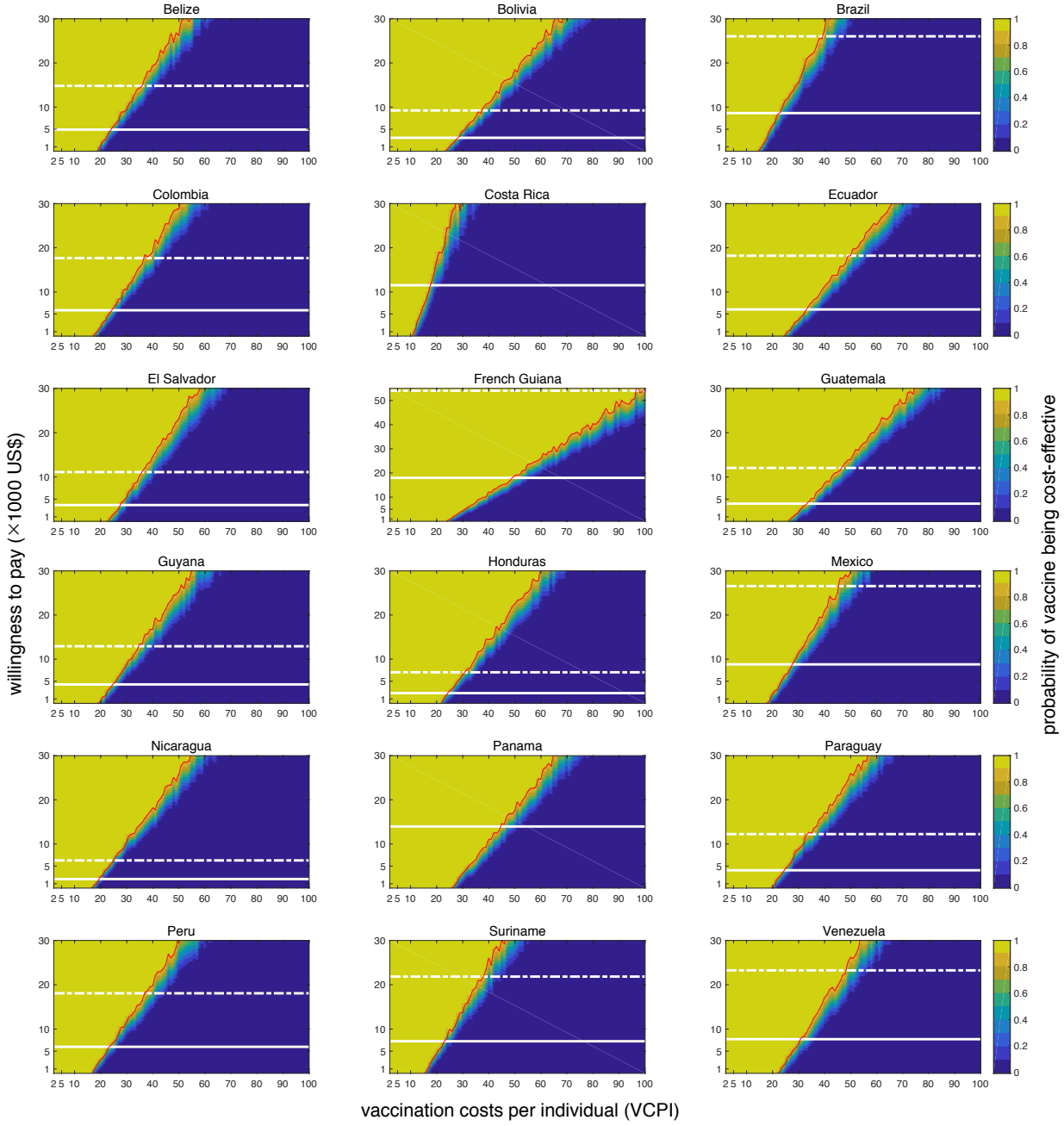


Figure 6.12: Probabilities of vaccine being cost-effective in 18 Latin American countries for a range of VCPI and willingness-to-pay. Solid line represents the willingness-to-pay threshold corresponding to the average of per capita GDP of each country in 2015 and 2016. Dashed line represents three times the average of per capita GDP of each country. The red curve represents the 90% probability of vaccine being cost-effective for a given VCPI

ranged between these values (Figure 6.14). Similarly, using three times the per-capita GDP, the

Table 6.6: Upper range of VCPI (US dollar) for a Zika vaccine candidate to be cost-saving (ICER<0), very cost-effective (threshold of the per capita GDP) or cost-effective (threshold of 3× the per capita GDP).

		Cost Saving	**Very Cost-Effective**		**Cost-Effective**	
Country	*Herd immunity*	VCPI	*GDP*	*VCPI*	*3×GDP*	*VCPI*
Belize	21%	$18	$4955	$23	$14865	$34
Bolivia	10%	$22	$3097	$27	$9291	$36
Brazil	18%	$14	$8694	$21	$26082	$38
Colombia	12%	$16	$5900	$23	$17700	$35
Costa Rica	2%	$10	$11563	$16	$34689	$29
Ecuador	8%	$24	$6084	$32	$18252	$48
El Salvador	16%	$22	$3719	$26	$11157	$34
French Guiana	18%	$23	$18036	$47	$54108	$96
Guatemala	14%	$25	$4032	$32	$12096	$45
Guyana	15%	$18	$4325	$23	$12975	$33
Honduras	14%	$21	$2358	$23	$7074	$29
Mexico	5%	$17	$8867	$26	$26601	$44
Nicaragua	17%	$16	$2109	$18	$6327	$24
Panama	15%	$25	$14009	$43	$42027	$82
Paraguay	17%	$19	$4094	$23	$12282	$32
Peru	4%	$16	$6042	$22	$18126	$35
Suriname	22%	$14	$7298	$21	$21894	$37
Venezuela	19%	$21	$7766	$29	$23298	$47

vaccine is still cost-effective for a VCPI up to $26 in Nicaragua and up to $98 in French Guiana (Figure 6.14). In the scenario with decreased attack rates, the results show that vaccination was very cost-effective with a VCPI of $4 or less in Mexico (mean incremental cost of $3054/DALY averted) and $41 or less in French Guiana (mean incremental cost of $15037/DALY averted), with other countries having an upper VCPI value in this range (Figure 6.15). The median percentage reduction of microcephaly in these scenarios exceeded 75% with vaccination (Technical Appendix, Figure A8). Summaries of the cost-effectiveness analysis for both scenarios of higher and lower attack rates are provided in Table 6.9.

Table 6.7: Mean ICER values with 95% confidence intervals corresponding to VCPI values under which vaccination program is at least 90% cost-effective in each country. The per capita GDP and three times the per capita GDP were used as thresholds for very cost-effective and cost-effective analysis, respectively. The dollar values in '()' indicates that the 95% CI extends to negative ICER values, which is considered cost-saving.

	Very cost-effective			**Cost-effective**		
Country	*VCPI*	*ICER*	*95% CI*	*VCPI*	*ICER*	*95% CI*
Belize	$23	$3,516	$144–$4,575	$34	$12,092	$7,379–$15,050
Bolivia	$27	$1,827	$(872)–$2,669	$36	$7,038	$4,249–$9,745
Brazil	$21	$6,356	$1,596–$7,223	$38	$21,725	$14,938–$27,441
Colombia	$23	$4,184	$1,284–$5,349	$35	$14,086	$9,447–$16,736
Costa Rica	$16	$7,352	$1,280–$9,234	$29	$29,061	$15,459–$30,561
Ecuador	$32	$4,451	$1,343–$5,560	$48	$15,581	$10,338–$17,576
El Salvador	$26	$1,379	$(1,884)–$2,826	$34	$8,177	$3,408–$9,785
French Guiana	$47	$14,475	$10,016–$16,653	$96	$49,934	$36,523–$53,661
Guatemala	$32	$2,544	$148–$3,944	$45	$9,786	$6,556–$11,859
Guyana	$23	$2,270	$(285)–$3,717	$33	$10,034	$5,884–$12,262
Honduras	$23	$892	$(1,711)–$1,705	$29	$4,992	$1,623–$6,142
Mexico	$26	$6,362	$2,564–$7,445	$44	$21,652	$14,717–$24,875
Nicaragua	$18	$595	$(1,465)–$1,231	$24	$4,829	$2,395–$6,068
Panama	$43	$11,001	$7,016–$13,486	$82	$37,247	$29,096–$43,898
Paraguay	$23	$2,348	$(305)–$3,332	$32	$9,903	$5,028–$10,670
Peru	$22	$4,332	$1,087–$4,870	$35	$14,028	$9,262–$16,432
Suriname	$21	$4,434	$1,505–$6,235	$37	$18,705	$12,714–$22,331
Venezuela	$29	$4,697	$623–$6,590	$47	$19,170	$13,160–$23,579

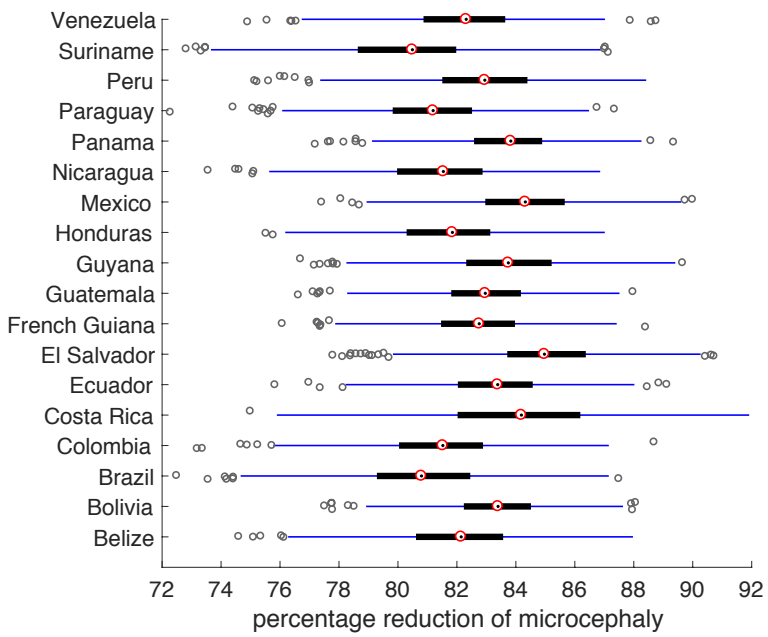


Figure 6.13: Boxplots for percentage reduction of microcephaly due to vaccination. The median is shown by the red circle.

Table 6.8: Attack rates for additional simulation scenarios. Attack rates were increased and decreased by 4%, with a lower bound of 1%. The model was calibrated to country-specific attack rate.

	Main Scenario	**Additional Scenarios**	
Country	**Attack rate (AR)**	**AR + 4%**	**AR − 4%**
Belize	21%	25%	17%
Bolivia	10%	14%	6%
Brazil	18%	22%	14%
Colombia	12%	16%	8%
Costa Rica	2%	6%	1%
Ecuador	8%	12%	4%
El Salvador	16%	20%	12%
French Guiana	18%	22%	14%
Guatemala	14%	18%	10%
Guyana	15%	19%	11%
Honduras	14%	18%	10%
Mexico	5%	9%	1%
Nicaragua	17%	21%	13%
Panama	15%	19%	11%
Paraguay	17%	21%	13%
Peru	4%	8%	1%
Suriname	22%	26%	18%
Venezuela	19%	23%	15%

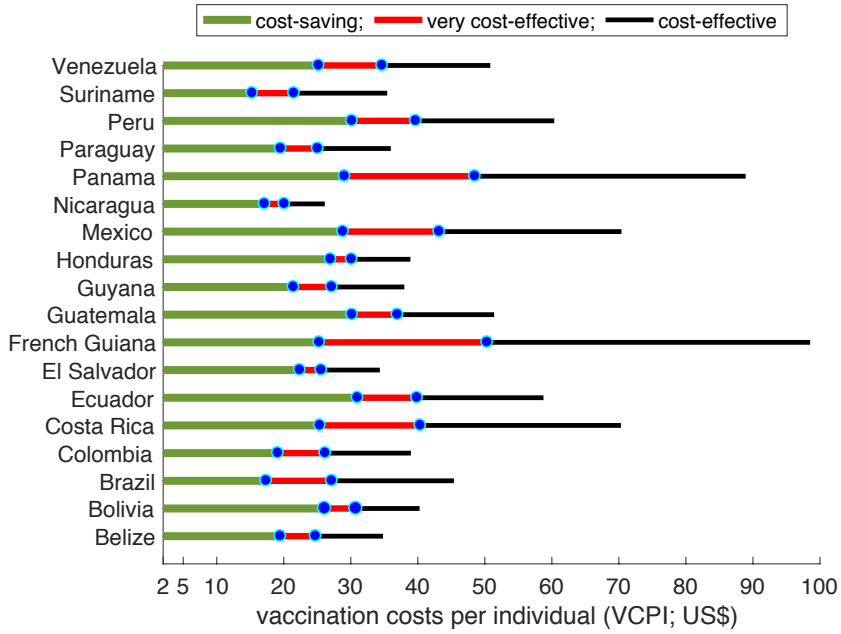


Figure 6.14: Upper range of VCPI (US dollar) for the scenarios of cost-saving (green), very cost-effective (red), and cost-effective (black). Estimates correspond to simulations calibrated to an increase of 4% in estimated attack rates for the 2015-2017 outbreaks.

