Mathematics of Planet Earth 7

Miranda I. Teboh-Ewungkem
Gideon Akumah Ngwa *Editors*

Infectious Diseases and Our Planet

Springer

Mathematics of Planet Earth

Volume 7

Springer's Mathematics of Planet Earth collection provides a variety of well-written books of a variety of levels and styles, highlighting the fundamental role played by mathematics in a huge range of planetary contexts on a global scale. Climate, ecology, sustainability, public health, diseases and epidemics, management of resources and risk analysis are important elements. The mathematical sciences play a key role in these and many other processes relevant to Planet Earth, both as a fundamental discipline and as a key component of cross-disciplinary research. This creates the need, both in education and research, for books that are introductory to and abreast of these developments.

Springer's MoPE series will provide a variety of such books, including monographs, textbooks, contributed volumes and briefs suitable for users of mathematics, mathematicians doing research in related applications, and students interested in how mathematics interacts with the world around us. The series welcomes submissions on any topic of current relevance to the international Mathematics of Planet Earth effort, and particularly encourages surveys, tutorials and shorter communications in a lively tutorial style, offering a clear exposition of broad appeal.

Responsible Editor(s):
Martin Peters, Heidelberg (martin.peters@springer.com)
Robinson dos Santos, São Paulo (robinson.dossantos@springer.com)
Additional Editorial Contacts:
Masayuki Nakamura, Tokyo (masayuki.nakamura@springer.com)

More information about this series at http://www.springer.com/series/13771

Miranda I. Teboh-Ewungkem
Gideon Akumah Ngwa
Editors

Infectious Diseases and Our Planet

Springer

Editors
Miranda I. Teboh-Ewungkem
Department of Mathematics
Lehigh University
Bethlehem, PA, USA

Gideon Akumah Ngwa
Department of Mathematics
University of Buea
Buea, Cameroon

ISSN 2524-4264 ISSN 2524-4272 (electronic)
Mathematics of Planet Earth
ISBN 978-3-030-50825-8 ISBN 978-3-030-50826-5 (eBook)
https://doi.org/10.1007/978-3-030-50826-5

Mathematics Subject Classification: 92-XX, 37-XX

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

This book is being dedicated to Professors Simon Levin of Princeton University, Louis Gross of the University of Tennessee and Thomas Hallam, Professor Emeritus of the University of Tennessee, for the seeds they planted with the Ecology related workshops they organized in the 1980s and 1990s in Trieste, Italy. Their work generated a new group of internationally based researchers and is making significant impacts.

Foreword

It is with delight that we gratefully acknowledge the efforts of the editors and contributors to this volume. We are flattered that the editors chose to dedicate the volume in honor of our efforts to direct the series of courses and workshops on mathematical ecology that were organized at what is now the Abdus Salam International Centre for Theoretical Physics (ICTP) in Trieste, Italy. This series began with a set of five courses and three workshops, sponsored at ICTP from 1982 to 2000, and was successful in introducing numerous developing-country scientists and mathematicians to current research that applies mathematical and computational methods to environmental problems. Many participants in these courses have proceeded to develop training programs in their own regions, including a wide array of courses and workshops in various countries in Latin America, Asia, and Africa. They have also subsequently mentored many new researchers who started work at these interfaces between quantitative and life sciences well after the initial series ended. In this sense, we consider our involvement in this series to have been one of the most significant contributions we have made to the development of interdisciplinary scientific expertise and connections across the world, and we echo the sentiments of Lord Robert May of Oxford who has stated that his involvement in this series has been "some of the most important teaching I have done." Through these courses, the ICTP was instrumental in fostering a worldwide appreciation for the benefits of mathematical analysis of environmental problems and provided the mechanism for numerous collaborations between researchers from different disciplines and countries.

It may be of interest to readers of this volume to be aware of some of the history of the series at ICTP. The series grew out of discussions between Professor Giovanni Vidossich of the Scuola Internazionale Superiore di Studi Avanzati (SISSA) and ICTP and Tom Hallam during Professor Vidossich's sabbatical at the University of Tennessee in 1980. With strong support from ICTP Director Abdus Salam, Professors Hallam, Levin, and Vidossich organized the first major gathering of researchers and instructors in mathematical ecology, to lecture extensively over a 2-week period in November 1982. This was followed by an intensive two-week International Research Symposium which brought to ICTP many of the world's leaders in the

field. There were 21 lecturers for the course, 140 participants from 32 developing countries and 14 developed countries, and 30 additional participants in the Research Symposium. The participants mainly had strong physics and mathematics training, though there were a few participants chosen who had mostly biological training. The plan was to cover sufficient background material during the formal course lectures so that participants would be able to follow with understanding the more advanced talks presented during the Research Symposium. Topics covered during the course included deterministic and stochastic population models, age-structured models, bioeconomics, epidemiological models, and diffusion models. Two books resulted from the 1982 experiences:

S.A. Levin and T.G. Hallam (editors). 1984. Mathematical Ecology, Proceedings, Trieste 1982. Lecture Notes in Biomathematics, Vol. 54. Springer-Verlag, Berlin. (A collection of papers presented during the Research Symposium.)

S.A. Levin and T.G. Hallam (editors). 1986. Mathematical Ecology. Biomathematics, Vol. 17. Springer-Verlag, Berlin. (A collection of the lectures presented during the course.)

Following on the success of the 1982 experiences, the three of us organized a second course, with attendant Research Symposium, in November–December 1986. This included four weeks of course lectures followed by a one-week Research Symposium. The attendees were chosen specifically to be more broadly representative of those carrying out research in mathematical ecology, with significantly higher numbers of participants chosen due to their biological backgrounds. Leading researchers were invited to present talks on current research throughout the latter part of the course as well as during the Research Symposium. There were 17 lecturers for the course, 175 participants from 44 developing countries and 20 developed countries, and 44 research lecturers. To foster interaction and potential collaborations among participants, all were invited to present brief talks on their research, and these talks were scheduled throughout the four weeks of the course. Topics covered during the courses included: population and community modeling, demography, ecosystem models, agricultural models, statistical ecology, bioeconomics, population genetics, evolutionarily stable strategies, infectious disease models, and environmental toxicology models. Thus, the topic coverage was considerably broader and more oriented towards applications of models than those during the 1982 course. This was the first course to integrate computer workshops with the formal course lectures. There were 25 research talks presented by participants from developing countries throughout the course. Two books resulted from the 1986 experiences:

T.G. Hallam, L.J. Gross and S.A. Levin (eds.). 1988. Mathematical Ecology: Proceedings, Trieste 1986. World Scientific Publishing Co., Singapore.

S.A. Levin, T.G. Hallam and L.J. Gross (eds.). 1989. Applied Mathematical Ecology. Biomathematics Series, Vol. 18. Springer-Verlag, Berlin. (a collection of the lectures presented during the Course).

The series continued with two types of activities: formal courses, including basic research background lectures similar to those in the 1982 and 1986 courses, alternated every two years with shorter, more in-depth workshops that focused on a few major research topics. The first workshop was held in November 1988, with

subsequent courses in 1990, 1994, and 2000 and workshops in 1992 and 1996. The total participants for each of these were considerably smaller than that of the earlier courses. Leading researchers were invited to give research lectures that were integrated throughout the courses. For all of these, computer workshops became an integral part of the programs. One of our great pleasures was that attendees at the later ICTP courses were often students of those who participated in earlier courses.

Formal research groups were implemented at the beginning of each course or workshop, based upon the interests of the participants. Each participant was required to be part of at least one of the research groups, some of which were organized around problems of regional concern. These research groups met throughout the course and made a formal presentation to all attendees at the end of the course. The objective was to foster collaborative learning and potentially create collaborations that would continue after the participants' time at ICTP ended. A small number of participants for each course and workshop were from developed countries, attending at their own expense, with the majority of these from Italy. One measure of the success of the sequence of courses is perhaps to note that 12 of the invited lecturers for the 1988–2000 courses and workshops were developing-country scientists who had previously attended a course or workshop as a participant, and one of these, Graciela Canziani of the National University of Central Buenos Aires in Argentina, joined us as a Director for the 1996 and 2000 courses. The 1988–2000 courses and workshops had 456 participants from developing countries, 92 from developed countries, and 112 lecturers many of whom were from developing countries. So, in total the series involved over 1000 participants, most of whom were from developing countries. The topic coverage of these ICTP activities directly aligns with those now considered part of the Mathematics of Planet Earth, contributing to the worldwide development of this major endeavor to apply mathematical methods to enhance the sustainability of our planet.

While it was a tremendous pleasure to have been given the opportunity to direct these activities over so many years, we have been even more honored by the many invitations we received to participate in activities all over the world led by those who had been participants in the initial series. While several follow-on activities were hosted at ICTP, led by our colleagues and former students, we have been thrilled by the growing capabilities to carry on these activities around the world due to a fine cadre of trained educators in mathematical ecology and computational biology, many of whom have been directly influenced by the sequence of ICTP courses. The current collection is evidence of how far the graduates from the courses and workshops have come in becoming leaders in their own right and a source of real gratification for us. As we confront the new challenges in assuring the sustainability of our planet and the new mathematics that will be necessary, our world is fortunate to have a new generation of researchers eager and able to confront the issues.

Knoxville, TN, USA — Louis J. Gross
Knoxville, TN, USA — Thomas G. Hallam
Princeton, NJ, USA — Simon A. Levin
July 2019

Preface

Book Focus The book is a collection of topical mathematical modeling works on current topics in indirectly and directly transmitted infectious diseases of humans, animals, and plants that characterize our planet.

Intended Audience

- *Graduate students in the early phases of their research*: Can serve as a source of topical works and ideas in modeling infectious.
- *Established researchers*: Serves a career development for these researchers.
- *Readers interested in understanding the dynamics of spread of infectious diseases*: It offers a broad range of models for these interested readers, who wish to further their understanding of the dynamics of spread of infectious diseases or gain insight into the modeling process.

Major Takeaways

- The book showcases the power of mathematical modeling in capturing the spread and interaction dynamics of infectious diseases of our planet without recourse to experimentation. For example, a laboratory experiment which detects that a mosquito has taken a blood meal may not be able to predict the impact of the blood meal on future population sizes of the mosquito.
- The book also showcases the power of mathematical models, by illustrating how different evolutionary processes can interact to enhance diseases spread with the goal being survival of the interacting populations.
- The book showcases the power of mathematical modeling of virulent infectious diseases such as the Ebola virus disease, whose methods are extendable to the currently spreading COVID-19 virus disease. The modeling process can inform key parameters that may affect the spread of these diseases and the role effective control might have in curbing transmission. An example is the role of media coverage in educating a population that may affect control and how the disease spreads.
- Control of the diseases that affect humans and animals in our planets requires an understanding of all processes and interactions related to the disease. This book showcases this well.

- There are challenges with control for diseases with multiple transmission pathways. Questions about which pathways might be dominant become important questions, and questions about whether less dominant pathways can tip a scale are questions worth exploring for the future. This book highlights these.

Bethlehem, PA, USA
Buea, Cameroon

Miranda I. Teboh-Ewungkem
Gideon Akumah Ngwa

Contents

Infectious Diseases and Our Planet

Miranda I. Teboh-Ewungkem and Gideon A. Ngwa

1 Introduction

The bio-diversity of our planet is incredibly rich and complex in its nature, comprising a huge collection of *bio-organisms*. We use the term *bio-organism* to refer to any biological entity that possess some or all the properties of *life*. The definition of life (or being alive) in biology is mostly descriptive, however, by *life*, we shall mean a *characteristic* for any entity on our planet that would enable preservation of that entity and reinforce its existence. In this light we can consider a bio-organism as having life if such an organism has the attribute of being organized (in cells, unicellular or multicellular) and in addition, the ability to perform some or all of the following functions: homeostasis (the ability to regulate its internal environment to maintain a constant state), metabolism (being able to feed and so can transform energy by decomposing organic matter into cellular components and vice versa), growth (being able to maintain a higher rate of anabolism/catabolism and be able to increase in size in all of its parts, rather than simply accumulate matter), adaptation (having the ability to change over time in response to the environment on our planet), responsive to stimuli (ability to react, say via motion, when exposed to external chemicals or others probes), and reproduction (the ability to produce new individual organisms, either asexually from a single parent organism or sexually from two parent organisms). These complex processes, called physiological functions, have underlying physical and chemical bases, as well as signaling and control mechanisms that are essential to maintaining

M. I. Teboh-Ewungkem (✉)
Department of Mathematics, Lehigh University, Bethlehem, PA, USA
e-mail: mit703@lehigh.edu

G. A. Ngwa
Department of Mathematics, University of Buea, Buea, Cameroon

M. I. Teboh-Ewungkem, G. A. Ngwa (eds.), *Infectious Diseases and Our Planet*,
Mathematics of Planet Earth 7, https://doi.org/10.1007/978-3-030-50826-5_1

life. A bio-entity that is capable of performing only one of the listed functions may not be classified as having life. We now discuss a few bio-entities under consideration here: A virus (giant molecule of nucleic acid coated with protein), for example, can replicate itself within a given environment, but because it neither metabolizes nor grows, it does not qualify as being alive. However, being a bio-entity, we would also consider virus in the discussion here. A fungus is another bio-entity that will be considered and could fall in either the living or non-living category, depending on the type. The last bio-entity to be considered is parasites, considered living.

Given the broad range of possibilities available to biological entities, we note that each bio-organism has its niche; the range of environmental conditions wherein each member of the given species can survive and reproduce. Survivorship and reproduction being key existential priorities for each of the bio-organisms that make up our planet's biosphere. Since there are several bio-organisms sharing the same limited-sized biosphere, issues of co-existence come into play. We can think of the many organisms co-existing in the same biosphere as being in some kind of equilibrium whereby the domain of dependence, and/or range of influence, of the activities of each bio-organism, even if overlapping, engenders some stability in the population numbers of the particular organism that the biosphere can *sustainably sustain*. Thus, this kind of equilibrium can be thought of as a scenario whereby each living organism will live in tandem with its kind and with members of other species in a co-existence framework. However, in some instances, there is a flare up of the population numbers of one species breaking the equilibrium and causing undesirable effects. For example, a parasite can conveniently survive and grow in a host if the parasite's numbers are within the host's carrying capacity.[1] It is important to think of sustainability in the sense that the environment sustains the life of the organism, without itself being destroyed. In the case where the parasite numbers become larger than the host's carrying capacity, the host then becomes sick (environmental degradation) and in the extreme situation dies because of the parasite burden. Incidentally, a dead host often goes down with the parasite, hence terminating the parasites' life cycle, such as in the case of malaria. In order for the parasite to survive and maintain a continuous lineage, it must adapt to its hosts system and find ways to function within the hosts evading the hosts' fighting mechanisms, such as its immune system, and live in tandem with the host, without killing the host, sometimes changing forms in the process. On the other hand, however, a virus that kills its hosts may terminate the transmission process in due time. However, even with the death of the host, the virus' potential to transmit may not be diminished but be quite high, as, for example, in the case of the Ebola virus disease (EVD) [1, 9], which is the disease discussed in the chapter entitled: *Modeling Ebola*

[1]The environmental carrying capacity, for any given biological species, is the maximum population size of that species that the specific environment can sustainably sustain. It is the limiting population size imposed by the available life sustaining determinants such as food, water, space, etc., attributable to the particular biological niche or environment.

Transmission Dynamics with Media Effects on Disease and Isolation Rates, by the authors Oduniyi, Gibbons, Oh, and Agusto, in [14]. In some cases, though an organism's biological make-up may be suitable for the growth of another, that is, it can serve as a host for the second organism, the carrying capacity of the host may be such that it cannot support the presence of the second invading organism. In this situation, inclusion of any quantity of the second organism will serve as a *disease* to the host. Since invading parasites can kill their hosts, to ensure continued survival of the species, the parasites have to find a way of transferring themselves from one host to another in a process that we now understand as *disease transmission.*

2 Transmittable Infectious Disease Classification

Parasitic bio-entities that have attained the status of disease agents often possess mechanism that enables their transfer from host to host. The process of transfer from one host to another is known as disease spread, and a disease agent that has the ability to spread within a given host population is said to be an *infectious disease agent*. Infectious diseases may be classified into two broad categories: Directly transmitted and indirectly transmitted infectious diseases. Directly transmitted infectious diseases will not require a secondary medium to transmit the infection from one host to another host. In this case it is sufficient that the susceptible and infected hosts have a physical contact that is long enough for the infection to pass from the infected host to the uninfected one. In this category, we can mention human diseases such as HIV/AIDS, Ebola, etc., among directly transmitted infectious diseases of humans (see [1, 9], as well as the chapters entitled: *Modeling Ebola Transmission Dynamics with Media Effects on Disease and Isolation Rates*, by Oduniyi, Gibbons, Oh, and Agusto, citation [14] and *Reducing the global HIV burden: The importance of uneven exposure to the results of HIV prevention trials*, by Moore, Boily, Donnell, and Dimitrov, as discussed in [7]). Some examples of directly transmitted diseases of animals are Pseudorabies as discussed in the chapter [6], entitled *Evidence for Multiple Transmission Routes for Pseudorabies in Wild Hogs*, by Levy, Lenhart, Collins, and Stiver, and the fungal pathogen *Batrachochytrium Salamandrivorans* as discussed in the chapter [5], entitled, *Identifying the dominant transmission pathway in a multi-stage infection model of the emerging fungal pathogen Batrachochytrium Salamandrivorans on the Eastern Newt*, by Islam, Gray, and Peace.

On the other hand, indirectly transmitted infectious diseases will require a secondary medium over which the infection can pass to the next host. These, that is, indirectly transmitted infectious diseases, can again be divided into two broad categories: Those that require a living bio-entity to help transmit the infection from host to host and those that required non-biological entities to help transmit the infectious pathogen from host to host. This classification becomes fuzzy when the infectious disease can be picked up by another host because of a visitation to a location that has been previously contaminated by an infectious host by leaving

behind infectious diseases agent, and the very disease can be transmitted from host to host simply because the infected and susceptible individuals are sharing a common space at the same time. Examples may include Ebola virus disease (as discussed in [14]), cholera, typhoid fever, and the recent Sars-Cov-2 virus, among infectious diseases of humans, and as discussed in [5] and [6], *Batrachochytrium Salamandrivorans* and Pseudorabies, among infectious diseases of animals. The latter disease, *Batrachochytrium Salamandrivorans*, is discussed in the chapter, entitled: *Identifying the dominant transmission pathway in a multi-stage infection model of the emerging fungal pathogen Batrachochytrium Salamandrivorans on the Eastern Newt*, by Islam, Gray, and Peace. It is worth mentioning that, in the context of our discussions, some of the directly transmittable diseases are also indirectly transmittable, transmitted via non-biological entity because they share same physical space in which contacts can occur, such as the Ebola virus disease (EVD) or the *Batrachochytrium Salamandrivorans* fungal infection, to name a few.

We shall also be interested in those infectious diseases that cannot be transmitted from host to host by mere touch or sharing of the same physical space but must require a second biological host within which the pathogen must grow and develop before it is in a form that can be transmitted to the next human, animal, or plant. In this category, we have malaria, yellow fever, leishmaniasis, etc., among infectious diseases of humans, which will require a secondary bio-entity called the disease vector, to help transmit the infection from one host to the other. See the chapters: [16] entitled, *A Multistage Mosquito-Centered Mathematical Model for Malaria Dynamics that Captures Mosquito Gonotrophic Cycle Contributions to its Population Abundance and Malaria Transmission*, by Teboh-Ewungkem, Ngwa, and Fomboh-Nforba; [15] entitled, *Charles Darwin meets Ronald Ross—A population-genetic framework for the evolutionary dynamics of malaria*, by Schneider; [12], entitled *The Effect of Demographic Variability and Periodic Fluctuations on Disease Outbreaks in a Vector-Host Epidemic Model* by Nipa and Allen and [2], entitled *Dynamic Regulation of T cell activation by coupled feedforward loops*, by Buri, Zelleke, and Ndifon. We also have the cassava mosaic virus diseases as plant diseases, as discussed in the chapter [3], entitled *Application of Mathematical Epidemiology to Crop Vector-Borne Diseases. The Cassava Mosaic Virus Disease Case*, by Chapwanya and Dumont. Because this form of transmission requires the interaction of at least two living organisms, say humans and mosquitoes, and development of the pathogen within these interacting living organisms, we will devote Sect. 4 to discuss further on vector-borne infections.

The pattern for disease transmission can be influenced by the spatial location/state of the organism to be infected. So the pattern would be different depending on whether we have a spatially distributed collection of host to be infected or we have a spatially distributed collection of hosts and diseases vectors to be infected and whether or not the infectious agent is spreading through the distributed hosts. Thus there are several categories of host–pathogen–disease vector mediated interactive mechanisms: In the one category, the host may be stationary over large time scales and the vector distributes the infectious disease agent from host to host as in most plants diseases, as discussed in the chapter [3]. In another category, both the host

and the vector are mobile and the pathogen is transferred from host to host by the vector as it follows the host from spatial location to spatial location, as is the case with human diseases such as malaria. In yet another instance, it is the infected host that carries the disease agent from one spatial location to other hosts in other spatial locations as could be the case with infectious diseases of humans such as cholera, HIV/AIDS, and Ebola virus disease to name but a few. Some of these classes of disease shall be examined in the chapters that lay ahead cited as [2, 3, 5–7, 12, 14–16]. As shall be seen, each type of transmission would require a specified type of modeling assumptions to get the transmission dynamics approximately correct.

Irrespective of the mode of transmission of infectious disease agents from one host to the other, it is important to note that each living organism has a natural defense mechanism that is hard-wired into its genetic make-up that allows it to fight invasion by parasites or death through exposure to toxins. These natural defense mechanism for single cellular organisms, including most parasites, is achieved via genetic mutations.[2]Mutations, when they occur can be advantageous or disadvantageous to the particular organism whose genetic make-up has mutated. Mutations are advantageous to the organism if such a mutation renders the bio-entity fitter and better to survive in the environment. This is probably the situation that has been observed for most pathogens that develop mutations that make them resistant to certain chemical substances, such as medicines that are usually useful for the elimination of the parasite within the host system. An example is the development of drug resistance by the malaria parasite as well as the development of resistance of insecticides by some mosquito species. This phenomenon of development of resistance through mutations is an important evolutionary survival pathway and therefore constitutes aspects to be considered when studying human diseases and their potential control mechanism through use of drugs. Thus understanding genetic mechanism, as discussed in the chapter [15], that can serve as plausible pathways towards the onset or development of drug resistance by parasites is an important idea in infectious disease modeling. Mutations that are detrimental to the particular living organism render such mutated organism less competitive in their biological niche, and in most extreme cases the mutated variants of the organism do not survive. This is Darwinism whereby the concept of survival of the fittest gets manifested through genetic mutations. We shall also consider at least one such model (see [15]) where Darwinism is considered towards a better understanding of the behavior of the malaria parasite.

[2]A mutation is a change that occurs in the DNA sequence of a living organism, either due to mistakes when the DNA is copied or as the result of environmental factors. In general, a mutation is recognizable as having occurred if there is an alteration of the nucleotide sequence in the genome of an organism, virus, or extrachromosomal DNA.

3 Defense Mechanisms Against Infection

Complex and higher order organisms will require more sophisticated defense mechanism to fight pathogens that can invade them, be it viruses, bacteria, or parasites, for example. This type of defense mechanism is triggered when the cells in the organism have the capacity to detect foreign objects in its system and/or activate or trigger mechanisms to produce other types of cells that can mitigate the invasion of a foreign object. The production of system cleansing cells (immune cells) that target invading pathogen cells in order to destroy them, falls broadly under the category of immune response. A properly functioning immune response system must have the ability to kick into action once a foreign organism invades the system (activation), the ability to produce killer cells that would target and kill the invading organism (active immunization), and the ability to switch off and go quiescent when the invading substance has been removed (deactivation). The understanding of the proper mechanisms for the functioning of the immune system in humans is still evolving. However, several hypotheses have been proposed for the better understanding of the action of the immune system. One plausible mathematical characterization would be to consider that the activation of the immune system happens in a two-step process: production of activator precursors followed by production of effector cells that then trigger the active immunization response. We shall consider some models that have this characteristic as well as in [2].

In some cases, the degree of infestation is so severe that the human immune system is overwhelmed such that in the absence of external assistance, be it pharmaceutically-related or just supportive management of the infected system(s), death is expected. For example, in malaria parasite infestation, if the parasite load in the human host is at manageable levels, the human's immune system (innate and adaptive) will effectively control the infection [10], however, in the case of high parasitemia, massive destruction of the red blood cells by the parasite would mean that if external interventions are not employed, the human would die of anemia. When it has been diagnosed that a malaria infection is present, anti-malarial drugs are normally prescribed to help reduce the parasite load in the patient to levels where the natural immune system can continue to offer protection to the human. For some viruses, for example, the Ebola virus disease, help is in the form of supportive care to help alleviate symptoms while allowing the body's natural mechanisms to fight. These supportive care, which may involve increased intake of electrolytes, can help boost the immune system, giving the affected system an opportunity to recover. Other supportive mechanisms involve oxygen therapy, treatment/management of secondary disease-related infections.

4 Vector-Borne Diseases: Vectors, Pathogen, and Transmission

Any living organism that facilitates the transfer of a pathogen from one living host to another living host (perhaps of the same species) will constitute a *disease vector*. By *disease transmission* we shall mean an effective transfer of the *disease pathogen* from one host to another. The disease pathogen is the infectious agent, usually another living organism, that when present within the living tissue of another host can cause sickness and possibly leading to the death of the infected host.

4.1 The Disease Vector

A living organism can serve as a possible (efficient or effective) disease vector if it has a complete biological and genetic make-up that requires that, from time to time, the organism will seek contact with the particular host (perhaps humans). This contact can be for several biological reasons including the need to derive determinants that it will require for growth and development. For example, the *Anopheles sp.* mosquito's biological make-up requires that the female interacts with vertebrate host to draw blood which she needs for the maturation of her eggs. The interaction between the disease vector and the host may therefore be driven by a physiological and biological need for growth and development. Though such interaction can also be to seek shelter, it can be safely assumed that where the interaction is successful, the disease vector's chances of survival as a species are improved as it may live to reproduce. A step towards the study of indirectly transmitted diseases could be to seek to understand the disease vectors themselves (as in [8, 11, 13], for example), and as such we may seek to understand and quantify the developmental and reproductive gains that would accrue to the vector after a successful interaction with the particular host. For example, a female *Anopheles sp.* mosquito that successfully acquires a blood meal from a vertebrate host (such as a human or animal) and also survives the resting period that she needs after this blood meal, will lay a batch of eggs that will hatch and eventually mature into more adult mosquitoes in the future. The reproductive gain in foraging for blood meals, against when weighed with risk of being killed by predators, is the successful maturation and subsequent laying of eggs. It is true that in the absence of disease, bio-organisms that have the appropriate biological and genetic make-up to feed on other organism will always seek to interact with those organisms. Those that require plants as a place to develop their next generation will also continue to interact with that plant.

Another step towards the study of indirectly transmitted diseases could be to seek to understand disease transmission process and possibly quantify the role of the vector in transmission and seek ways to prevent their success. The chapters by Teboh-Ewungkem et al. [16], Chapwanya and Dumont [3] and [12], address these. While [16] focuses on the malaria transmitting vectors' interaction with humans,

Table 1 Examples of some disease vectors and the diseases they carry. Note that other interacting populations may exist, even though not mentioned

Disease vectors	Species	Diseases transmitted	Pathogen transmitted	Example Interacting population
Mosquito	*Anopheles sp.*	Malaria	*Plasmodium sp.*	Humans and animals
	Aedes aegypti	Dengue	Dengue virus	Humans
	Aedes aegypti	Zika	Zika virus	Humans
	Aedes sp. & *Mansonia sp.*	Lymphatic filariasis (elephantiasis)	Parasitic filarial worms (*Wuchereria bancrofti*)	Humans
Tse-Tse fly	*Glossina sp.*	Sleeping sickness	*Trypanosoma brucei*	Humans
Cockcroach	*Periplaneta sp.*	Bacteria and viruses	Salmonella, Polio virus	Humans
Sand fly	*Phlebotominae sp.*	Leishmaniasis	Leishmania parasites	Humans
White fly	*Bemisia tabaci*	Cassava mosaic disease	Cassava mosaic virus	Cassava plants

[3] looks at that of the white-flies, their interaction and disease transmission with the cassava plants. Nipa and Allen in [12] study a general vector–host interaction model. Table 1 summarizes some disease transmitting vectors and the diseases they can transmit.

4.2 The Disease Transmission

Disease transmission for indirectly transmitted infections occurs when the pathogen is passed from the host to the disease vector and again when it is passed from the disease vector to the host. So, for these kinds of diseases, infection occurs twice: In the first instance, the vector gets infected by the human host and the disease then has time to mature within the vector and in the second case a second human is infected by the vector. For such an infection to occur, there must be a successful interaction between the disease vector and the host, as well as a successful transfer of the pathogen from host to vector or from vector to host. Such double infections would have to occur over distinct points in time that are separated by the length of time equivalent to the maturation period of the infection in both organisms. In this regard, one may regard the pathogen as an opportunist that only exploits the life style of the disease vector, and see the vector as the active conduit that helps transfer the infection from one human to another. The second step then, towards a proper understanding of indirectly transmitted diseases of humans would then

be to seek to understand the indices of transmissibility; that is the quantifiable measurements of quantities that can be used to ascertain whether or not the disease will spread between and within the different populations upon introduction of at least one index case and factors that may facilitate this process, be it environmental and/or evolutionary, as in the chapters [3, 12] and [16].

4.3 The Disease Pathogen

The disease pathogens in some instances are living organism while others are not, and, normally could take advantage of the life style of the disease vector and then could divide its life cycle and developmental pathway so that part of it is in the host and the other part is in the disease vector. This type of division of life cycle of the organism can, from an evolutionary standpoint, be beneficial to the pathogen since at each time it has part of its progeny in two different biological entities, thereby making extinction difficult. For example, the pathogen that causes the disease malaria in humans has divided its life cycle such that one part of the life cycle is in the human and the other part in the mosquito. One major characteristic of a disease pathogen that has divided its life cycle to be able to survive in different organism is that only certain forms of the pathogen can grow in the different hosts. So there is a need to develop in essentially different ways, so as to produce the forms of the organism that can be transferred from host to the disease vector and vice versa. For example, the form of the malaria parasite than can be transmitted from humans to mosquitoes and begin development within the mosquito are called gametocytes and they are produced in the human. On the other hand, the form that can be transmitted to humans by mosquitoes and begin development in the human are called sporozoites and these are produced in the mosquito. Each of the hosts, humans and vectors, offers a different biological niche to the disease pathogen and the different parasite forms within these hosts introduce delays in the development and maturation times of the infection both in the human and in the vector. A second step, therefore, in the study of indirectly transmitted diseases would be to seek to understand how the disease pathogen weaves its own life style into the life style of its vector. In this case we may wish to understand how the disease agent affects the *behavior* of both the disease vector and the host as well as factors that allow them to survive and thrive, as in the chapters [2] and [15].