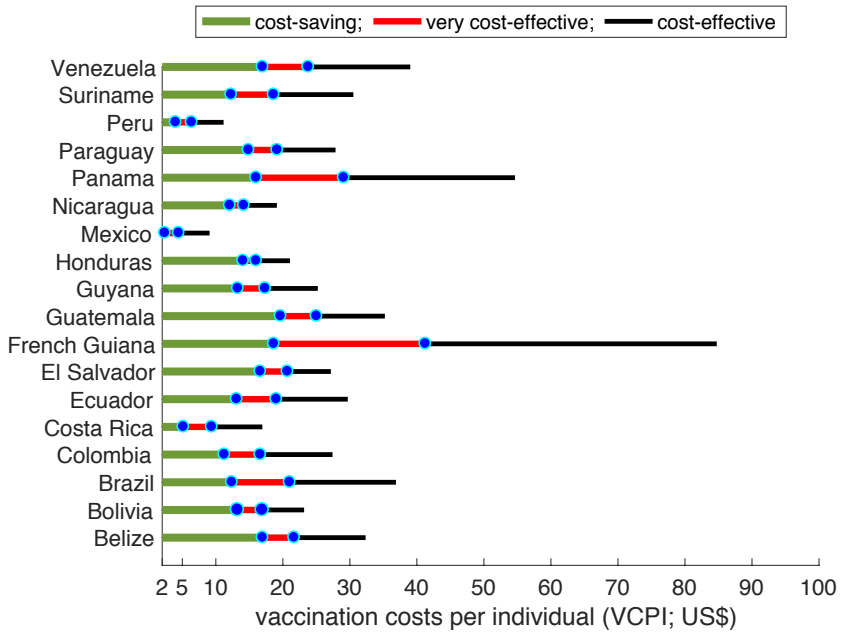


Figure 6.15: Upper range of VCPI (US dollar) for the scenarios of cost-saving (green), very cost-effective (red), and cost-effective (black). Estimates correspond to simulations calibrated to an decrease of 4% in estimated attack rates for the 2015-2017 outbreaks.

Table 6.9: Mean ICER values with 95% confidence intervals corresponding to VCPI values under which vaccination program is at least 90% cost-effective in each country. The per capita GDP were used as thresholds for cost-effective analysis. The dollar values in '()' indicates that the 95% CI extends to negative ICER values, which is considered cost-saving.

Simulations calibrated to an increase of 4% in estimated attack rates.

	Very cost-effective			**Cost-effective**		
Country	*VCPI*	*ICER*	*95% CI*	*VCPI*	*ICER*	*95% CI*
Belize	$24	$2689	$(1194)–$3773	$34	$10892	$6136–$14590
Bolivia	$30	$1189	$(1334)–$2164	$40	$6920	$4089–$8997
Brazil	$27	$6589	$2553–$7720	$45	$21841	$15274–$26353
Colombia	$26	$4181	$1458–$5539	$38	$13721	$9371–$16483
Costa Rica	$40	$9072	$4951–$12098	$70	$30013	$22287–$33976
Ecuador	$39	$3618	$973–$5276	$58	$15088	$10878–$18087
El Salvador	$25	$1098	$(2753)–$2733	$34	$7545	$2781–$10230
French Guiana	$50	$14914	$10328–$18865	$98	$49466	$34961–$53192
Guatemala	$36	$2197	$(200)–$3521	$51	$10076	$6620–$11936
Guyana	$27	$2691	$(250)–$4032	$37	$9665	$5907–$11632
Honduras	$30	$1078	$(1445)–$1723	$38	$5439	$2953–$6806
Mexico	$43	$7099	$4304–$8866	$70	$23159	$18270–$27829
Nicaragua	$20	$1067	$(757)–$1789	$26	$5069	$2673–$6063
Panama	$48	$10427	$6843–$13151	$88	$34894	$27744–$42041
Paraguay	$25	$2662	$5–$3705	$35	$9702	$5960–$11045
Peru	$39	$4398	$1577–$5465	$60	$15565	$11540–$17911
Suriname	$21	$4820	$798–$6335	$35	$17716	$11223 –$21123
Venezuela	$34	$4820	$1944–$7838	$51	$19982	$11823–$21092

Simulations calibrated to an increase of 4% in estimated attack rates.

	Very cost-effective			**Cost-effective**		
Country	*VCPI*	*ICER*	*95% CI*	*VCPI*	*ICER*	**95% CI**
Belize	$21	$2344	$(812)–$3581	$32	$12128	$7102–$14, 544
Bolivia	$16	$909	$(1459)–$2077	$23	$7207	$3196–$8, 751
Brazil	$21	$6720	$2642–$8089	$36	$20704	$14484–$24808
Colombia	$16	$3465	$266–$4008	$27	$14476	$9076–$17082
Costa Rica	$9	$6661	$(741)–$8037	$18	$25476	$12133–$38507
Ecuador	$19	$4241	$688–$5265	$29	$13608	$8413–$16282
El Salvador	$20	$1183	$(1846)–$2852	$27	$8404	$3222–$9843
French Guiana	$41	$15037	$10339–$17905	$84	$48232	$37689–$57894
Guatemala	$25	$2445	$(447)–$3601	$35	$9639	$5399–$11411
Guyana	$17	$2130	$(1099)–$3429	$25	$10149	$5292–$13610
Honduras	$16	$946	$(1896)–$1676	$21	$5276	$1658–$7219
Mexico	$4	$3054	$(5722)–$2798	$9	$19550	$3620–$23927
Nicaragua	$14	$802	$(1638)–$1335	$19	$4798	$2246–$6295
Panama	$29	$11311	$5967–$13785	$54	$34281	$24242–$41282
Paraguay	$19	$2627	$(29)–$3344	$27	$8492	$5258–$10724
Peru	$6	$2594	$(2114)–$2779	$11	$13487	$3063–$17903
Suriname	$18	$5057	$1164–$6269	$30	$16836	$10634–$20560
Venezuela	$23	$4915	$808–$6501	$39	$19481	$12735–$23902

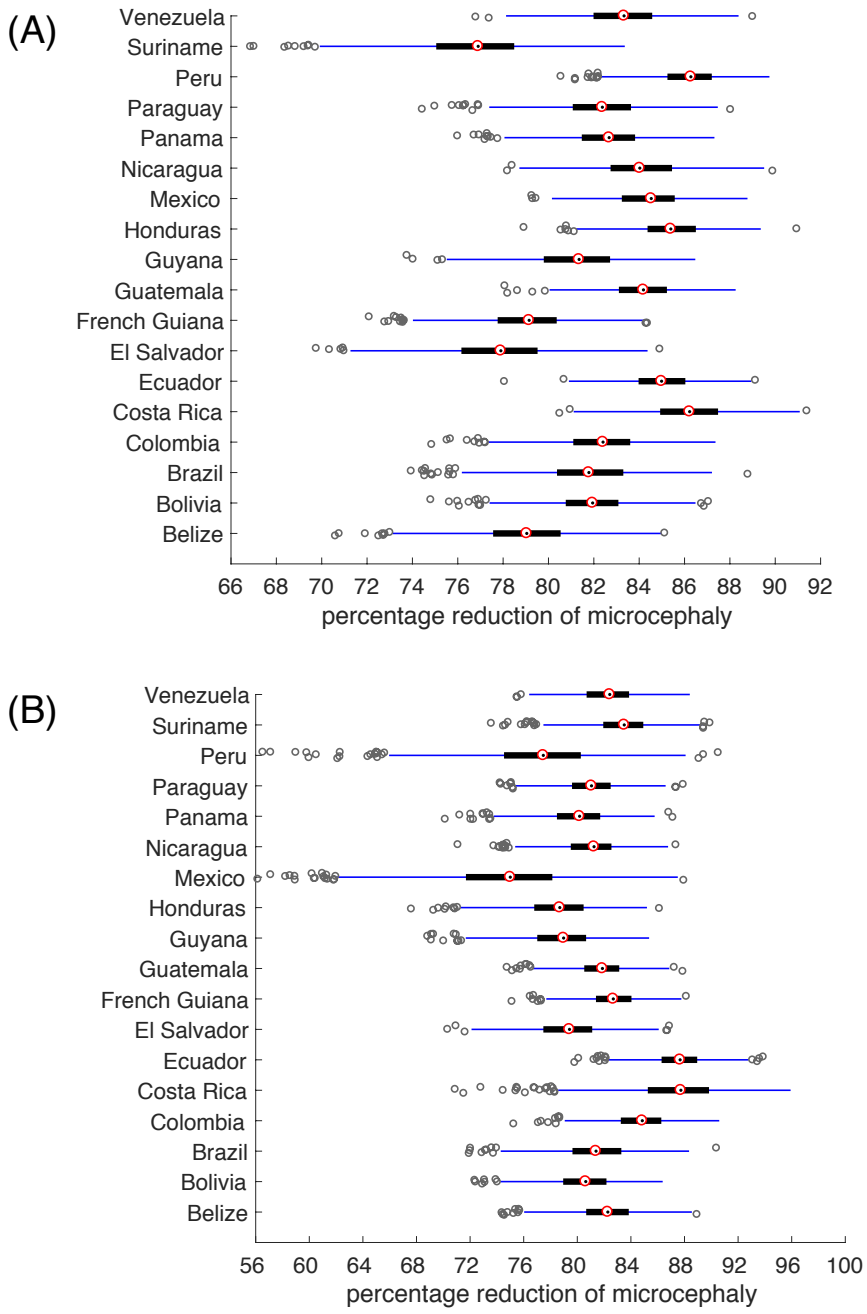


Figure 6.16: Boxplots for the percentage reduction of microcephaly due to vaccination for scenarios: (A) an increase of 4%; and (B) a decrease of 4% to estimated attack rates for the 2015–2017 outbreaks. The median is shown by the red circle.

6.5 Discussion

In this chapter, we evaluated the cost-effectiveness of a Zika vaccine candidate from a government perspective under a number of plausible scenarios. We utilized the comprehensive ABM developed in Chapter 5 by extending it to include vaccination dynamics. Our analysis considered: (i) in a Colombian population setting where the model was calibrated to estimated basic reproduction number $\mathcal{R}_0$ of Colombia, taking into account existing herd-immunity from previous outbreaks and relative transmission of asymptomatic infection, and (ii) 17 other countries in the Americas where the model was calibrated to attack rates estimated for the 2015-2017 outbreaks. Our analysis determined a range of VCPI within which vaccination is cost-saving, very cost-effective, and cost-effective for these countries. Although a number of factors (e.g., the level of pre-existing herd immunity, attack rate, costs associated with the management of Zika infection and its outcomes, and the WTP) are critical in determining VCPI for cost-effectiveness, our results show that targeted vaccination of women of reproductive age would be cost-effective, and even cost-saving, in all countries studied here if VCPI is sufficiently low. Furthermore, vaccination with a protection efficacy in the range 60% – 90% significantly reduces the incidence of microcephaly, with a median percentage reduction that exceeds 75% in simulated scenarios.