5 Discussion and Conclusion

The manuscripts published, herein, address different aspects of the diseases discussed. The results highlight the intricate interconnection between our planet and infectious diseases. Mother nature plays a role as it can drive seasonal fluctuations of diseases pathogens, as well as human factors. In [12], Nipa et al. used a stochastic vector–host indirectly transmissible framework to show how seasonal patterns and

variability in the demography of the vectors and host affect disease outbreaks for a general vector-borne disease.

With malaria as a specific example, Teboh-Ewungkem et al. in [16] highlighted one aspect of indirectly transmitted diseases, which has often been neglected in mathematical models. They highlighted the fact that for an indirectly transmitted infectious disease to spread within human populations each disease vector must interact with two different humans at two different time periods in the following way: In the first instance, the vector interacts with human A and may pick up the infection. Then the vector, if it successfully picked the infection, survives through the incubation period and becomes infectious and then interacts with human $B \neq A$, it can then pass on the infection to the human B. So, since it is the vector that actively seeks the humans, the infected vector must interact with at least two different humans at two different points in time (the length of this time difference will be compared with the length of the incubation period in the vector to infer infectiousness of the vector) to propagate the infection into the human population. On the other hand, many vectors can pick up the infection from one human simultaneously, which is an important aspect to consider when writing down models for indirectly transmitted diseases. The authors noted that any realistic mathematical model should take this aspect into consideration, as well as the fact that transmission of the infection is contingent on effective interactive contact. Thus, the reasons why disease vectors of humans are efficient vectors for human diseases are partly because the need to interact with humans is hard-wired into the physiological make of the disease vector as the interaction is tied to the survivability of the vectors next generation. For example, the mosquito, the vector that transmits malaria interacts with humans because the female adult needs to harvest blood from humans that she needs for the maturation of her eggs. Understanding and capturing these aspects as malaria continues to affect millions and kill many, especially children.

Instead of the human malaria disease, Chapwanya and Dumont [3] investigated the cassava mosaic virus disease that is also transmitted by a vector, in this case the white fly. The disease affects the cassava plant, a plant that forms an integral perennial crop serving as an important staple food for millions in the African continent on a regular basis and as food security against famine, since its roots can last a long time in the ground. Thus, understanding the interactions between the vectors that transmit the disease, the disease causing virus and the cassava plants are essential to the health and survival status of a significant population on our planet, those that rely on the cassava plant for nutrition. Thus controlling the disease and minimizing infection transmission is important.

Diseases with multiple transmission pathways can introduce challenges especially with control. Thus questions about which pathway might be dominant are important because they will impact how control measures are applied. Moreover, how these transmission pathways are adapted to the environment, climate, and our planet is important because a non-dominant pathway may potentially still serve as a source allowing for the potentials of increased epidemic frequencies and posing a challenge to disease control. In the chapter [6], where Levy et al. studied the viral disease, Pseudorabies, in Wild Hog populations in the Smokey

Mountains, they illustrated that the disease has multiple transmission routes: In particular, they highlighted four transmission routes, namely: direct route though density dependent contacts between wild hogs, or in the process of mating or during nursing, from mothers to piglets, and the fourth being as a result of stress, with carriers becoming re-infected. Notice that the first three routes are tied to the life style and evolutionary need for survival and protection of the next generation. While the fourth route, linked to stress factors, can be attributed to both natural causing and man-made factors since stress could be generated as a result of many factors including overcrowding, raised water levels, food shortage, etc. Regardless of the transmission route, pseudorabies in wild hogs impacts other domestic and wild animals in the regions inhabited by the diseased animals and has implications to our planet. Hence finding ways to curtail the spread of this disease is important for the ecosystems that interact with these wild hog populations.

In order to find effective and sustainable control measures for diseases with multiple transmission pathways, we must first understand the pathways and understand what factors are the greatest propagators of the diseases for specific pathways or how these pathways may be intricately tied to the demography of the populations affected. Islam et al., in [5], examined the transmission dynamics *Batrachochytrium salamandrivorans* (Bsal), an emerging fungal pathogen that affects the North American salamander population. As noted in [4], salamanders play a significant role in our ecosystem as they perform various ecological functions and provide ecosystem services that benefit the human race. Noted examples include: their service as pest controllers as they feed on mosquitoes; their service as nutrition for other larger animals, hence maintaining a balance in the ecosystem; their service as health indicators of the ecosystem due to their vulnerability and susceptibility to environmental toxic substances and drought; their role in carbon cycling and hence link to climate change, and their service as pets for humans, to name a few. Thus, maintaining a healthy salamander population is tied to the health of our planet. Thus models to investigate how to prevent and control any potential *Batrachochytrium salamandrivorans* (Bsal) infection is important. In this light, the authors Islam et al., in [5], investigated the invasion potential of the Bsal fungal infection in a population of Eastern Newts and showed that population density was a factor in the form of transmission that was dominant. In particular, for small population densities, they showed that the dominant transmission pathway was the direct host-to-host contact transmissions, meanwhile for larger population densities, environmental transmission was the dominant transmission form. Hence, the form of control should be tied to population density information in order to achieve a greater success of inhibiting invasion.

Control of the diseases that affect the humans and animals in our planets requires an understanding of all aspects tied to the disease. For malaria, for example, one has to view the disease as a three component problem: the parasite component, the human component, and the vector component, all interacting. The chapter in [16] captures all three components, with the vector as diseases drivers and the human serving as components infected by the malaria parasite. The chapters in [2] and [15] focus on the parasite component in the human. When a parasite infects a human,

it transforms and change forms with the end goal being survival in the human, leading to the production of the transmissible forms of the parasites, the mature gametocytes. However, during such an infection the body establishes a defense mechanism, with T cells being part of the defense mechanism cells. For malaria in particular, two types of immune response are reported—innate and adaptive, developed due to repeated exposure to malaria. In [2], Buri et al. studied the principles involved in the activation of T cells after a pathogen invasion. Their focus was on the adaptive immune system T cells and its activation, a process that requires that at least two signals be received before the T cells can be activated.

Within human host malaria parasites can respond positively to effective pharmaceutical control measures, if there is no drug resistance to the parasite. This is a primary method of disease control in a malaria-infected and sick patient. If, however, the parasite form is resistant to the administered drugs, then there is a problem and disease control and or eradication is not achieved. Schneider, in the chapter [15], investigated the dynamics of resistance—conferring mutations and their resulting impacts on genetic mutation. There, the author shows that the multiplicity of infection (MOI), defined as the number of super infections during the course of an infection, mediates the interplay between selection and recombination.

For diseases such as HIV/AID and Ebola, human factors are an integral aspects of control, especially with no actual cure (even though for HIV, there are medications that can decrease transmissibility). In [7], the authors highlighted that with the goal to reduce the global HIV burden, uneven exposure between trial participants in randomized controlled trials much be considered as they can affect the effectiveness of the trial. In particular, their results showed that effectiveness decreased with HIV exposure rate and trial length. In [14], Odunyi et al. studied media effects on an Ebola model that incorporates isolated and non-isolated individuals, as well as sexually infectious individuals. They showed that increased media effect is correlated with lower Ebola disease epidemic peak and this peaks lags behind the epidemic peak without media coverage.

In all, several factors contribute to the transmission of diseases within our planet. Some aspects are enhanced by human factors like deforestation, etc., while others by natural factors such as the climatic aspects of the regions affected by the specific disease. Overall, understanding how to control these diseases, as well as the actions we must take, are vital and require an understanding of the underlying process.

References

1. Agusto, F.B., Teboh-Ewungkem, M.I., Gumel, A.B.: Mathematical assessment of the effect of traditional beliefs and customs on the transmission dynamics of the 2014 ebola outbreaks. BMC Med. **13**(1), 96 (2015)
2. Buri, G., Zelleke, G.M., Ndifon, W.: Dynamic regulation of T cell activation by coupled feedforward loops. In: Teboh-Ewungkem, M.I., Ngwa, G.A. (eds.) Mathematics of Planet Earth: Infectious Diseases and Our Planet. Springer, Berlin, 2020

3. Chapwanya, M., Dumont, Y.: Application of mathematical epidemiology to crop vector-borne diseases. The cassava mosaic virus disease case. In: Teboh-Ewungkem, M.I., Ngwa, G.A. (eds.) Mathematics of Planet Earth: Infectious Diseases and our Planet. Springer, 2020
4. Gray, M.J., Lewis, J.P., Nanjappa, P., Klocke, B., Pasmans, F., Martel, A., Stephen, C., Olea, G.P., Sacerdote-Velat, A., Smith, S.A., Christman, M.R., Williams, J.M., Olson, D.H.: *Batrachochytrium salamandrivorans*: The North American response and a call for action. PLoS Pathog. **11**(12), e1005251 (2015)
5. Islam, M.R., Gray, M.J., Peace, A.: Identifying the dominant transmission pathway in a multi-stage infection model of the emerging fungal pathogen *Batrachochytrium Salamandrivorans* on the Eastern Newt. In: Teboh-Ewungkem, M.I., Ngwa, G.A. (eds.) Mathematics of Planet Earth: Infectious Diseases and our Planet. Springer, Berlin, 2020
6. Levy, B., Lenhart, S., Collins, C., Stiver, W.: Evidence for multiple transmission routes for pseudorabies in wild hogs. In: Teboh-Ewungkem, M.I., Ngwa, G.A. (eds.) Mathematics of Planet Earth: Infectious Diseases and our Planet. Springer, Berlin, 2020
7. Moore, J.R., M.-C. Boily, Donnell, D.J., Dimitrov, D.: Reducing the global HIV burden: the importance of uneven exposure to the results of HIV prevention trials. In: Teboh-Ewungkem, M.I., Ngwa, G.A. (eds.) Mathematics of Planet Earth: Infectious Diseases and our Planet. Springer, Berlin, 2020
8. Ngwa, G.A.: On the population dynamics of the malaria vector. Bull. Math. Biol. **68**(8), 2161–2189 (2006)
9. Ngwa, G.A., Teboh-Ewungkem, M.I.: A mathematical model with quarantine states for the dynamics of ebola virus disease in human populations. Comput. Math. Methods Med. **2016**, 9352725 (2016)
10. Ngwa, G.A., Woldegerima, W.A. Teboh-Ewungkem, M.I.: A mathematical study of the implict role of innate and adaptive immune responses on within-humna plasmodium falciparum parasite levels. J. Biol. Syst. **28**(02), 377–429
11. Ngwa, G.A., Wankah, T.T., Fomboh-Nforba, M.Y., Ngonghala, C.N., Teboh-Ewungkem, M.I.: On a reproductive stage-structured model for the population dynamics of the malaria vector. Bull. Math. Biol. **76**, 2476–2516 (2014)
12. Nipa, K.F., Allen, L.J.S.: The effect of demographic variability and periodic fluctuations on disease outbreaks in a vector-host epidemic model. In: Teboh-Ewungkem, M.I., Ngwa, G.A. (eds.) Mathematics of Planet Earth: Infectious Diseases and our Planet. Springer, Berlin, 2020
13. Nourridine, S., Teboh-Ewungkem, M.I., Ngwa, G.A.: A mathematical model of the population dynamics of disease transmitting vectors with spatial consideration. J. Biol. Dyn. **5**(4), 335–365 (2011)
14. Oduniyi, E., Gibbons, B., Oh, M., Agusto, F.B.: Modeling ebola transmission dynamics with media effects on disease and isolation rates. In: Teboh-Ewungkem, M.I., Ngwa, G.A. (eds.) Mathematics of Planet Earth: Infectious Diseases and our Planet. Springer, 2020
15. Schneider, K.A.: Charles Darwin meets Ronald Ross: a population-genetic framework for the evolutionary dynamics of malaria. In: Teboh-Ewungkem, M.I., Ngwa, G.A. (eds.) Mathematics of Planet Earth: Infectious Diseases and our Planet. Springer, Berlin, 2020
16. Teboh-Ewungkem, M.I., Ngwa, G.A., Fomboh-Nforba, M.Y.: A multistage mosquito-centered mathematical model for malaria dynamics that captures mosquito gonotrophic cycle contributions to its population abundance and malaria transmission. In: Teboh-Ewungkem, M.I., Ngwa, G.A. (eds.) Mathematics of Planet Earth: Infectious Diseases and our Planet. Springer, Berlin (2020)

The Effect of Demographic Variability and Periodic Fluctuations on Disease Outbreaks in a Vector–Host Epidemic Model

Kaniz Fatema Nipa and Linda J. S. Allen

1 Introduction

Seasonal changes often impact disease incidence due to the effect on the dynamics of the pathogen or the host. This is especially true for vector-transmitted diseases such as malaria, dengue, Zika, Lyme disease, chikungunya, or leishmaniasis, but it is also true for human and avian influenza, where changes in contacts such as in school settings or seasonal changes in temperature impact transmission [4, 6, 10, 12, 23, 27, 28]. How climate change will impact seasonal variations in disease prevalence raises some important public health questions, especially for vector-transmitted diseases [9, 18]. Mathematical modeling is an important tool to help address some of these questions.

Deterministic and stochastic epidemic models capture seasonal contact or other types of seasonal behavior via periodic coefficients [4–6, 8, 11, 12, 16, 19, 20, 23, 26–29]. Predicting the occurrence of disease outbreaks with periodic environments is more complex than in a constant environment. Recent studies on deterministic and stochastic epidemic models with periodic environments have shown that it is not sufficient to consider the average of the basic reproduction number [4–6, 26]. In this investigation, we extend some of these stochastic studies to a stochastic periodic vector–host epidemic model, a nonhomogeneous stochastic model, to investigate how seasonality and demographic variability impact the probability of a disease outbreak. In particular, a multitype branching process approximation is used to calculate the probability of a disease outbreak [5]. The approximation for probability of a disease outbreak is computed from a system of ordinary differential equations which is derived from the backward Kolmogorov differential

K. F. Nipa · L. J. S. Allen (✉)
Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX, USA
e-mail: linda.j.allen@ttu.edu

M. I. Teboh-Ewungkem, G. A. Ngwa (eds.), *Infectious Diseases and Our Planet*,
Mathematics of Planet Earth 7, https://doi.org/10.1007/978-3-030-50826-5_2

equations. The approximation shows that the risk of a disease outbreak is periodic and depends on the particular time at which either an infected vector or an infected host is introduced into entirely susceptible vector and susceptible host populations. Numerical examples with periodic transmission rates for vector and host illustrate the times at which there is the greatest probability of an outbreak and also demonstrate how these times are related to the peak transmission rates for vector or host.

2 Vector–Host ODE Model

We develop methods for a simple vector–host model. The host population is divided into susceptible and infected individuals, where infected individuals either die or recover. We assume there are no disease-related deaths. The vector population is also divided into susceptible and infected individuals with births and deaths but with no recovery and no disease-related deaths. The host model is an SIS model, with the total population size equal to $N_s = S + I$, and vector population is an SI model, where the variables for healthy and infected vectors are denoted by M and V, respectively, with a total population size $N_m = M + V$. The host and vector demography is modeled with parameter b equal to the birth and the natural death rate of hosts, parameter c is equal to the birth and natural death rate of the vectors, and γ is the recovery rate of the host. The transmission rate from host to vector is $\beta_m(t)$ and from vector to host is $\beta_s(t)$ which are time-dependent periodic functions with the same period. That is,

$$\beta_i(t) = \beta_i(t + p), \quad i = m, s, \quad t \in (-\infty, \infty), \tag{1}$$

with period $p > 0$. The following compartmental diagram (Fig. 1) illustrates the relation between vector and host.

Written as a system of ordinary differential equations (ODEs), the vector–host model takes the following form:

$$\text{Infected} \begin{cases} \dot{I} &= \frac{\beta_s(t)}{N_s} SV - \gamma I - bI \\ \dot{V} &= \frac{\beta_m(t)}{N_m} MI - cV \end{cases} \tag{2}$$

$$\text{Healthy} \begin{cases} \dot{S} &= -\frac{\beta_s(t)}{N_s} SV + bN_s - bS + \gamma I \\ \dot{M} &= -\frac{\beta_m(t)}{N_m} MI + cN_m - cM. \end{cases} \tag{3}$$

The vector and host population sizes are constant, N_s and N_m, respectively, and therefore the disease-free equilibrium (DFE) for host and vector is $\bar{S} = N_s$ and $\bar{M} = N_m$. The existence and stability of a periodic endemic equilibrium could also

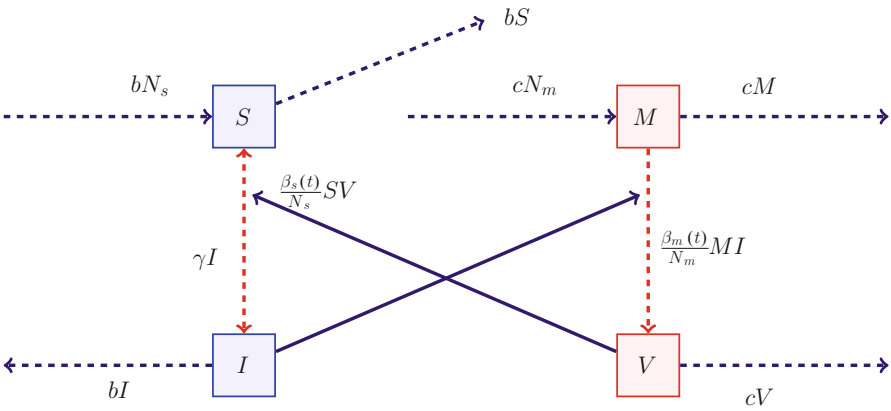


Fig. 1 Compartmental diagram for the vector–host model

be investigated as in [11] but we focus on the dynamics near the DFE to study the probability of an outbreak in the more general stochastic setting.

In the special case when the transmission rates are constant, $\beta_m = \bar{\beta}_m$ and $\beta_s = \bar{\beta}_s$, it is straightforward to compute the basic reproduction number from the next generation matrix as follows:

$$\bar{\mathcal{R}}_0 = \sqrt{\frac{\bar{\beta}_m \bar{\beta}_s}{c(b+\gamma)}}. \tag{4}$$

(See e.g. [25].) When this threshold parameter exceeds 1, a disease outbreak occurs. A variety of alternate forms for the basic reproduction number have also been defined known as type or target reproduction numbers [21, 22, 24]. For example, the square of the basic reproduction number in (4), $[\bar{\mathcal{R}}_0]^2$, is defined as the type reproduction number when control measures are applied only to the host or to the vector populations [13, 21]. Control measures that reduce the basic reproduction number also reduce the probability of an outbreak.

For nonautonomous systems of differential equations that model epidemics with time-periodic coefficients, Wang and Zhao [26] and Bacaër and Guernaoui [6] derived sufficient conditions for existence of the threshold parameter $\mathcal{R}_0$ (conditions (A1)–(A7) in the Appendix). These conditions depend on the linearized differential system for I and V at the DFE and the monodromy matrix (fundamental matrix) [26]. The linearized periodic differential equation for I and V satisfies

$$\begin{aligned} \dot{I} &= \beta_s(t)V - \gamma I - bI \\ \dot{V} &= \beta_m(t)I - cV. \end{aligned} \tag{5}$$

For $X = (I, V)^T$, the linear system can be expressed as $\dot{X} = [\mathcal{F}(t) - \mathcal{V}(t)]X$, where

$$\mathcal{F}(t) = \begin{bmatrix} 0 & \beta_s(t) \\ \beta_m(t) & 0 \end{bmatrix} \text{ and } \mathcal{V}(t) = \begin{bmatrix} \gamma + b & 0 \\ 0 & c \end{bmatrix}. \tag{6}$$

The systems (2), (3), and (5) satisfy the assumptions (A1)-(A7) (see Appendix). Only in special cases, does there exist an explicit solution for $\mathcal{R}_0$ [16, 17, 26, 29]. However, the value of $\mathcal{R}_0$ for the periodic system can be numerically approximated by considering the system

$$\dot{X} = [\mathcal{F}(t)/\lambda - \mathcal{V}(t)]X. \tag{7}$$

The spectral radius of the fundamental matrix solution of this system, evaluated at $t = p$, is unity when $\lambda = \mathcal{R}_0$, i.e., the dominant Floquet multiplier of the monodromy matrix is unity. In particular, if $\Phi_{\mathcal{F}/\lambda-\mathcal{V}}(t)$ is the monodromy matrix and the spectral radius $\rho(\Phi_{\mathcal{F}/\lambda-\mathcal{V}}(p)) = 1$, then $\lambda = \mathcal{R}_0$ (Theorem 2.1 p. 704, [26]). To find a numerical approximation for $\mathcal{R}_0$, we use an increasing sequence λ_i, $i = 1, 2, \ldots$ with $\lambda_{i+1} - \lambda_i < 0.001$ until a particular value λ_i is reached whereby a nontrivial periodic solution is obtained satisfying $|X(t) - X(t + p)| < \epsilon$ for ϵ sufficiently small and t sufficiently large. We apply this numerical method in Sect. 5. Another numerical method to compute $\mathcal{R}_0$ was proposed by Posny and Wang [20]. This method approximates the operator eigenvalue problem by a matrix eigenvalue problem. Klausmeier proposed using a computer algebra system to compute the Floquet multipliers [14].

3 Nonhomogeneous Process

The ODE model serves as a framework to formulate the infinitesimal transition probabilities for a time-nonhomogeneous stochastic process. The variables are discrete-valued and time is continuous, $t \in [0, \infty)$,

$$S(t), I(t) \in \{0, 1, 2, 3, \ldots, N_s\} \text{ and } M(t), V(t) \in \{0, 1, 2, \ldots, N_m\}.$$

The ODE rates are used to define the infinitesimal transition probabilities and they are summarized in Table 1. Denote the changes in the healthy and infected states for a small interval of time as $(\Delta I(t), \Delta V(t), \Delta S(t), \Delta M(t))$, e.g., $\Delta I(t) = I(t + \Delta t) - I(t)$. Let $\Sigma(t)$ be the sum of all the transition rates,

$$\Sigma(t) = \frac{\beta_s(t)}{N_s} S(t)V(t) + b(I(t) + N_s + S(t)) + \gamma I(t) + \frac{\beta_m(t)}{N_m} M(t)I(t) + c(V(t) + N_m + M(t)).$$

The transmission rates are time-periodic, as in (1), with period $p > 0$. Due to the seasonal variation in the transmission rates, the process is nonhomogeneous in time. Let $Y(t) = (I(t), V(t), S(t), M(t))$ and $\Delta Y(t) = Y(t + \Delta t) - Y(t)$. For example, the probability of a host infection in a small period of time Δt is

$$\mathbb{P}\{\Delta Y(t) = (1, 0, -1, 0)|Y(t)\} = \frac{\beta_s(t)}{N_s} S(t)V(t)\Delta t + o(\Delta t).$$

4 Branching Process Approximation

Approximation of the nonhomogeneous stochastic process near the DFE leads to a nonhomogeneous multitype branching process in two variables, $I(t)$ and $V(t)$. The events 1–5 in Table 1 when $S(t) = N_s$, $M(t) = N_m$, define a multitype branching process for $I(t)$ and $V(t)$. The transition probabilities are linear in $I(t)$ and $V(t)$ with periodic parameters $\beta_m(t)$ and $\beta_s(t)$. For the branching process, we will let the initial time be τ and the final time be t. Then the transition probability from state (i, v) at time τ to state (j, k) at time t is defined as follows:

$$p_{(i,v),(j,k)}(\tau, t) = \mathbb{P}\{(I(t), V(t)) = (j, k)|(I(\tau), V(\tau)) = (i, v)\}, \tag{8}$$

where $\tau < t$.

Bacaër and Dads recently verified that the basic reproduction number $\mathcal{R}_0$ from the nonautonomous ODE model with periodic coefficients serves as a threshold for disease extinction in the multitype branching process [5]. In particular, Bacaër and Dads proved a general result for a multitype branching process consisting of n types. Here, we state their assumptions and their results in terms of our system (5)–(6).

Let $\mathbb{P}_{ext}$ be the asymptotic probability of disease extinction. Bacaër and Dads showed that $\mathbb{P}_{ext}$ is a p-periodic function of τ, where τ is the time at which an infected individual is introduced into the population.

Theorem 4.1 (Proposition, p. 36 [5]) *Assume the linear periodic system (5)–(6) satisfies*

Table 1 Infinitesimal transition probabilities for the stochastic vector–host model

Event	Description	$(\Delta I(t), \Delta V(t), \Delta S(t), \Delta M(t))$	Probabilities
1	Host infection	$(1, 0, -1, 0)$	$\frac{\beta_s(t)}{N_s} S(t)V(t)\Delta t + o(\Delta t)$
2	Infected host death	$(-1, 0, 0, 0)$	$bI(t)\Delta t + o(\Delta t)$
3	Host recovery	$(-1, 0, 1, 0)$	$\gamma I(t)\Delta t + o(\Delta t)$
4	Vector infection	$(0, 1, 0, -1)$	$\frac{\beta_m(t)}{N_m} M(t)I(t)\Delta t + o(\Delta t)$
5	Infected vector death	$(0, -1, 0, 0)$	$cV(t)\Delta t + o(\Delta t)$
6	Healthy host birth	$(0, 0, 1, 0)$	$bN_s\Delta t + o(\Delta t)$
7	Healthy vector birth	$(0, 0, 0, 1)$	$cN_m\Delta t + o(\Delta t)$
8	Healthy host death	$(0, 0, -1, 0)$	$bS(t)\Delta t + o(\Delta t)$
9	Healthy vector death	$(0, 0, 0, -1)$	$cM(t)\Delta t + o(\Delta t)$
10	No change	$(0, 0, 0, 0)$	$1 - \Sigma(t)\Delta t + o(\Delta t)$

(H1) $\mathcal{F}(t)$ *is nonnegative with at least one entry strictly positive for* $t \geq 0$.
(H2) $K(t) = \mathcal{F}(t) - \mathcal{V}(t)$ *is irreducible for all* $t \geq 0$.
(H3) $\mathcal{F}(t)$ *and* $\mathcal{V}(t)$ *are piecewise continuous and* $p-$*periodic for* $t \geq 0$.
(H4) $\mathcal{V}(t) = (v_{ij}(t))$ *has nonpositive off-diagonal elements,* $v_{ij}(t) \leq 0$, $i \neq j$, $t \geq 0$ *and has positive diagonal elements,* $v_{ii}(t) \geq A > 0$, $t \geq 0$.

If $\mathcal{R}_0 \leq 1$, *then the multitype branching process with the same periodic rates has an asymptotic probability of extinction* $\mathbb{P}_{ext} = 1$ *and if* $\mathcal{R}_0 > 1$, *then the multitype branching process has an asymptotic probability of extinction that is a* $p-$*periodic function* $\mathbb{P}_{ext}$ *satisfying* $0 < \mathbb{P}_{ext} < 1$.

The assumptions (H1)-(H4) hold for the linear periodic system (5)–(6). Most important, in their proof of this result, Bacaër and Dads verified when $\mathcal{R}_0 > 1$ that the solution $\mathbb{P}_{ext}(t) = p_{(i,v),(0,0)}(\tau, t)$ for the multitype branching process converges to a unique nontrivial periodic solution $\mathbb{P}_{ext}$ of period $p > 0$ [5].

In verifying Theorem 4.1, Bacaër and Dads applied the forward Kolmogorov differential equations. The backward Kolmogorov differential equations are often used in studying the extinction process as they provide a simpler method for numerically computing the probability of an outbreak. We apply the events 1–5 in Table 1 with $S(t) = N_s$ and $M(t) = N_m$ to derive the backward Kolmogorov differential equations for the two infectious groups I and V.

Let the initial time be τ with $\Delta\tau > 0$ a small time step and apply the rates from Table 1. The following relation follows for the infinitesimal transition probabilities:

$$\begin{aligned} p_{(i,v),(j,k)}(\tau - \Delta\tau, t) &= \beta_s(\tau - \Delta\tau)vp_{(i+1,v),(j,k)}(\tau, t)\Delta\tau \\ &\quad + (\gamma + b)ip_{(i-1,v),(j,k)}(\tau, t)\Delta\tau \\ &\quad + \beta_m(\tau - \Delta\tau)ip_{(i,v+1),(j,k)}(\tau, t)\Delta\tau + cvp_{(i,v-1),(j,k)}(\tau, t)\Delta\tau \\ &\quad + [1 - (\beta_s(\tau - \Delta\tau)v + (\gamma + b)i + \beta_m(\tau - \Delta\tau)i + cv)\Delta\tau] \\ &\quad \times \left[p_{(i,v),(j,k)}(\tau, t)\right] + o(\Delta\tau). \end{aligned}$$

Subtracting $p_{(i,v),(j,k)}(\tau, t)$ from both sides, dividing by $\Delta\tau$, and letting $\Delta\tau \to 0$, the backward Kolmogorov differential equations are

$$\begin{aligned} -\frac{\partial p_{(i,v),(j,k)}(\tau, t)}{\partial \tau} &= \beta_s(\tau)vp_{(i+1,v),(j,k)}(\tau, t) + (\gamma + b)ip_{(i-1,v),(j,k)}(\tau, t) \\ &\quad + \beta_m(\tau)ip_{(i,v+1),(j,k)}(\tau, t) + cvp_{(i,v-1),(j,k)}(\tau, t) \\ &\quad - [\beta_s(\tau)v + (\gamma + b)i + \beta_m(\tau)i + cv]p_{(i,v),(j,k)}(\tau, t). \end{aligned} \tag{9}$$

We also define offspring probability generating functions (pgfs) for an infected host or an infected vector and a more general pgf for the entire population of infected hosts and infected vectors at the DFE. Probability generating functions are power series representations of random variables. For example, an offspring pgf for each

of the two types of infectious stages, I and V, is a power series in the variables u_1 and u_2. It has the general form,

$$f_k(u_1, u_2) = \sum_{i=0}^{\infty} \sum_{v=0}^{\infty} \mathbb{P}_k(i, v) u_1^i u_2^v, \quad u_i \in [0, 1], \quad i = 1, 2,$$

where $k = I, V$ are the two types. For the random variable I, the value of $\mathbb{P}_I(i, v)$ is the probability that in a small period of time one I individual transmits the infection to i individuals of the same type and to v individuals of type V. The infected individual transmitting the disease is also counted as a new infection, as the individual remains infectious after transmitting the infection. A similar definition applies to the random variable V. In our model, the probabilities $\mathbb{P}_k(i, v)$ are time-dependent. Also an infected host only transmits the infection to a vector and an infected vector only transmits to a host. The probabilities come from the assumptions in Table 1, where $S \approx N_s$, $M \approx N_m$, and initially either $I = 1$ or $V = 1$.