Cost-effectiveness analysis was based on using direct medical cost estimates associated with the treatment of symptomatic Zika infection, GBS cases, and long-term neurological sequelae caused by microcephaly condition. Although the likelihood of cost-effectiveness was shown to be sensitive to willingness-to-pay and vaccination costs, the largest range of VCPI for cost-effectiveness corresponded to scenarios in which the population is fully susceptible or the effect of other interventions to blunt ZIKV transmission is relatively low. However, non-pharmaceutical measures (including vector control programs), increased access to contraception [229], and pre-existing herd effects as a result of naturally acquired immunity in previous outbreaks could decrease the range of VCPI for cost-effectiveness, requiring a significantly higher willingness-to-pay for vaccination to prove cost-effective. Previous work suggests that a prophylactic vaccine with a protection efficacy of 75% reduces the incidence of prenatal infections by at least 94% if 90% of women of reproductive age are vaccinated [224] These estimates are higher than what our model predicts (with a median percentage reduction between 75% and 88%) in similar scenarios, which is expected given the deterministic nature of model used in the previous study [224]. Nevertheless, the findings indicate that targeted vaccination is an important preventive measure for mitigating the impact of ZIKV infection in future outbreaks.

The strength of our study relies on the evaluation of cost-effectiveness for countries affected by Zika with estimated attack rates exceeding 2% within a single modelling framework. Our analysis was based on an individual-level stochastic approach, accounting for parameter uncertainty and heterogeneities in disease transmission. Due to its dynamic nature, the simulation model also takes into account the accruing herd immunity during the epidemic that results from the indirect protection effects of naturally acquired immunity in the population.

The results presented in this chapter should be considered within the context of study limitations. First, we note that our analysis was based on estimates of attack rates during the 2015-2017 ZIKV outbreaks in the Latin and South American countries [52, 217, 227], which were regarded as the level of pre-existing herd immunity in the simulations. Should this level fall at the time of vaccine availability in future outbreaks, the expected changes in the VCPI range for cost-effectiveness require further analysis. Second, although the initial phase of clinical trials indicates high levels of neutralizing antibodies [215, 216], the range of vaccine efficacy is not ascertained, and our estimates relied on the assumption that a single dose of vaccine would provide a protection efficacy of 60% to 90%. The efficacy data can also provide information on the number of vaccine doses required, which would affect the vaccination costs per individual. We also assumed that the risk of microcephaly is independent of vaccine-induced immunity in a vaccinated pregnant woman if infection occurred. In the absence of pre-existing immunity, clinical and epidemiological studies indicate that a significant portion (up to 80%) of ZIKV-infected individuals experience asymptomatic infection without presenting clinical symptoms. We assumed that vaccine-induced immunity further reduces the chance of clinical manifestation (if infection occurred), and therefore considered infection following vaccination to be asymptomatic.

We assumed that during the epidemic, pregnant women are vaccinated (with a coverage of 80%) early in their first trimester that is associated with the highest risk of microcephaly. Yet, we understand that due to various factors, including access to healthcare resources and late recognition of pregnancy, vaccination may not occur prior to any potential ZIKV infection during pregnancy. The risk of microcephaly following vaccination was not altered if infection occurred but the disease was considered to be asymptomatic. However, in terms of costs associated with microcephaly (which dominate), we expect the results of cost-effectiveness analysis to hold because we did not alter the risk of microcephaly in the presence of vaccine-induced immunity in pregnant women. Without the outcomes of clinical trials, our model did not consider the possible adverse side effects of vaccination and their associated costs. Although, other neurological disorders have been reported in association with ZIKV infection (including encephalitis, meningoencephalitis,

myelitis, and optical neuritis), we considered only microcephaly and GBS outcomes. Finally, in the context of cost-effectiveness analysis from a government perspective, our analysis excluded indirect costs such as loss of productivity and earnings in families inflicted by microcephaly and GBS, yet we understand that the lifetime indirect costs related to the care of children with microcephaly could be substantial. The validation of the above assumptions requires efficacy data from clinical trials, which are currently lacking.

Despite these limitations that merit further investigation as relevant information and data become available, we have provided estimates for Zika vaccine cost-effectiveness to inform decision makers for the implementation of a targeted vaccination program. The findings suggest that a vaccine has the potential to significantly reduce the health and economic burden of ZIKV infection in at-risk populations.

Chapter 7

Closing Remarks and Future Directions

Cost-effective analysis, as part of health economics, is widely used in many settings across the world to evaluate the potential impact of healthcare interventions and technologies, often in the face of budget constraints, scarce resources, and even ethical considerations. Traditional economic evaluations often utilize clinical trials and observational studies which are inadequate in many aspects to support decision-making processes, largely because of limited follow-up time-horizon and small sample sizes. In order to address these limitations for the long-term consequences of health intervention strategies, mathematical and computational models are needed to extrapolate results to a larger scale and a longer timeframe. The choice on the type of model, including explicit and implicit assumptions on system dynamics and parameters, depends on the context of the intervention or technology [230]. Incorrect assumptions or an inappropriate choice of model can lead to suboptimal decisions at best, with potential negative consequences in the quality of life or wasted resources. We found that most economic evaluations are conducted using aggregate frameworks such as state-transition (Markov) models or compartmental models with differential equations [231]. While these approaches span a vast literature, they often impose restrictive limitations such as homogeneous populations and linear interactions, which hinders their application to communicable disease dynamics. Moreover, aggregate models often adopt assumptions that are inaccurate or inadequate to faithfully represent the underlying system dynamics and may generate spurious results.

In this thesis, we described a more comprehensive modelling framework in which traditional cost-effectiveness analysis is integrated with an ABM computational system to overcome the limitations of other methodologies. ABM can adopt less stringent assumptions and have a "bottom-up" approach in which system dynamics are generated as a result of modelling at the individual level. For example, ABM can capture indirect effects such as herd immunity without explicitly having to model them. This is an important characteristic that is overlooked in many Markov models with cohort structure for evaluating vaccination effects. As a result, the data-generating process of an ABM captures non-linear interactions among agents and feedback loops. In our framework, we

integrate this data-generating process into cost-effectiveness analysis, thus avoiding the inherent difficulties of traditional methods, such as censored data and homogeneity.

In order to test the robustness of our framework, we considered two case studies: (i) a human-to-human infection transmission (i.e., Haemophilus infuenzae) model and (ii) a vector-borne disease (i.e., Zika) model. In each case, we developed a disease-specific agent-based model to determine how a potential vaccine candidate may affect the epidemic dynamics, but also present their first cost-effectiveness analysis and implications for vaccination strategies in different population settings. For a routine vaccination against Haemophilus infuenzae serotype 'a', our analysis suggests a significant reduction in costs by the tenth year of the vaccination program, with significant decreases across all cost categories, including immediate hospitalization and long-term disability. The total costs of the immunization program are significantly lower than the costs required to provide life-time care of severely debilitated survivors of invasive Hia disease. Similarly, for a potential vaccination against Zika virus, our analysis shows that targeted vaccination of women of reproductive age would be cost-effective (and even cost-saving) for sufficiently low vaccine costs per individual in many countries in the Americas affected by previous Zika outbreaks. We also found that a Zika vaccine can provide substantial benefits by reducing the incidence of microcephaly, which leads to life-long sequelae with reduced quality of life.

Given the results of our case studies, we believe this framework advances the research efforts in understating the mechanisms of disease processes and evaluating the potential impact of health intervention strategies in a more systematic manner, and with a higher level of realism and flexibility. Our results show that the empirical reliability generated by our model can be far more informative than traditional aggregate models. As part of further studies in near-term research efforts, this framework can be extended to include a number of other important investigations in health economics of interventions or technologies, such as:

- Cost-benefit analysis, which examines both costs and consequences in monetary terms.
- Cost-utility analysis, which examines costs and a single consequence in the form of a health-related quality of life measure.
- Cost-minimization analysis, which examines the least costly consequence among alternatives with equivalent consequences.
- Cost-consequence analysis, which examines the costs and multiple consequences in their natural units without aggregation into a single consequence.

- Input cost analysis, which examines the costs of all alternatives but not their consequences.
- Cost-related outcome analysis, which examines the consequences of all alternatives in monetary terms but not the input costs incurred.

Despite the growing application of ABM in various fields of research beyond health economics, a number of limitations exists that need to be addressed in future work. For example, while the computational aspects of ABM have been detailed in existing literature, the underlying theoretical basis has rarely been used in its construction. As a consequence, the advantage of ABM to capture realistic features of real-world phenomena is undermined by the lack of dynamical systems tools for their analysis. In §2.4, we introduced a brief mathematical formalism in terms of recursive functions, but little is gained in terms of analytical power without methods like bifurcation and stability analyses that are well-established for investigating differential equation-based models. Recently, however, it has been established that the basic mathematical nature of many agent-based models can be derived from a sequential dynamical system. This characterization makes ABM amenable to powerful symbolic analysis, in which the language and tools of mathematical theory are used to examine the system. Although it is beyond the scope of this thesis, the theory of sequential dynamical systems and their application to ABM will be a subject of future studies within the framework proposed here.

Bibliography

[1] Thomas C. Schelling. "Dynamic Models of Segregation". In: *The Journal of Mathematical Sociology* 1.2 (July 1971), pp. 143–186. ISSN: 0022-250X, 1545-5874. DOI: 10.1080/0022250X.1971.9989794.

[2] Shu-Heng Chen, Chia-Ling Chang, and Ye-Rong Du. "Agent-Based Economic Models and Econometrics". In: *The Knowledge Engineering Review* 27.02 (June 2012), pp. 187–219. ISSN: 0269-8889, 1469-8005. DOI: 10.1017/S0269888912000136.

[3] Michael J. North et al. "Multiscale Agent-Based Consumer Market Modeling". In: *Complexity* (2010), NA–NA. ISSN: 10762787, 10990526. DOI: 10.1002/cplx.20304.

[4] M. Cristelli, L. Pietronero, and A. Zaccaria. "Critical Overview of Agent-Based Models for Economics". In: (Jan. 10, 2011). arXiv: 1101.1847 [physics, q-fin]. URL: http://arxiv.org/abs/1101.1847 (visited on 02/01/2019).

[5] Donald L. DeAngelis and Volker Grimm. "Individual-Based Models in Ecology after Four Decades". In: *F1000Prime Reports* 6 (June 2, 2014). ISSN: 20517599. DOI: 10.12703/P6-39.

[6] Judith A. Effken et al. "Simulating Nursing Unit Performance With OrgAhead: Strengths and Challenges". In: *CIN: Computers, Informatics, Nursing* 30.11 (Nov. 2012), pp. 620–626. ISSN: 1538-2931. DOI: 10.1097/NXN.0b013e318261f1bb.

[7] Michael W. Macy and Robert Willer. "From Factors to Actors: Computational Sociology and Agent-Based Modeling". In: *Annual Review of Sociology* 28 (2002), pp. 143–166. ISSN: 03600572, 15452115. JSTOR: 3069238.

[8] Daniel G. Brown et al. "Path Dependence and the Validation of Agent-based Spatial Models of Land Use". In: *International Journal of Geographical Information Science* 19.2 (Feb. 2005), pp. 153–174. ISSN: 1365-8816, 1362-3087. DOI: 10.1080/13658810410001713399.

[9] William A. Brock and Cars H. Hommes. "Heterogeneous Beliefs and Routes to Chaos in a Simple Asset Pricing Model". In: *Journal of Economic Dynamics and Control* 22.8-9 (July 1998), pp. 1235–1274. ISSN: 01651889. DOI: 10.1016/S0165-1889(98)00011-6.

[10] Andrew Ilachinski. *Irreducible Semi-Autonomous Adaptive Combat (ISAAC): An Artificial-Life Approach to Land Warfare.* Aug. 1997. URL: http://www.dtic.mil/dtic/tr/fulltext/u2/a362371.pdf (visited on 07/19/2018).

[11] James Moffat, Josephine Smith, and Susan Witty. "Emergent Behaviour: Theory and Experimentation Using the MANA Model". In: *Journal of Applied Mathematics and Decision Sciences* 2006 (Oct. 2, 2006), pp. 1–13. ISSN: 1173-9126, 1532-7612. DOI: 10.1155/JAMDS/2006/54846.

[12] Raymond R. Hill, Lance E. Champagne, and Joseph C. Price. "Using Agent-Based Simulation and Game Theory to Examine the WWII Bay of Biscay U-Boat Campaign". In: *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* 1.2 (Apr. 2004), pp. 99–109. ISSN: 1548-5129, 1557-380X. DOI: 10.1177/875647930400100204.

[13] Michael J North. "A Theoretical Formalism for Analyzing Agent-Based Models". In: *Complex Adaptive Systems Modeling* 2.1 (2014), p. 3. ISSN: 2194-3206. DOI: 10.1186/2194-3206-2-3.

[14] E. Bonabeau. "Agent-Based Modeling: Methods and Techniques for Simulating Human Systems". In: *Proceedings of the National Academy of Sciences* 99 (Supplement 3 May 14, 2002), pp. 7280–7287. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.082080899.