Let τ be the initial time and let $(I(\tau), V(\tau)) = (1, 0)$ or $(0, 1)$, denoted as e_1 or e_2, respectively. At the DFE, for an infected host, there are only three events that occur in a small period of time $\Delta\tau$, either recovery or death of the host or infection of a vector (other events have probability of order $o(\Delta\tau)$ in Table 1). The offspring pgf for the host I, given $(I(\tau), V(\tau)) = e_1$, is

$$f_I(u_1, u_2, \tau) = \frac{\beta_m(\tau) u_1 u_2 + \gamma + b}{\beta_m(\tau) + \gamma + b}, \quad u_i \in [0, 1], \quad i = 1, 2. \tag{10}$$

That is, one infected host recovers with probability $\gamma/(\beta_m(\tau) + \gamma + b)$ or dies with probability $b/(\beta_m(\tau) + \gamma + b)$ or infects a vector with probability $\beta_m(\tau)/(\beta_m(\tau) + \gamma + b)$. Note the term $u_1 u_2$ means one infected host generates one infected vector (u_2 raised to the power one) and remains infectious (u_1 raised to the power one). Similarly, the offspring pgf for a vector V, given there is one infected vector at time τ, $(I(\tau), V(\tau)) = e_2$, is

$$f_V(u_1, u_2, \tau) = \frac{\beta_s(\tau) u_1 u_2 + c}{\beta_s(\tau) + c}, \quad u_i \in [0, 1], \quad i = 1, 2. \tag{11}$$

That is, one infected vector dies with probability $c/(\beta_s(\tau) + c)$ or infects a host with probability $\beta_s(\tau)/(\beta_s(\tau) + c)$.

We define another generating function for all hosts I and vectors V at the DFE, to obtain differential equations for the probability of extinction. The generating function for I and V, given $(I(\tau), V(\tau)) = (i, v)$, is the following expectation:

$$G_{(i,v)}(u_1, u_2, \tau, t) = \mathbb{E}\left[u_1^{I(t)} u_2^{V(t)} | (I(\tau), V(\tau)) = (i, v)\right]. \tag{12}$$

The generating function can be expressed in terms of the transition probability in (8),

$$G_{(i,v)}(u_1, u_2, \tau, t) = \sum_{j,k\geq 0} p_{(i,v),(j,k)}(\tau, t)u_1^j u_2^k. \tag{13}$$

For $j = k = 0$, the coefficient of the preceding sum is the probability of extinction, $p_{(i,v),(0,0)}(\tau, t)$ which is also equal to $G_{(i,v)}(0, 0, \tau, t)$.

We assume independence of the random variables I and V in the branching process approximation. This is a reasonable assumption when the size of the total population is large and the number of infected hosts and vectors is small. Mathematically, if initially $(I(\tau), V(\tau)) = (i, v)$, then the generating function $G_{(i,v)}$ is the product of i and v generating functions beginning from e_1 and e_2, respectively,

$$G_{(i,v)}(u_1, u_2, \tau, t) = [G_{e_1}(u_1, u_2, \tau, t)]^i [G_{e_2}(u_1, u_2, \tau, t)]^v. \tag{14}$$

For initial data $(I(\tau), V(\tau)) = (i, v)$, the probability of extinction at time t is given by

$$p_{(i,v),(0,0)}(\tau, t) = [p_{e_1,(0,0)}(\tau, t)]^i [p_{e_2,(0,0)}(\tau, t)]^v. \tag{15}$$

The backward Kolmogorov differential equations lead to a simple system of differential equations that can be solved numerically to compute the probability of ultimate extinction as $t \to \infty$.

For simplicity, let $G_{(i,v)} \equiv G_{(i,v)}(u_1, u_2, \tau, t)$ and $p_{(i,v),(j,k)} \equiv p_{(i,v),(j,k)}(\tau, t)$. Taking the derivative of $G_{(i,v)}$ with respect to τ in (14) and applying the identities $G_{(i,0)} = G_{e_1}^i$ and $G_{(0,v)} = G_{e_2}^v$ yield

$$\frac{\partial G_{e_1}}{\partial \tau} = \frac{1}{i[G_{(i-1,0)}]} \sum_{j,k} \frac{\partial p_{(i,0),(j,k)}}{\partial \tau} u_1^j u_2^k$$

$$\frac{\partial G_{e_2}}{\partial \tau} = \frac{1}{v[G_{(0,v-1)}]} \sum_{j,k} \frac{\partial p_{(0,v),(j,k)}}{\partial \tau} u_1^j u_2^k.$$

Substituting the backward Kolmogorov differential equations (9) into the right side, applying the identity (14), the offspring pgfs (10) and (11) and simplifying, the differential equation for G_{e_1} is given by

$$\begin{aligned}
\frac{\partial G_{e_1}}{\partial \tau} &= -\frac{1}{G_{(i-1,0)}} \sum_{j,k} \Big[(\gamma + b) p_{(i-1,0),(j,k)} u_1^j u_2^k + \beta_m(\tau) p_{(i,1),(j,k)} u_1^j u_2^k \\
&\quad - (\gamma + b + \beta_m(\tau)) p_{(i,0),(j,k)} u_1^j u_2^k \Big] \\
&= -\gamma - b - \beta_m(\tau) G_{e_1} G_{e_2} + [\gamma + b + \beta_m(\tau)] G_{e_1} \\
&= -(\beta_m(\tau) + b + \gamma)[f_I(G_{e_1}, G_{e_2}, \tau) - G_{e_1}].
\end{aligned} \tag{16}$$

Similarly, the differential equation for G_{e_2} has the form

$$\frac{\partial G_{e_2}}{\partial \tau} = -(\beta_s(\tau) + c)[f_V(G_{e_1}, G_{e_2}, \tau) - G_{e_2}]. \tag{17}$$

The differential equations in (16) and (17) are also referred to as the backward Kolmogorov differential equations for the branching process approximation for I and V [15]. Differential equations for the time-homogeneous branching processes are derived in a similar manner (see e.g., [1, 3]).

The differential equations in (16) and (17) do not depend explicitly on u_1, u_2 or t. If each of these variables is set to zero, then the probability of extinction from (13) and (14) in a time period of τ can also be written as

$$p_{(i,v),(0,0)}(\tau, 0) = [p_{e_1,(0,0)}(\tau, 0)]^i [p_{e_2,(0,0)}(\tau, 0)]^v. \tag{18}$$

The expression (18) is equivalent to (15). Therefore, the asymptotic probability of extinction can be approximated from the backward Kolmogorov differential equations (16) and (17).

The partial differential equations in (16)–(17) are ordinary differential equations when $u_1 = u_2 = 0$ and $t = 0$. Denote $G_{e_j}(0, 0, \tau, 0)$ as $P_j(\tau)$, $j = 1, 2$, respectively,

$$\frac{dP_1}{d\tau} = -(\beta_m(\tau) + b + \gamma)[f_I(P_1, P_2, \tau) - P_1] = F_1(P_1, P_2, \tau),$$
$$\frac{dP_2}{d\tau} = -(\beta_s(\tau) + c)[f_V(P_1, P_2, \tau) - P_2] = F_2(P_1, P_2, \tau).$$

These equations are solved backward in time. Let $\tau_k = -k\Delta\tau$, where $\Delta\tau > 0$. For example, Euler's method yields the approximation

$$P_j(\tau_{k+1}) \approx P_j(\tau_k) - \Delta\tau F_j(P_1(\tau_k), P_2(\tau_k), \tau_k), \quad P_j(0) = 0, \quad j = 1, 2.$$

For a large positive integer n, $\tau = -np \ll 0$ and small $\Delta\tau$, the solution for P_j converges to a periodic solution on $[-np, (1-n)p]$. To graph this solution forward in time, a change of variable is made to the interval $[0, p]$. That is, the value $P_j(\tau + p) = \mathbb{P}_{ext}(0)$ and the value $P_j(\tau) = \mathbb{P}_{ext}(p)$, so that the ultimate probability of extinction when $(I(t), V(t)) = e_j$ and $t \in [0, p]$ is approximated by $P_j(\tau + p - t)$ and the probability of an outbreak by $1 - P_j(\tau + p - t)$. Then for $(I(t), V(t)) = (i, v)$, the probability of an outbreak is approximated by

$$\mathbb{P}_{\text{outbreak}}(t) = 1 - [P_1(\tau + p - t)]^i [P_2(\tau + p - t)]^v. \tag{19}$$

In the special case of constant transmission rates, where the periodic transmission rates $\beta_j(t)$ are replaced by their time average, there exists an explicit formula for the

probability of an outbreak [2, 7]. It can be expressed in terms of the average basic reproduction number [2], that is,

$$\bar{\mathbb{P}}_{\text{outbreak}} = 1 - (q_I)^i (q_V)^v, \tag{20}$$

where the values of q_I and q_V are

$$q_I = \frac{\bar{\beta}_m}{\bar{\beta}_m + b + \gamma} \left(\frac{1}{\bar{\mathcal{R}}_0} \right)^2 + \frac{b + \gamma}{\bar{\beta}_m + b + \gamma}, \tag{21}$$

$$q_V = \frac{\bar{\beta}_s}{\bar{\beta}_s + c} \left(\frac{1}{\bar{\mathcal{R}}_0} \right)^2 + \frac{c}{\bar{\beta}_s + c}. \tag{22}$$

In this special case, the stochastic process is a time-homogeneous Markov chain.

5 Numerical Examples

Several numerical examples illustrate the results in the preceding sections. Three sets of periodic transmission rates are assumed,

$$\begin{aligned} \beta_s(t) &= \bar{\beta}_s \left(1 + \delta_s \cos\left(\frac{2\pi t}{p} \right) \right), \\ \beta_m(t) &= \bar{\beta}_m \left(1 + \delta_m \cos\left(\frac{2\pi t}{p} \right) \right), \end{aligned} \tag{23}$$

or

$$\begin{aligned} \beta_s(t) &= \bar{\beta}_s \left(1 + \delta_s \cos\left(\frac{2\pi t}{p} \right) \right), \\ \beta_m(t) &= \bar{\beta}_m \left(1 + \delta_m \sin\left(\frac{2\pi t}{p} \right) \right), \end{aligned} \tag{24}$$

or

$$\begin{aligned} \beta_s(t) &= \bar{\beta}_s \left(1 + \delta_s \sin\left(\frac{2\pi t}{p} \right) \right), \\ \beta_m(t) &= \bar{\beta}_m \left(1 + \delta_m \cos\left(\frac{2\pi t}{p} \right) \right), \end{aligned} \tag{25}$$

where the period $p = 20$, $\delta_m = \delta_s = 0.9$, and the average transmission rates are $\bar{\beta}_s = 0.04$ and $\bar{\beta}_m = 2$. The period represents a length of 1 year. In the numerical examples, the effects of vector death rate c and host recovery rate γ on probability of

Table 2 Parameter values for the vector–host model

Parameter	Description	Value
N_s	Total number of hosts	100
N_m	Total number of vectors	1000
$\bar{\beta}_s$	Transmission rate from vector to host	0.04
$\bar{\beta}_m$	Transmission rate from host to vector	2.0
γ	Recovery rate of the host	0.05–0.4
b	Birth and death rate of hosts	0.02
c	Birth and death rate of vectors	0.05–0.4
δ_s	Amplitude of transmission $\beta_s(t)$	0.9
δ_m	Amplitude of transmission $\beta_m(t)$	0.9
p	Period of transmission rate	20

an outbreak are considered. A summary of the parameter values for the numerical examples is given in Table 2. The specific parameter values are hypothetical but biologically meaningful. They are chosen for illustration purposes. For example, a recovery rate $\gamma = 0.2$ implies an average of 5 units of time which is equal to 3 months. The birth and death rates in the host are assumed smaller than in the vector, $b \leq c$, because the population turnover is faster in the vector than in the host. The average transmission rate from host to vector is chosen to be greater than the average transmission rate from vector to host, $\bar{\beta}_m > \bar{\beta}_s$. Also, we use sine and cosine functions to represent the periodic transmission rates. Such types of functions have been used in the literature for vector-transmitted diseases [11, 20, 26]. The phase shift between the two transmission rates in (24) and (25) is chosen to help identify which of the two transmission rates have the greatest impact and how it affects the probability of an outbreak. The value of $\mathcal{R}_0$ is computed from the solution of (7).

For each numerical example, we compare the multitype branching process estimate $\mathbb{P}_{\text{outbreak}}(t)$ in (19) with the probability of an outbreak in the full nonlinear nonhomogeneous stochastic model. To simulate the full nonlinear process, a Monte Carlo approach is used with a time step Δt chosen sufficiently small such that during each time step only one of the 10 events in Table 1 occurs. The Monte Carlo simulation was checked using a sequence of progressively smaller time steps to ensure accuracy. For $\Sigma(t)\Delta t < 1$, each of the events occurs with the probability given in Table 1, e.g., the probability of host infection is $\beta_s(t)S(t)V(t)\Delta t/N_s$ and the probability of no change is $1 - \Sigma(t)\Delta t$. To approximate the probability of an outbreak in the full nonhomogeneous stochastic process, 5000 sample paths are simulated for a given set of initial conditions, $(I(t), V(t)) = (I_0, V_0)$, $t \in [0, p]$. Each sample path continues until a time $T > t$ is reached when either $I(T) + V(T) = 0$ or $I(T) + V(T) = 50$. If the total infected population reaches 50, then it is assumed that there is an outbreak. The proportion of sample paths q out of 5000 that hit zero is an estimate for the probability of extinction of both I and V and the remaining proportion $1 - q$ that reach 50 is an estimate of the probability of an outbreak. Other threshold values larger than or smaller than 50 were also checked. Larger threshold values can be applied if the total population sizes are sufficiently

Table 3 The average basic reproduction number and the probabilities of an outbreak $\bar{\mathbb{P}}_{\text{outbreak}} = 1 - (q_I)^i (q_V)^v$ are computed for the time-homogeneous Markov process. The probability of an outbreak depends on the initial number of infected hosts and infected vectors, $(i, v) = (1, 0)$, $(0, 1)$ or $(2, 2)$ but not on the time of introduction. The average transmission parameters are $\bar{\beta}_s = 0.04$, $\bar{\beta}_m = 2$, and $b = 0.02$

Parameters			$\bar{\mathbb{P}}_{\text{outbreak}}$		
c	γ	$\bar{\mathcal{R}}_0$	$1 - q_I$	$1 - q_V$	$1 - (q_I)^2(q_V)^2$
0.05	0.1	3.65	0.87	0.41	0.99
0.1	0.1	2.58	0.80	0.24	0.98
0.2	0.1	1.83	0.66	0.12	0.91
0.4	0.1	1.29	0.38	0.036	0.64
0.1	0.05	3.38	0.88	0.26	0.99
0.1	0.2	1.91	0.65	0.21	0.92
0.1	0.4	1.38	0.39	0.14	0.72
0.05	0.05	4.78	0.92	0.425	0.998
0.2	0.2	1.35	0.41	0.075	0.698

large and $\mathcal{R}_0$ is not too close to 1. However, if both vector and host population sizes are small on the order of 100 or less or when $\mathcal{R}_0$ is close to 1, then the branching process approximation may have poor agreement with the numerical simulations.

The probability of an outbreak in the full nonhomogeneous stochastic process is also compared to the time-homogeneous Markov process, where the transmission parameters $\beta_j(t)$ are replaced by their time average, i.e., formula (20). Table 3 is a summary of the average basic reproduction number $\bar{\mathcal{R}}_0$ and probabilities of an outbreak, $\bar{\mathbb{P}}_{\text{outbreak}}$, for the time-homogeneous Markov process for a range of values for the vector death rate c and the host recovery rate γ.

In the first example, the time-periodic transmission rates $\beta_s(t)$ and $\beta_m(t)$ are modeled with cosine functions that have the same phase, Eqs. (23). Other parameter values are in Table 2 with $c = 0.1 = \gamma$ and a basic reproduction number $\mathcal{R}_0 = 2.58$. Graphed in Fig. 2 are the ODE solution and two sample paths of the nonhomogeneous stochastic process. In the nonhomogeneous stochastic process, the probability of an outbreak depends on the time of introduction of an infected host, either $t = 0$ or $t = 10$:

$$\mathbb{P}_{\text{outbreak}}(0) = 0.87 \text{ and } \mathbb{P}_{\text{outbreak}}(10) = 0.63.$$

The effect of the transmission rates $\beta_s(t)$ and $\beta_m(t)$ on the probability of an outbreak is illustrated in Figs. 3, 4, and 5. In Fig. 3, the transmission rates have the same phase, Eqs. (23). In Fig. 4, there is a phase shift in host to vector transmission, Eqs. (24) and in Fig. 5, there is a phase shift in vector to host transmission, Eqs. (25). In each figure, the transmission rates are graphed in the top two panels and the periodic probabilities of an outbreak $\mathbb{P}_{\text{outbreak}}(t)$ in the bottom two panels. The branching process estimate is verified by checking five different introductions of either one infected host or one infected vector at times $t = 0, 5, 10, 15, 20$ through

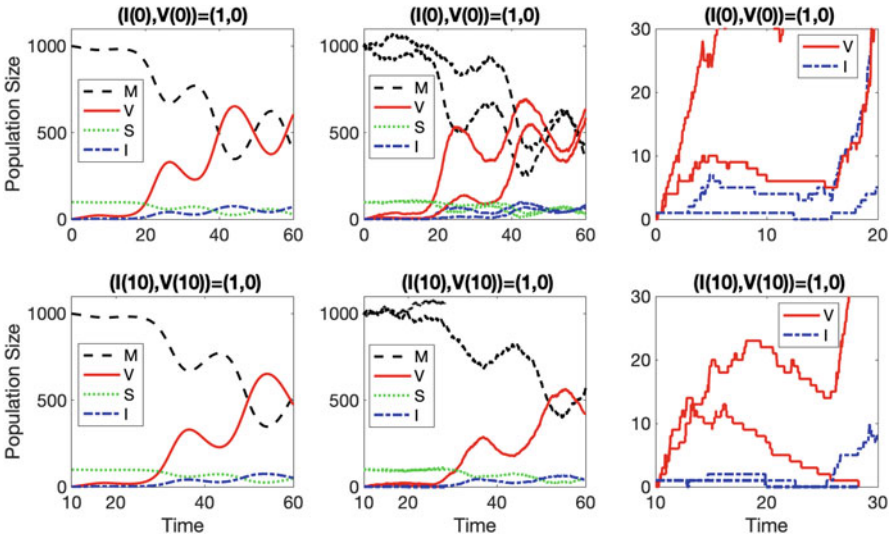


Fig. 2 Graphs of the ODE solution are in the first column and two sample paths of the nonhomogeneous stochastic process are in the second and third columns. The third column is a close-up view of the sample paths in the second column. The initial conditions depend on the time an infected host is introduced, either $t = 0$ in the top row or $t = 10$ in the bottom row. Three of the sample paths illustrate outbreaks but one sample path for the initial condition at $t = 10$ illustrates no outbreak. The transmission rates are given in (23) and parameter values are in Table 2 with $c = 0.1 = \gamma$. The basic reproduction number for the ODE model is $\mathcal{R}_0 = 2.58$ and the branching approximation estimates are $\mathbb{P}_{\text{outbreak}}(0) = 0.87$ and $\mathbb{P}_{\text{outbreak}}(10) = 0.63$

a Monte Carlo simulation of 5000 sample paths of the nonhomogeneous stochastic process. The results from the Monte Carlo simulations show good agreement with the branching process approximation.

Evident in Fig. 3 is the change in shape of the $\mathbb{P}_{\text{outbreak}}(t)$ as compared to the transmission rates. What might be expected is that the time of the lowest probability of an outbreak would correspond to the time of the smallest transmission rate and similarly, the time of the greatest probability of an outbreak with the time of the largest transmission rate. Instead the extrema of the probability of an outbreak are shifted left of the extrema of the transmission rates. The graphs in Figs. 4 and 5 also illustrate a shift left that is primarily dependent on whether the host or the vector initiates the outbreak. If an infected host initiates the outbreak, then the shift left depends on the transmission rate from host to vector and if an infected vector initiates the outbreak, then the shift depends on the transmission rate from vector to host.

In the remaining examples, the host recovery rate γ and the vector death rate c are varied (Fig. 6 and Tables 4, 5, and 6). Often vector control measures increase the vector death rate c whereas host prophylactic treatments increase the recovery rate γ. Figure 6 and Tables 4, 5, and 6 illustrate that increasing c or γ decreases the value of $\mathcal{R}_0$ and the probability of an outbreak, $\mathbb{P}_{\text{outbreak}}(t)$. Increasing the number

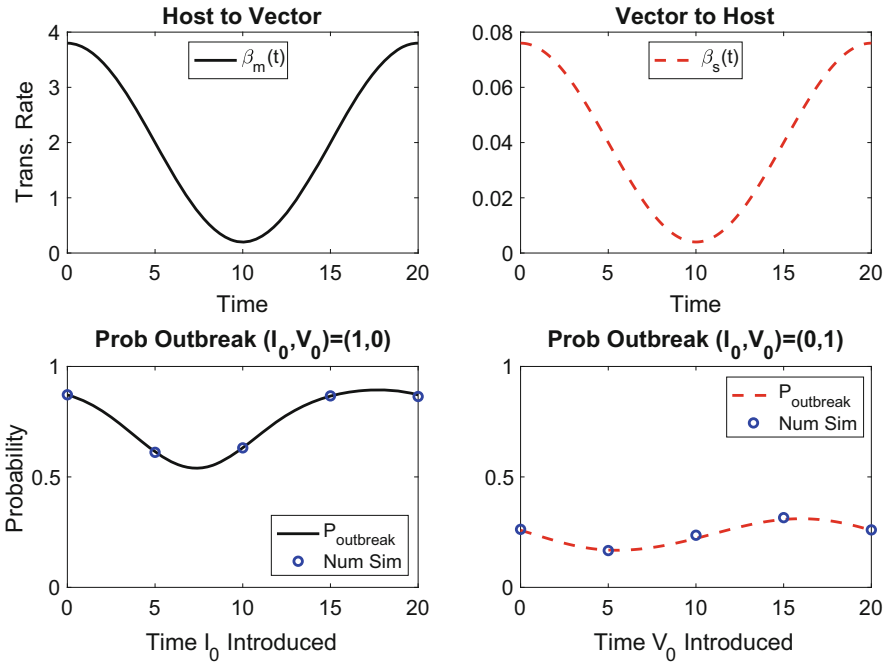


Fig. 3 The periodic transmission rates in Eqs. (23) are graphed in the top two panels and the probabilities of an outbreak $\mathbb{P}_{\text{outbreak}}(t)$ in the bottom two panels. Five different introductions of one infected host or one infected vector at times $t = 0, 5, 10, 15, 20$ are checked using the Monte Carlo simulation of the full nonhomogeneous process (Num Sim, circles). The solid black curves represent host to vector transmission and the probability of an outbreak for initial infection in the host, whereas the dashed red curves represent vector to host transmission and probability of an outbreak for initial infection in the vector. The initial conditions are either $(I_0, V_0) = (1, 0)$ or $(0, 1)$. The parameter values are the same as in Fig. 2. The basic reproduction number is $\mathcal{R}_0 = 2.58$

of infected hosts or vectors that are introduced, also increases the probability of an outbreak (Fig. 6 bottom panel).

The graphs of the periodic probabilities of an outbreak $\mathbb{P}_{\text{outbreak}}(t)$ in Fig. 6 can be compared to the constant probabilities of an outbreak $\bar{\mathbb{P}}_{\text{outbreak}}$. In Table 3, the values of $\bar{\mathbb{P}}_{\text{outbreak}} = 1 - (q_i)^i (q_V)^v$ are given for the time-homogeneous Markov process with initial conditions $(i, v) = (I_0, V_0) = (1, 0)$ and $(0, 1)$.

In Tables 4, 5, and 6, the average values of the periodic probability of an outbreak $\mathbb{P}_{\text{outbreak}}(t)$,

$$\widehat{P}_{\text{outbreak}} = \frac{1}{p} \int_0^p \mathbb{P}_{\text{outbreak}}(t)\, dt,$$

are recorded for a range of values of c and γ and for the transmission rates in Eqs. (23)–(25). The maximum and minimum values of $\mathbb{P}_{\text{outbreak}}(t)$ and the basic reproduction numbers $\mathcal{R}_0$ are also recorded. Comparing the values of $\mathcal{R}_0$ and

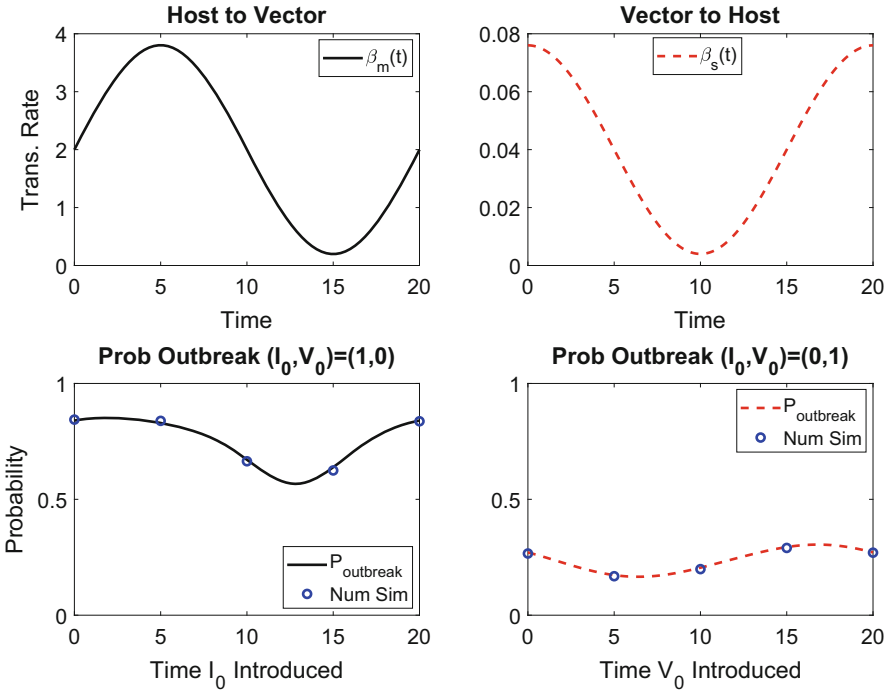


Fig. 4 The same parameters as in Fig. 3 are assumed, except that the transmission rates are given in (24) with a phase shift in host to vector transmission. The basic reproduction number is $\mathcal{R}_0 = 2.53$

$\widehat{P}_{\text{outbreak}}$, Tables 4, 5, and 6 columns 3–5, to the corresponding values in the time-homogeneous Markov process, Table 3 columns 3–5, there are slight differences. These differences increase when there is a phase shift in the time-dependent transmission rates, in Tables 5 and 6 versus Table 3. For all transmission rates, however, the average probabilities of an outbreak $\widehat{P}_{\text{outbreak}}$ are less than or equal to probabilities of an outbreak in the homogeneous process. Also, the extrema for the periodic probability of an outbreak $\mathbb{P}_{\text{outbreak}}(t)$ are shifted left of the corresponding extrema of the transmission rates. For example, in Table 4 the times of the extrema for the probability of an outbreak range from $t_M = 15.5{-}18.6$ and from $t_m = 4.6{-}8.0$, prior to the times of the extrema for the transmission rates at $t_M = 20$ and $t_m = 10$.

Some general observations can be made from Tables 4, 5, and 6. The times at which the extrema occur in the probabilities of an outbreak are in closer agreement to those of the transmission rates when $\mathcal{R}_0$ is large and when infection begins with the host, $I_0 = 1$. This can be seen by comparing the extrema of $\mathbb{P}_{\text{outbreak}}(t)$ to the host to vector transmission rate $\beta_m(t)$. For example, in Tables 4 and 6 for $\mathcal{R}_0 = 4.78$ and 4.68 and $I_0 = 1$, the extrema of $\mathbb{P}_{\text{outbreak}}(t)$ occur at $t_M = 18.6$ and 19.5 and $t_m = 7.9$ and 8.0, respectively, while the extrema of $\beta_m(t)$ occur at $t_M = 20$ and

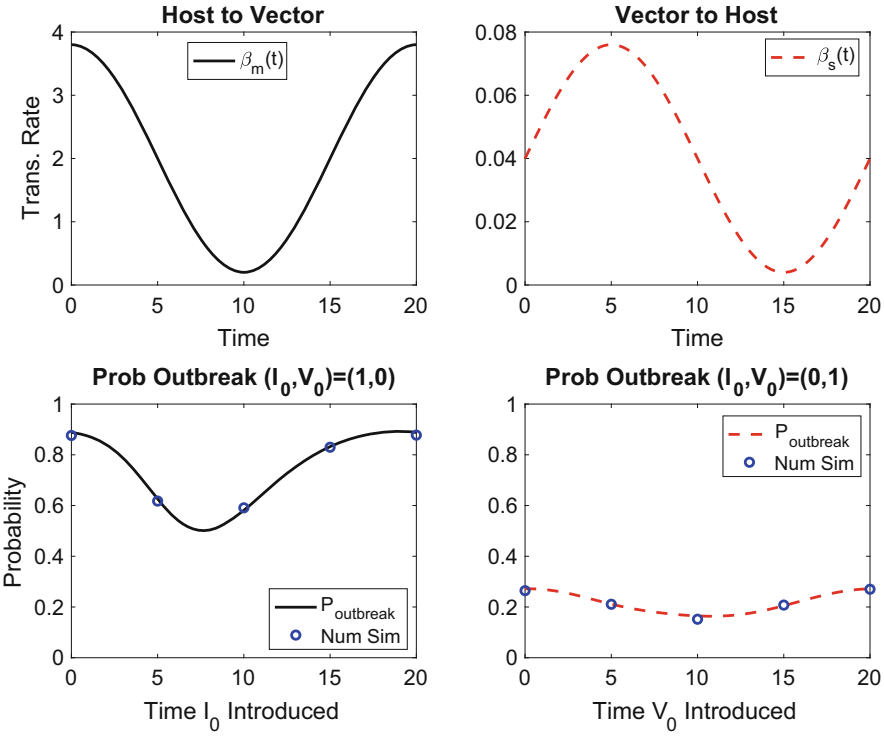


Fig. 5 The same parameters as in Fig. 3 are assumed, except that the transmission rates are given in (25) with a phase shift in vector to host transmission. The basic reproduction number is $\mathcal{R}_0 = 2.49$

$t_m = 10$. In Table 5 for $\mathcal{R}_0 = 4.79$ and $I_0 = 1$, the extrema for $\mathbb{P}_{\text{outbreak}}(t)$ occur at $t_M = 3.7$ and $t_m = 13.1$ while the extrema of $\beta_m(t)$ occur at $t_M = 5$ and $t_m = 15$.

Another observation is related to the effect of increasing the parameters c or γ on the probability of an outbreak. In Tables 5 and 6, there is a switch in the location of the extrema when either c or γ increases. As the vector birth and death rate c increases but γ is fixed, the extrema of $\mathbb{P}_{\text{outbreak}}(t)$ switch from dominance by host to vector transmission to vector to host transmission. A similar switch takes place if γ increases but c is fixed, but the dominance switches from vector to host transmission to host to vector transmission. For example, in Table 5, when $c = 0.4$ the maxima of $\mathbb{P}_{\text{outbreak}}(t)$ occur at $t_M = 18.5$ or 17.5 for $I_0 = 1$ or $V_0 = 1$, respectively. This is a shift left of the maximum of $\beta_s(t)$ which occurs at $t_M = 20$. A similar type of switch occurs in Table 6 when $\gamma = 0.4$. The maxima of $\mathbb{P}_{\text{outbreak}}(t)$ occur at $t_M = 18.2$ or 18.7 for $I_0 = 1$ or $V_0 = 1$, respectively, a shift left of the maximum of $\beta_m(t)$ which occurs at $t_M = 20$.

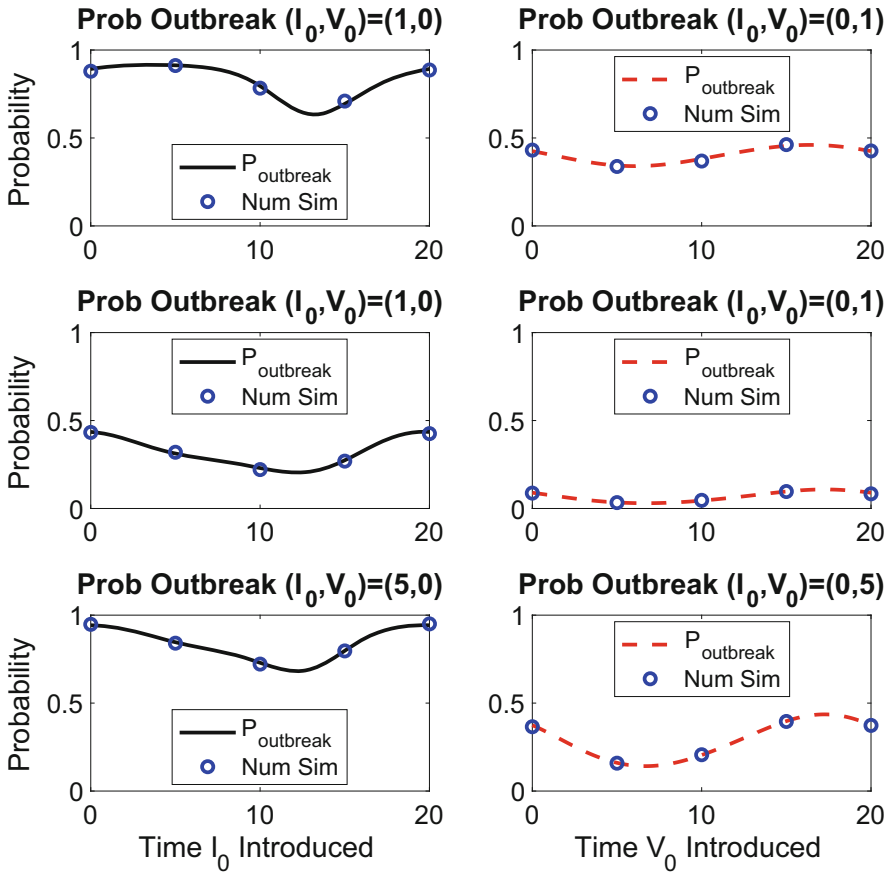


Fig. 6 The probability of an outbreak $\mathbb{P}_{\text{outbreak}}(t)$ in (19) and the full nonhomogeneous process show good agreement when introduction of hosts or vectors at $t = 0, 5, 10, 15, 20$, based on 5000 sample paths (Num Sim, circles). In the top two panels, $c = 0.05$, $\gamma = 0.1$, and $\mathcal{R}_0 = 3.73$. In the middle and bottom two panels, $c = 0.2 = \gamma$ and $\mathcal{R}_0 = 1.27$. These panels differ in number of infected vectors or infected hosts that are introduced. Initial number of infected hosts or vectors is either $(I_0, V_0) = (1, 0)$, $(0, 1)$, $(5, 0)$ or $(0, 5)$, respectively. The transmission rates are given in (24) with a phase shift in host to vector transmission and other parameter values are given in Table 2

6 Conclusion

For a nonhomogeneous, time-periodic stochastic vector–host model, we have shown that the backward Kolmogorov differential equations for the nonhomogeneous multitype branching process approximation can be used to derive a system of differential equations to approximate the asymptotic probability of an outbreak, $\mathbb{P}_{\text{outbreak}}(t)$ for $t \in [0, p]$ in (19). The periodic transmission rates lead to a periodic probability of an outbreak which depends on the number of infected vectors or

Table 4 The basic reproduction number $\mathcal{R}_0$, the average probability of outbreak $\widehat{P}_{\text{outbreak}}$, the maximum (t_M, M) and the minimum (t_m, m) of the probability of an outbreak $\mathbb{P}_{\text{outbreak}}(t)$ for $t \in (0, 20]$ for initial conditions $(I_0, V_0) = (1, 0)$ and $(I_0, V_0) = (0, 1)$, respectively. The transmission rates are defined in Eqs. (23) and the remaining parameter values are in Table 2

Parameters			$\widehat{P}_{\text{outbreak}}$		Max (t_M, M)		Min (t_m, m)	
c	γ	$\mathcal{R}_0$	$I_0 = 1$	$V_0 = 1$	$I_0 = 1$	$V_0 = 1$	$I_0 = 1$	$V_0 = 1$
0.05	0.1	3.68	0.82	0.41	(18.6, 0.93)	(15.8, 0.47)	(8.0, 0.62)	(5.4, 0.35)
0.1	0.1	2.58	0.75	0.24	(17.7, 0.89)	(16.0, 0.31)	(7.4, 0.54)	(5.5, 0.17)
0.2	0.1	1.84	0.62	0.11	(16.6, 0.81)	(16.2, 0.18)	(5.9, 0.41)	(6.0, 0.06)
0.4	0.1	1.36	0.41	0.039	(15.6, 0.58)	(16.3, 0.07)	(4.6, 0.25)	(7.2, 0.012)
0.1	0.05	3.39	0.84	0.26	(17.8, 0.94)	(16.1, 0.33)	(7.2, 0.69)	(5.8, 0.18)
0.1	0.2	1.94	0.59	0.21	(17.6, 0.80)	(15.8, 0.28)	(7.6, 0.33)	(5.1, 0.14)
0.1	0.4	1.47	0.38	0.15	(17.3, 0.60)	(15.5, 0.21)	(8.0, 0.14)	(4.7, 0.10)
0.05	0.05	4.78	0.88	0.42	(18.6, 0.96)	(15.8, 0.48)	(7.9, 0.75)	(5.6, 0.36)
0.2	0.2	1.35	0.38	0.074	(16.1, 0.59)	(15.6, 0.12)	(6.1, 0.19)	(5.2, 0.035)

Table 5 Similar to Table 4, except that the transmission rates are defined in Eqs. (24) with a phase shift in host to vector transmission

Parameters			$\widehat{P}_{\text{outbreak}}$		Max (t_M, M)		Min (t_m, m)	
c	γ	$\mathcal{R}_0$	$I_0 = 1$	$V_0 = 1$	$I_0 = 1$	$V_0 = 1$	$I_0 = 1$	$V_0 = 1$
0.05	0.1	3.73	0.82	0.40	(3.6, 0.92)	(16.4, 0.46)	(13.2, 0.63)	(6.0, 0.34)
0.1	0.1	2.53	0.74	0.24	(1.9, 0.85)	(16.8, 0.31)	(12.8, 0.57)	(6.4, 0.17)
0.2	0.1	1.70	0.58	0.11	(19.9, 0.71)	(17.3, 0.17)	(12.0, 0.46)	(6.8, 0.05)
0.4	0.1	1.48	0.22	0.024	(18.5, 0.30)	(17.5, 0.05)	(6.5, 0.17)	(7.5, 0.006)
0.1	0.05	3.26	0.84	0.25	(1.9, 0.91)	(16.6, 0.33)	(12.8, 0.72)	(6.4, 0.18)
0.1	0.2	1.91	0.58	0.20	(1.7, 0.73)	(17.0, 0.26)	(13.0, 0.36)	(6.4, 0.14)
0.1	0.4	1.40	0.33	0.13	(1.5, 0.48)	(17.3, 0.18)	(13.2, 0.14)	(6.4, 0.09)
0.05	0.05	4.79	0.89	0.42	(3.7, 0.95)	(16.2, 0.48)	(13.1, 0.76)	(6.0.0.36)
0.2	0.2	1.27	0.31	0.066	(19.7, 0.44)	(17.3, 0.11)	(12.2, 0.21)	(6.7, 0.030)

infected hosts introduced into a susceptible population and on the time at which they are introduced. It is also shown that the estimate for the probability of an outbreak $\mathbb{P}_{\text{outbreak}}(t)$ is in good agreement with Monte Carlo simulations of the full nonlinear nonhomogeneous stochastic model. The results of the analysis and the numerical simulations indicate that the time at which the host or vector transmission is highest does not coincide with the greatest risk for infection. The times of the extrema, t_M and t_m, of the probability of an outbreak are shifted left of the extrema of the transmission rates. In our model, the shift is smallest when the basic reproduction number $\mathcal{R}_0$ is large and when infection begins in the host.

Some implicit assumptions in the multitype branching process approximation may limit some of the applications of this method. The assumption of independence of initial number infected restricts these results to a small number of initial infections and a large population size. If $\mathcal{R}_0 < 1$, the branching process approximation is subcritical and all solutions approach zero (extinction). But if $\mathcal{R}_0 > 1$, the branching process is supercritical, and solutions either approach zero (extinction) or they

Table 6 Similar to Table 4, except that the transmission rates are defined in Eqs. (25) with a phase shift in vector to host transmission

Parameters			$\widehat{P}_{\text{outbreak}}$		Max (t_M, M)		Min (t_m, m)	
c	γ	$\mathcal{R}_0$	$I_0 = 1$	$V_0 = 1$	$I_0 = 1$	$V_0 = 1$	$I_0 = 1$	$V_0 = 1$
0.05	0.1	3.46	0.81	0.38	(19.4, 0.93)	(0.15, 0.44)	(8.1, 0.60)	(10.3, 0.33)
0.1	0.1	2.49	0.73	0.21	(19.0, 0.89)	(0.19, 0.27)	(7.6, 0.50)	(10.8, 0.16)
0.2	0.1	1.80	0.60	0.096	(18.4, 0.79)	(0.27, 0.14)	(7.1, 0.38)	(11.6, 0.058)
0.4	0.1	1.30	0.36	0.03	(17.8, 0.52)	(0.49, 0.049)	(6.6, 0.22)	(12.7, 0.012)
0.1	0.05	3.37	0.83	0.24	(19.2, 0.94)	(0.61, 0.30)	(7.6, 0.67)	(11.0, 0.18)
0.1	0.2	1.75	0.55	0.16	(18.7, 0.78)	(19.5, 0.21)	(7.7, 0.28)	(9.6, 0.13)
0.1	0.4	1.23	0.26	0.078	(18.2, 0.46)	(18.7, 0.10)	(8.0, 0.074)	(6.7, 0.058)
0.05	0.05	4.68	0.88	0.41	(19.5, 0.96)	(0.43, 0.46)	(8.0, 0.74)	(10.6, 0.35)
0.2	0.2	1.26	0.31	0.047	(17.7, 0.50)	(19.3, 0.071)	(7.0, 0.13)	(11.2, 0.032)

grow without bound (outbreak). For the supercritical case, which is of interest, the branching process approximation requires a sufficiently large population size. For our numerical examples, the host population size $N_s = 100$ and vector population size $N_m = 1000$ are sufficiently large to apply the branching process approximation. When host and vector population sizes are both less than 100, then the branching process approximation may not yield a good approximation.

These results have applications to other nonhomogeneous stochastic epidemic models with periodic transition rates. Seasonality plays an important role in many infectious diseases, including malaria, dengue, Zika, Lyme disease, chikungunya, leishmaniasis, and influenza [4, 6, 10, 12, 23, 27, 28]. The roles of host, vector, and pathogen responses to seasonal cues in the environment are important in defining the periodic transition rates which in turn can be used to compute the time-dependent risk of an outbreak by applying a branching process approximation. The effect of climate change on seasonal variations and prevalence of vector-transmitted diseases are important to public health [9, 18]. The models and methods presented here, as well as more complex models, are useful in testing hypotheses and designing experiments that will help identify the times at which vector control or prophylactic measures are the most effective.

Acknowledgments We thank the two referees and the editors for their helpful suggestions.

Appendix

The following assumptions are taken from Wang and Zhao [26]. The ODE system can be expressed as the following system:

$$\dot{x}_i = \mathcal{F}_i(t, x) - \mathcal{V}_i(t, x) \equiv f_i(t, x), \quad i = 1, \ldots, 4,$$

where $x = (x_1, x_2, x_3, x_4) = (I, V, S, M)$ and $\mathcal{V}_i = \mathcal{V}_i^+ - \mathcal{V}_i^-$.

(A1) The functions $\mathcal{F}_i(t,x)$, $\mathcal{V}_i^+(t,x)$, and $\mathcal{V}_i^-(t,x)$ are nonnegative and continuous on $\mathbb{R}\times\mathbb{R}_4^+$ and continuously differential with respect to x.
(A2) The functions $\mathcal{F}_i(t,x)$, $\mathcal{V}_i^+(t,x)$, and $\mathcal{V}_i^-(t,x)$ are p-periodic in t, $p>0$.
(A3) If $x_i=0$, then $\mathcal{V}_i^-(t,x)=0$.
(A4) $\mathcal{F}_i(t,x)=0$ for $i=3,4$.
(A5) If $I=V=0$, then $\mathcal{F}_i(t,x)=V_i^+(t,x)=0$ for $i=1,2$.
(A6) The DFE is stable if $I=0=V$.
(A7) The spectral radius of matrix V is less than one.

Assumptions (A1)–(A7) hold for the ODE system (2) and (3).

References

1. Allen, L.J.S.: A primer on stochastic epidemic models: formulation, numerical simulation, and analysis. Infect. Dis. Model. **2**(2), 128–142 (2017)
2. Allen, L.J.S., Lahodny, G.E., Jr.: Extinction thresholds in deterministic and stochastic epidemic models. J. Biol. Dyn. **6**(2), 590–611 (2012)
3. Athreya, K.B., Ney, N.E.: Branching Processes. Dover Publications, Mineola, New York (2004)
4. Bacaër, N.: Approximation of the basic reproduction number R0 for vector-borne diseases with a periodic vector population. Bull. Math. Biol. **69**(3), 1067–1091 (2007)
5. Bacaër, N., Dads, E.H.A.: On the probability of extinction in a periodic environment. J. Math. Biol. **68**(3), 533–548 (2014)
6. Bacaër, N., Guernaoui, S.: The epidemic threshold of vector-borne diseases with seasonality: the case of cutaneous leishmaniasis in Chichaoua, Morocco. J. Math. Biol. **53**(3), 421–436 (2006)
7. Bartlett, M.S.: The relevance of stochastic models for large-scale epidemiological phenomena. J. R. Stat. Soc. C **13**(1), 2–8 (1964)
8. Billings, L., Forgoston, E.: Seasonal forcing in stochastic epidemiology models. Ricerche Mat. **67**(1), 27–47 (2018)
9. Campbell-Lendrum, D., Manga, L., Bagayoko, M., Sommerfeld, J.: Climate change and vector-borne diseases: what are the implications for public health research and policy? Philos. Trans. R. Soc. B **370**(1665), 20130552 (2015)
10. Fox, S.J., Miller, J.C., Meyers, L.A.: Seasonality in risk of pandemic influenza emergence. PLoS Comput. Biol. **13**(10), e1005749 (2017)
11. Gao, D., Lou, Y., Ruan, S.: A periodic Ross–Macdonald model in a patchy environment. Discrete Contin. Dynam. Systems B **19**(10), 3133–3145 (2014)
12. Grassly, N.C., Fraser, C.: Seasonal infectious disease epidemiology. Proc. R. Soc. Lond. B: Biol. Sci. **273**(1600), 2541–2550 (2006)
13. Heffernan, J.M., Smith, R.J., Wahl, L.M.: Perspectives on the basic reproductive ratio. J. R. Soc. Interface **2**(4), 281–293 (2005)
14. Klausmeier, C.A.: Floquet theory: a useful tool for understanding nonequilibrium dynamics. Theor. Ecology **1**(3), 153–161 (2008)
15. Klein, B. Macdonald, P.D.M.: The multitype continuous-time Markov branching process in a periodic environment. Adv. Appl. Prob. **12**(1), 81–93 (1980)
16. Ma, J., Ma, Z.: Epidemic threshold conditions for seasonally forced SEIR models. Math. Biosci. Eng. **3**(1), 161–172 (2006)
17. Mitchell, C., Kribs, C.: A comparison of methods for calculating the basic reproductive number for periodic epidemic systems. Bull. Math. Biol. **79**(8), 1846–1869 (2017)

18. Parham, P.E., Michael, E.: Modeling the effects of weather and climate change on malaria transmission. Environ. Health Perspect. **118**(5), 620–626 (2010)
19. Parham, P.E., Michael, E.: Outbreak properties of epidemic models: The roles of temporal forcing and stochasticity on pathogen invasion dynamics. J. Theor. Biol. **271**(1), 1–9 (2011)
20. Posny, D., Wang, J.: Computing the basic reproductive numbers for epidemiological models in nonhomogeneous environments. Appl. Math. Comput. **242**, 473–490 (2014)
21. Roberts, M.G., Heesterbeek, J.A.P.: A new method for estimating the effort required to control an infectious disease. Proc. R. Soc. Lond. B: Biol. Sci. **270**(1522), 1359–1364 (2003)
22. Shuai, Z., Heesterbeek, J.A.P., van Den Driessche, P.: Extending the type reproduction number to infectious disease control targeting contacts between types. J. Math. Biol. **67**(5), 1067–1082 (2013)
23. Suparit, P., Wiratsudakul, A., Modchang, C.: A mathematical model for Zika virus transmission dynamics with a time-dependent mosquito biting rate. Theor. Biol. Med. Model. **15**(1), 1–11 (2018)
24. van den Driessche, P.: Reproduction numbers of infectious disease models. Infect. Dis. Model. **2**(3), 288–303 (2017)
25. van den Driessche, P., Watmough, J.: Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math. Biosci. **180**(1–2), 29–48 (2002)
26. Wang, W., Zhao, X.-Q.: Threshold dynamics for compartmental epidemic models in periodic environments. J. Dyn. Differ. Equ. **20**(3), 699–717 (2008)
27. Wang, X., Zhao, X.-Q.: Dynamics of a time-delayed Lyme disease model with seasonality. SIAM J. Appl. Dyn. Syst. **16**(2), 853–881 (2017)
28. Wang, X., Zhao, X.-Q.: A malaria transmission model with temperature-dependent incubation period. Bull. Math. Biol. **79**(5), 1155–1182 (2017)
29. Wesley, C.L., Allen, L.J.S.: The basic reproduction number in epidemic models with periodic demographics. J. Biol. Dyn. **3**(2–3), 116–129 (2009)

Evidence for Multiple Transmission Routes for Pseudorabies in Wild Hogs

Benjamin Levy, Suzanne Lenhart, Charles Collins, and William Stiver

1 Introduction

Invasive species can be defined as living organisms that have established residence in a non-native ecosystem. Typically an invasive species enters an ecosystem via translocation, meaning the given species was entered into an ecosystem in which it did not evolve which in some cases can result in a catastrophe. For instance, the emerald ash borer is an invasive species in North America that has dramatically decreased the total population of ash trees on the continent [8]. All taxa, geographic locations, and even entire classes of animals such as reptiles can be threatened by invasive species [18, 39]. To make matters worse, the impacts and costs associated with an invasive species increase as the species becomes established in an area [16]. With this in mind invasive species are among the world's most pressing environmental concerns [33].

European Wild Boar were (*Sus scrofa*) was brought to the United States in the sixteenth century and poses a significant environmental threat [45]. A number of studies have detailed the ecological and economic impacts of the species in the United States [13–15, 33]. Here we focus on a specific sub-population that exists in Great Smoky Mountains National Park (GSMNP). The population originated from a hunting preserve in Hoopers Bald, North Carolina where escaped boars bred

B. Levy (✉)
Fitchburg State University, Fitchburg, MA, USA
e-mail: blevy1@fitchburgstate.edu

S. Lenhart · C. Collins
University of Tennessee, Knoxville, TN, USA
e-mail: slenhart@utk.edu; collins@utk.edu

W. Stiver
National Park Service, Great Smoky Mountains National Park, Gatlinburg, TN, USA
e-mail: bill_stiver@nps.gov

M. I. Teboh-Ewungkem, G. A. Ngwa (eds.), *Infectious Diseases and Our Planet*,
Mathematics of Planet Earth 7, https://doi.org/10.1007/978-3-030-50826-5_3

with pigs of domestic ancestry to produce a hybrid population [21]. We refer to this hybrid population as wild hogs. This hybrid population exists throughout the Park and surrounding area, and is the only established population of wild swine in GSMNP.

The wild hog population has caused significant damage since entering GSMNP. They are in direct competition with native species, and their rooting behavior is extremely destructive resulting in altered nutrient cycles and disrupted forest succession patterns [6, 20]. Moreover, the population is a vessel for a number of diseases including porcine brucellosis, parvovirus, leptospirosis, toxoplasmosis, and pseudorabies [9, 17, 32]. Some are a concern for wildlife and pets, such as parvovirus and pseudorabies, while others are a concern for humans. Toxoplasmosis is a parasitic infection that is the leading cause of death attributed to foodborne illness in the United States [11]. Both brucellosis and leptospirosis are bacteria that can spread to humans that have been in contact with wildlife [9, 10]. Since these diseases impact humans and can even lead to death, understanding how they spread in reservoir populations such as wild hogs is important.

Because of their negative impacts, GSMNP started a control program in the late 1950s. Through the control program traps are set within Park boundaries to collect wild hogs, and they are also actively sought out by Park rangers in the backcountry. In either case, once a ranger locates a hog they are shot and left to decay in the Park. Throughout the paper we refer to these activities as "hunting," and thousands of hogs have been eliminated from GSMNP in this manner. This removal program has successfully limited the population and produced a significant amount of useful data [24]. One result of the control program was the formation of a disease monitoring program for wild hogs in the Park [9, 37].

Wild hogs spread diseases that have serious implications for a large number of domestic and wild animals throughout the region, especially the pseudorabies virus (PRV). Clinical signs of PRV and impacts have been well documented in commercial hog facilities but are less understood in wild hogs [28]. Since North Carolina is one of the largest producers of pork in the United States, commercial pig farmers and the state government are concerned about the potential presence of pseudorabies in the state. Beginning in 2001, GSMNP partnered with the North Carolina Department of Agriculture and Consumer Services (NCDACS), the Tennessee Department of Agriculture (TDA), and the US Department of Agriculture, Animal and Plant Health Inspection Service, Veterinary Services (APHIS) to begin monitoring wild hogs in the Park for disease. Specifically, 42.5% of hogs that were shot by rangers from 2001–2013 were tested for pseudorabies and brucellosis, as they pose the greatest threat to humans and the domestic pig industry [9]. Though no hogs have tested positive for brucellosis, the population has become a reservoir for pseudorabies. From 2001–2004, all blood samples taken from hogs shot in GSMNP tested negative for pseudorabies. However, hogs began testing positive for PRV starting in 2005 with the prevalence increasing steadily reaching as high as 56.9% (see Fig. 1 and Table 1) [9, 30, 37]. The total number of hogs that were caught as part of the control program and tested for PRV ranged from as few as 58 in 2012 to

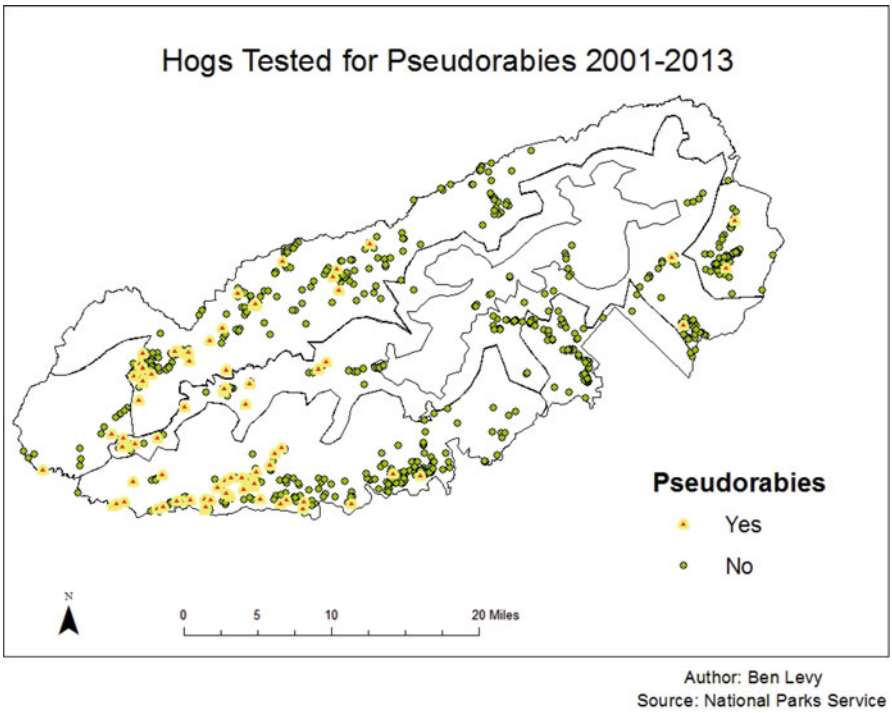


Fig. 1 Map depicting the 42.5% of hogs shot as part of the control program that were tested for pseudorabies from 2001–2013. Although the disease is concentrated in the Western half of the Park, a small pocket of disease is present in the far eastern regions

Table 1 Results from ongoing disease survey of wild hogs in GSMNP from 2005–2013. Though a relatively small number of samples were taken each year, a general trend of increasing prevalence can be seen in the data [30]

Year	2005	2006	2007	2009	2009	2010	2011	2012	2013
Blood samples	150	208	64	106	155	105	90	58	69
Positive for pseudorabies	2	4	10	9	9	4	19	33	20
Prevalence	0.013	0.019	0.156	0.085	0.058	0.038	0.211	0.569	0.290

as many as 208 in 2006. The highest percentage of hogs that were tested was 82% in 2006, while the lowest percentage was 22% in 2007.

Since PRV was first identified in the early twentieth century, a considerable amount of research has been conducted to study the impact of the disease on pigs [19, 27, 34–36, 43]. Results from such studies have shown that venereal transmission is the primary way the disease spreads in domestic pig populations [29, 36]. Several modeling studies have considered the experimental information. For instance, a stochastic susceptible–infected–recovered (SIR) model was used to consider the spread of PRV in domestic pigs, the results of which confirmed the transmission

routes posed by experimental work in [44]. While disease dynamics in domestic pigs is well understood, important disease characteristics in wild hogs remain unclear. One thing that is known about the transmission of pseudorabies in wild hogs is that venereal transmission alone cannot account for the observed prevalence of the disease in the United States [41]. This implies more than one transmission route may exist in wild hogs. While some of the aforementioned studies will be discussed further in the next section, we mention them here to highlight the need for additional research.

Disease in wildlife is only one feature in the broader One Health concept. The health of our planet is linked with our environmental, wildlife, domestic animal, and human health as well as the use of resources. The health (and ultimate survival) of any one group is connected to the health of the others. Studying a disease in wild hogs is an issue for the health of our planet. Mathematical models can aid in understanding and managing biological systems.

Mathematical models can be used to better understand a biological system including providing support for transmission routes of infectious diseases. For example, [4] used an ODE model to show that aerosol transmission is the more dominant transmission route in pandemic influenza compared to contact transmission, which can be used to reduce, and guide control of, future outbreaks. In a second example researchers were able to understand spatial patterns of the Ebola virus disease in humans only after incorporating bat movement patterns in their model [23]. In another study, collected data were compared to various model structures for cholera and the resulting information was used to determine the most accurate model formulation [22]. When considering transmission routes for lumpy skin disease in cattle, a mathematical model was used to determine that indirect transmission is more likely to exist than direct transmission [26]. Another example of modeling aiding the understanding of disease dynamics is when researchers in [7] found that empirical results related to avian influenza could only be reproduced after adding indirect transmission via environmental contamination, providing support for this transmission route. Finally, modeling was used to show that preferential sexual transmission could not reproduce observed pseudorabies prevalence in the United States, which has implications for the spread of the disease in domestic pig populations [41]. While the mathematical formulation, available data, and goal of each model varied, they all shared the commonality of using mathematical modeling to better understand transmission routes and observed disease dynamics of an infectious disease. It is the aim of this study to model pseudorabies in the wild hog population within GSMNP in order to better understand transmission routes, consider how pseudorabies is spreading throughout GSMNP, and improve the general understanding of disease dynamics in wild hogs, which is vitally important to the viability of domestics pig production [29]. To this end, we extend an existing population model for wild hogs that is discrete in time and space to include disease dynamics to address these issues.

Each of the models mentioned in the previous paragraph is continuous in time and/or space. In fact, most ecological models found in the literature are continuous in time and/or space, as seen in the review presented by Arino and Van den Driessche

[3] and further supported in [2]. Populations that have distinct movement, birth, or growth patterns in different times of the year may be represented more clearly with models that are discrete in time and possibly in space. For an introduction to population modeling in discrete time, see the books by Allen and Edelstein-Keshet [1, 12]. Note that in a model with continuous time (ordinary differential equations), all of the events or actions occur at the same time, but the order of events is crucial in models with discrete time [5]. In the modeling in this paper, the order of events for actions of hogs is important, and data in different regions of the Park dictate the use of discrete spatial components.

In addition to better understanding transmission routes and important dynamics of pseudorabies, this study also contributes to the literature on discrete models. In the sections to follow we describe what is known about disease dynamics of Pseudorabies in wild hogs, extend a discrete population model to include disease transmission, explain how we estimated parameters for the model, and discuss results that support the existence of multiple transmission routes for PRV in wild hogs.

2 Disease Dynamics

Pseudorabies, also known as Aujeszky's Disease, was first identified in 1902. This herpes viral infection is highly contagious in domestic pigs as it leads to respiratory illness in adults and high mortality rate for piglets [29]. After an initial shedding period of about 7 days, adults recover but carry the disease for life in a latent form. Stress can occasionally cause recovered individuals to begin shedding the disease and therefore become infectious once again [29]. Thus all hosts who contract PRV will test positive for life. The decrease in birthrate caused by the disease poses a serious threat to the domestic pig industry. Since the US pig industry became PRV-free in 2004, wild hogs represent the reservoir for the disease which means there is vested interest in limiting the spread of the PRV in wild hog populations [29, 41]. Although the disease has been well studied in domestic pigs due to its impact on production, there have been few studies on the effects in wild hogs. What is clear is that the disease has less of an impact on wild hogs as it does not decrease birthrates and instead causes benign symptoms [27, 41]. However, in addition to the impact on domestic pigs, the disease can be contracted by other animals such as bear and coyotes, and is especially deadly in canines [9].

The primary transmission route in domestic pigs is venereal transmission, though it is possible that several other less significant transmission routes exist [29, 36]. However, the disease acts differently in populations of wild swine. Possible transmission routes for pseudorabies in wild hog populations include nose-to-nose or direct contact, venereal transmission, reinfection of a recovered carrier due to stress during which individuals become infectious again, and transmission to piglets during nursing [19, 29, 36, 41]. While the specific transmission route(s) are unknown, a recent study determined that density-dependent transmission can

account for the reported seroprevalence found in the United States, but venereal transmission alone cannot [41]. Because of this, we consider each aforementioned possible transmission route in our model of the disease and use available data to determine which transmission route(s) are most likely to exist in the population. Throughout the paper we use the term "direct transmission" in place of "density-dependent transmission." Density-dependent transmission uses the concept that contact rates will increase with population density and thus so will transmission rates. Direct transmission, therefore, encapsulates several transmission types including nose-to-nose and venereal transmission.