[15] C M Macal and M J North. "Tutorial on Agent-Based Modelling and Simulation". In: *Journal of Simulation* 4.3 (Sept. 2010), pp. 151–162. ISSN: 1747-7778, 1747-7786. DOI: 10.1057/jos.2010.3.

[16] Sander van der Hoog. "Deep Learning in (and of) Agent-Based Models: A Prospectus". In: (June 20, 2017). arXiv: 1706.06302 [q-fin]. URL: http://arxiv.org/abs/1706.06302 (visited on 02/10/2019).

[17] Alfred J. Lotka. "Contribution to the Theory of Periodic Reactions". In: *The Journal of Physical Chemistry* 14.3 (Jan. 1909), pp. 271–274. ISSN: 0092-7325, 1541-5740. DOI: 10.1021/j150111a004.

[18] W. O. Kermack and A. G. McKendrick. "A Contribution to the Mathematical Theory of Epidemics". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 115.772 (Aug. 1, 1927), pp. 700–721. ISSN: 1364-5021, 1471-2946. DOI: 10.1098/rspa.1927.0118.

[19] Joshua M. Epstein. "Remarks on the Foundations of Agent-Based Generative Social Science". In: *Handbook of Computational Economics*. Vol. 2. Elsevier, 2006, pp. 1585–1604. ISBN: 978-0-444-51253-6. DOI: 10.1016/S1574-0021(05)02034-4. URL: http://linkinghub.elsevier.com/retrieve/pii/S1574002105020344 (visited on 08/01/2018).

[20] Domenico Delli Gatti et al., eds. *Agent-Based Models: A Toolkit*. 1st ed. Cambridge University Press, Mar. 22, 2018. ISBN: 978-1-108-41499-9 978-1-108-22727-8 978-1-108-40004-6. DOI: 10.1017/9781108227278. URL: https://www.cambridge.org/core/product/identifier/9781108227278/type/book (visited on 01/06/2019).

[21] Matteo Richiardi. "Agent-Based Models as Recursive Systems". In: *Agent-Based Models*. Ed. by Domenico Delli Gatti et al. 1st ed. Cambridge University Press, Mar. 22, 2018, pp. 33–42. ISBN: 978-1-108-41499-9 978-1-108-22727-8 978-1-108-40004-6. DOI: 10.1017/9781108227278.004. URL: https://www.cambridge.org/core/product/identifier/CBO9781108227278A027/type/book_part (visited on 01/21/2019).

[22] Reinhard Laubenbacher et al. "A Mathematical Formalism for Agent-Based Modeling". In: *arXiv preprint arXiv:0801.0249* (2007).

[23] Alan Veliz-Cuba, Abdul Salam Jarrah, and Reinhard Laubenbacher. "Polynomial Algebra of Discrete Models in Systems Biology". In: *Bioinformatics* 26.13 (July 1, 2010), pp. 1637–1643. ISSN: 1460-2059, 1367-4803. DOI: 10.1093/bioinformatics/btq240.

[24] Sven Banisch and Ricardo Lima. "Markov Chain Aggregation for Simple Agent-Based Models on Symmetric Networks: The Voter Model". In: (Sept. 18, 2012). arXiv: 1209.3902 [nlin, physics:physics]. URL: http://arxiv.org/abs/1209.3902 (visited on 01/21/2019).

[25] Sven Banisch, Ricardo Lima, and Tanya Araújo. "Agent Based Models and Opinion Dynamics as Markov Chains". In: (Aug. 8, 2011). arXiv: 1108.1716 [physics]. URL: http://arxiv.org/abs/1108.1716 (visited on 01/16/2019).

[26] Luis R. Izquierdo et al. "Techniques to Understand Computer Simulations: Markov Chain Analysis". In: *Journal of Artificial Societies and Social Simulation* 12.1 (2009), pp. 1–6. URL: https://ideas.repec.org/a/jas/jasssj/2008-19-2.html.

[27] Herbert Gintis. "Markov Models of Social Dynamics: Theory and Applications". In: *ACM Transactions on Intelligent Systems and Technology* 4.3 (June 1, 2013), p. 1. ISSN: 21576904. DOI: 10.1145/2483669.2483686.

[28] Thomas Lux. "Estimation of Agent-Based Models Using Sequential Monte Carlo Methods". In: *Journal of Economic Dynamics and Control* 91 (June 2018), pp. 391–408. ISSN: 01651889. DOI: 10.1016/j.jedc.2018.01.021.

[29] Volker Grimm et al. “A Standard Protocol for Describing Individual-Based and Agent-Based Models”. In: *Ecological Modelling* 198.1-2 (Sept. 2006), pp. 115–126. ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2006.04.023.

[30] Robert Axtell. “1 Why Agents ? On the Varied Motivations for Agent Computing in the Social Sciences”. In: 2000.

[31] Jim Doran. “Agent Design for Agent-Based Modelling”. In: *Agent-Based Computational Modelling*. Ed. by Francesco C. Billari et al. Heidelberg: Physica-Verlag, 2006, pp. 215–223. ISBN: 978-3-7908-1640-2. DOI: 10.1007/3-7908-1721-X_11. URL: http://link.springer.com/10.1007/3-7908-1721-X_11 (visited on 12/26/2018).

[32] Matteo G. Richiardi. “Agent-Based Computational Economics: A Short Introduction”. In: *The Knowledge Engineering Review* 27.02 (June 2012), pp. 137–149. ISSN: 0269-8889, 1469-8005. DOI: 10.1017/S0269888912000100.

[33] Andrew T. Crooks and Alison J. Heppenstall. “Introduction to Agent-Based Modelling”. In: *Agent-Based Models of Geographical Systems*. Ed. by Alison J. Heppenstall et al. Dordrecht: Springer Netherlands, 2012, pp. 85–105. ISBN: 978-90-481-8926-7 978-90-481-8927-4. DOI: 10.1007/978-90-481-8927-4_5. URL: http://www.springerlink.com/index/10.1007/978-90-481-8927-4_5 (visited on 12/14/2018).

[34] Marek Laskowski and Seyed M. Moghadas. “A General Framework for Agent-Based Modelling with Applications to Infectious Disease Dynamics.” In: *Biomat 2013*. International Symposium on Mathematical and Computational Biology. World Scientific, May 2014, pp. 318–339. ISBN: 978-981-4602-21-1 978-981-4602-22-8. DOI: 10.1142/9789814602228_0019.

[35] Dirk Helbing, ed. *Social Self-Organization*. Understanding Complex Systems. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. ISBN: 978-3-642-24003-4 978-3-642-24004-1. DOI: 10.1007/978-3-642-24004-1. URL: http://link.springer.com/10.1007/978-3-642-24004-1 (visited on 01/17/2019).

[36] Alexis Drogoul, Fabien Michel, and Jacques Ferber. “Multi-Agent Systems and Simulation: A Survey from the Agent Community’s Perspective”. In: *Multi-Agent Systems*. Ed. by Danny Weyns and Adelinde Uhrmacher. Vol. 20091861. CRC Press, June 3, 2009, pp. 3–51. ISBN: 978-1-4200-7023-1 978-1-4200-7024-8. DOI: 10.1201/9781420070248.pt1. URL: http://www.crcnetbase.com/doi/abs/10.1201/9781420070248.pt1 (visited on 01/17/2019).

[37] Davide Secchi. "Agent-Based Models of Bounded Rationality". In: *Team Performance Management: An International Journal* 23.1/2 (Mar. 14, 2017), pp. 2–12. ISSN: 1352-7592. DOI: 10.1108/TPM-12-2016-0052.

[38] Stewart Robinson. *Simulation: The Practice of Model Development and Use*. Chichester, West Sussex, England ; Hoboken, NJ: John Wiley & Sons, Ltd, 2004. 316 pp. ISBN: 978-0-470-84772-5.

[39] Robert Ernest Marks. "Validating Simulation Models: A General Framework and Four Applied Examples". In: *Computational Economics* 30.3 (Sept. 20, 2007), pp. 265–290. ISSN: 0927-7099, 1572-9974. DOI: 10.1007/s10614-007-9101-7.

[40] Giorgio Fagiolo, Paul Windrum, and Alessio Moneta. "Empirical Validation of Agent-Based Models: Alternatives and Prospects". In: *Journal of Artificial Societies and Social Simulation* 10 (May 2007).

[41] Matteo Richiardi. "Estimation of Agent-Based Models". In: *Agent-Based Models*. Ed. by Domenico Delli Gatti et al. 1st ed. Cambridge University Press, Mar. 22, 2018, pp. 183–221. ISBN: 978-1-108-41499-9 978-1-108-22727-8 978-1-108-40004-6. DOI: 10.1017/9781108227278.010. URL: https://www.cambridge.org/core/product/identifier/CBO9781108227278A064/type/book_part (visited on 01/30/2019).

[42] Jakob Grazzini and Matteo Richiardi. "Estimation of Ergodic Agent-Based Models by Simulated Minimum Distance". In: *Journal of Economic Dynamics and Control* 51 (Feb. 2015), pp. 148–165. ISSN: 01651889. DOI: 10.1016/j.jedc.2014.10.006.

[43] Jakob Grazzini, Matteo G. Richiardi, and Mike Tsionas. "Bayesian Estimation of Agent-Based Models". In: *Journal of Economic Dynamics and Control* 77 (Apr. 2017), pp. 26–47. ISSN: 01651889. DOI: 10.1016/j.jedc.2017.01.014.

[44] Robert G. Sargent. "An Overview of Verification and Validation of Simulation Models". In: *Proceedings of the 19th Conference on Winter Simulation - WSC '87*. The 19th Conference. Atlanta, Georgia, United States: ACM Press, 1987, pp. 33–39. ISBN: 978-0-911801-32-3. DOI: 10.1145/318371.318379.

[45] Jakob Grazzini. "Analysis of the Emergent Properties: Stationarity and Ergodicity". In: *Journal of Artificial Societies and Social Simulation* 15.2 (2012). ISSN: 1460-7425. DOI: 10.18564/jasss.1929.

[46] Jakob Grazzini, Matteo Richiardi, and Lisa Sella. "The Agent-Based Experiment". In: *Agent-Based Models*. Ed. by Domenico Delli Gatti et al. 1st ed. Cambridge University Press, Mar. 22, 2018, pp. 143–162. ISBN: 978-1-108-41499-9 978-1-108-22727-8 978-1-108-40004-6. DOI: 10.1017/9781108227278.008. URL: https://www.cambridge.org/core/product/identifier/CBO9781108227278A053/type/book_part (visited on 01/28/2019).

[47] Peter C. B. Phillips and Zhijie Xiao. "A Primer on Unit Root Testing". In: *Journal of Economic Surveys* 12.5 (Dec. 1998), pp. 423–470. ISSN: 09500804. DOI: 10.1111/1467-6419.00064.

[48] Jennifer Badham et al. "Developing Agent-Based Models of Complex Health Behaviour". In: *Health & Place* 54 (Nov. 2018), pp. 170–177. ISSN: 13538292. DOI: 10.1016/j.healthplace.2018.08.022.

[49] Affan Shoukat, Thomas Vilches, and Seyed M. Moghadas. "Cost-Effectiveness of a Potential Zika Vaccine Candidate: A Case Study for Colombia". In: *BMC Medicine* 16.1 (Dec. 2018). ISSN: 1741-7015. DOI: 10.1186/s12916-018-1091-x.

[50] Affan Shoukat, Robert Van Exan, and Seyed M. Moghadas. "Cost-Effectiveness of a Potential Vaccine Candidate for Haemophilus Influenzae Serotype 'a'". In: *Vaccine* 36.12 (Mar. 2018), pp. 1681–1688. ISSN: 0264410X. DOI: 10.1016/j.vaccine.2018.01.047.

[51] Mehdi Najafi et al. "The Effect of Individual Movements and Interventions on the Spread of Influenza in Long-Term Care Facilities". In: *Medical Decision Making* 37.8 (2017), pp. 871–881.

[52] Qian Zhang et al. "Spread of Zika Virus in the Americas". In: *Proceedings of the National Academy of Sciences* 114.22 (May 30, 2017), E4334–E4343. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1620161114.

[53] Seyed M. Moghadas et al. "Asymptomatic Transmission and the Dynamics of Zika Infection". In: *Scientific Reports* 7.1 (Dec. 2017). ISSN: 2045-2322. DOI: 10.1038/s41598-017-05013-9.

[54] Srinivasan Venkatramanan et al. "Using Data-Driven Agent-Based Models for Forecasting Emerging Infectious Diseases". In: *Epidemics* 22 (2018), pp. 43–49.