3 Modeling Pseudorabies

Our approach is to extend an existing metapopulation model to include disease dynamics for wild hogs in GSMNP. The metapopulation model is described in detail in [24]. Here we summarize the model and highlight key features. Since oak acorn abundance is believed to be the main driver of wild hog population dynamics, detailed overstory vegetation information from [25] plays an important role in the model. The data were used to partition GSMNP into 8 discrete spatial regions and were integrated with yearly oak acorn data to populate each region with food in the form of oak mast and other soft mast such as tubers and roots [31]. We use the general term mast to refer to the sum of hard mast (acorns) as well as soft mast (tubers and roots) that exist in our model. Calories available in each region decline throughout the year, the specific values of which are used to adjust caloric-dependent parameters such as survival rate, movement rate, and birthrate. Since Park rangers actively hunt the population and have detailed records of this activity, we also model this activity. Of the 8 regions in the model, 6 of them are inside Park boundaries where rangers actively hunt, while 2 represent exterior regions where hogs will not be subject to the control program. The model itself is also discrete in time with 1-month time steps. After each time step hogs move between regions based on the given season and food availability. Parameter values were estimated using an optimization scheme that identified the set of values that most closely matched the hunting data in [30]. The resulting model integrates what is known about the population with available data to accurately represent population-level wild hog births, death, and movement in GSMNP, and is therefore well suited to model the spread of pseudorabies in the Park. The order of events in the spatially and temporally discrete metapopulation model is described in more detail below.

Since the disease has an infectious period of 1 week, we first reduced the metapopulation model described previously from a 1-month time step to a 1-week time step. Since the movement rates cannot simply be scaled from a monthly value to a weekly value, the authors decided to re-estimate all population parameters in the same manner described in [24], the only difference being the size of the time step and implementation of the hunting rate. In [24], authors used a 1-month time step and varied hunting rates by region and year, whereas here we use constant

Table 2 A description of estimated parameters found in the pseudorabies model with a time step of 1 week

Name	Value	Description
$Surv0$	0.96	Survival factor if there is no mast
$SurvMax$	0.99	Survival factor as mast approaches a maximum level
$BR0$	0.10	Percent of population that give birth and whose piglets survive given no mast
$BRMax$	0.13	Percent of population that give birth and whose piglets survive as mast approaches a maximum level
$Move0$	0.14	Percent of feral hogs moving with no available mast
$MoveMax$	0.01	Percent of feral hogs moving as mast approaches a maximum level
H_H	0.09	On-season hunting rate, from January through May
H_L	0.06	Off-season hunting rate, from June through December

on- and off-season hunting rates each year and a 1-week time step. Parameters estimated for the model with a 1-week time step are shown in Table 2. The resulting model with a time step of 1 week almost identically mimics the model with a time step of 1 month described in [24].

We consider three classes within the population: Susceptibles (S), Infected (I), and Carriers (C). Susceptible individuals have never contracted the disease and are vulnerable to infection. Infected individuals are those who are symptomatic and infectious. Carriers have recovered from the disease and are no longer infectious but, since they still carry the virus in a latent form, could re-enter the infected class again at rate ϕ, at which point they would once again be infectious and able to transmit the disease [42]. Since each region has a different total area and will contain different numbers of individuals in a given time step, we use a different transmission rate in each region, denoted by β_r. These β_r values account for all direct transmission between infecteds and susceptibles. Since β_r includes venereal transmission, we increase this rate by factor γ during mating season under the assumption that there is increased contact between individuals during this time [36].

We carry out the same order of events as are described in [24], with the addition of partitioning the population into three classes and a disease transmission event taking place after births and before movement. The model from [24], therefore, takes the following form:

1. Update the mast for the month since many of the parameters that govern feral hog dynamics in GSMNP are driven by hard mast availability [38, 40]. The total mast $M_{r,t}$ in region r at time t is a function of the hard mast ($HM_{r,t}$) and soft mast (SM_r) in the given region. Specifically, hard mast is set yearly in August by the mast index value for the given year ($MI_{r,y}$) and then declines throughout the year as a result of decay (δ) and consumption (C_P) by the population in the given region ($P_{r,t}$):

$$HM_{r,t+1} = \begin{cases} MI_{r,y} & m = 8 \\ ((1-\delta)HM_{r,t} - C_P P_{r,t})^+ & m \neq 8, \end{cases}$$

$$M_{r,t} = HM_{r,t} + SM_r. \tag{1}$$

Note that the rate of decay (δ) was scaled by 1/4 to account for the new weekly time step.

2. Hogs are hunted in each region and time step ($H_{r,t}$) at a rate determined by the season. The Park applies a higher rate early in the year (H_H) and a lower rate later in the year (H_L):

$$H_{r,t} = \begin{cases} H_H \cdot P_{r,t} & m = 1-5, \\ H_L \cdot P_{r,t} & m = 6-12. \end{cases}$$

3. Compute the portion of the post-hunted population that survives in each region ($Surv_{r,t}$). We do this before adding births because only surviving adults can reproduce. We record the number of surviving susceptibles (SS), surviving infected (SI), and surviving carriers (SC) that move on to the next event:

$$\begin{aligned} Surv_{r,t} &= F(M_{r,t}, Surv0, SurvMax, M_h), \\ SS_{r,t} &= S_{r,t} \cdot (1 - Hrate) \cdot Surv_{r,t}, \\ SI_{r,t} &= I_{r,t} \cdot (1 - Hrate) \cdot Surv_{r,t}, \\ SC_{r,t} &= C_{r,t} \cdot (1 - Hrate) \cdot Surv_{r,t}, \end{aligned}$$

where the function F scales the given parameter based on food availability in the given region and $Hrate$ represents either the higher level hunting rate or lower level hunting rate depending on the month.

4. If the month is January, we compute the number of births ($BR_{r,t}$) based on the percent of the population that is a mature female (B_F), average litter size (L_A), surviving population, and mast supplies. A percent of piglets (α) will contract a latent form of the infection via nursing and become members of the carrier class:

$$BR_{r,t} = \begin{cases} B_F \cdot L_A \cdot F(M_{r,t}, BR0, BRMax, M_h) & m = 1 \\ 0 & m \neq 1, \end{cases}$$

$$\begin{aligned} SS_{r,t} &= [S_{r,t} \cdot (1 - Hrate) \cdot Surv_{r,t} \cdot (1 + BR_{r,t})] \\ &\quad + [(I_{r,t} + C_{r,t}) \cdot (1 - \alpha) \cdot (1 - Hrate) \cdot Surv_{r,t} \cdot BR_{r,t}], \\ SI_{r,t} &= I_{r,t} \cdot (1 - Hrate) \cdot Surv_{r,t}, \\ SC_{r,t} &= C_{r,t} \cdot (1 - Hrate) \cdot Surv_{r,t} \\ &\quad + [(I_{r,t} + C_{r,t}) \cdot \alpha \cdot (1 - Hrate) \cdot Surv_{r,t} \cdot BR_{r,t}]. \end{aligned}$$

5. Disease transmission:
   Wild hogs in the Park mate during September which means there will be more sexual contacts during this month resulting in an increased contact-based transmission rate by a factor of γ:

$$\beta_r = \begin{cases} \hat{\beta}_r & m \neq 9 \\ \gamma \hat{\beta}_r & m = 9 \end{cases}$$

   Additionally, carriers become reinfected at rate ϕ and hogs can contract the disease as a result of contact with infected individuals:

$$\begin{aligned} SS_{r,t} &= SS_{r,t} e^{-\beta_r SI_{r,t}}, \\ SI_{r,t} &= SS_{r,t}(1 - e^{-\beta_r SI_{r,t}}) + SC_{r,t}\phi, \\ SC_{r,t} &= SC_{r,t}(1 - \phi) + SI_{r,t}. \end{aligned}$$

6. Perform movement using either general movement or seasonal movement, which is dependent upon the time of year. During seasonal movement hogs moving up in elevation during the warmer months, down in elevation during the colder month. In the absence of seasonal movement, hogs are more likely to leave a region if there is inadequate food and less likely to leave a region if food is abundant. This is carried out in our disease model exactly as described in [24] by applying movement rates equally to each class. The total population of each class that exist in each region at the beginning of the following time step ($S_{r,t+1}$, $I_{r,t+1}$, and $C_{r,t+1}$) are recorded at the end of this event as determined by the number of individuals that move into each region.

Pseudorabies was first detected in GSMNP in region 4 in February of 2005. We, therefore, initialize the model in 2004, introduce 5 infected individuals into region 4 in February 2005, and run the model until 2013, which is the most recent year of which we have disease data. The pseudorabies-related data and disease parameter estimation process are detailed in the next section.

4 Parameter Estimation

Our pseudorabies model contains the following parameters that do not have known values : β_1, β_2, β_3, β_4, β_5, β_6, α, γ, and ϕ. Regions 7 and 8 border regions 4 and 5, are exterior to the Park, and their purpose is to mimic reality by providing areas where the hog population will not be subject to the control program. Since these regions were created using the same area and overstory composition as regions 4 and 5, we assume that $\beta_7 = \beta_4$ and $\beta_8 = \beta_5$. The list of all unknown parameters is described below in Table 3.

Table 3 A list and description of new parameters and variables found in the pseudorabies model

Parameter	Description
β_r	Direct transmission rate in region r for $r = 1, \ldots, 6$
α	Percent of piglets that become carriers as a result of nursing from mothers that carry the virus
γ	Percent increase in transmission rate during mating season
ϕ	Percent of carrier class that becomes reinfected at each time step

To estimate parameters we formulate an optimization problem to provide parameters that best match available prevalence data from [30]. Moreover, since the specific transmission routes that exist in the population are unknown as discussed previously, we use the available disease data to determine which transmission routes are most likely to exist in the population. We achieve this by testing all possible combinations of transmission routes in the model and seeing which combination best matches the data.

Park officials began sampling hogs that were shot as part of the control program for disease in 2001 and the first case of pseudorabies was detected in region 4 in February 2005 [9]. The specific number of samples and geographic location of tests varied greatly each year. The total number of individuals tested in a given year range from as little as 64 to as many as 208 (see Table 1), and the vast majority of tests were conducted on hogs found in regions 4 and 5 with very few hogs tested from other regions (see Table 4).

To obtain the parameter values that produce hunting prevalence levels that best match the available hunting prevalence data, we use data from 2005 through 2013 and implement an optimization scheme that relates model output with available data. We only include data from regions 4 and 5 as the majority of tests were conducted on hogs found in these regions. Additionally, since a consistent number of disease tests were not conducted year-to-year within regions 4 and 5, we weight prevalence values by the number of disease samples taken in each region. Multiplying prevalence by the number of observations produces the estimated number of seropositive hogs that were shot in from each region. With this in mind the optimization procedure is essentially trying to reproduce the number of hogs that were shot in each region that tested positive for pseudorabies. By only including data from regions 4 and 5 while also weighting prevalence values by the number of samples taken, our parameter estimates will best approximate the most reliable data.

Let s_r represent a vector containing the number of hogs sampled for disease in region r, let HI_r represent the vector containing yearly prevalence of infected hogs in the model hunted from region r, and let HI_r^* represent the vector containing yearly prevalence of infected hogs from region r in the hunting data. Each of these vectors is 1×9 as the data we are using are from 2005 through 2013. Since both recovered and actively infected hogs will test positive for the disease, the prevalence data include individuals from both the infected and carrier classes. Thus,

Table 4 Blood samples, detected cases, and resulting prevalence for pseudorabies in each region of GSMNP

Disease samples from Region 1									
	2005	2006	2007	2008	2009	2010	2011	2012	2013
Total blood samples	0	3	1	1	6	1	7	5	1
Detected cases	0	0	0	0	0	0	1	2	1
Prevalence	0	0	0	0	0	0	0.14	0.4	1
Disease samples from Region 2									
	2005	2006	2007	2008	2009	2010	2011	2012	2013
Total blood samples	0	0	0	0	0	0	0	33	10
Detected cases	0	0	0	0	0	0	0	1	0
Prevalence	0	0	0	0	0	0	0	0.03	0
Disease Samples from Region 3									
	2005	2006	2007	2008	2009	2010	2011	2012	2013
Total blood samples	0	0	0	0	0	25	24	0	4
Detected cases	0	0	0	0	0	1	1	0	0
Prevalence	0	0	0	0	0	0.03	0.04	0	0
Disease samples from Region 4									
	2005	2006	2007	2008	2009	2010	2011	2012	2013
Total blood samples	19	49	71	39	26	36	53	63	35
Detected cases	2	4	6	2	1	3	11	15	12
Prevalence	0.11	0.082	0.085	0.052	0.038	0.083	0.21	0.24	0.34
Disease samples from Region 5									
	2005	2006	2007	2008	2009	2010	2011	2012	2013
Total blood samples	0	48	53	34	78	19	31	50	19
Detected cases	0	1	2	2	8	1	4	16	5
Prevalence	0	0.021	0.038	0.059	0.10	0.053	0.13	0.32	0.26
Disease samples from Region 6									
	2005	2006	2007	2008	2009	2010	2011	2012	2013
Total blood samples	10	13	11	35	26	16	18	22	21
Detected cases	0	1	2	7	4	2	0	1	2
Prevalence	0	0.077	0.18	0.2	0.15	0.13	0	0.045	0.095

$$HI_{r,t} = I_{r,t} + C_{r,t}.$$

The optimization problem can be stated as

$$\underset{x}{\text{Minimize}} \quad J(x) = \frac{||s_4 \cdot HI_4 - s_4 \cdot HI_4^*||_2}{||s_4 \cdot HI_4^*||_2} + \frac{||s_5 \cdot HI_5 - s_5 \cdot HI_5^*||_2}{||s_5 \cdot HI_5^*||_2}, \tag{2}$$

where x represents all possible values of $\beta_1,\ \beta_2,\ \beta_3,\ \beta_4,\ \beta_5,\ \beta_6,\ \alpha,\ \gamma$, and ϕ that are included in the model, and each term represents the relative error in the given region.

Additionally, we require that the parameters fall within a reasonable range. As a result, the above problem is also restricted by the following linear constraints:

$$\begin{aligned} 0 \leq & \alpha \leq 1 \\ 0 \leq & \phi \leq 1 \\ & \gamma \geq 1. \end{aligned} \tag{3}$$

As stated our model contains at least 6 unknown direct transmission rates (β_r), which can increase to as many as 9 if all transmission routes are included in the model. To reduce the number of parameters used in the estimation process we also consider scaling transmission rates obtained from region 4 for use in other regions by scaling β_4 by the relative area of other regions. We chose to use the value estimated for region 4 as this region contains the most removals and therefore has the most reliable data. Specifically,

$$\beta_r = \beta_4 \frac{A_4}{A_r},$$

for $r = [1, 2, 3, 6, 7, 8]$, where A_r represents the area of region r. This relationship will produce larger direct transmission rates in regions with a smaller area than region 4. This approach allowed us to use a unique transmission rate for each region while avoiding overfitting the model by reducing the number of parameters to be estimated.

To solve the above optimization problem we used the Global Optimization Toolbox from MATLAB. Specifically, we used the fmincon local solver in coordination with the MultiStart Algorithm. The fmincon solver is a derivative-free solver that accepts smooth, nonlinear objective functions, and enforces the linear constraints given in (3). The MultiStart Algorithm generates uniformly distributed random starting points within the given bounds and passes them one-by-one to the local solver, fmincon, which attempts to find a local basin of attraction relative to each given start values. Since fmincon only converges on local minimums, the MultiStart Algorithm allowed us to test a large number of evenly distributed starting points to ensure we identified the global minimum. In all instances we used 100 starting points.

5 Results and Discussion

We estimated parameters by implementing the aforementioned optimization scheme with the objective function in (2), which only includes data from regions 4 and 5. We also consider all combinations of transmission routes. This process included estimating 6 transmission rates (one for each region), as well as estimating transmission rates for just regions 4 and 5 and scaling β_4 for use in other regions. To

Table 5 Summary of parameters estimation results where different combinations of transmission routes are considered. The objective function value used in the optimization scheme is given in the column labeled "Regions 4 and 5," while the performance of the model on the entire dataset is seen in the right most column. The top half of the table displays results where a transmission rate was estimated for each region. The bottom half shows results where transmission rates were only estimated in regions 4 and 5 and were scaled for use in other regions, resulting in fewer total parameters

Number of betas		Trans. routes included	Total # of variables	Relative error	
6 (per region)	2 (scaled)			Regions 4 and 5	Aggregate all data
✓		D	6	0.99	1.03
✓		D , I	7	1.02	1.17
✓		D, V	7	0.62	0.69
✓		D, R	7	0.96	0.72
✓		D, I, V	8	0.62	0.69
✓		D, I, R	8	0.32	0.63
✓		D, V, R	8	1.16	0.90
✓		D, I, R, V	9	0.32	0.65
	✓	D	2	1.96	1.05
	✓	D, I	3	1.06	1.05
	✓	D, V	3	1.18	0.88
	✓	D, R	3	0.84	0.67
	✓	D, I, V	4	1.06	1.05
	✓	D, I, R	4	0.76	0.65
	✓	D, V, R	4	0.44	0.50
	✓	D, I, R, V	5	0.44	0.50

Transmission routes: D direct transmission (included in all cases), *I* increased contact during mating season, *V* vertical transmission, *R* carriers becoming reinfected

view how well the estimated parameters match the entire dataset we also calculated the relative error in the entire Park. We did this for each parameter combination by aggregating available data in all regions and using the calculation seen in each of the terms in (2). All results are presented in Table 5. Notice that there are instances where the relative error in the entire Park is lower than in just regions 4 and 5, such as in rows 4 and 7. This can be explained by the fact that data from all regions were aggregated in these calculations. In these cases the number of infections in one region was too low while the total infecteds in another region were too high so that when the data were combined across all regions they balanced out to produce a lower relative error.

The first characteristic we notice is how direct transmission alone cannot account for the prevalence seen in the Park as evidenced by the objective function value for both instances of D only in Table 5. This is also seen visually in Fig. 2 where model output is plotted against the data using parameters estimated by only considering direct transmission in each region. The right column in Fig. 2 displays model output plotted against the data, both of which have been weighted by the

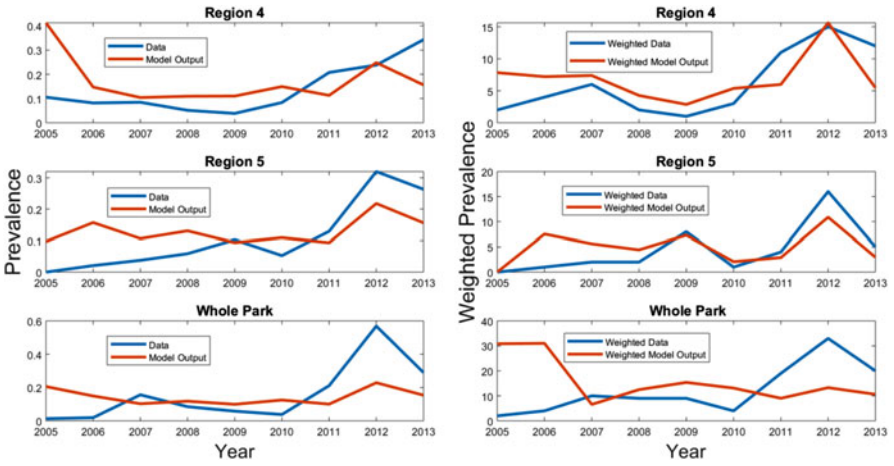


Fig. 2 Model output from estimating a direct transmission rate in each region and no additional transmission routes. The right column displays model output and data that have been weighted by the number of observations in the given year. The left column displays unweighted data and model output

number of observations in the given year. The left-hand column displays unweighted model output and data. Although the weighted model output matches a few of the intermediate points resulting in a moderate fit, the earlier and later years do not fit well and the prevalence in the model does not have the increasing trend seen in the data. As a result, the overall fit of the data is poor in this case. This is especially evident in the unweighted output and both graphs representing the fit relative to the whole park. These results do not agree with previous findings that direct density-dependent contact alone could account for seroprevalence of pseudorabies seen in wild swine populations throughout the USA [41]. This discrepancy could be explained by the scale at which the two studies were conducted, which implies local population dynamics may play an important role in determining disease dynamics.

Including each subsequent transmission route individually does generally reduce the objective function value compared to just the contact-based transmission. Although the reinfection transmission route does not perform well with just contact-based transmission, the objective function value decreases significantly when reinfection is paired with at least one additional transmission route. Furthermore, the lowest objective function values include reinfection as a transmission route. As a result, individuals becoming reinfected due to stress may be the most viable additional transmission route that exists in this wild hog population, which agrees with the relatively sparse pseudorabies literature [29]. However, we see conflicting results when it comes to the transmission routes of vertical transmission and increased contact during mating season. While including an increase in the transmission rate during mating season with reinfection produces the lowest objective function value when estimating six transmission rates, when estimating two transmission rates the

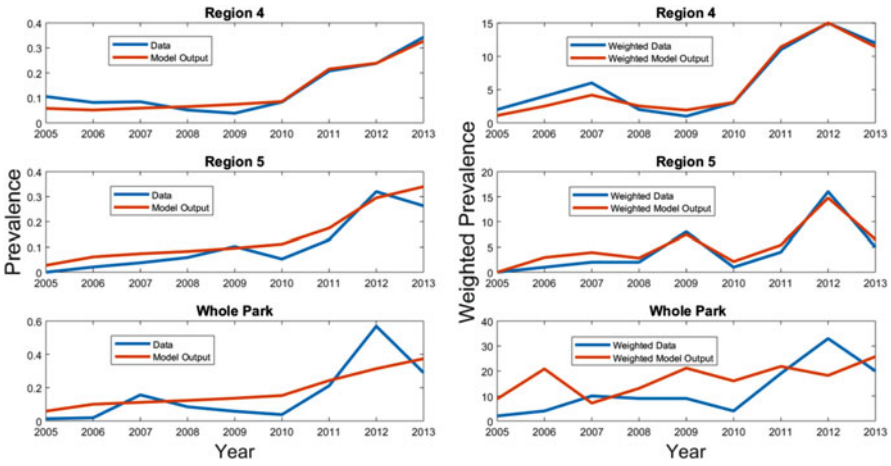


Fig. 3 Model output from estimating a different transmission rate in each region, and including increased contact during mating season, some individuals becoming reinfected, and vertical transmission from mothers to piglets

smallest objective function value is obtained when including vertical transmission with reinfection. This implies that there may be an important interaction between transmission routes since the impact of some transmission routes in spreading the disease appears to be dependent on the inclusion of others in our model. Until this inconsistency is resolved we assume that each transmission route plays some role in spreading pseudorabies. Objective function values when all transmission routes were included in the model are seen in Table 5, rows 8 and 16, and model output using these parameters is shown in Figs. 3 and 4. Parameter estimates used in these two figures are shown in Tables 6 and 7. We note that in Table 6 the vertical transmission rate α is essentially 0. This may be because vertical transmission is a less relevant pathway for the disease, but we acknowledge that it also may be due to having a larger number of parameters in that scenario.

While we obtain the lowest objective function value in the case with the most variables included in the model, the decreases cannot be attributed to overfitting as similar results are obtained with far fewer parameters by scaling the transmission rate from region 4. These facts can be seen by comparing rows 8 and 16 in Table 5 as well as Figs. 3 and 4. In other words, the decrease in objective function value is a result of the transmission routes included in the model rather than simply the number of parameters. There are several differences in results from fitting a transmission rate in each region compared to fitting just two transmission rates. First, notice in Figs. 3 and 4 how we obtain more dynamics in the unweighted output when we fit a parameter value in each region compared to the relatively smooth curves produced by fitting just two transmission rates. Additionally, estimating two transmission rates and scaling β_4 for use in other regions performs better on the entire park.

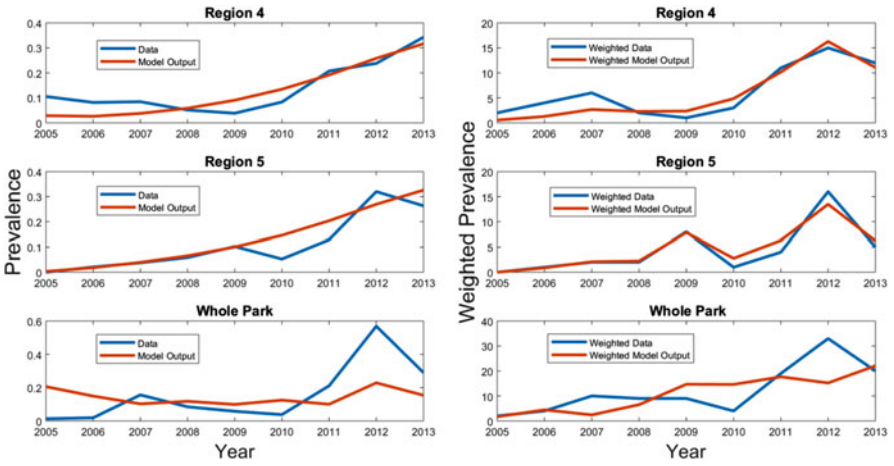


Fig. 4 Model output from estimating two transmission routes and scaling them for use in other regions, and including increased contact during mating season, some individuals becoming reinfected, and vertical transmission from mothers to piglets

Table 6 Parameter estimates obtained from including a different transmission rate in each region, and including increased contact during mating season, some individuals becoming reinfected, and vertical transmission from mothers to piglets

Parameter	β_1	β_2	β_3	β_4	β_5	β_6	α	γ	ϕ
Value	0.0010	0.0017	0.0686	0.0026	0.0019	0.0010	6.10×10^{-8}	1.042	0.0040

Table 7 Parameter estimates obtained from using two transmission rates and scaling them for use in other regions, and including increased contact during mating season, some individuals becoming reinfected, and vertical transmission from mothers to piglets

Parameter	β_1	β_2	α	γ	ϕ
Value	0.0019	0.0016	0.99	4.99	0.0044

A final outcome from this work is further development of the discrete spatial disease modeling literature. By considering the performance of the model with a different transmission rate in each region as well as the model that scales transmission rates for use in other regions, we are able to evaluate the method of scaling disease transmission rates by relative area. The performance of the model on the entire park is most relevant to this discussion, which is located in the right most column in Table 5. We can see that scaling parameter values by relative area is comparable to all corresponding instances where a different parameter value was fit in each region, and in some cases the scaled values performed better. It is also worth noting that in each scenario when β_4 is scaled for use in region 5, the values are comparable to those estimated for use in region 5. This implies that contact-based transmission of pseudorabies scales with the area of the region being modeled, which can be used to scale pseudorabies transmission rates estimated from data

obtained in one location for use in another location. Additionally, since this method of scaling parameter values is found to be useful here, it can be considered as a possible feature in any number of subsequent discrete spatial disease studies.

6 Conclusions

Transmission and dynamics of pseudorabies have been studied in domestic pigs but are not well understood in wild hogs. Specifically, transmission routes for the disease in wild populations are unknown [41]. In this project we use our pseudorabies model to analyze potential transmission routes that exist in the wild hog population in GSMNP. To achieve this, a metapopulation model that was previously developed for the population was adapted to a 1-week time step to accommodate a compartmental disease model for pseudorabies [24]. Parameters for the model were estimated using the pseudorabies prevalence data sampled from the Park from 2005–2013. In addition to basic contact-based transmission, we considered all combinations of increased transmission during mating season, transmission from mothers to piglets during nursing, and the possibility that carriers become reinfected and therefore able to transmit the disease. In each instance, we estimated different transmission rates for each region, as well as estimated different rates for the two regions containing the most prevalence data and scaled one to use in all other regions.

When parameters for a discrete spatial model are estimated from data, their values only apply to the spatial scale of the data from which they were estimated. This limits their use as they cannot reliably be applied to model dynamics in a different location. By scaling the transmission rate obtained in region 4 by relative area for use in other regions, we contribute to the relatively undeveloped literature on discrete spatial disease modeling. Since we obtain comparable results when we scaled transmission rates versus estimating a transmission rate in each region, this method can be used to apply rates to future pseudorabies scenarios as well as other discrete disease models in general.

Our work shows that our current knowledge of how pseudorabies spreads in a wild hog population is incomplete. While our results provide evidence that increased transmission during mating season, transmission from mothers to piglets, and carriers becoming reinfected may all exist as a transmission route for pseudorabies within GSMNP, the strongest evidence for an additional transmission route is in the form of carriers becoming reinfected due to stress, as our objective function values saw the largest decline when this transmission route was included in the model. Since we see dramatic reductions in objective function values when at least three transmission routes were included in the model, there may be a dependence of some transmission routes on the existence of others. Such support for the existence of numerous transmission routes for pseudorabies in wild hog populations encourages empirical tests to clearly define transmission routes so we can better understand disease dynamics, make future predictions, and consider intervention strategies.

Our model with discrete spatial and temporal features provides results that encourage further work in discrete modeling. Discrete modeling can contribute greatly to studies of various systems on our planet, but further work is needed on how to estimate transmission rates in distinct spatial regions, taking into account the sizes, the boundaries, and resources that exist in each region.

References

1. Allen, L.J.: Introduction to Mathematical Biology. Pearson/Prentice Hall (2007)
2. Allen, L.J., Brauer, F., Van den Driessche, P., Wu, J.: Mathematical Epidemiology, vol. 1945. Springer, Berlin (2008)
3. Arino, J., Van den Driessche, P.: Disease spread in metapopulations. Fields Inst. Commun. **48**(1), 1–13 (2006)
4. Atkinson, M.P., Wein, L.M.: Quantifying the routes of transmission for pandemic influenza. Bull. Math. Biol. **70**(3), 820–867 (2008)
5. Bodine, E.N., Gross, L.J., Lenhart, S.: Order of events matter: Comparing discrete models for optimal control of species augmentation. J. Biol. Dyn. **6**(sup2), 31–49 (2012)
6. Bratton, S.P.: The effect of the European wild boar (Sus scrofa) on the high-elevation vernal flora in Great Smoky Mountains National Park. Bull. Torrey Botanical Club **101**(4), 198–206 (1974)
7. Breban, R., Drake, J.M., Stallknecht, D.E., Rohani, P.: The role of environmental transmission in recurrent avian influenza epidemics. PLoS Comput. Biol. **5**(4), e1000346 (2009)
8. Cappaert, D., McCullough, D.G., Poland, T.M., Siegert, N.W.: Emerald ash borer in North America: a research and regulatory challenge. Am. Entomol. **51**(3), 152–165, 51(3)
9. Cavendish, T., Stiver, W., Delozier, E.K.: Disease surveillance of wild hogs in Great Smoky Mountains National Park—a focus on pseudorabies. In: National Conference on Feral Hogs, p. 7
10. Center for Disease Control and Prevention (CDC): Leptospirosis (2010). https://www.cdc.gov/leptospirosis/index.html
11. Center for Disease Control and Prevention (CDC): Toxoplasmosis (toxoplasma infection) (2018). https://www.cdc.gov/parasites/toxoplasmosis/index.html
12. Edelstein-Keshet, L.: Mathematical Models in Biology. SIAM (2005)
13. Engeman, R.M., Smith, H.T., Severson, R., Severson, M.A., Woolard, J., Shwiff, S.A., Constantin, B., Griffin, D.: Damage reduction estimates and benefit-cost ratios for feral swine control from the last remnant of a basin marsh system in Florida. Environ. Conserv. **31**(3), 207–211 (2004)
14. Engeman, R.M., Smith, H.T., Shwiff, S.A., Constantin, B., Woolard, J., Nelson, M., Griffin, D.: Prevalence and economic value of feral swine damage to native habitat in three Florida state parks. Environ. Conserv. **30**(4), 319–324 (2003)
15. Engeman, R.M., Woolard, J., Smith, H.T., Bourassa, J., Constantin, B.U., Griffin, D.: An extraordinary patch of feral hog damage in Florida before and after initiating hog removal. Hum. Wildl. Conflicts **1**(2), 271–275 (2007)
16. Epanchin-Niell, R.S. and Hastings, A.: Controlling established invaders: integrating economics and spread dynamics to determine optimal management. Ecol. Lett. **13**(4), 528–541 (2010)
17. Gerhold, R.W., Saraf, P., Chapman, A., Zou, X., Hickling, G., Stiver, W.H., Houston, A., Souza, M., Su, C.: *Toxoplasma gondii* seroprevalence and genotype diversity in select wildlife species from the southeastern United States. Parasit. Vectors **10**(1), 508 (2017)
18. Gibbon, J.W., Scott, D.E., Ryan, T.J., Buhlmann, K.A., Tuberville, T.D., Metts, B.S., Greene, J.L., Mills, T., Leiden, Y., Poppy, S., et al.: The global decline of reptiles, déjà vu amphibians: reptile species are declining on a global scale. six significant threats to reptile populations

are habitat loss and degradation, introduced invasive species, environmental pollution, disease, unsustainable use, and global climate change. BioSci **50**(8), 653–666 (2000)
19. Hahn, E., Page, G., Hahn, P., Gillis, K., Romero, C., Annelli, J., Gibbs, E.: Mechanisms of transmission of Aujeszky's disease virus originating from feral swine in the USA. Vet. Microbiol. **55**(1–4):123–130 (1997)
20. Howe, T.D. and Bratton, S.P.: Winter rooting activity of the European wild boar in the Great Smoky Mountains National Park. Castanea **41**(3), 256–264 (1976)
21. Jones, P.: A historical study of the European wild boar in North Carolina. Master of Arts in Education Thesis, Appalachian State Teachers College, Boone, North Carolina (1957)
22. Lee, E.C., KellyJr, M.R., Ochocki, B.M., Akinwumi, S.M., Hamre, K.E., Tien, J.H., Eisenberg, M.C.: Model distinguishability and inference robustness in mechanisms of cholera transmission and loss of immunity. J. Theor. Biol. **420**, 68–81 (2017)
23. Leroy, E.M., Epelboin, A., Mondonge, V., Pourrut, X., Gonzalez, J.-P., Muyembe-Tamfum, J.-J., Formenty, P.: Human ebola outbreak resulting from direct exposure to fruit bats in Luebo, Democratic Republic of Congo, 2007. Vector Borne Zoonotic Dis. 9(6):723–728 (2009)
24. Levy, B., Collins, C., Lenhart, S., Madden, M., Corn, J., Salinas, R.A., Stiver, W.: A metapopulation model for feral hogs in Great Smoky Mountains National Park. Nat. Resour. Model. **29**(1), 71–97 (2016)
25. Madden, M., Welch, R., Jordan, T., Jackson, P., Seavey, R., Seavey, J.: Digital Vegetation Maps for the Great Smoky Mountains National Park. The University of Georgia, Department of Geography, Athens (2004)
26. Magori-Cohen, R., Louzoun, Y., Herziger, Y., Oron, E., Arazi, A., Tuppurainen, E., Shpigel, N.Y., Klement, E.: Mathematical modelling and evaluation of the different routes of transmission of lumpy skin disease virus. Vet. Res. **43**(1), 1 (2012)
27. Mengeling, W., Pirtle, E.: Susceptibility of feral swine to acute and latent infections with pseudorabies virus. In: Proceedings of the Feral Pig Symposium, Orlando, pp. 37–38 (1989)
28. Miller, R.S., Farnsworth, M.L., Malmberg, J.L.: Diseases at the livestock–wildlife interface: status, challenges, and opportunities in the United States. Prev. Vet. Med. **110**(2), 119–132 (2013)
29. Müller, T., Hahn, E., Tottewitz, F., Kramer, M., Klupp, B., Mettenleiter, T., Freuling, C.: Pseudorabies virus in wild swine: a global perspective. Arch. Virol. **156**(10), 1691–1705 (2011)
30. National-Park-Service: Great Smoky Mountains National Park feral hog harvest data from 1980–2013 (1980). Accessed 30 Aug 2011
31. National-Park-Service: Great Smoky Mountains National Park hard mast index from 1981–2010 (1981). Accessed 30 Aug 2011
32. New Jr, J.C., Delozier, K., Barton, C.E., Morris, P.J., Potgieter, L.N.: A serologic survey of selected viral and bacterial diseases of European wild hogs, Great Smoky Mountains National Park, USA. J. Wildl. Dis. **30**(1), 103–106 (1994)
33. Olson, L.J., et al.: The economics of terrestrial invasive species: a review of the literature. Agric. Resour. Econ. Rev. **35**(1), 178 (2006)
34. Pirtle, E.C., Sacks, J.M., Nettles, V.F., Rollor, I.: Prevalence and transmission of pseudorabies virus in an isolated population of feral swine. J. Wildl. Dis. **25**(4), 605–607 (1989)
35. Pomeranz, L.E., Reynolds, A.E., Hengartner, C.J.: Molecular biology of pseudorabies virus: impact on neurovirology and veterinary medicine. Microbiol. Mol. Biol. Rev. **69**(3), 462–500 (2005)
36. Romero, C.H., Meade, P., Santagata, J., Gillis, K., Lollis, G., Hahn, E.C., Gibbs, E.P.J.: Genital infection and transmission of pseudorabies virus in feral swine in Florida, USA. Vet. Microbiol. **55**(1–4), 131–139 (1997)
37. Sandfoss, M.R., DePerno, C.S., Betsill, C.W., Palamar, M.B., Erickson, G., Kennedy-Stoskopf, S.: A serosurvey for Brucella suis, classical swine fever virus, porcine circovirus type 2, and pseudorabies virus in feral swine (Sus scrofa) of eastern North Carolina. J. Wildl. Dis. **48**(2), 462–466 (2012)

38. Scott, C., Pelton, M.: Seasonal food habits of the European wild hog in the Great Smoky Mountains National Park. Proc. Southeast. Assoc. Game Fish Commissioners **29**, 585–593 (1975)
39. Simberloff, D.: Conservation Biology for All, chapter 7. In: Invasive Species, pp. 131–152. Oxford University Press, Oxford (2010)
40. Singer, F.J.: Wild pig populations in the national parks. Environ. Manag. **5**(3), 263–270 (1981)
41. Smith, G.: Preferential sexual transmission of pseudorabies virus in feral swine populations may not account for observed seroprevalence in the USA. Prev. Vet. Med. **103**(2), 145–156 (2012)
42. Torremorell, M.: Viral causes of infertility and abortion in swine. In: Current Therapy in Large Animal Theriogenology, 2nd edn. Elsevier, Amsterdam (2006)
43. Tozzini, F., Poli, A., Croce, G.D.: Experimental infection of European wild swine (Sus scrofa l.) with pseudorabies virus. J. Wildl. Dis. **18**(4), 425–428 (1982)
44. Van Nes, A.: Epidemiology: Mathematical modelling of pseudorabies virus (syn. Aujeszky's disease virus) outbreaks aids eradication programmes: a review. Vet. Q **23**(1), 21–26 (2001)
45. Witmer, G.W., Sanders, R.B., Taft, A.C.: Feral swine—are they a disease threat to livestock in the United States? USDA National Wildlife Research Center-Staff Publications, p. 292. (2003)

Application of Mathematical Epidemiology to Crop Vector-Borne Diseases: The Cassava Mosaic Virus Disease Case

Michael Chapwanya and Yves Dumont

1 Introduction

Agriculture began from around 9000 BC independently in different places around the world. Since then, humans have developed their knowledge to produce food more efficiently and to increase crop yields, while promoting the development of the human community. However, this is possible as long as Mother Earth has the capacity to sustain all developments induced by humans. Looking ahead to the year 2050, approximately 9 billion people are expected to live in places where 10,000 years ago, only a few hundred thousands were living. In essence, there is a challenge to provide food for people on (more and more) limited area of (arable) land. Even if crop yields have drastically increased during the last century, crops have to face new dangers in the form of diseases and pests. For a number of years, pesticides were developed as a solution to combat the spread of pests and diseases. Unfortunately, we now know that most of these pesticides have caused a lot of damage (like cancer, loss in the biodiversity, etc.). Consequently, there is need to develop sustainable approaches to maintain yields and reduce the use of chemical products as much

M. Chapwanya
Department of Mathematics and Applied Mathematics, University of Pretoria, Pretoria, South Africa
e-mail: m.chapwanya@up.ac.za

Y. Dumont (✉)
CIRAD, UMR AMAP, Pôle de Protection des Plantes, F-97410 Saint-Pierre, La Réunion, France

AMAP, Univ Montpellier, CIRAD, CNRS, INRA, IRD, Montpellier, France

Department of Mathematics and Applied Mathematics, University of Pretoria, Pretoria, South Africa

EPITAG, LIRIMA, Paris, France
e-mail: yves.dumont@cirad.fr

M. I. Teboh-Ewungkem, G. A. Ngwa (eds.), *Infectious Diseases and Our Planet*, Mathematics of Planet Earth 7, https://doi.org/10.1007/978-3-030-50826-5_4

as possible in order to protect the biodiversity, decrease the risk of cancer or other diseases, and, also, to protect our Earth for the future generations.

While the first mathematical model on inoculation against smallpox was developed by Daniel Bernoulli around 1760, it is mostly acknowledged that Mathematical Epidemiology and the first compartmental models were initiated, not by mathematicians, but by public health physicians, like Sir R.A. Ross, W.H. Hamer, A.G. McKendrick, and W.O. Kermack between 1900 and 1935, [6] (see also [4]). Since then, various epidemiological models have been developed and studied mathematically. However, while mathematical models for vector-borne diseases in humans and animals are well documented in the literature, very little work has been done in modeling vector-borne diseases in plants.

The aim of this chapter is to present an example of crop–vector–virus interaction, using cassava crop because it has become a very important crop for many people in Africa. After rice and maize, it is now the third-largest source of food and carbohydrates [17]. We will begin here in the introduction by giving a brief historical background of crop diseases to emphasize that crop diseases are as old as human diseases like malaria, and probably started with the neolithic revolution when humans moved from a lifestyle of hunting and gathering to one of agriculture and settlement. We then explain the importance of cassava and the impact of its main disease, the cassava mosaic disease, and summarize previous modeling works.

1.1 Crop Diseases in the Past

Crop diseases are known since the ancient time (see, for instance, [32] and also [5]). Plant diseases have been recorded in a few Egyptian writings, on Sumerian clay tablets (1200 BC), and also in the Old Testament and in the Talmud. Even in the far East, plant diseases were recognized, in particular, in India, Vedic literature (1500–500 BC), and also some two thousands old works are available in China. In ancient Greece, some diseases of trees, cereals, and vegetables have been reported by Theophrastus (371–287 BC), also known as the "father of Botany." He was student, collaborator, and successor of Aristotle. Theophrastus and Pliny the Elder (23–79 AD) became the references in Europe until the end of the Middle Age. However, between the eighth and thirteenth century, the Arab agricultural revolution occurred. Ibn al-Awwam (Seville, late 12th) wrote one of the most important medieval book on Agriculture and also about the symptoms of tree and vine diseases as well as (more or less realistic) methods to cure. Aztec, Inca, and Mayas in the Americas also faced diseases on maize and potatoes, that will be imported by conquistadors, first, in Spain and in Portugal (in 1493, for maize, when Christopher Colombus went back to Spain [20]; potatoes were only recorded in the literature in 1537 after the Spaniards discovered high Andes of South America [22]), before their spreading throughout Europe along the sixteenth–seventeenth centuries. They became major crops in many countries in the "Old World." Particularly potatoes, leading to one of the most notable historical impacts of plant disease, the Irish potato famine of 1845,

that was caused by late blight fungus, *Phytophthora infestans*. Approximately one million people perished from starvation; around one million and a half more are believed to have left Ireland and immigrated to the USA. The French vine industry was devastated in 1878 by another fungal pathogen, the powdery mildew. This is the first time where disease management was able to find a solution and produce the first fungicide ever, the Bordeaux mixture (made with copper sulfate and lime mixture). However, for the vine industry, the real disaster came with the phylloxera, a sap sucking insect, that first devastated most of the vineyards in the continent before spreading to the rest of the world, except some places, like Chile. In fact, the number of diseases (and also pest attacks) has increased drastically since the nineteenth century, with the increase in population mobility and also with the increase of trade exchanges, including the importation of alien plants/plant products containing diseases or insects (vectors of diseases). What is unusual with crop vector-borne diseases is that in general one vector (aphids, for instance) is able to transmit several to hundreds of viruses on many crops. So that several crops can be impacted on at the same time by several virus causing more damages and eventually death. Since the discoveries of new lands, the importation of new crops, the increase of trades and population displacements, diseases and vectors have invaded or are invading areas where plants are fully susceptible, such that, in the recent decades, pests and diseases have become a major problem in many crops around the world, and, especially in Africa.

1.2 About Cassava, Cassava Mosaic Disease, and Its Vector

This chapter focuses on cassava (*Manihot esculenta* Crantz) and the cassava mosaic virus. Cassava was domesticated in the southwestern Amazon of Brazil and Bolivia some 8,000–10,000 years ago (see [25] for further details). It was introduced from the Americas onto Africa by Portuguese traders in the sixteenth century. Since then it has spread on the African continent, such that, after rice and corn, it has become a very important staple food for over 500 millions people in Africa alone. Cassava can grow in places where other crops cannot. It is tolerant to drought and can grow on soils with a low nutrient capacity. Cassava is a perennial crop. However the storage roots can be harvested from 6 to 24 months after planting, depending on the cultivar and the growing conditions. Moreover, the roots can be left in the ground without harvesting for a long period of time, making it a very useful crop as a security against famine. That is why cassava is considered as a food security crop. Its study and its protection against bio-aggressors are of main importance. The current average yields per hectare is around 12.5 tonnes, while the optimal yields can be as high as 80 tonnes under optimal conditions. This shows that there is potential for increase in production. More than 1 million of small-scale farmers around the world are producing cassava. See [18] for further information.

Cassava mosaic geminiviruses (CMGs) are the causal agents of cassava mosaic disease (CMD), which is one of the most widespread and devastating diseases of cassava in Africa. For instance, in the 1990s a severe form of the CMD lead to yield

losses of 80 to 100% in Uganda and Kenya (FAO, 2015). Currently, nine distinct CMG species (seven African and two Indian) have been found to infect cassava [40].

CMD is now prevalent in many parts of Africa [42]. The geminiviruses and their whitefly vector (*Bemisia tabaci*) have been studied extensively and much attention has been given to possible control measures and their deployment since the 1990s. CMD was first described in Tanzania in 1894. Till then, CMD became prevalent in most of the countries of sub-Saharan Africa due to the continuous cultivation of susceptible cultivars and the indiscriminate use of diseased propagation material. As highlighted in [24] (and references therein), CMD is also propagated routinely from stem cuttings coming from infected plants, such that the virus can be maintained within the plant waiting for vectors to spread. Early studies found that some varieties of cassava were less affected than others by CMD, so that resistance programs were launched in the 1930s–1940s. In the 1980s and 1990s, many projects aimed at controlling CMD were launched in Kenya, Ivory Coast, and Uganda, such that CMD received more attention than any other crop disease and this also explains why CMD is always so prominent in the literature on cassava. However, whatever the research efforts done in the last thirty years, CMD is still a problem.

Cassava mosaic geminiviruses are transmitted in a persistent manner by whiteflies, *B. tabaci* (see [33] for a recent review). This was first showed in 1932 in the Democratic Republic of Congo. Further studies showed that whiteflies are the sole vectors. There are some vectors that can acquire the virus but are unable to transmit it. The minimum time between acquisition and becoming infected is around 7 h. However, the transmission efficiency is very low (0.15–1.7% for insect collected in the field [14], to moderate (4–13%) for laboratory-reared insects [35]. The CMGs spreading is mainly related to the susceptibility/resistance of the cassava varieties, the phytosanitation measures, and if CMGs are systemic within the infected plants. However, clearly, the mobility of the whiteflies is a key factor and in general higher incidence due to severe CMD is mainly due to a high level of infestation by *B. tabaci*. Last but not least *B. tabaci* also causes direct damages to cassava by feeding on phloem or deposition of honeydew, which acts as a substrate for sooty molds (a black, non-parasitic, superficial growth of fungi on plant surfaces), that reduces both respiration and photosynthesis [38]. Surprisingly, as highlighted in [33], the research effort on the ecology of these insects was very little compared to all work done on the virus and resistant cassava cultivars. However, *B. tabaci* is a polyphagous herbivore that can potentially use a wide range of different host plants (more than 500) in cassava production landscape so that, in general, intercropping cassava with other plants (maize, coffee, sweet potato, bean, etc.) may result in a lower *B. tabaci* population abundance. However, the mechanisms resulting in the decay are not necessarily related to host-preference. Coming back to the biology of *B. tabaci*, published information suggests that its development period from egg to adult emergence is between 19 and 29 days, and the species goes through four nymphal instars before entering a pupal phase. However, in this paper, we will not enter into such details since we mainly focus on *B. tabaci* as vector of the CMD. Note also that *B. tabaci* is able to transmit more than 150 viruses, such that it has

become a very serious pest, not only for cassava. In particular, another virus that is really detrimental to cassava, also transmitted by *B. tabaci*, is the cassava brown streak virus (CBSV) [31].

The symptoms of CMD are very well known. Common symptoms include misshapen leaves, chlorosis, mottling, and mosaic. Less severe symptoms are ill-defined mosaic patterns, green mosaic with slight or absence of leaf distortion. Thus, in general, symptoms range from barely perceptible mosaic to stunting and general decline of plants. The more severe the symptoms, the lower the yield. In general young plants are more severely infected than old ones. Last but not least, the symptoms are even more severe in the case of co-infection by two cassava strains, like ACMV (African Cassava Mosaic Virus) and EACMV (Eastern-Africa Cassava Mosaic Virus). There is a synergistic interaction between the two viruses.

Since the early CMD epidemics, phytosanitary strategies have been developed. Of course, these control strategies should be sustainable, and preferably with no use of pesticides. Following [42], three approaches to decrease losses due to a virus disease have been proposed:

(i) decrease the proportion of infected plants;
(ii) delay infection to such a late stage of crop growth so that losses become unimportant;
(iii) decrease the severity of damage sustained after infection in diverse ways.

These objectives can be achieved using phytosanitation, disease-resistant varieties, good cultural practices, vector control, and mild-strain protection. All previous control measures, combined or not, have been more or less studied. In this work, we will particularly explore roguing (the removal of diseased/infected plants), that is a very well-known means of virus disease control. It is now clear that whiteflies are not easy to control by insecticides, such that vector control seems to be ineffective. However, we will also explore other control approaches that are used by small-scale farmers for other crops against whiteflies that are not necessarily used to protect cassava.

1.3 On the Usefulness of Mathematical Modeling of Cassava Mosaic Disease

The previous summary on Cassava and Cassava Mosaic Virus shows that despite years of study, field investigations, and field experiments, the risk of severe epidemics like in the 1990s cannot be excluded. We believe it is not enough to launch new research programs based only on field experiments/observations and laboratory experiments. We need to consider additional tools such as mathematical modeling and computer simulations. Mathematical modeling offers a cheaper, convenient, and powerful tool to explore the Cassava-CMD system and go further in order to better understand the possible dynamics that can occur and to possibly point out new ways worth investigating. Some mathematical models on CMD have already been developed, see for instance [15, 24, 26], but they were mostly studied using

numerical simulations, without providing the qualitative analysis of the proposed mathematical models. Of course, here we distinguish compartmental models from statistical models, which are, in general, based on data sets. The paper from Holt et al. [24] provides many interesting results. However their models are mainly based on the fact that *B. tabaci* population is limited by cassava individuals, which, in an African context, is not really true since in the introduction, we emphasize that the host range for *B. tabaci* is large and that in general cassava is inter-cropped with other plants. Thus the vector population is modeled using a logistic-like equation in the susceptible compartment, assuming some kind of skip-oviposition in the birth-rate functional, but this is not clearly explained. They also assumed constant replanting and additional mortality rate for infected plants. Both crop and vector interact using mass-action principle. Since, it has been highlighted that *B. tabaci* population size can be very large, the use of a simple mass-action function could be questionable: in fact it is not the case here since the vector population is limited by the number of cassava individuals (which is a restriction, as explained above). However, the reader has to be aware that this is the way the modeling works: you consider and/or give modeling assumptions to build your model(s), such that for the same biological problem you can derive many different models, see for example the discussion in [27]. Using their models [24], Holt and collaborators obtained various results, mainly based on numerical simulations. Their study showed a range of dynamics including cycles of infection. In their conclusions, clearly the selection of cuttings and roguing are the main tools to lower the incidence of the disease. In [46] and more recently in [23], authors included vector aggregation (and also dispersal in [23]) in their models, through nonlinear acquisition and inoculation rates. These previous works highlight the wide diversity and complexity of plant-vector-virus modeling. However, among all these models and also many others in crop epidemiology, plants are always considered as mature individuals, ignoring their growth and the effect of vectors on them before and during outbreaks [9].

The structure of the chapter is as follows: in Sect. 2, we build the mathematical model related to Cassava-CMVD interactions taking into account plant biomass growth. In Sect. 3, a qualitative analysis is provided. In Sect. 4, a specific study about the permanence of the disease is done. Finally, in Sect. 5, a sensitivity analysis followed by numerical simulations are given to illustrate the theoretical results. In particular, we discuss the usefulness of roguing and eventually other control options that could be tested with our approach.

2 The Mathematical Model

The model is formulated by considering two compartments: the total plant biomass P, and the vector population density V. The total biomass is further subdivided into two disjoint epidemiological states: the susceptible healthy biomass H_p, and the infectious biomass I_p, so that $P = H_p + I_p$. Similarly, the vector population is divided into two disjoint analogous classes, the susceptible vectors S_v, and the infectious vector I_v so that $V = S_v + I_v$. Based on the above knowledge, we propose

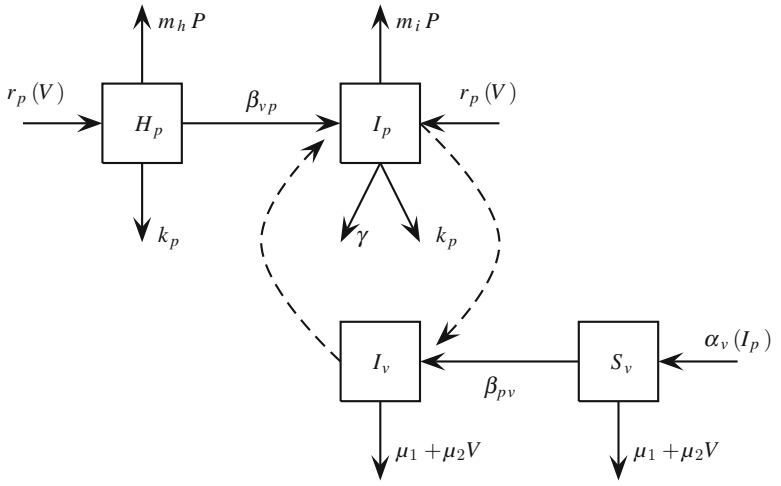


Fig. 1 Schematic diagram of the model

the epidemiological structure represented by the compartmental diagram presented in Fig. 1, p. 63, which leads to the following system of equations:

$$\begin{cases} \dfrac{dH_p}{dt} = r_p(V)H_p - m_h P H_p - \beta_{vp} H_p I_v - k_p H_p, \\ \dfrac{dI_p}{dt} = r_p(V)I_p + \beta_{vp} H_p I_v - m_i P I_p - \gamma I_p - k_p I_p, \\ \dfrac{dS_v}{dt} = \alpha_v(I_p)V - (\mu_1 + \mu_2 V)\, S_v - \beta_{pv} S_v I_p, \\ \dfrac{dI_v}{dt} = \beta_{pv} S_v I_p - (\mu_1 + \mu_2 V)\, I_v, \end{cases} \tag{1}$$

with initial conditions

$$H_p(0) = H_p^0 \geq 0, \quad I_p(0) = I_p^0 \geq 0, \quad S_v(0) = S_v^0 \geq 0, \quad I_v(0) = I_v^0 \geq 0,$$

where all the parameters are positive. Let us now describe the formulation of model (1) in detail.

Following [30] (see also [39]), the crop biomass growth is assumed to be a nonlinear process, influenced by maintenance (respiration, etc.) and intra-specific competition, while the biomass production (through photosynthesis, reproduction, etc.) is assumed to be a linear process of the biomass. Thus, the total biomass P follows a logistic-like equation

$$\frac{dP}{dt} = (r_p(V) - k_p)P - m_h(P)P,$$

where $r_p(V)$ is the biomass production rate, k_p is the harvest rate, and the function $m_h(P)$ represents the losses of biomass due to respiration, transpiration, and maintenance. It is a positive increasing function of the biomass [30]. According to [28], where the authors fit cassava (dry) biomass growth with a logistic equation, we set $m_h(P) = m_h P$. Our crop growth model is very "simple" compared to previous process-based models (PBM) (like, for instance, GUMCAS [36], GUMCAS-DSSAT [19], and SIMANIHOT [41]), but sufficient to explore the dynamics of the disease, while keeping the model mathematically tractable. Last but not least, the use of the previous PBM needs to estimate all the physiological parameters not only on healthy plants but also on diseased plants. These data are in general never available.

The function $r_p(V)$ is assumed to be a decreasing and bounded function of V, i.e., $r_p'(V) \leq 0$. This makes sense since the increase in the vector population may negatively impact the production of new biomass.

According to whiteflies biology, and like in [9], we assume that the vector population follows a logistic growth with $\alpha_v - \mu_1$ being the net growth rate, and μ_2 being the death rate due to the density effect. Contrary to [24], we do not rely the dynamics on the presence/absence of cassava, since in the African context, many other cultivated crops can host whiteflies. In fact, we follow [26], where the authors revisited the Holt Cassava Model by considering a logistic equation to model the whiteflies dynamic. However, we take into account that whiteflies have a negative effect on Cassava growth rate, such that the growth rate function is decreasing with increasing V, through a parameter ϕ, related to sap sucking. We also assume that all vector newborns are susceptible and, according to the mean lifespan of whiteflies, we assume that all infected vectors stay infected for the rest of their lives (the impact of the viruliferous period can be very important within a spatio-temporal model [9]).

The presence of the virus in the host and in the vector populations has several effects. We model only two known effects that were not necessarily developed in previous models:

(i) First, infected plants grow less than susceptible plants: the maintenance of the infected biomass, m_i, is larger than m_h, the maintenance parameter of healthy plant, due to plant defense mechanisms.
(ii) Second, the vector growth rate, $\alpha_v - \mu_1$, can also be impacted positively by the presence of diseased plants (see [10]), so that we consider $\alpha_v(I_p) - \mu_1$.

Infection between the plants and vectors is modeled by the mass-action principle. The parameter β_{vp} represents the contact rate between infectious vectors and susceptible plants. Similarly, β_{pv} represents the contact rate between infectious plants and susceptible vectors. The assumption is that the environment is uniform, homogeneous, and randomly mixed. In the model we also introduce the roguing parameter γ, i.e., the rate at which infected biomass is removed. In addition, k_p is the rate at which the cassava is harvested for their tuberous roots. The model assumes both the susceptible and infected biomass can be harvested. A summary (and values) of all parameters is provided in Table 1, p. 65. Detailed explanations of the given range of values is provided at the beginning of Sect. 5, p. 81.

Table 1 Range of values for the parameters of system (1)

Parameters	Name	Range	Unit	Source
r	Biomass maximal growth rate	[0.01, 0.06]	day^{-1}	Estimated from [28]
ϕ	Impact of sap sucking per vector	[0, 0.02]	vector^{-1}	Assumed
m_h	Maintenance rate per unit of biomass	[0.003, 0.01]	$\text{biomass}^{-1}\text{day}^{-1}$	Estimated from [28]
$m_i \geq m_h$	Maintenance rate per unit of biomass	[0.01, 0.1]	$\text{biomass}^{-1}\text{day}^{-1}$	Assumed
	when the biomass is infected			
β_{vp}	Infection rate	[0.002, 0.032]	$\text{vector}^{-1}\text{day}^{-1}$	[24]
γ	Roguing rate	[0, 1]	day^{-1}	Assumed
α	Whitefly rate of increase	[0.1, 1]	day^{-1}	Estimated from [26]
μ_1	Average death rate	[0.05, 0.2]	day^{-1}	Estimated from [26]
μ_2	Average density death rate	[0.0001, 0.01]	$\text{vector}^{-1}\text{day}^{-1}$	Estimated from [26]
β_{pv}	Acquisition rate	[0.01, 0.1]	$\text{biomass}^{-1}\text{day}^{-1}$	Assumed
k_p	Mean harvest rate	[0.0025, 0.0072]	day^{-1}	Estimated from [24]
δ	Boost parameter related to I_p	[0, 1]	–	Assumed

Following (1), when CMV is circulating, P and V satisfy the following equations:

$$\frac{dP}{dt} = (r_p(V) - k_p)P - (m_h H_p + m_i I_p)P - \gamma I_p, \tag{2}$$

and

$$\frac{dV}{dt} = \alpha_v(I_p)V - (\mu_1 + \mu_2 V)V. \tag{3}$$

3 Qualitative Analysis

In this section we take $r_p(V) = r/(1 + \phi V)$, which satisfies the assumption above for $\phi, r > 0$ with $r_p(V) \leq r_p(0) = r$. In addition, a possible choice for α_v is $\alpha_v(I_p) = \alpha\left(1 + \frac{\delta I_p}{1 + I_p}\right)$ which satisfies $\alpha_v'(I_p) > 0$ and $\alpha \leq \alpha_v(I_p) \leq \alpha(1 + \delta)$ for all I_p. We start with the well-posedness result which we state as follows.

Theorem 1 *Assume $\alpha_v(0) > \mu_1$ and $r_p(0) > k_p$. The system defines a dynamical system on the biologically feasible region*

$$\Omega = \left\{(H_p, I_p, S_v, I_v) \in \mathbb{R}^4_+ : H_p + I_p \leq P^\dagger,\ S_v + I_v \leq V^\dagger\right\},$$

where $V^\dagger = \dfrac{\alpha(1+\delta) - \mu_1}{\mu_2}$ *and* $P^\dagger = \dfrac{r_p(0) - k_p}{m_h}$.