[55] Melissa Tracy, Magdalena Cerdá, and Katherine M. Keyes. "Agent-Based Modeling in Public Health: Current Applications and Future Directions". In: *Annual Review of Public Health* 39.1 (Apr. 2018), pp. 77–94. ISSN: 0163-7525, 1545-2093. DOI: 10.1146/annurev-publhealth-040617-014317.

[56] Peter Zweifel, Friedrich Breyer, and Mathias Kifmann. *Health Economics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. ISBN: 978-3-540-27804-7 978-3-540-68540-1. DOI: 10.1007/978-3-540-68540-1. URL: http://link.springer.com/10.1007/978-3-540-68540-1 (visited on 02/11/2019).

[57] Gianluca Baio. *Bayesian Methods in Health Economics*. Chapman & Hall/CRC Biostatistics Series. OCLC: ocn835887421. Boca Raton: CRC Press, Taylor & Francis Group, 2013. 225 pp. ISBN: 978-1-4398-9555-9.

[58] William S. Weintraub and David J. Cohen. "The Limits of Cost-Effectiveness Analysis". In: *Circulation: Cardiovascular Quality and Outcomes* 2.1 (Jan. 2009), pp. 55–58. ISSN: 1941-7713, 1941-7705. DOI: 10.1161/CIRCOUTCOMES.108.812321.

[59] Shainoor J. Ismail et al. "Canada's National Advisory Committee on Immunization (NACI): Evidence-Based Decision-Making on Vaccines and Immunization". In: *Vaccine* 28 (Apr. 2010), A58–A63. ISSN: 0264410X. DOI: 10.1016/j.vaccine.2010.02.035.

[60] National Institute for Health and Care Excellence. *Guide to the Processes of Technology Appraisal*. Apr. 2018. URL: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/technology-appraisal-processes-guide-apr-2018.pdf.

[61] Joseph C. Gardiner, Cathy J. Bradley, and Marianne Huebner. "The Cost-Effectiveness Ratio in the Analysis of Health Care Programs". In: *Handbook of Statistics*. Vol. 18. Bioenvironmental and Public Health Statistics. Elsevier, Jan. 1, 2000, pp. 841–869. DOI: 10.1016/S0169-7161(00)18030-7. URL: http://www.sciencedirect.com/science/article/pii/S0169716100180307 (visited on 02/12/2019).

[62] Louise B. Russell et al. "The Role of Cost-Effectiveness Analysis in Health and Medicine". In: *JAMA* 276.14 (Oct. 9, 1996), pp. 1172–1177. ISSN: 0098-7484. DOI: 10.1001/jama.1996.03540140060028.

[63] Herbert E. Klarman, Jhon O??S. Francis, and Gerald D. Rosenthal. "Cost Effectiveness Analysis Applied to the Treatment of Chronic Renal Disease:" in: *Medical Care* 6.1 (Jan. 1968), pp. 48–54. ISSN: 0025-7079. DOI: 10.1097/00005650-196801000-00005.

[64] G. W. Torrance, W. H. Thomas, and D. L. Sackett. "A Utility Maximization Model for Evaluation of Health Care Programs". In: *Health Services Research* 7.2 (1972), pp. 118–133. ISSN: 0017-9124. pmid: 5044699.

[65] S. Fanshel and J. W. Bush. "A Health-Status Index and Its Application to Health-Services Outcomes". In: *Operations Research* 18.6 (Dec. 1970), pp. 1021–1066. ISSN: 0030-364X, 1526-5463. DOI: 10.1287/opre.18.6.1021.

[66] null Miyamoto. "Quality-Adjusted Life Years (QALY) Utility Models under Expected Utility and Rank Dependent Utility Assumptions". In: *Journal of Mathematical Psychology* 43.2 (June 1999), pp. 201–237. ISSN: 0022-2496. DOI: 10.1006/jmps.1999.1256. pmid: 10366516.

[67] Milton C. Weinstein, George Torrance, and Alistair McGuire. "QALYs: The Basics". In: *Value in Health* 12 (Mar. 2009), S5–S9. ISSN: 10983015. DOI: 10.1111/j.1524-4733.2009.00515.x.

[68] F. Sassi. "Calculating QALYs, Comparing QALY and DALY Calculations". In: *Health Policy and Planning* 21.5 (July 28, 2006), pp. 402–408. ISSN: 0268-1080, 1460-2237. DOI: 10.1093/heapol/czl018.

[69] Sarah J. Whitehead and Shehzad Ali. "Health Outcomes in Economic Evaluation: The QALY and Utilities". In: *British Medical Bulletin* 96.1 (Dec. 1, 2010), pp. 5–21. ISSN: 0007-1420. DOI: 10.1093/bmb/ldq033.

[70] Gunnar Nemeth. "Health Related Quality of Life Outcome Instruments". In: *European Spine Journal* 15.S1 (Jan. 2006), S44–S51. ISSN: 0940-6719, 1432-0932. DOI: 10.1007/s00586-005-1046-8.

[71] World Bank. *World Development Report 1993: Investing in Health.* New York: Oxford University Press, 1993.

[72] Christopher J. L. Murray, ed. *The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020 ; Summary.* Global Burden of Disease and Injury Series 1. OCLC: 832847878. Cambridge: Harvard School of Public Health [u.a.], 1996. 43 pp. ISBN: 978-0-9655466-0-7.

[73] World Health Organisation. *About the Global Burden of Disease (GBD) Project.* URL: https://www.who.int/healthinfo/global_burden_disease/about/en/.

[74] Christopher J. Murray. "Quantifying the Burden of Disease: The Technical Basis for Disability-Adjusted Life Years." In: *Bulletin of the World health Organization* 72.3 (1994), p. 429.

[75] A. Mehrez and A. Gafni. “Quality-Adjusted Life Years, Utility Theory, and Healthy-Years Equivalents”. In: *Medical Decision Making: An International Journal of the Society for Medical Decision Making* 9.2 (1989 Apr-Jun), pp. 142–149. ISSN: 0272-989X. DOI: 10.1177/0272989X8900900209. pmid: 2501627.

[76] Luis Prieto and José A. Sacristán. “Problems and Solutions in Calculating Quality-Adjusted Life Years (QALYs)”. In: *Health and Quality of Life Outcomes* 1.1 (Dec. 19, 2003), p. 80. ISSN: 1477-7525. DOI: 10.1186/1477-7525-1-80.

[77] Gerard Duru et al. “Limitations of the Methods Used for Calculating Quality-Adjusted Life-Year Values:” in: *PharmacoEconomics* 20.7 (2002), pp. 463–473. ISSN: 1170-7690. DOI: 10.2165/00019053-200220070-00004.

[78] S. Anand and K. Hanson. “Disability-Adjusted Life Years: A Critical Review”. In: *Journal of Health Economics* 16.6 (Dec. 1997), pp. 685–702. ISSN: 0167-6296. pmid: 10176779.

[79] Trude Arnesen and Erik Nord. “The Value of DALY Life: Problems with Ethics and Validity of Disability Adjusted Life Years”. In: *BMJ : British Medical Journal* 319.7222 (Nov. 27, 1999), pp. 1423–1425. ISSN: 0959-8138. pmid: 10574867. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1117148/ (visited on 02/11/2019).

[80] Sudhir Anand and Kara Hanson. “DALYs: Efficiency versus Equity”. In: *World Development* 26.2 (Feb. 1, 1998), pp. 307–310. ISSN: 0305-750X. DOI: 10.1016/S0305-750X(97)10019-5.

[81] World Health Organization. *Cost-Effectiveness Threshold Values*. URL: https://www.who.int/choice/costs/en/.

[82] Elliot Marseille et al. “Thresholds for the Cost-Effectiveness of Interventions: Alternative Approaches”. In: *Bulletin of the World Health Organization* 93.2 (Feb. 1, 2015), pp. 118–124. ISSN: 1564-0604. DOI: 10.2471/BLT.14.138206. pmid: 25883405.

[83] A. M. Garber and C. E. Phelps. “Economic Foundations of Cost-Effectiveness Analysis”. In: *Journal of Health Economics* 16.1 (Feb. 1997), pp. 1–31. ISSN: 0167-6296. pmid: 10167341.

[84] Louis R Eeckhoudt and James K Hammit. “Background Risks and the Value of a Statistical Life”. In: *Journal of Risk and Uncertainty* 23.3 (2001), pp. 261–279. ISSN: 08955646, 15730476. JSTOR: 41761049.

[85] Andrew R. Willan and Andrew H. Briggs. *Statistical Analysis of Cost-Effectiveness Data: Willan/Statistical Analysis of Cost-Effectiveness Data*. Chichester, UK: John Wiley & Sons, Ltd, July 14, 2006. ISBN: 978-0-470-85628-4 978-0-470-85626-0 978-0-470-85627-7. DOI: 10.1002/0470856289. URL: http://doi.wiley.com/10.1002/0470856289 (visited on 02/13/2019).

[86] B. A. van Hout et al. "Costs, Effects and C/E-Ratios alongside a Clinical Trial". In: *Health Economics* 3.5 (1994 Sep-Oct), pp. 309–319. ISSN: 1057-9230. pmid: 7827647.

[87] Peter Wakker and Marc P. Klaassen. "Confidence Intervals for Cost/Effectiveness Ratios". In: *Health economics* 4.5 (1995), pp. 373–381.

[88] E. M. Laska, M. Meisner, and C. Siegel. "Statistical Inference for Cost-Effectiveness Ratios". In: *Health Economics* 6.3 (1997 May-Jun), pp. 229–242. ISSN: 1057-9230. pmid: 9226141.

[89] Gianluca Baio. "Bayesian Models for Cost-Effectiveness Analysis in the Presence of Structural Zero Costs". In: (July 19, 2013). arXiv: 1307.5243 [math, stat]. URL: http://arxiv.org/abs/1307.5243 (visited on 02/13/2019).

[90] Simon G. Thompson and Richard M. Nixon. "How Sensitive Are Cost-Effectiveness Analyses to Choice of Parametric Distributions?" In: *Medical Decision Making* 25.4 (July 2005), pp. 416–423. ISSN: 0272-989X, 1552-681X. DOI: 10.1177/0272989X05276862.

[91] A O'Hagan and J W Stevens. "Bayesian Methods for Design and Analysis of Cost-Effectiveness Trials in the Evaluation of Health Care Technologies". In: *Statistical Methods in Medical Research* 11.6 (Dec. 2002), pp. 469–490. ISSN: 0962-2802, 1477-0334. DOI: 10.1191/0962280202sm305ra.

[92] A. O'Hagan, J. W. Stevens, and J. Montmartin. "Inference for the Cost-Effectiveness Acceptability Curve and Cost-Effectiveness Ratio". In: *PharmacoEconomics* 17.4 (Apr. 2000), pp. 339–349. ISSN: 1170-7690. DOI: 10.2165/00019053-200017040-00004. pmid: 10947489.

[93] Daniel F. Heitjan, Alan J. Moskowitz, and William Whang. "Bayesian Estimation of Cost-Effectiveness Ratios from Clinical Trials". In: *Health Economics* 8.3 (1999), pp. 191–201. ISSN: 1099-1050. DOI: 10.1002/(SICI)1099-1050(199905)8:3<191::AID-HEC409>3.0.CO;2-R.

[94] A. O'Hagan, J. W. Stevens, and J. Montmartin. "Bayesian Cost-Effectiveness Analysis from Clinical Trial Data". In: *Statistics in Medicine* 20.5 (Mar. 15, 2001), pp. 733–753. ISSN: 0277-6715. DOI: 10.1002/sim.861. pmid: 11241573.

[95] David J. Spiegelhalter and Nicola G. Best. "Bayesian Approaches to Multiple Sources of Evidence and Uncertainty in Complex Cost-Effectiveness Modelling". In: *Statistics in medicine* 22.23 (2003), pp. 3687–3709.

[96] A. O'Hagan and J. W. Stevens. "A Framework for Cost-Effectiveness Analysis from Clinical Trial Data". In: *Health Economics* 10.4 (June 2001), pp. 303–315. ISSN: 1057-9230. DOI: 10.1002/hec.617. pmid: 11400253.

[97] Daniel Polsky et al. "Confidence Intervals for Cost-Effectiveness Ratios: A Comparison of Four Methods". In: *Health economics* 6.3 (1997), pp. 243–252.