Proof We want to show that for any non-negative initial data, the system possesses for all $t \geq 0$, a unique solution which lies in the region Ω. We will provide the proof in two steps.

First we want to show that Ω is a positively invariant set. In particular, we want to show that no trajectory leaves Ω by crossing one of its faces, [7]. On the contrary, let us assume there exist $t_1 > 0$ such that $H_p(t_1) = 0$, $H_p'(t_1) < 0$ with $H_p(t) > 0,\ I_p(t) > 0,\ S_v(t) > 0,\ I_v(t) > 0$ for all $t \in (0, t_1)$. The first equation gives

$$\frac{dH_p(t_1)}{dt} = 0,$$

which is a contradiction. Therefore $H_p(t) \geq 0$ for all $t \geq 0$. Similarly for $t_2 > 0$ with $I_p(t_2) = 0$, $I_p'(t_2) < 0$ with $H_p(t) > 0,\ S_v(t) > 0,\ I_v(t) > 0$ for all $t \in (0, t_2)$, we have

$$\frac{dI_p(t_2)}{dt} = \beta_{vp} H_p I_v > 0,$$

and $t_3 > 0$ with $S_v(t_3) = 0$, $S_v'(t_3) < 0$ with $H_p(t) > 0,\ I_p(t) > 0,\ I_v(t) > 0$ for all $t \in (0, t_3)$, we have

$$\frac{dS_v(t_3)}{dt} = \alpha_v(I_p) I_v > 0,$$

and $t_4 > 0$ with $I_v(t_4) = 0$, $I_v'(t_4) < 0$ with $H_p(t) > 0,\ I_p(t) > 0,\ S_v(t) > 0$ for all $t \in (0, t_4)$, we have

$$\frac{dI_v(t_4)}{dt} = \beta_{pv} S_v I_p > 0.$$

Therefore, in all cases the solution remains in Ω for any initial data in Ω.

In the second step, we use the a prior estimate derived below together with the fact that the right-hand side of the system is a locally Lipschitz function.

From Eq. (2), assuming $I_p \geq 0$ we have the inequality

$$\frac{dP}{dt} \leq (r_p(V) - k_p) - P(m_h H_p + m_i I_p) \leq (r_p(0) - k_p) - m_h P^2,$$

so that $\lim_{t\to\infty} \sup P(t) \leq \dfrac{r_p(0) - k_p}{m_h}$, where we have used the assumption that $r_p(V) > k_p$. Similarly, from Eq. (3) we have $\lim_{t\to\infty} \sup V(t) \leq \dfrac{\alpha(1+\delta) - \mu_1}{\mu_2}$ provided $\alpha_v(0) > \mu_1$. That is all the solutions are uniformly bounded. Combining the above two steps we conclude that (1) defines a dynamical system on Ω. □

3.1 Computation of $\mathcal{R}_0$

In this section we use the next generation matrix approach described in [44] (see also [12]), to compute the basic reproduction number $\mathcal{R}_0$, of the system. In our case, this is defined as the number of secondary infections that an infectious crop or vector could produce when introduced in a healthy population.

The system (1) has a disease-free equilibrium $E_{\text{DFE}} = (P^*, 0, V^*, 0)$. We consider only the equations where the infection progress, namely

$$\frac{dX}{dt} = \mathcal{F}(X) - \mathcal{V}(X), \tag{4}$$

where, in this case $X = (I_p, I_v)^t$, and

$$\mathcal{F}(X) = \begin{pmatrix} r_p(V)I_p + \beta_{vp}H_pI_v, \\ \beta_{pv}S_vI_p \end{pmatrix} \quad \text{and} \quad \mathcal{V}(X) = \begin{pmatrix} \left(m_iP + \gamma + k_p\right) I_p \\ \left(\mu_1 + \mu_2 V\right) I_v \end{pmatrix}.$$

Equation (4) distinguishes the rate of appearance of new infections $\mathcal{F}(X)$, from the transfer into or out of the population by any other means, $\mathcal{V}(X)$. Then

$$J_F(X) = \begin{pmatrix} r_p(V) & -r_v^2(V)\phi/rI_p + \beta_{vp}H_p \\ \beta_{pv}S_v & 0 \end{pmatrix}$$

and

$$J_V(X) = \begin{pmatrix} m_iI_p + m_iP + \gamma + k_p & 0 \\ 0 & \mu_1 + \mu_2V + \mu_2I_v \end{pmatrix}.$$

We set

$$F = \begin{pmatrix} r_p(V^*) & \beta_{vp}P^* \\ \beta_{pv}V^* & 0 \end{pmatrix}$$

and

$$V = \begin{pmatrix} m_iP^* + \gamma + k_p & 0 \\ 0 & \mu_1 + \mu_2V^* \end{pmatrix} = \begin{pmatrix} (m_i - m_h)P^* + \gamma + r_p(V^*) & 0 \\ 0 & \alpha_v(0) \end{pmatrix},$$

where $P^* = \dfrac{r_p(V^*) - k_p}{m_h}$ and $V^* = \dfrac{\alpha_v(0) - \mu_1}{\mu_2}$. Then the next generation matrix K is

$$FV^{-1} = \begin{pmatrix} r_p(V^*) & \beta_{vp}P^* \\ \beta_{pv}V^* & 0 \end{pmatrix} \begin{pmatrix} \dfrac{1}{(m_i - m_h)P^* + \gamma + r_p(V^*)} & 0 \\ 0 & \dfrac{1}{\alpha_v(0)} \end{pmatrix},$$

that is

$$FV^{-1} = \begin{pmatrix} \dfrac{r_p(V^*)}{(m_i - m_h)P^* + \gamma + r_p(V^*)} & \dfrac{\beta_{vp}P^*}{\alpha_v(0)} \\ \dfrac{\beta_{pv}V^*}{(m_i - m_h)P^* + \gamma + r_p(V^*)} & 0 \end{pmatrix}.$$

From which we deduce that

$$\mathcal{R}_0 = \frac{1}{2}\left(\mathcal{V}_0 + \sqrt{\mathcal{V}_0^2 + 4\mathcal{H}_0}\right), \tag{5}$$

where

$$\mathcal{V}_0 = \frac{r_p(V^*)}{(m_i - m_h)P^* + \gamma + r_p(V^*)},$$

represent a vertical transmission (transmission through the production of new plant biomass), and

$$\mathcal{H}_0 = \frac{\beta_{pv}\beta_{vp}V^*P^*}{\alpha_v(0)[(m_i - m_h)P^* + \gamma + r_p(V^*)]} = \frac{\beta_{pv}\beta_{vp}V^*P^*}{\alpha_v(0)[m_iP^* + \gamma + k_p]}$$

represent a horizontal transmission (through vectors).

Remark 1 Note also that when $(m_i - m_h)P^* + \gamma = 0$, that is when there is no roguing, i.e., $\gamma = 0$, and $m_i = m_h$, we have $\mathcal{V}_0 = 1$, such that we always have $\mathcal{R}_0 > 1$.

Remark 2 Instead of the complex expression given in (5), for the basic reproduction ratio, it would be biologically more natural to consider $\mathcal{V}_0 + \sqrt{\mathcal{H}_0}$, with both ways of transmission: the one-phase infectious process, from plant to plant, and the two-phase infectious process, from vector to plant and from plant to vector. Indeed, $\mathcal{V}_0 + \sqrt{\mathcal{H}_0} < 1$ implies $\mathcal{R}_0 < 1$, and $\mathcal{R}_0 > 1$ implies $\mathcal{V}_0 + \sqrt{\mathcal{H}_0} > 1$.

Finally, following [44], we deduce the following result

Theorem 2 *When $\mathcal{R}_0 < 1$ the DFE is locally asymptotically stable, while unstable when $\mathcal{R}_0 > 1$.*

3.2 Trivial Equilibria and Their Stabilities/Unstabilities

Several equilibria exist for model (1). In this section we focus on the following equilibria:

$$E_+ = (0, 0, V^*, 0), \qquad E_{++} = (P^\dagger, 0, 0, 0),$$

with $P^\dagger = \dfrac{r_p(0) - k_p}{m_h} > P^* = \dfrac{r_p(V^*) - k_p}{m_h}$; the disease-free equilibrium

$$E_{\text{DFE}} = \left(P^*, 0, V^*, 0\right).$$

Computing the Jacobian of system (1) leads to

$$J = \begin{pmatrix} f(H_p, I_p, S_v, I_v) & -m_h H_p & -\dfrac{r\phi H_p}{(1+\phi V)^2} & -\dfrac{r\phi H_p}{(1+\phi V)^2} - \beta_{vp} H_p \\ \beta_{vp} I_v - m_i I_p & g(H_p, I_p, S_v, I_v) & -\dfrac{r\phi I_p}{(1+\phi V)^2} & -\dfrac{r\phi I_p}{(1+\phi V)^2} + \beta_{vp} H_p \\ 0 & \alpha_v'(I_p)V - \beta_{pv} S_v & h(H_p, I_p, S_v, I_v) & \alpha_v(I_p) - \mu_2 S_v \\ 0 & \beta_{pv} S_v & \beta_{pv} I_p - \mu_2 I_v & -\mu_2(I_v + V) - \mu_1 \end{pmatrix},$$

where $f(H_p, I_p, S_v, I_v) = r_p(V) - m_h(P + H_p) - \beta_{vp} I_v - k_p$, $g(H_p, I_p, S_v, I_v) = r_p(V) - m_i(P + I_p) - \gamma - k_p$, and $h(H_p, I_p, S_v, I_v) = \alpha_v(I_p) - (\mu_1 + \mu_2 V) - \beta_{pv} I_p - \mu_2 S_v$. Thus at E_+ we have

$$J_{E_+} = \begin{pmatrix} r_p(V^*) - k_p & 0 & 0 & 0 \\ 0 & r_p(V^*) - \gamma - k_p & 0 & 0 \\ 0 & \alpha_v'(0)V^* - \beta_{pv} V^* & -\alpha_v(0) - \mu_2 V^* & \alpha_v(0) - \mu_2 V^* \\ 0 & \beta_{pv} V^* & 0 & -\mu_2 V^* - \mu_1 \end{pmatrix}. \tag{6}$$

Expanding in terms of the first column, it is clear that at least one of the eigenvalues is strictly positive, which implies that E_+ is unstable. At E_{++} we have

$$J_{E_{++}} = \begin{pmatrix} r_p(0) - k_p & -m_h P^\dagger & -r\phi P^\dagger & -P^\dagger(r\phi + 1) \\ 0 & -\gamma - k_p & 0 & \beta_{vp} P^\dagger \\ 0 & 0 & \alpha_v(0) - \mu_1 & \alpha_v(0) \\ 0 & 0 & 0 & -\mu_1 \end{pmatrix}. \tag{7}$$

$J_{E_{++}}$ being an upper triangular matrix, the eigenvalues are the diagonal terms. Since $\alpha_v(0) - \mu_1$ is strictly positive, E_{++} is unstable.

At the disease-free equilibrium, we have

$$J_{\text{DFE}}=\begin{pmatrix} -(r_p(V^*)-k_p) & -m_hP^* & -\dfrac{\phi P^* r_p^{*2}}{r} & -\dfrac{\phi P^* r_p^{*2}}{r}-\beta_{vp}P^* \\ 0 & -(m_i-m_h)P^*-\gamma & 0 & \beta_{vp}P^* \\ 0 & \alpha_v'(0)-\beta_{pv}V^* & -\mu_2 V^* & \alpha_v(0)-\mu_2 V^* \\ 0 & \beta_{pv}V^* & 0 & -\mu_2 V^*-\mu_1 \end{pmatrix}.$$

Considering $|J_{\text{DFE}}-\lambda I|$, we expand in terms of the first column and this gives $\lambda_1=-(r_p(V^*)-k_p)<0$ since $r_p(V^*)>k_p$. We then expand the remaining matrix in terms of the second column and $\lambda_2=-\mu_2 V^*$. The remaining two eigenvalues are eigenvalues of the matrix

$$\begin{pmatrix} -(m_i-m_h)P^*-\gamma & \beta_{vp}P^* \\ \beta_{pv}V^* & -\mu_1-\mu_2 V^* \end{pmatrix},$$

or

$$\begin{pmatrix} -\left((m_i-m_h)P^*+\gamma\right) & \beta_{vp}P^* \\ \beta_{pv}V^* & -\alpha_v(0) \end{pmatrix}$$

after applying the definition of P^* and V^* to the diagonal entries of the reduced matrix. Clearly the trace of this matrix is negative and the requirement for positive determinant simplifies to $\left((m_i-m_h)P^*+\gamma\right)\alpha_v(0)-\beta_{pv}\beta_{vp}P^*V^*>0$, that is if

$$\mathcal{J}_0=\frac{\beta_{pv}\beta_{vp}P^*V^*}{\left((m_i-m_h)P^*+\gamma\right)\alpha_v(0)}<1. \tag{8}$$

We summarize the previous results as follows

Proposition 1 *Consider model* (1)*:*

- *The equilibria E_+ and E_{++} are always unstable.*
- *Assuming $(m_i-m_h)P^*+\gamma>0$. The disease-free equilibrium, DFE, is locally asymptotically stable when $\mathcal{J}_0<1$, and unstable otherwise.*

Remark 3 When $(m_i-m_h)P^*+\gamma=0$, according to the previous section, and Remark 1, p. 68, we have $\mathcal{R}_0>1$, such that DFE is always unstable.

Remark 4 A remark is necessary here with regards to the basic reproduction numbers defined in (5) and (8). They are mathematically equivalent: as indicated in [11] it will depend on the choice of $\mathcal{F}$. It is however important to have a biologically relevant decomposition. Using the next generation matrix approach, if we choose $r_p(V)I_p$ as part of $\mathcal{V}$, then $\mathcal{R}_0^2=\mathcal{J}_0$. However a direct computation shows that since $\mathcal{R}_0=1$ so is $\mathcal{J}_0$. Indeed, following (5), $\mathcal{R}_0=1$ is equivalent to

$$\sqrt{\mathcal{V}_0+4\mathcal{H}_0}=2-\mathcal{V}_0,$$

or equivalently $\mathcal{H}_0 = 1 - \mathcal{V}_0$, i.e.,

$$\frac{\beta_{pv}\beta_{vp}P^*V^*}{[(m_i - m_h)P^* + \gamma + r_p(V^*)]\alpha_v(0)} = 1 - \frac{r_p(V^*)}{(m_i - m_h)P^* + \gamma + r_p(V^*)},$$

or

$$\frac{\beta_{pv}\beta_{vp}P^*V^*}{\alpha_v(0)} = (m_i - m_h)P^* + \gamma,$$

that is exactly equivalent to have $\mathcal{J}_0 = 1$. More precisely, using the previous result, we can show the following relationships:

$$\mathcal{R}_0 \leq \mathcal{J}_0 \leq 1, \qquad \text{and} \qquad 1 \leq \mathcal{J}_0 \leq \mathcal{R}_0.$$

3.3 Global Stability of the DFE

Next we improve on the previous result and show that the disease-free equilibrium is globally asymptotically stable (GAS) when $\mathcal{R}_0 \leq 1$. Following the work of [8], we rewrite our system by splitting the uninfected compartments $x = (H_p, S_v)$ from the infected compartments $y = (I_p, I_v)$. That is

$$\begin{cases} \dfrac{dx}{dt} = f(x, y), \\ \dfrac{dy}{dt} = g(x, y), \end{cases}$$

such that $g(x, 0) = 0$. Let us consider the system

$$\frac{dx}{dt} = f(x, 0) = \begin{pmatrix} \dfrac{rH_p}{1 + \phi S_v} - m_h H_p^2 \\ \alpha_v(0)S_v - (\mu_1 + \mu_2 S_v)\, S_v \end{pmatrix}. \tag{9}$$

The second equation is the well-known logistic equation, for which we know that V^* is GAS. Then, we deduce that P^* is GAS for the first equation. Altogether the DFE $x^* = (P^*, V^*)$ is GAS for system (9). Now we consider

$$g(x, y) = \begin{pmatrix} r_p(V)I_p + \beta_{vp}H_pI_v - \left(m_iP + \gamma + k_p\right) I_p \\ \beta_{pv}S_vI_p - (\mu_1 + \mu_2 V)\, I_v \end{pmatrix}$$

such that

$$J_g(x, y) = \begin{pmatrix} r_p(V) - \left(m_iP + m_iI_p + \gamma + k_p\right) & -\dfrac{r\phi I_p}{(1 + \phi V)^2} + \beta_{vp}H_p \\ \beta_{pv}S_v & -(\mu_1 + \mu_2 V + \mu_2 I_v) \end{pmatrix},$$

and

$$J_g(\text{dfe}, 0) = \begin{pmatrix} r_p(V^*) - \left(m_i P^* + \gamma + k_p\right) & \beta_{vp} P^* \\ \beta_{pv} V^* & -\left(\mu_1 + \mu_2 V^*\right) \end{pmatrix}.$$

Then

$$g(x, y) = \begin{pmatrix} -\left((m_i - m_h) P^* + \gamma\right) & \beta_{vp} P^* \\ \beta_{pv} V^* & -\alpha_v(0) \end{pmatrix} y - \begin{pmatrix} \beta_{vp}\left(P^* - H_p\right) \\ \beta_{pv}(V^* - S_v) \end{pmatrix}.$$

Now we have to show that $\hat{g}(x, y) \geq 0$. This is true if $S_v \leq V^*$ (always verified), and if $H_p \leq P^*$ (always verified).

In addition,

$$A = \begin{pmatrix} -\left((m_i - m_h) P^* + \gamma\right) & \beta_{vp} P^* \\ \beta_{pv} V^* & -\alpha_v(0) \end{pmatrix}$$

is a Metzler Matrix. Thus using the result from [8], we have the following result.

Theorem 3 *Assume $\gamma + (m_i - m_h) P^* > 0$. The disease-free equilibrium, (DFE) is globally asymptotically stable if $\mathcal{J}_0 \leq 1$ and unstable for $\mathcal{J}_0 > 1$.*

In the next section we investigate the permanence of system (1), p. 63, and, potentially, the existence and stability of endemic equilibria.

4 Permanence of System (1): Endemic Equilibrium

Studying the existence (and the stability/instability) of endemic equilibria for system (1) is not an easy task when $\alpha \equiv \alpha_v(I_p)$. However, in Annexe A, p. 91, we show that two types of endemic equilibria may exist (under conditions): a full endemic equilibrium (when the whole biomass is infected, i.e., $H_p = 0$), and an endemic equilibrium (when healthy and infected biomasses co-exist).

Since $\alpha_v(I_p) \geq \alpha_v(0)$, for all $I_p \geq 0$, we would like to show that a direct study is not always necessary and that it is possible to consider the following subsystem of (1) to derive interesting facts about the dynamics of system (1):

$$\begin{cases} \dfrac{dH_p}{dt} = r_p(V) H_p - m_h H_p P - \beta_{vp} H_p I_v - k_p H_p, \\ \dfrac{dI_p}{dt} = r_p(V) I_p + \beta_{vp} H_p I_v - m_i P I_p - \gamma I_p - k_p I_p, \\ \dfrac{dS_v}{dt} = \alpha_v(0) V - \left(\mu_1 + \mu_2 V\right) S_v - \beta_{pv} S_v I_p, \\ \dfrac{dI_v}{dt} = \beta_{pv} S_v I_p - \left(\mu_1 + \mu_2 V\right) I_v, \end{cases} \tag{10}$$

Using the previous results, it is straightforward to show that system (10) admits a unique non-negative solution, $\underline{x}$. Let x be the solution of (1). Then, assuming that $x(0) = \underline{x}(0)$, and using the comparison Theorem (see, for instance, Theorem 3.2, p. 209, in [45]), we can deduce that $\underline{x}$ is a lower solution of system (1), that is $x(t) \geq \underline{x}(t)$, for all $t > 0$.

Thus, any sufficient conditions to show permanence of (10), i.e., $\underline{x(t)} > 0$, will ensure permanence of (1), $x(t) > 0$ for all $t > 0$.

Assuming the dynamics of the vector to be very fast compared to the dynamics of the biomass, we assume quasi-steady state approximation such that $V = V^*$. Then, system (10) reduces to

$$\begin{cases} \dfrac{dH_p}{dt} = r^* H_p - m_h P H_p - \beta_{vp} H_p I_v - k_p H_p, \\ \dfrac{dI_p}{dt} = r^* I_p + \beta_{vp} H_p I_v - m_i P I_p - \gamma I_p - k_p I_p, \\ \dfrac{dI_v}{dt} = \beta_{pv}(V^* - I_v) I_p - \alpha_v I_v, \end{cases} \tag{11}$$

where $r^* = r/(1 + \phi V^*)$. For the equilibria we need to solve

$$\begin{aligned} r^* H_p^* - m_h P^* H_p^* - \beta_{vp} H_p^* I_v^* - k_p H_p^* &= 0, \\ r^* I_p + \beta_{vp} H_p^* I_v^* - m_i P^* I_p^* - \gamma I_p^* - k_p I_p^* &= 0, \\ \beta_{pv}(V^* - I_v^*) I_p^* - \alpha_v I_v^* &= 0. \end{aligned} \tag{12}$$

Notice that the definition of P^* here is different from earlier results. Thus looking for equilibria in Eq. (12) leads to

$$\left(r^* - k_p - m_h H_p - m_i I_p\right) P^* = \gamma I_p, \tag{13}$$

that is

$$\left(r^* - k_p - m_h P^* - (m_i - m_h) I_p^*\right) P^* = \gamma I_p^*, \tag{14}$$

so that

$$\frac{r^* - k_p - m_h P^*}{(m_i - m_h) P^* + \gamma} P^* = I_p^*, \tag{15}$$

assuming either $\gamma > 0$ or $m_i > m_h$. Thus, if $I_p = 0$ then $P_{\text{dfe}} = \dfrac{r^* - k_p}{m_h}$, and thus $I_v = 0$ and $H_{p,\text{dfe}} = P_{\text{dfe}}$.

Now assume that $0 \leq \gamma + k_p < r^*$, and $H_p = 0$, then from the second equation and third equation we deduce,

$$P^{\#} = I_p^{\#} = \frac{r^* - \gamma - k_p}{m_i}, \qquad I_v^{\#} = \frac{\beta_{pv} P^{\#}}{\alpha_v + \beta_{pv} P^{\#}} V^*.$$

The basic reproduction number for the reduced system (11) becomes

$$\mathcal{R}_0 = \frac{1}{2}\Bigg(\frac{r^*}{(m_i - m_h) P_{\text{dfe}} + \gamma + r^*} + \sqrt{\left(\frac{r^*}{(m_i - m_h) P_{\text{dfe}} + \gamma + r^*} \right)^2 + 4 \frac{\beta_{pv} \beta_{vp} P^* V^*}{\left((m_i - m_h) P_{\text{dfe}} + \gamma + r^* \right) \alpha_v}} \Bigg).$$

Assume now that $H_p > 0$, Then, using the first equation, we have

$$\frac{r^* - k_p - m_h P^*}{\beta_{vp}} = I_v.$$

Replacing I_v in the third equation leads to

$$\beta_{pv}\left(V^* - I_v\right) I_p - \alpha_v I_v = \beta_{pv}\left(V^* - \frac{r^* - k_p - m_h P^*}{\beta_{vp}} \right) I_p - \alpha_v \frac{r^* - k_p - m_h P^*}{\beta_{vp}} = 0.$$

Replacing I_p by $\dfrac{r^* - k_p - m_h P^*}{(m_i - m_h) P + \gamma} P$ leads to

$$\beta_{pv} \left(V^* - \frac{r^* - k_p - m_h P^*}{\beta_{vp}} \right) \frac{r^* - k_p - m_h P^*}{(m_i - m_h) P^* + \gamma} P^* = \alpha_v \frac{r^* - k_p - m P^*}{\beta_{vp}}.$$

Assuming that $\left(r^* - k_p - m P^*\right) \neq 0$, we can simplify, such that

$$\beta_{pv} \left(\beta_{vp} V^* - r^* + k_p + m_h P^* \right) P^* = \alpha_v \left((m_i - m_h) P^* + \gamma \right),$$

that is

$$\beta_{pv} m_h P^{*2} - \beta_{pv} \left(r^* - k_p - \beta_{vp} V^* - \alpha_v (m_i - m_h) \right) P^* - \gamma \alpha_v = 0,$$

which is a quadratic equation in P. The discriminant of this equation is

$$\Delta = \left(\beta_{pv} (r^* - k_p - \beta_{vp} V^*) - \alpha_v (m_i - m_h) \right)^2 + 4 \gamma \alpha_v \beta_{pv} m_h > 0,$$

such that there exists only one non-negative solution

$$P^* = \frac{1}{2 \beta_{pv} m_h} \left(\beta_{pv} \left(r^* - k_p - \beta_{vp} V^* - \alpha_v (m_i - m_h) \right) + \sqrt{\Delta} \right).$$

Then we can deduce

$$I_p^* = \frac{r^* - k_p - m_h P^*}{(m_i - m_h)P^* + \gamma} P^*, \qquad H_p^* = P^* \left(1 - \frac{r^* - k_p - m_h P^*}{(m_i - m_h)P^* + \gamma}\right)$$

and

$$I_v^* = \frac{r^* - k_p - m_h P^*}{\beta_{vp}}.$$

We summarize the previous results as follows

Proposition 2 *System* (11) *always admits the following three (non-zero) equilibria:*

- *The disease-free equilibrium* $DFE = (P_{dfe}, 0, 0)$*, whatever* $\gamma \geq 0$.
- *When* $0 \leq \gamma + k_p < r^*$*, a Full Disease Equilibrium,* $FDE = (0, P^{\#}, I_v^{\#})$.
- *When* $\gamma > 0$*, an endemic equilibrium* $EE = (H_p^*, I_p^*, I_v^*)$.

One way, to investigate the local (asymptotic) stability, is to compute the Jacobian

$$J(X) = \begin{pmatrix} r^* - k_p - m_h I_p - 2m_h H_p - \beta_{vp} I_v & -m_h H_p & -\beta_{vp} H_p \\ \beta_{vp} I_v - m_i I_p & r^* - m_i P - m_i I_p - \gamma - k_p & \beta_{vp} H_p \\ 0 & \beta_{pv}(V^* - I_v) & -\beta_{pv} I_p - \alpha_v \end{pmatrix}$$

$$= \begin{pmatrix} r^* - k_p - m_h I_p - 2m_h H_p - \beta_{vp} I_v & -m_h H_p & -\beta_{vp} H_p \\ r^* - k_p - m_h P - m_i I_p & r^* - m_i P - m_i I_p - \gamma - k_p & \beta_{vp} H_p \\ 0 & \beta_{pv}(V^* - I_v) & -\beta_{pv} I_p - \alpha_v \end{pmatrix}.$$

The stability of the disease-free equilibrium follows from the definition of $\mathcal{R}_0$. However, in the particular case of $FDE = (0, P^{\#}, I_v^{\#})$, we obtain the following Jacobian Matrix:

$$J(FDE) = \begin{pmatrix} r^* - k_p - m_h P^{\#} - \beta_{vp} V^{\#} & 0 & 0 \\ r^* - k_p - (m_i + m_h) P^{\#} & r^* - 2m_i P^{\#} - \gamma - k_p & 0 \\ 0 & \beta_{pv}\left(V^* - V^{\#}\right) & -\left(\beta_{pv} P^{\#} + \alpha_v\right) \end{pmatrix},$$

that is

$$J(FDE) = \begin{pmatrix} r^* - k_p - m_h \dfrac{r^* - \gamma - k_p}{m_i} - \beta_{vp} V^{\#} & 0 & 0 \\ \gamma - m_h P^{\#} & -r^* + \gamma + k_p & 0 \\ 0 & \beta_{pv}\left(V^* - V^{\#}\right) & -\beta_{pv} P^{\#} - \alpha_v \end{pmatrix},$$

or

$$J(FDE) = \begin{pmatrix} (r^* - k_p)\left(1 - \dfrac{m_h}{m_i}\right) + \gamma\dfrac{m_h}{m_i} - \beta_{vp}V^{\#} & 0 & 0 \\ \gamma - m_h P^{\#} & -r^* + \gamma + k_p & 0 \\ 0 & \beta_{pv}\left(V^* - V^{\#}\right) & -\beta_{pv}P^{\#} - \alpha_v \end{pmatrix}$$

Clearly, the eigenvalues are the terms in the diagonal. They are negative if $\gamma + k_p < r^*$ (this is the condition to have existence of FDE), and $\gamma + P^*(m_i - m_h) < \dfrac{m_i}{m_h}\beta_{vp}V^{\#}$. After manipulation, and using (8), the last condition is equivalent to $\mathcal{J}_0 > \dfrac{P^*}{P^{\#}}\left(1 + \dfrac{\beta_{vp}P^{\#}}{\alpha_v}\right) > 1$. In other words, when $\mathcal{J}_0$ is sufficiently large, the whole crop can become infected.