[98] Gianluca Baio, Andrea Berardi, and Anna Heath. *Bayesian Cost-Effectiveness Analysis with the R Package BCEA*. Use R! Cham: Springer International Publishing, 2017. ISBN: 978-3-319-55716-8 978-3-319-55718-2. DOI: 10.1007/978-3-319-55718-2. URL: http://link.springer.com/10.1007/978-3-319-55718-2 (visited on 10/21/2018).

[99] Z. Jin et al. "Haemophilus Influenzae Type a Infection and Its Prevention". In: *Infection and Immunity* 75.6 (June 1, 2007), pp. 2650–2654. ISSN: 0019-9567. DOI: 10.1128/IAI.01774-06.

[100] Paul King. "Haemophilus Influenzae and the Lung (Haemophilus and the Lung)". In: *Clinical and Translational Medicine* 1.1 (2012), p. 10. ISSN: 2001-1326. DOI: 10.1186/2001-1326-1-10.

[101] Kenneth Todar. *Todar's Online Textbook of Bacteriology*.

[102] James P. Watt et al. "Burden of Disease Caused by Haemophilus Influenzae Type b in Children Younger than 5 Years: Global Estimates". In: *The Lancet* 374.9693 (2009), pp. 903–911.

[103] Dept. of Vaccines and Biologicals. *Haemophilus Influenzae Type b (Hib) Meningitis in the Pre-Vaccine Era : A Global Review of Incidence, Age Distributions, and Case-Fatality Rates*. World Health Organization, 2002. URL: http://www.who.int/iris/handle/10665/67572.

[104] Anders Hviid and Mads Melbye. "Impact of Routine Vaccination with a Conjugate Haemophilus Influenzae Type b Vaccine". In: *Vaccine* 22.3-4 (Jan. 2004), pp. 378–382. ISSN: 0264410X. DOI: 10.1016/j.vaccine.2003.08.001.

[105] "Haemophilus influenzae type b (Hib) Vaccination Position Paper – July 2013". In: *Releve Epidemiologique Hebdomadaire* 88.39 (Sept. 27, 2013), pp. 413–426. ISSN: 0049-8114. pmid: 24143842.

[106] K. Galil et al. “Reemergence of Invasive Haemophilus Influenzae Type b Disease in a Well-Vaccinated Population in Remote Alaska”. In: *The Journal of Infectious Diseases* 179.1 (Jan. 1999), pp. 101–106. ISSN: 0022-1899. DOI: 10.1086/314569. pmid: 9841828.

[107] Michael L. Jackson et al. “Modeling Insights into Haemophilus Influenzae Type b Disease, Transmission, and Vaccine Programs”. In: *Emerging Infectious Diseases* 18.1 (Jan. 2012), pp. 13–20. ISSN: 1080-6040, 1080-6059. DOI: 10.3201/eid1801.110336.

[108] S. Y. Oh et al. “School-Aged Children: A Reservoir for Continued Circulation of Haemophilus Influenzae Type b in the United Kingdom”. In: *The Journal of Infectious Diseases* 197.9 (May 1, 2008), pp. 1275–1281. ISSN: 0022-1899. DOI: 10.1086/586716. pmid: 18422439.

[109] Raymond S.W. Tsang and Marina Ulanova. “The Changing Epidemiology of Invasive Haemophilus Influenzae Disease: Emergence and Global Presence of Serotype a Strains That May Require a New Vaccine for Control”. In: *Vaccine* 35.33 (July 2017), pp. 4270–4275. ISSN: 0264410X. DOI: 10.1016/j.vaccine.2017.06.001.

[110] Raymond S.W. Tsang et al. “Characteristics of Invasive Haemophilus Influenzae Serotype a (Hia) from Nunavik, Canada and Comparison with Hia Strains in Other North American Arctic Regions”. In: *International Journal of Infectious Diseases* 57 (Apr. 2017), pp. 104–107. ISSN: 12019712. DOI: 10.1016/j.ijid.2017.02.003.

[111] Raymond S. W. Tsang et al. “Laboratory Characterization of Invasive *Haemophilus Influenzae* Isolates from Nunavut, Canada, 2000–2012”. In: *International Journal of Circumpolar Health* 75.1 (Jan. 2016), p. 29798. ISSN: 2242-3982. DOI: 10.3402/ijch.v75.29798.

[112] Y. A. Li et al. “Invasive Bacterial Diseases in Northern Canada, 2006-2013”. In: *Canada Communicable Disease Report* 42.4 (2016), p. 74.

[113] R. S. W. Tsang et al. “A Review of Invasive Haemophilus Influenzae Disease in the Indigenous Populations of North America”. In: *Epidemiology and Infection* 142.07 (July 2014), pp. 1344–1354. ISSN: 0950-2688, 1469-4409. DOI: 10.1017/S0950268814000405.

[114] Jenny L. Rotondo et al. “The Epidemiology of Invasive Disease Due to Haemophilus Influenzae Serotype a in the Canadian North from 2000 to 2010”. In: *International Journal of Circumpolar Health* 72.1 (Jan. 2013), p. 21142. ISSN: 2242-3982. DOI: 10.3402/ijch.v72i0.21142.

[115] Naushaba Degani et al. “Invasive Bacterial Diseases in Northern Canada”. In: *Emerging infectious diseases* 14.1 (2008), p. 34.

[116] Michael G. Bruce et al. “Epidemiology of Haemophilus Influenzae Serotype a, North American Arctic, 2000–2005”. In: *Emerging infectious diseases* 14.1 (2008), p. 48.

[117] Vic Eton et al. “Epidemiology of Invasive Pneumococcal and Haemophilus Influenzae Diseases in Northwestern Ontario, Canada, 2010–2015”. In: *International Journal of Infectious Diseases* 65 (Dec. 2017), pp. 27–33. ISSN: 12019712. DOI: 10.1016/j.ijid.2017.09.016.

[118] Marina Ulanova and Raymond SW Tsang. “Haemophilus Influenzae Serotype a as a Cause of Serious Invasive Infections”. In: *The Lancet infectious diseases* 14.1 (2014), pp. 70–82.

[119] Elisabeth E. Adderson et al. “Invasive Serotype a Haemophilus Influenzae Infections with a Virulence Genotype Resembling Haemophilus Influenzae Type b: Emerging Pathogen in the Vaccine Era?” In: *Pediatrics* 108.1 (2001), e18–e18.

[120] Marina Ulanova. “Global Epidemiology of Invasive *Haemophilus Influenzae* Type a Disease: Do We Need a New Vaccine?” In: *Journal of Vaccines* 2013 (2013), pp. 1–14. ISSN: 2090-7974, 2090-7990. DOI: 10.1155/2013/941461.

[121] A. D. Cox et al. “Developing a Vaccine for Haemophilus Influenzae Serotype a: Proceedings of a Workshop”. In: *Canada Communicable Disease Report* 43.5 (2017), p. 89.

[122] L. Barreto et al. “The Emerging Haemophilus Influenzae Serotype a Infection and a Potential Vaccine: Implementation Science in Action”. In: *DISEASE REPORT* 43.5 (2017), p. 85.

[123] Andrew D. Cox et al. “Investigating the Candidacy of a Capsular Polysaccharide-Based Glycoconjugate as a Vaccine to Combat Haemophilus Influenzae Type a Disease: A Solution for an Unmet Public Health Need”. In: *Vaccine* 35.45 (Oct. 2017), pp. 6129–6136. ISSN: 0264410X. DOI: 10.1016/j.vaccine.2017.09.055.

[124] P. G. Coen et al. “Mathematical Models of Haemophilus Influenzae Type b”. In: *Epidemiology & Infection* 120.3 (1998), pp. 281–295.

[125] K. Auranen et al. “Modelling Transmission, Immunity and Disease of Haemophilus Influenzae Type b in a Structured Population”. In: *Epidemiology and Infection* 132.5 (Oct. 2004), pp. 947–957. ISSN: 0950-2688, 1469-4409. DOI: 10.1017/S0950268804002493.

[126] Angjelina Konini et al. “Dynamics of Naturally Acquired Antibody against Haemophilus Influenzae Type a Capsular Polysaccharide in a Canadian Aboriginal Population”. In: *Preventive Medicine Reports* 3 (June 2016), pp. 145–150. ISSN: 22113355. DOI: 10.1016/j.pmedr.2016.01.004.

[127] Angjelina Konini and Seyed M. Moghadas. "Modelling the Impact of Vaccination on Curtailing Haemophilus Influenzae Serotype 'a'". In: *Journal of Theoretical Biology* 387 (Dec. 2015), pp. 101–110. ISSN: 00225193. DOI: 10.1016/j.jtbi.2015.09.026.

[128] T. Leino et al. "Dynamics of Natural Immunity Caused by Subclinical Infections, Case Study on Haemophilus Influenzae Type b (Hib)". In: *Epidemiology and Infection* 125.3 (Dec. 2000), pp. 583–591. ISSN: 0950-2688. pmid: 11218209.

[129] David L Heymann. *Control of Communicable Diseases Manual*. OCLC: 903098348. 2015. ISBN: 978-0-87553-018-5.

[130] *Patient Cost Estimator*. Canadian Institute for Health Information.

[131] Ora Kendall et al. *Bacterial Meningitis In Canada: Hospitalizations*. Dec. 1, 2005. URL: https://www.canada.ca/en/public-health/services/reports-publications/canada-communicable-disease-report-ccdr/monthly-issue/2005-31/bacterial-meningitis-canada-hospitalizations-1994-2001.html.

[132] Charlotte Jackson et al. "Effectiveness of Haemophilus Influenzae Type b Vaccines Administered According to Various Schedules: Systematic Review and Meta-Analysis of Observational Data". In: *The Pediatric Infectious Disease Journal* 32.11 (Nov. 2013), pp. 1261–1269. ISSN: 0891-3668. DOI: 10.1097/INF.0b013e3182a14e57.

[133] U. K. Griffiths et al. "Dose-Specific Efficacy of Haemophilus Influenzae Type b Conjugate Vaccines: A Systematic Review and Meta-Analysis of Controlled Clinical Trials". In: *Epidemiology and Infection* 140.08 (Aug. 2012), pp. 1343–1355. ISSN: 0950-2688, 1469-4409. DOI: 10.1017/S0950268812000957.

[134] Nadia A. Charania and Seyed M. Moghadas. "Modelling the Effects of Booster Dose Vaccination Schedules and Recommendations for Public Health Immunization Programs: The Case of Haemophilus Influenzae Serotype b". In: *BMC Public Health* 17.1 (Dec. 2017). ISSN: 1471-2458. DOI: 10.1186/s12889-017-4714-9.

[135] Statistics Canada. *Childhood National Immunization Coverage Survey*. 5185. Statistics Canada, June 28, 2017.

[136] T. Lagergård and P. Branefors. "Nature of Cross-Reactivity between Haemophilus Influenzae Types a and b and Streptococcus Pneumoniae Types 6A and 6B". In: *Acta Pathologica, Microbiologica, Et Immunologica Scandinavica. Section C, Immunology* 91.6 (Dec. 1983), pp. 371–376. ISSN: 0108-0202. pmid: 6201038.

[137] Anikó Fekete et al. “Synthesis of Octa- and Dodecamers of D-Ribitol-1-Phosphate and Their Protein Conjugates”. In: *Carbohydrate Research* 341.12 (Sept. 4, 2006), pp. 2037–2048. ISSN: 0008-6215. DOI: 10.1016/j.carres.2005.10.023. pmid: 16458277.

[138] Elizabeth C. Briere et al. *Prevention and Control of Haemophilus Influenzae Type b Disease: Recommendations of the Advisory Committee on Immunization Practices (ACIP).* Feb. 28, 2014, 63(RR01), 1–14.

[139] Jeffrey C. Kwong et al. *Ontario Burden of Infectious Disease Study (ONBOIDS): An OAHP-P/ICES Report.* Toronto: Ontario Agency for Health Protection and Promotion, Institute for Clinical Evaluative Sciences, 2010.

[140] G. Samoukovic et al. “Transporting a Critically Ill Patient from the Canadian North - Lessons Learned from Almost a Decade of SkyService Medevac Experience”. In: *International Journal of Infectious Diseases* 14 (Mar. 2010), e132–e133. ISSN: 12019712. DOI: 10.1016/j.ijid.2010.02.1778.

[141] Rob Van Exan. *Heamophilus Influenzae Disease in Canada: A Disease Model for Evaluating Vaccine Programs.* Technical Report. Cannaught Laboratories Limited, 1986.

[142] *Statistics Canada. Consumer Price Index. Catalogue No. 62-001-XWE.* Statistics Canada.

[143] Christopher Murray, David B Evans, and Arnab Acharya. “Development of WHO Guidelines on Generalized Cost-Effectiveness Analysis”. In: *Health Economics* 9 (2000), 235:251.

[144] Center for Disease Control and Prevention. *Vaccines for Children Program (VFC).* Sept. 1, 2017. URL: https://www.cdc.gov/vaccines/programs/vfc/awardees/vaccine-management/price-list/index.html (visited on 08/01/2017).

[145] Karen Edmond et al. “Global and Regional Risk of Disabling Sequelae from Bacterial Meningitis: A Systematic Review and Meta-Analysis”. In: *The Lancet. Infectious Diseases* 10.5 (May 2010), pp. 317–328. ISSN: 1474-4457. DOI: 10.1016/S1473-3099(10)70048-7. pmid: 20417414.

[146] Ulla Kou Griffiths, Andrew Clark, and Rana Hajjeh. “Cost-Effectiveness of Haemophilus Influenzae Type b Conjugate Vaccine in Low- and Middle-Income Countries: Regional Analysis and Assessment of Major Determinants”. In: *The Journal of Pediatrics* 163.1 (July 2013), S50–S59.e9. ISSN: 00223476. DOI: 10.1016/j.jpeds.2013.03.031.

[147] Joshua A. Salomon et al. “Common Values in Assessing Health Outcomes from Disease and Injury: Disability Weights Measurement Study for the Global Burden of Disease Study 2010”. In: *The Lancet* 380.9859 (2012), pp. 2129–2143.

[148] E. Dubé et al. “Vaccine Acceptance, Hesitancy and Refusal in Canada: Challenges and Potential Approaches”. In: *Canada Communicable Disease Report* 42.12 (2016), p. 246.

[149] Center for Disease Control and Prevention. *Diphtheria, Tetanus, and Pertussis Vaccine Safety*. Centers for Disease Control and Prevention, Oct. 27, 2017. URL: https://www.cdc.gov/vaccinesafety/vaccines/dtap-tdap-vaccine.html.

[150] S. A. Halperin et al. “Safety and Immunogenicity of Haemophilus Influenzae-Tetanus Toxoid Conjugate Vaccine given Separately or in Combination with a Three-Component Acellular Pertussis Vaccine Combined with Diphtheria and Tetanus Toxoids and Inactivated Poliovirus Vaccine for the First Four Doses”. In: *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 28.5 (May 1999), pp. 995–1001. ISSN: 1058-4838. DOI: 10.1086/514741. pmid: 10452624.

[151] J. Eskola et al. “Randomised Trial of the Effect of Co-Administration with Acellular Pertussis DTP Vaccine on Immunogenicity of Haemophilus Influenzae Type b Conjugate Vaccine”. In: *Lancet (London, England)* 348.9043 (Dec. 21, 1996–28), pp. 1688–1692. ISSN: 0140-6736. DOI: 10.1016/S0140-6736(96)04356-5. pmid: 8973430.

[152] D. P. Greenberg et al. “Immunogenicity of a Haemophilus Influenzae Type B-Tetanus Toxoid Conjugate Vaccine When Mixed with a Diphtheria-Tetanus-Acellular Pertussis-Hepatitis B Combination Vaccine”. In: *The Pediatric Infectious Disease Journal* 19.12 (Dec. 2000), pp. 1135–1140. ISSN: 0891-3668. pmid: 11144372.

[153] Angjelina Konini, Mingsong Kang, and Seyed M. Moghadas. “Simulating Immune Interference on the Effect of a Bivalent Glycoconjugate Vaccine against *Haemophilus Influenzae* Serotypes “a” and “b””. In: *Canadian Journal of Infectious Diseases and Medical Microbiology* 2016 (2016), pp. 1–8. ISSN: 1712-9532, 1918-1493. DOI: 10.1155/2016/5486869.

[154] Kathleen Whyte et al. “Recurrent Invasive Haemophilus Influenzae Serotype a Infection in an Infant”. In: *Microbiology Discovery* 3.1 (2015), p. 4. ISSN: 2052-6180. DOI: 10.7243/2052-6180-3-4.

[155] World Health Organization. *Observed Rate of Vaccine Reactions. Haemophilus Influenzae Type B.* Information Sheet. 2012. URL: http://www.who.int/vaccine_safety/initiative/tools/HiB_Vaccine_rates_information_sheet.pdf?ua=1.

[156] Robert S. Lanciotti et al. “Genetic and Serologic Properties of Zika Virus Associated with an Epidemic, Yap State, Micronesia, 2007”. In: *Emerging Infectious Diseases* 14.8 (Aug. 2008), pp. 1232–1239. ISSN: 1080-6040, 1080-6059. DOI: 10.3201/eid1408.080287.

[157] Enny S. Paixao et al. “History, Epidemiology, and Clinical Manifestations of Zika: A Systematic Review”. In: *American Journal of Public Health* 106.4 (Apr. 2016), pp. 606–612. ISSN: 0090-0036, 1541-0048. DOI: 10.2105/AJPH.2016.303112.

[158] Byung-Hak Song et al. “Zika Virus: History, Epidemiology, Transmission, and Clinical Presentation”. In: *Journal of Neuroimmunology* 308 (July 2017), pp. 50–64. ISSN: 01655728. DOI: 10.1016/j.jneuroim.2017.03.001.

[159] Michael A. Johansson et al. “Zika and the Risk of Microcephaly”. In: *New England Journal of Medicine* 375.1 (July 7, 2016), pp. 1–4. ISSN: 0028-4793, 1533-4406. DOI: 10.1056/NEJMp1605367.

[160] Alice Panchaud et al. “Emerging Role of Zika Virus in Adverse Fetal and Neonatal Outcomes”. In: *Clinical Microbiology Reviews* 29.3 (July 2016), pp. 659–694. ISSN: 1098-6618. DOI: 10.1128/CMR.00014-16. pmid: 27281741.

[161] Fabienne Krauer et al. “Zika Virus Infection as a Cause of Congenital Brain Abnormalities and Guillain–Barré Syndrome: Systematic Review”. In: *PLOS Medicine* 14.1 (Jan. 3, 2017). Ed. by Lorenz von Seidlein, e1002203. ISSN: 1549-1676. DOI: 10.1371/journal.pmed.1002203.

[162] Shamez N Ladhani et al. “Outbreak of Zika Virus Disease in the Americas and the Association with Microcephaly, Congenital Malformations and Guillain–Barré Syndrome”. In: *Archives of Disease in Childhood* 101.7 (July 2016), pp. 600–602. ISSN: 0003-9888, 1468-2044. DOI: 10.1136/archdischild-2016-310590.

[163] L. D. Kramer and G. D. Ebel. “Dynamics of Flavivirus Infection in Mosquitoes”. In: *Adv. Virus Res.* 60 (2003). DOI: 10.1016/S0065-3527(03)60006-0.

[164] Deckard, D. Trew. *Male-to-Male Sexual Transmission of Zika Virus—Texas.* Jan. 2016.

[165] G. Venturi. “An Autochthonous Case of Zika Due to Possible Sexual Transmission, Florence, Italy, 2014”. In: *Euro. Surveill.* 21 (2016). DOI: 10.2807/1560-7917.ES.2016.21.8.30148.

[166] B. D. Foy. “Probable Non-Vector-Borne Transmission of Zika Virus, Colorado, USA”. In: *Emerg. Infect. Dis.* 17 (2011). DOI: 10.3201/eid1705.101939.

[167] Daozhou Gao et al. “Prevention and Control of Zika as a Mosquito-Borne and Sexually Transmitted Disease: A Mathematical Modeling Analysis”. In: *Scientific Reports* 6.1 (Sept. 2016). ISSN: 2045-2322. DOI: 10.1038/srep28070.

[168] Reuters. “Brazil Reports Zika Infection from Blood Transfusions”. In: (Februrary 4, 2016). URL: https://www.reuters.com/article/us-health-zika-brazil-blood/brazil-reports-zika-infection-from-blood-transfusions-idUSKCN0VD22N (visited on 12/13/2016).

[169] World Health Organization. “Zika Situation Report”. In: (). URL: http://www.who.int/emergencies/zika-virus/situation-report/en/ (visited on 08/20/2018).

[170] World Health Organization. “Fifth Meeting of the Emergency Committee under the International Health Regulations (2005) Regarding Microcephaly, Other Neurological Disorders and Zika Virus”. In: ().

[171] Isaac I Bogoch et al. “Anticipating the International Spread of Zika Virus from Brazil”. In: *The Lancet* 387.10016 (Jan. 2016), pp. 335–336. ISSN: 01406736. DOI: 10.1016/S0140-6736(16)00080-5.

[172] Centre for Disease Control and Prevention. “Areas with Zika”. In: (). URL: http://www.cdc.gov/zika/geo/ (visited on 08/06/2016).

[173] L. Dinh et al. “Estimating the Subcritical Transmissibility of the Zika Outbreak in the State of Florida, USA, 2016”. In: *Theoret. Biol. Med. Modelling.* 9 (2016). DOI: 10.1186/s12976-016-0046-1.

[174] Government of Canada. *Surveillance of Zika Virus*. Oct. 17, 2018. URL: http://www.healthycanadians.gc.ca/diseases-conditions-maladies-affections/disease-maladie/zika-virus/surveillance-eng.php?id=zikacases#s1 (visited on 12/01/2016).

[175] Pan American Health Organization. *Zika Epidemiological Update*. July 26, 2017. URL: https://reliefweb.int/sites/reliefweb.int/files/resources/2017-jul-26-phe-epi-update-zika-virus.pdf (visited on 08/01/2017).

[176] Mark R. Duffy et al. “Zika Virus Outbreak on Yap Island, Federated States of Micronesia”. In: *New England Journal of Medicine* 360.24 (June 11, 2009), pp. 2536–2543. ISSN: 0028-4793, 1533-4406. DOI: 10.1056/NEJMoa0805715.

[177] Gerardo Chowell et al. “Using Phenomenological Models to Characterize Transmissibility and Forecast Patterns and Final Burden of Zika Epidemics”. In: *PLoS Currents* (2016). ISSN: 2157-3999. DOI: 10.1371/currents.outbreaks.f14b2217c902f453d9320a43a35b9583.

[178] H. Nishiura et al. “Preliminary Estimation of the Basic Reproduction Number of Zika Virus Infection during Colombia Epidemic, 2015–2016”. In: *Travel Med. Infect. Dis.* 14 (2016). DOI: 10.1016/j.tmaid.2016.03.016.

[179] United Nations Statistics Division. *Demographic and Social Statistics*. URL: https://unstats.un.org/unsd/demographic-social/products/dyb/index.cshtml.

[180] W. L. Walker. “Zika Virus Disease Cases—50 States and the District of Columbia, January 1-July 31, 2016”. In: *MMWR Morb Mortal Wkly Rep.* 65 (2016). DOI: 10.15585/mmwr.mm6536e5.

[181] T. A. Perkins et al. “Model-Based Projections of Zika Virus Infections in Childbearing Women in the Americas”. In: *Nat. Microbiol* 1 (2016). DOI: 10.1038/nmicrobiol.2016.126.

[182] L. Yakob et al. “Low Risk of a Sexually-Transmitted Zika Virus Outbreak”. In: *Lancet Infect Dis.* 16 (2016). DOI: 10.1016/S1473-3099(16)30324-3.

[183] Michael Reece et al. “Sexual Behaviors, Relationships, and Perceived Health Among Adult Men in the United States: Results from a National Probability Sample”. In: *The Journal of Sexual Medicine* 7 (Oct. 2010), pp. 291–304. ISSN: 17436095. DOI: 10.1111/j.1743-6109.2010.02009.x.

[184] Debby Herbenick et al. “Sexual Behaviors, Relationships, and Perceived Health Status Among Adult Women in the United States: Results from a National Probability Sample”. In: *The Journal of Sexual Medicine* 7 (Oct. 2010), pp. 277–290. ISSN: 17436095. DOI: 10.1111/j.1743-6109.2010.02010.x.

[185] L. M. Styer et al. “Mosquitoes Do Senesce: Departure from the Paradigm of Constant Mortality”. In: *Am. J. Trop. Med. Hyg.* 76 (2007).

[186] M. Trpis and W. Hausermann. “Dispersal and Other Population Parameters of Aedes Aegypti in an African Village and Their Possible Significance in Epidemiology of Vector-Borne Diseases”. In: *Am. J. Trop. Med. Hyg.* 35 (1986). DOI: 10.4269/ajtmh.1986.35.1263.

[187] J. L. Putnam and T. W. Scott. “Blood-Feeding Behavior of Dengue-2 Virus-Infected Aedes Aegypti”. In: *Am. J. Trop. Med. Hyg.* 52 (1995). DOI: 10.4269/ajtmh.1995.52.225.

[188] H. J. Wearing and P. Rohani. “Ecological and Immunological Determinants of Dengue Epidemics”. In: *Proc Natl. Acad. Sci. USA* 103 (2006). DOI: 10.1073/pnas.0602960103.

[189] P. Gallian. “Zika Virus in Asymptomatic Blood Donors in Martinique”. In: *Blood.* 129 (2017). DOI: 10.1182/blood-2016-09-737981.

[190] Michel Jacques Counotte et al. “Sexual Transmission of Zika Virus and Other Flaviviruses: A Living Systematic Review”. In: *PLOS Medicine* 15.7 (July 24, 2018). Ed. by Lorenz von Seidlein, e1002611. ISSN: 1549-1676. DOI: 10.1371/journal.pmed.1002611.

[191] Barry Atkinson et al. “Presence and Persistence of Zika Virus RNA in Semen, United Kingdom, 2016”. In: *Emerging Infectious Diseases* 23.4 (Apr. 2017), pp. 611–615. ISSN: 1080-6040, 1080-6059. DOI: 10.3201/eid2304.161692.

[192] H. Nishiura and S. B. Halstead. “Natural History of Dengue Virus (DENV) – 1 and DENV – 4 Infections: Reanalysis of Classic Studies”. In: *J. Infect Dis.* 195 (2007). DOI: 10.1086/511825.

[193] L. B. Carrington and C. P. Simmons. “Human to Mosquito Transmission of Dengue Viruses”. In: *Front Immunol.* 5 (2014). DOI: 10.3389/fimmu.2014.00290.

[194] D. J. Gubler et al. “Viraemia in Patients with Naturally Acquired Dengue Infection”. In: *Bull World Health Org.* 59 (1981).

[195] Miranda Chan and Michael A. Johansson. “The Incubation Periods of Dengue Viruses”. In: *PLoS ONE* 7.11 (Nov. 30, 2012). Ed. by Nikos Vasilakis, e50972. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0050972.

[196] D. J. Gubler et al. “Viraemia in Patients with Naturally Acquired Dengue Infection”. In: *Bulletin of the World Health Organization* 59.4 (1981), pp. 623–630. ISSN: 0042-9686. pmid: 6976230.

[197] Maimuna S Majumder et al. “Estimating a Feasible Serial Interval Range for Zika Fever”. In: *Bulletin of the World Health Organization* (Feb. 9, 2016). ISSN: 0042-9686. DOI: 10.2471/BLT.16.171009.

[198] S. Towers. “Estimate of the Reproduction Number of the 2015 Zika Virus Outbreak in Barranquilla, Colombia, and Estimation of the Relative Role of Sexual Transmission”. In: *Epidemics* 17 (2016). DOI: 10.1016/j.epidem.2016.10.003.

[199] N. M. Ferguson. “Countering the Zika Epidemic in Latin America”. In: *Science* 353 (2016). DOI: 10.1126/science.aag0219.

[200] S. Riley. “Epidemiology: Making High-Res Zika Maps”. In: *Nat. Microbiol.* 1 (2016). DOI: 10.1038/nmicrobiol.2016.157.

[201] K. A. Liebman. "Determinants of Heterogeneous Blood Feeding Patterns by Aedes Aegypti in Iquitos, Peru". In: *PLoS Negl. Trop. Dis.* 8 (2014). DOI: 10.1371/journal.pntd.0002702.

[202] D. Champredon and S. M. Moghadas. "Quantifying the Contribution of Asymptomatic Infection to the Cumulative Incidence". In: *Epidemiology and Infection* 145.06 (Apr. 2017), pp. 1256–1258. ISSN: 0950-2688, 1469-4409. DOI: 10.1017/S0950268817000115.

[203] V. Duong. "Asymptomatic Humans Transmit Dengue Virus to Mosquitoes". In: *Proc Natl. Acad. Sci. USA* 112 (2015). DOI: 10.1073/pnas.1508114112.

[204] J. R. Coura, M. Suárez-Mutis, and S. Ladeia-Andrade. "A New Challenge for Malaria Control in Brazil: Asymptomatic Plasmodium Infection–a Review". In: *Memórias do Instituto Oswaldo Cruz.* 101 (2006). DOI: 10.1590/S0074-02762006000300001.

[205] N. M. Nguyen. "Host and Viral Features of Human Dengue Cases Shape the Population of Infected and Infectious Aedes Aegypti Mosquitoes". In: *Proc Natl. Acad. Sci. USA* 110 (2013). DOI: 10.1073/pnas.1303395110.

[206] D. J. Gubler et al. "Epidemic Dengue 3 in Central Java, Associated with Low Viremia in Man". In: *Am. J. Trop. Med. Hyg.* 30 (1981). DOI: 10.4269/ajtmh.1981.30.1094.

[207] D. J. Gubler et al. "Epidemiologic, Clinical, and Virologic Observations on Dengue in the Kingdom of Tonga". In: *Am. J. Trop. Med. Hyg* 27 (1978). DOI: 10.4269/ajtmh.1978.27.581.

[208] Megan R. Reynolds et al. "Vital Signs: Update on Zika Virus–Associated Birth Defects and Evaluation of All U.S. Infants with Congenital Zika Virus Exposure — U.S. Zika Pregnancy Registry, 2016". In: *MMWR. Morbidity and Mortality Weekly Report* 66.13 (Apr. 7, 2017), pp. 366–373. ISSN: 0149-2195, 1545-861X. DOI: 10.15585/mmwr.mm6613e1.

[209] World Health Organization. *WHO Global Consultation on Research Related to Zika Virus Infection.* Mar. 2016.

[210] World Health Organization. "Zika Virus (ZIKV) Vaccine Target Product Profile (TPP): Vaccine to Protect against Congenital Zika Syndrome for Use during an Emergency". In: (). URL: http://www.who.int/immunization/research/development/WHO_UNICEF_Zikavac_TPP_Feb2017.pdf.

[211] Anna Durbin and Annelies Wilder-Smith. "An Update on Zika Vaccine Developments". In: *Expert Review of Vaccines* 16.8 (Aug. 3, 2017), pp. 781–787. ISSN: 1476-0584, 1744-8395. DOI: 10.1080/14760584.2017.1345309.

[212] Francisco A. Lagunas-Rangel, Martha E. Viveros-Sandoval, and Arturo Reyes-Sandoval. "Current Trends in Zika Vaccine Development". In: *Journal of Virus Eradication* 3.3 (July 1, 2017), pp. 124–127. ISSN: 2055-6640. pmid: 28758019.

[213] Ralph A. Tripp and Ted M. Ross. "Development of a Zika Vaccine". In: *Expert Review of Vaccines* 15.9 (Sept. 2016), pp. 1083–1085. ISSN: 1476-0584, 1744-8395. DOI: 10.1080/14760584.2016.1192474.

[214] Dan H. Barouch, Stephen J. Thomas, and Nelson L. Michael. "Prospects for a Zika Virus Vaccine". In: *Immunity* 46.2 (Feb. 2017), pp. 176–182. ISSN: 10747613. DOI: 10.1016/j.immuni.2017.02.005.

[215] Martin R Gaudinski et al. "Safety, Tolerability, and Immunogenicity of Two Zika Virus DNA Vaccine Candidates in Healthy Adults: Randomised, Open-Label, Phase 1 Clinical Trials". In: *The Lancet* 391.10120 (Feb. 2018), pp. 552–562. ISSN: 01406736. DOI: 10.1016/S0140-6736(17)33105-7.

[216] Alan D. T. Barrett. "Current Status of Zika Vaccine Development: Zika Vaccines Advance into Clinical Evaluation". In: *npj Vaccines* 3.1 (Dec. 2018). ISSN: 2059-0105. DOI: 10.1038/s41541-018-0061-9.

[217] Felipe J. Colón-González et al. "After the Epidemic: Zika Virus Projections for Latin America and the Caribbean". In: *PLOS Neglected Tropical Diseases* 11.11 (Nov. 1, 2017). Ed. by Uwem Friday Ekpo, e0006007. ISSN: 1935-2735. DOI: 10.1371/journal.pntd.0006007.

[218] Margaret A. Honein et al. "Birth Defects Among Fetuses and Infants of US Women With Evidence of Possible Zika Virus Infection During Pregnancy". In: *JAMA* 317.1 (Jan. 3, 2017), p. 59. ISSN: 0098-7484. DOI: 10.1001/jama.2016.19006.

[219] V. M. Cao-Lormeau et al. "Guillain-Barré Syndrome Outbreak Associated with Zika Virus Infection in French Polynesia: A Case-Control Study". In: *Lancet (London, England)* 387.10027 (Apr. 9, 2016), pp. 1531–1539. ISSN: 1474-547X. DOI: 10.1016/S0140-6736(16)00562-6. pmid: 26948433.

[220] Jorge A. Alfaro-Murillo et al. "A Cost-Effectiveness Tool for Informing Policies on Zika Virus Control". In: *PLOS Neglected Tropical Diseases* 10.5 (May 20, 2016). Ed. by Hélène Carabin, e0004743. ISSN: 1935-2735. DOI: 10.1371/journal.pntd.0004743.

[221] W. N. Nembhard et al. "Patterns of First-Year Survival among Infants with Selected Congenital Anomalies in Texas, 1995-1997". In: *Teratology* 64.5 (Nov. 2001), pp. 267–275. ISSN: 0040-3709. DOI: 10.1002/tera.1073. pmid: 11745833.

[222] Simon Cauchemez et al. “Association between Zika Virus and Microcephaly in French Polynesia, 2013–15: A Retrospective Study”. In: *The Lancet* 387.10033 (May 2016), pp. 2125–2132. ISSN: 01406736. DOI: 10.1016/S0140-6736(16)00651-6.

[223] Joshua A Salomon et al. “Disability Weights for the Global Burden of Disease 2013 Study”. In: *The Lancet Global Health* 3.11 (Nov. 2015), e712–e723. ISSN: 2214109X. DOI: 10.1016/S2214-109X(15)00069-8.

[224] David P. Durham et al. “Evaluating Vaccination Strategies for Zika Virus in the Americas”. In: *Annals of Internal Medicine* 168.9 (May 1, 2018), p. 621. ISSN: 0003-4819. DOI: 10.7326/M17-0641.

[225] Joe Parkin Daniels. “Tackling Teenage Pregnancy in Colombia”. In: *The Lancet* 385.9977 (Apr. 2015), pp. 1495–1496. ISSN: 01406736. DOI: 10.1016/S0140-6736(15)60738-3.

[226] Diego Cuellar. *Estimacion Del Numero de Gestantes En Colombia (Estimation of the Number of Pregnant Women in Colombia)*. Direccion de Epidemiologıa y Demografıa., Oct. 2013. URL: https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/GCFI/1.2%20Estimaciones%20GestantesNacimientos.pdf.

[227] United Nations Development Programme. “A Socio-Economic Impact Assessment of the Zika Virus in Latin America and the Caribbean: With a Focus on Brazil, Colombia and Suriname”. In: (Mar. 2017). URL: https://www.undp.org/content/undp/en/home/librarypage/hiv-aids/a-socio-economic-impact-assessment-of-the-zika-virus-in-latin-am.html.

[228] Wu Zeng et al. “Cost-Effectiveness of Dengue Vaccination in Ten Endemic Countries”. In: *Vaccine* 36.3 (Jan. 2018), pp. 413–420. ISSN: 0264410X. DOI: 10.1016/j.vaccine.2017.11.064.

[229] Rui Li et al. “Cost-Effectiveness of Increasing Access to Contraception during the Zika Virus Outbreak, Puerto Rico, 2016”. In: *Emerging Infectious Diseases* 23.1 (Jan. 2017), pp. 74–82. ISSN: 1080-6040, 1080-6059. DOI: 10.3201/eid2301.161322.

[230] Alan Brennan, Stephen E. Chick, and Ruth Davies. “A Taxonomy of Model Structures for Economic Evaluation of Health Technologies”. In: *Health Economics* 15.12 (Dec. 2006), pp. 1295–1310. ISSN: 10579230, 10991050. DOI: 10.1002/hec.1148.

[231] Sun-Young Kim and Sue J. Goldie. “Cost-Effectiveness Analyses of Vaccination Programmes”. In: *Pharmacoeconomics* 26.3 (2008), pp. 191–215.