Now we consider $EE = (H_p^*, I_p^*, I_v^*)$. We obtain

$$J(EE) = \begin{pmatrix} -m_h H_p^* & -m_h H_p^* & -\beta_{vp}H_p^* \\ r^* - k_p - m_h P^* - m_i I_p & r^* - m_i P^* - m_i I_p - \gamma - k_p & \beta_{vp}H_p^* \\ 0 & \beta_{pv}\left(V^* - I_v^*\right) & -\beta_{pv}I_p^* - \alpha_v \end{pmatrix}.$$

In that case, we know that $\gamma + k_p > r^*$, such that the characteristic polynomial becomes

$$p(x) = x^3 + a_2x^2 + a_1x + a_0, \tag{16}$$

where, clearly,

$$a_2 = -\text{trace}(J(X^*)) = m_h H_p^* + m_i P^* + m_i I_p^* - r^* + \gamma + k_p + \beta_{pv}I_p^* + \alpha_v > 0.$$

The coefficient a_1 is given by

$$\begin{aligned} a_1 &= \sum_i i < j \begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix} \\ &= \begin{vmatrix} -m_h H_p^* & -m_h H_p^* \\ r^* - k_p - m_h P^* - m_i I_p & r^* - m_i P^* - m_i I_p - \gamma - k_p \end{vmatrix} \\ &\quad + \begin{vmatrix} -m_h H_p & -\beta_{vp}H_p \\ 0 & -\beta_{pv}I_p - \alpha_v \end{vmatrix} + \begin{vmatrix} r^* - m_i P^* - m_i I_p - \gamma - k_p & \beta_{vp}H_p \\ \beta_{pv}\left(V^* - I_v\right) & -\beta_{pv}I_p - \alpha_v \end{vmatrix}. \end{aligned}$$

It is clear that

$$\begin{vmatrix} -m_h H_p^* & -m_h H_p^* \\ r^* - k_p - m_h P^* - m_i I_p^* & r^* - m_i P^* - m_i I_p - \gamma - k_p \end{vmatrix} = m H_p^* ((m_i - m_h) P^* + \gamma) > 0,$$

$$\begin{vmatrix} -m_h H_p & -\beta_{vp} \\ 0 & -\beta_{pv} I_p - \alpha_v \end{vmatrix} = m_h H_p^* \left(\beta_{pv} I_p + \alpha_v\right) > 0,$$

and

$$\begin{vmatrix} r^* - m_i P^* - m_i I_p - \gamma - k_p & \beta_{vp} H_p^* \\ \beta_{pv} \left(V^* - I_v\right) & -\beta_{pv} I_p - \alpha_v \end{vmatrix}$$
$$= \left(m_i P^* + m_i I_p + \gamma + k_p - r^*\right) \left(\beta_{pv} I_p^* + \alpha_v\right) - \beta_{pv} \beta_{vp} H_p^* \left(V^* - I_v^*\right).$$

However, from (12), p. 73 we have

$$\beta_{pv} \beta_{vp} H_p V^* = \left(m_i P^* + \gamma - r^* - k_p\right) \left(\alpha_v + \beta_{pv} I_p^*\right)$$

such that

$$\begin{vmatrix} r^* - m_i P^* - m_i I_p - \gamma - k_p & \beta_{vp} H_p^* \\ \beta_{pv} \left(V^* - I_v\right) & -\beta_{pv} I_p - \alpha_v \end{vmatrix} = m_i I_p^* (\beta_{pv} I_p^* + \alpha_v) + \beta_{pv} \beta_{vp} H_p^* I_v^* > 0.$$

In fact

$$a_1 = m_h H_p^* ((m_i - m_h) P^* + \gamma) + (m_h H_p^* + m_i I_p^*)(\beta_{pv} I_p^* + \alpha_v) + \beta_{pv} \beta_{vp} H_p^* I_v^* > 0.$$

Thus we deduce that $a_1 > 0$. Finally

$$a_0 = -\begin{vmatrix} -m_h H_p & -m_h H_p & -\beta_{vp} H_p \\ r^* - k_p - m_h P^* - m_i I_p^* & r^* - m_i P^* - m_i I_p - \gamma - k_p & \beta_{vp} H_p \\ 0 & \beta_{pv} \left(V^* - I_v\right) & -\beta_{pv} I_p - \alpha_v \end{vmatrix}$$
$$= m_h H_p^* \begin{vmatrix} r^* - m_i P^* - m_i I_p - \gamma - k_p & \beta_{vp} H_p \\ \beta_{pv} \left(V^* - I_v\right) & -\beta_{pv} I_p - \alpha_v \end{vmatrix}$$
$$+ \left(r^* - k_p - m_h P^* - m_i I_p^*\right) \begin{vmatrix} \beta_{pv} \left(V^* - I_v\right) & -\beta_{pv} I_p - \alpha_v \\ -m H_p & -\beta_{vp} H_p \end{vmatrix}.$$

We have, as given above,

$$m_h H_p^* \begin{vmatrix} r^* - m_i P^* - m_i I_p - \gamma - k_p & \beta_{vp} H_p \\ \beta_{pv}(V^* - I_v) & -\beta_{pv} I_p - \alpha_v \end{vmatrix}$$
$$= m_h H_p^* \left(m_i I_p^* \left(\beta_{pv} I_p^* + \alpha_v \right) + \beta_{vp} H_p^* \beta_{pv} I_v^* \right) > 0,$$

and

$$\left(r^* - k_p - m_h P^* - m_i I_p^* \right) \begin{vmatrix} \beta_{pv}(V^* - I_v) & -\beta_{pv} I_p - \alpha_v \\ -m_h H_p & -\beta_{vp} H_p \end{vmatrix}$$
$$= -\left(m_h P^* + m_i I_p^* - r^* + k_p \right) \left(\beta_{pv} \beta_{vp} H_p \left(V^* - I_v \right) + m_h H_P^* \left(\beta_{pv} I_p + \alpha_v \right) \right),$$

with

$$m_h H_p^* \left(m_i I_p^* \left(\beta_{pv} I_p^* + \alpha_v \right) \right) - \left(m_h P^* + m_i I_p^* - r^* + k_p \right) m_h H_P^* \left(\beta_{pv} I_p + \alpha_v \right)$$
$$= -\left(m_h P^* - r^* + k_p \right) m_h H_P^* \left(\beta_{pv} I_p + \alpha_v \right)$$

such that

$$a_0 = m_h H_p^* \beta_{vp} H_p^* \beta_{pv} I_v^* - \left(m_h P^* - r^* + k_p \right) m_h H_P^* \left(\beta_{pv} I_p + \alpha_v \right)$$
$$- \left(m_h P^* + m_i I_p^* - r^* + k_p \right) \beta_{pv} \beta_{vp} H_p \left(V^* - I_v \right)$$
$$a_0 = \beta_{pv} \beta_{vp} H_p I_v \left(m_h H_p^* + m_h P^* + m_i I_p^* - r^* + k_p \right)$$
$$- \left(m_h P^* - r^* + k_p \right) m_h H_P^* \left(\beta_{pv} I_p + \alpha_v \right)$$
$$- \left(m_h P^* + m_i I_p^* - r^* + k_p \right) \beta_{pv} \beta_{vp} H_p^* V^*.$$

However, using the fact that

$$r^* - k_p - m_h P^* = \beta_{vp} I_v^*,$$

and

$$H_p^* \left(\beta_{vp} I_v - m_i I_p \right) = \left(m_i I_p + \left[\gamma + k_p - r^* \right] \right) I_p > 0,$$

we have

$$a_0 = \beta_{pv} \beta_{vp} H_p^* I_v^* \left(m_h H_p^* + m_h P^* + m_i I_p^* - r^* + k_p \right)$$
$$+ \beta_{vp} I_v m_h H_p^* \left(\beta_{pv} I_p + \alpha_v \right) + \beta_{pv} \beta_{vp} H_p^* V^* \left(\beta_{vp} I_v - m_i I_p^* \right),$$

i.e.,

$$a_0 = \beta_{pv}\beta_{vp}H_pI_v\left(m_hH_p^* + m_hP^* + m_iI_p^* - r^* + k_p\right) + \beta_{vp}I_vm_hH_P^*\left(\beta_{pv}I_p + \alpha_v\right) + \beta_{pv}\beta_{vp}V^*\left(m_iI_p + \gamma - r^* + k_p\right)I_p > 0.$$

Thus the first assumption of Rough–Hurwitz is verified: $a_i > 0$.

We now consider γ, the roguing parameter, as a bifurcation parameter. Let us define

$$\begin{aligned}\Delta(\gamma) = a_1a_2 - a_0 &= [m_hH_p^*((m_i - m_h)P^* + \gamma) + (m_hH_p^* + m_iI_p^*)\beta_{pv}I_p^* + \alpha_v) \\ &+ \beta_{pv}\beta_{vp}H_p^*I_v^*] \times [m_hH_p^* + m_iP^* + m_iI_p^* - r^* + \gamma + k_p + \beta_{pv}I_p^* + \alpha_v] \\ &- \beta_{pv}\beta_{vp}H_p^*I_v^*\left(m_hH_p^* + m_hP^* + m_iI_p^* - r^* + k_p\right) \\ &+ \beta_{vp}I_vm_hH_p^*\left(\beta_{pv}I_p + \alpha_v\right) + \beta_{pv}\beta_{vp}H_p^*V^*\left(\beta_{vp}I_v - m_iI_p^*\right).\end{aligned}$$

In order to show LAS of the endemic equilibrium, we need to show that $\Delta(\gamma) > 0$. However, due to the complexity of the formula, this will be investigated numerically. In addition, it is well known that if $\Delta(\gamma) = 0$ for some values of γ, it means that the polynomial (16) has pure imaginary conjugate roots and one real root. However, it is straightforward to show that this real root is simply $-a_2$ (because $a_1a_2 - a_0 = 0$).

In fact, if the polynomial defined in (16) has a pair of complex conjugate roots, $a \pm bi$ where $a, b \in \mathbb{R}$, which cross the real axis as γ passes through γ^*, then $\Delta(\gamma)$ changes sign as γ passes through γ^*. This can be showed very easily because we have the following relationship:

$$\Delta(\gamma) = -2a\left(b^2 + r^2\right) - 2a^3. \tag{17}$$

Similarly, if we assume that $\Delta(\gamma)$ changes sign, then according to the previous formula $a(\gamma)$ will change sign too. However, in order to derive a Hopf Bifurcation (see, for instance, Theorem 3.4.2 in [21]), we have an additional property to verify, the so-called transversality condition that indicates that the eigenvalues cross the x-axis with a non-zero velocity. In other words

$$\frac{da}{d\gamma}(\gamma^*) = a'(\gamma^*) \neq 0. \tag{18}$$

However, using (17), we have that

$$\Delta'(\gamma_i) = -2a'(\gamma^*)(b^2 + r^2)$$

Table 2 Parameters values

r	ϕ	m_h	m_i	β_{vp}	γ	α_v	μ_1	μ_2	β_{pv}	k_p	δ
0.04	0	0.01	0.01	0.008	[0, 1]	0.2	0.12	0.0002	0.02	0.005	0

Table 3 Hopf bifurcation points coordinates and first Lyapunov coefficient estimate for several harvesting rates (based on Table 1, p. 65)

k_p		H_i	I_p	I_v	γ_i^*	l_1
0.0025	H_2	0.561818	0.097525	3.863322	0.208952	$-2.953532e^{-05}$
	H_1	0.030780	0.113949	4.506589	0.045791	$-1.590897e^{-05}$
0.005	H_2	0.524541	0.090960	3.605623	0.195185	$-3.069432e^{-05}$
	H_1	0.031325	0.106193	4.203102	0.043544	$-1.620631e^{-05}$
0.0072	H_2	0.490470	0.085230	3.380376	0.182667	$-1.659246e^{-05}$
	H_1	0.031934	0.099374	3.935864	0.041605	$-1.659246e^{-03}$

such that verifying (18) is equivalent to verifying $\Delta'(\gamma^*) \neq 0$. Finally, we deduce that a Hopf Bifurcation related to the roguing parameter γ may occur at (EE, γ^*) if

- there exists a value γ^* such that $\Delta(\gamma^*) = 0$, such the Jacobian J_{EE} has a simple pair of pure imaginary eigenvalues and one real negative eigenvalue, $-a_2$,
- and $\Delta'(\gamma^*) \neq 0$.

In addition, the stability of the periodic solutions is given by the sign of the first Lyapunov coefficient of the dynamics, $l_1(EE, \gamma^*)$. If $l_1(EE, \gamma^*) < 0$, then these solutions are stable limit cycles and the Hopf bifurcation is supercritical, while if $l_1(EE, \gamma^*) > 0$ the Hopf bifurcation is subcritical [29]. l_1 being very difficult to obtain theoretically, we may use MatCont software [13] to estimate it.

We consider the parameters values provided in Table 2, p. 80, to illustrate Hopf Bifurcations. These values were chosen thanks to the (range of) values given in Table 1, p. 65.

In Table 3, p. 80, we present Hopf bifurcation results related to different values of k_p. In fact, it is clear that if $k_p \geq k_p^*$, then there is no Hopf Bifurcation.

To illustrate the Hopf Bifurcation property, we focus on $k_p = 0.005$, in Table 3, when $\delta = 0$, for simplicity. In Fig. 2, p. 81, we show the bifurcation analysis of system (11) done with MatCont [13], when $\gamma \in [0.02, 0.12]$: we show that two values, γ_1^* and γ_2^*, exist where a Hopf bifurcation occurs.

However, these two values can be obtained by solving $\Delta(\gamma^*) = 0$. This is illustrated in Figs. 3, p. 81, and 9, p. 88.

Last but not least, since we know that the Hopf bifurcation points are supercritical, we can estimate the periods for different values of $\gamma \in [\gamma_1^*, \gamma_2^*]$, using MatCont. See Fig. 4, p. 82. Surprisingly, the period belongs to a large interval, i.e., in [130, 445]. This period might certainly be linked to the crop growth parameters. The period is large except for values of γ closed to γ_2^*. However, the maximum period, 445 days, is reached at $\gamma = 0.05$. This period is almost 15 months, barely impossible to detect it in the field.

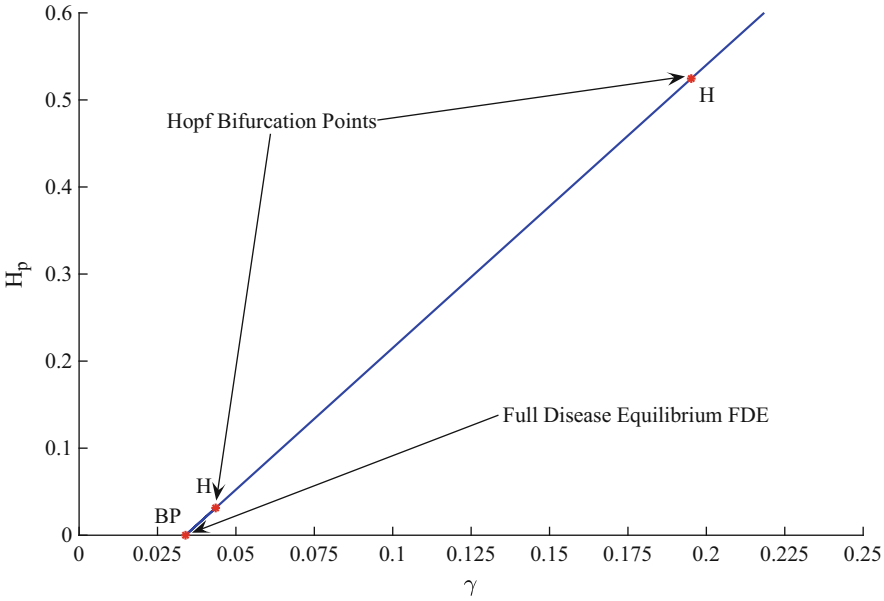
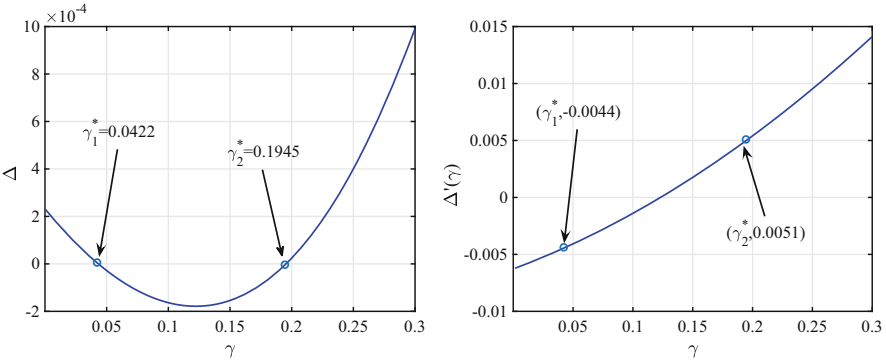


Fig. 2 Bifurcation analysis made with MatCont—$k_p = 0.005$



(a) Δ with γ. (b) Δ' with γ.

Fig. 3 The Hopf Bifurcation values γ_1^* and γ_2^*, obtained with MatCont (Fig. 2), are recovered by solving $\Delta(\gamma) = 0$ such that $\Delta'(\gamma_i^*) \neq 0$ - $k_p = 0.005$

5 Numerical Simulations

For the sensitivity analysis and the numerical simulations, we first discuss the parameters values given in Table 1, p. 65. Some of them are based on the values used in [24, 26, 28], and some of them are estimated based on the available knowledge on the virus and the cassava crops. In fact, despite so much work and published

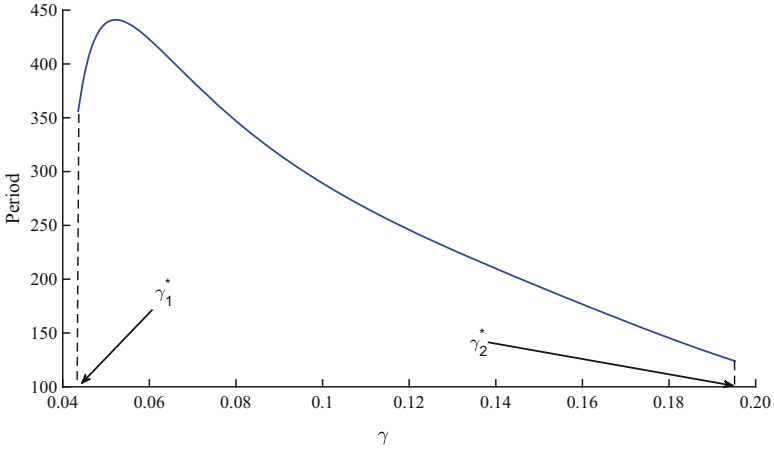


Fig. 4 Period estimate of the periodic solution according to $\gamma \in [\gamma_1^*, \gamma_2^*]$, using MatCont, for $k_p = 0.005$

papers on cassava and CMV, parameter estimation has always been a critical issue. Thus, after carefully checking the literature on cassava and CMD, we have tried to provide "realistic" estimates for all parameters used in our model. It is important to note that here we consider cassava biomass and not individual plants, hence some adjustments to the literature values were necessary.

In fact, it is not easy to get information about cassava growth: among all the work done on cassava, surprisingly, only one study from Australia was found about cassava biomass growth, [28]. These are not exactly the same environmental conditions found in Africa, but in any case, at the plants level, this study shows that a logistic equation can represent well the (dry) biomass growth. We will consider the values obtained in [28] to estimate the range of values of r and m_h. In Africa, the range of planting density can be rather large. However, for the sake of simplicity, we will consider a standard density of 10,000 plants per ha, or 1 plant per m^2.

According to (FAOSTAT2019), the average productivity between 2010 and 2017 is around 9 tons per ha in Africa, this leads to a mean productivity of 0.9 kg per m^2. Then assuming a harvest index (root weight divided by total plant weight) between 0.5 and 0.7, these leads to a total biomass weight per square meter between 1.3kg and 1.8 kg. Typically, harvest can start 8 months (250 days) after planting, so that the harvest rate k_p ranges from $1.3/250$ to $1.8/250$, that is $k_p \in [0.0052, 0.0072]$ roughly. If the harvest starts later, say 500 days after planting, then $k_p \in [0.0026, 0.0036]$, so that we can choose the interval $[0.0025, 0.0072]$ for the harvesting rate in our model.

Last, but not least, the estimate of acquisition and inoculation rates is also another issue. Only two old studies have been done (see, for instance, [16]) on the transmission efficiency and it is in general very low (from 0.17% to 2%) for the vectors. Surprisingly, we did not find information related to the number of

plants visited by vectors, or the number of feeding bites, such that estimate of acquisition and inoculation rates is rather difficult. According to the literature, it seems that the transmission rate, from the infected vector to the susceptible host, β_{vp}, is rather low; in contrary, thanks to the fact that the feeding time is long (couple of hours), the inoculation rate from the infected plant to the susceptible vector, β_{pv}, can be high. Thus, without detailed information available in the literature, we will use the estimates of k_1 and k_2 given in [24], to estimate β_{vp} and β_{pv}. However, due to the fact that they have been estimated for individual plants and not for biomass we have to divide by the total biomass weight per square meter given above, since we consider a density of one individual per square meter. Thus $\beta_{vp} \in [0.0011, 0.0246]$, while for β_{pv}, we will not consider the same range, but a larger range, say $[0.0056, 0.23]$.

5.1 Sensitivity Analysis of Model (1)

To gain insight into the correct strategies for control of the crop vector-borne disease as described by model (1), we perform a sensitivity analysis.

The most important parameter is the basic reproduction number, $\mathcal{J}_0$. In the current context it represents the amount of new infections per unit of plant biomass (vector) by the introduction of one unit of infected plant biomass (vector). In this section we aim to find out how the basic reproduction ratio responds to changes in the selected parameters. Mathematically, the sensitivity of $\mathcal{J}_0$ with respect to a parameter p is given by

$$\mathscr{F}_p^{\mathcal{J}_0} = \frac{\partial \mathcal{J}_0}{\partial p}.$$

However, we will consider the sensitivity index. This measures the change in $\mathcal{J}_0$ with respect to the percentage change in parameter p. The normalized sensitivity index of a variable $\mathcal{J}_0$, that depends differentiably on a parameter p, is defined as

$$\mathscr{E}_p^{\mathcal{J}_0} = \frac{\partial \mathcal{J}_0}{\partial p} \frac{p}{\mathcal{J}_0}.$$

The sensitivity index of $\mathcal{J}_0$ with respect to the parameter p is positive if $\mathcal{J}_0$ is increasing with respect to p and negative if $\mathcal{J}_0$ is decreasing with respect to p. For convenience, and following the discussion in Remark 4, we present the local sensitivity analysis on $\mathcal{J}_0$. Straightforward calculation leads to the following results:

$$\mathscr{E}_r^{\mathcal{J}_0} = \frac{\gamma r}{[(m_i - m_h)P^* + \gamma]P^* m_h(1 + \phi V^*)},$$

$$\mathscr{E}_\phi^{\mathcal{J}_0} = -\frac{\gamma \phi}{[(m_i - m_h)P^* + \gamma]P^* m_h(1 + \phi V^*)^2},$$

$$\mathscr{E}_{k_p}^{\mathcal{J}_0} = -\frac{\gamma \phi}{[(m_i - m_h)P^* + \gamma]P^* m_h}, \quad \mathscr{E}_{m_h}^{\mathcal{J}_0} = -\frac{\gamma \phi (r_p(V^*) - k_p)}{[(m_i - m_h)P^* + \gamma]P^* m_h^2},$$

$$\mathscr{E}_{m_i}^{\mathcal{J}_0} = -\frac{P^* m_i}{(m_i - m_h)P^* + \gamma}, \quad \mathscr{E}_\alpha^{\mathcal{J}_0} = \frac{\mu_1}{\mu_2 V^*}, \quad \mathscr{E}_{\mu_1}^{\mathcal{J}_0} = -\frac{\mu_1}{\mu_2 V^*},$$

$$\mathscr{E}_{\mu_2}^{\mathcal{J}_0} = -1, \quad \mathscr{E}_{\beta_{pv}}^{\mathcal{J}_0} = 1, \quad \mathscr{E}_{\beta_{vp}}^{\mathcal{J}_0} = 1,$$

$$\mathscr{E}_\gamma^{\mathcal{J}_0} = -\frac{\gamma}{(m_i - m_h)P^* + \gamma}.$$

However, the sensitivity index gives us only partial information because we consider only the sensitivity for one parameter only and also only on $\mathcal{J}_0$. We will now consider an additional sensitivity analysis considering that all parameters are changing. We will also focus on the system's variables H_p, I_p, S_v, and I_v. Using Table 1, p. 65, we derive some global sensitivity analysis using two well-known methods: the eFast and the LHS-PRCC methods. The eFast method given in Fig. 5, p. 85, highlights first-order effects (main effects) and total effects (main and all interaction effects) of the parameters on the Model Outputs. We also derive a LHS-PRCC sensitivity analysis given in Fig. 6, p. 86. LHS-PRCC stands for Latin Hypercube Sampling and PRCC for Partial rank correlation coefficient. These two methods give complementary information. Indeed the PRCC method provides mainly information about how the outputs are impacted if we increase (or decrease) the inputs of a specific parameter while the eFast indicates which parameter uncertainty has the greatest impact on the output variability (see, for instance, [34] for further explanations). Clearly here, the LHS-PRCC method provides the most interesting sensitivity analysis in terms of the contribution of each parameter that may depend on the chosen variable. However, some parameters, related to the whiteflies dynamics (α_v, μ_2) or the transmission (β_{vp} and β_{pv} or the plant growth (r, ϕ, γ) may have a strong effect on the system dynamics. The values for the intervals used for the sensitivity analysis are given in Table 1, p. 65.

In Figs. 7, p. 87, and 8, p. 87, we show the LHS-PRCC sensitivity analysis of both Basic reproduction numbers, $\mathcal{R}_0$ and $\mathcal{J}_0$. Since they are mathematically equivalent, it make sense that their sensitivity analysis are almost similar. The results are in agreement with what we expect: when some parameters increase, then either $\mathcal{R}_0$ or $\mathcal{J}_0$ increases or decreases. The thresholds are more sensitive to some parameters, like β_{vp}, β_{pv}, r_p, or ϕ. However, we can see that the roguing parameter, γ, has a strong negative effect, like k_p, the harvest rate. As we have seen earlier, very little is known about the values taken by β_{vp} and β_{pv}, such that additional studies would be welcome to conduct.

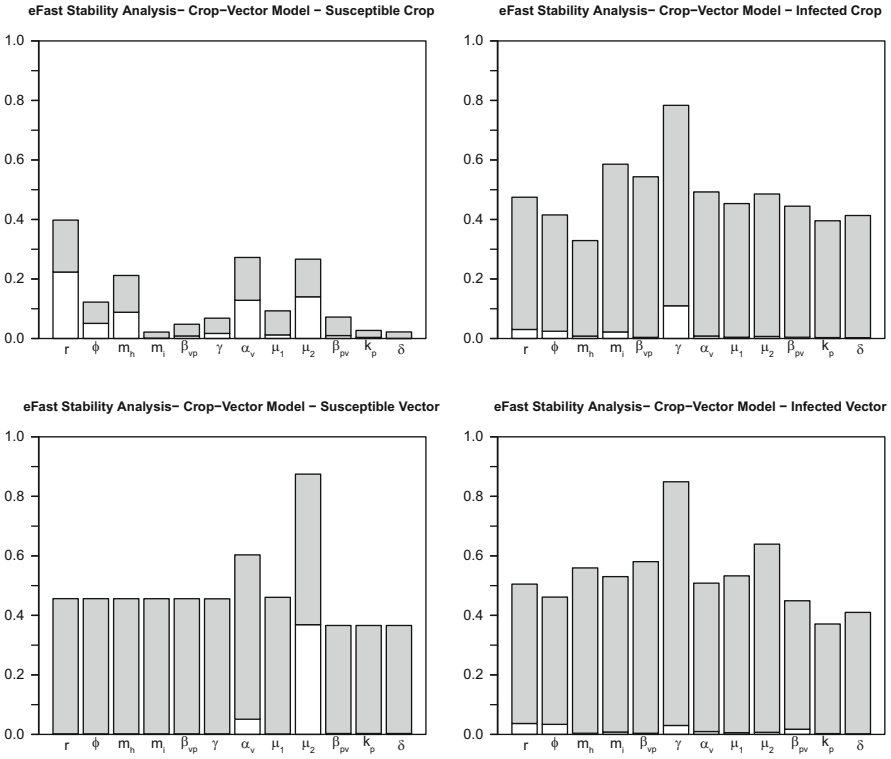


Fig. 5 eFAST sensitivity analysis. White bar:first-order effects; Sum of white and gray bars: total effect

Table 4 Parameters values with $\delta \geq 0$

r	ϕ	m_h	m_i	β_{vp}	γ	α_v	δ	μ_1	μ_2	β_{pv}	k_p
0.04	0	0.01	0.01	0.008	[0, 1]	0.2	[0, 1]	0.12	0.0002	0.02	0.005

5.2 *Numerical Simulations*

The above system of ordinary differential equations is highly nonlinear, hence in this section we present numerical simulations to support the results of the previous sections. In addition, we also present simulations to illustrate the model behavior with respect to the given parameters. We begin with the basic model where all parameters are chosen according to the baseline values provided in Table 1, p. 65. The system is integrated using MatLab's *ode solvers* and we adjust the solver's *RelTol* and *AbsTol* until the simulations have converged. The rest of the simulation were then performed using these baseline tolerance values.

The following simulations are done using parameters values given in Table 4, p. 85, and chosen using Table 1, p. 65.

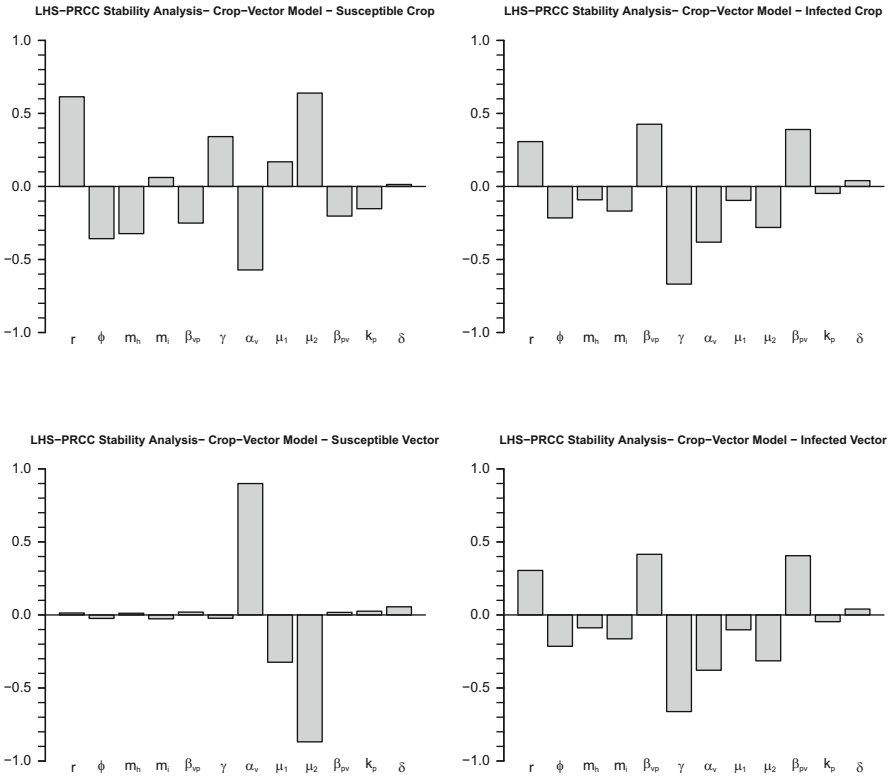


Fig. 6 LHS-PRCC sensitivity analysis

In particular, when $\delta = 0$, Fig. 9, p. 88, explores the impact of the roguing parameter, γ and also of $\mathcal{R}_0$ on the Healthy crop biomass, through the Hopf bifurcation. Thus, we consider either γ (Fig. 9a) or $\mathcal{R}_0$ (Fig. 9b) as bifurcation parameter and the vertical axis shows the values of the Healthy plant biomass. In Fig. 9a, an increase in γ, correspond to a large amount of healthy plant biomass at equilibrium. Then, with $\mathcal{R}_0$ we obtain the opposite: when $\mathcal{R}_0$ is large, then $H_p = 0$ and the Full Disease equilibrium is reached.

Clearly, thanks to the plot of the stability function Δ in Fig. 3, the endemic equilibrium loses stability as $\mathcal{R}_0$ decreases when oscillatory solutions appear.

In Fig. 10, p. 88, we provide the same simulations with $\delta > 0$, namely $\delta = 0.2$. Clearly, Hopf bifurcation occurs for smaller and larger values of γ, compared to the case when $\delta = 0$. In Fig. 11, p. 89, we consider the case when $\delta \geq 0$, for a given value of γ, here 0.1, where Hopf bifurcation occurs. Here we need to solve the full system as given in (1). Interestingly, MatLab solvers struggled to solve the full system and adjustments of *RelTol* and *AbsTol* were again necessary. First, in Fig. 11, p. 89, with (a), $\gamma = 0.1$, and (b), $\gamma = 0.2$, we show that the oscillatory behavior is amplified as δ grows. Thus $\delta > 0$ not only increases the interval, for

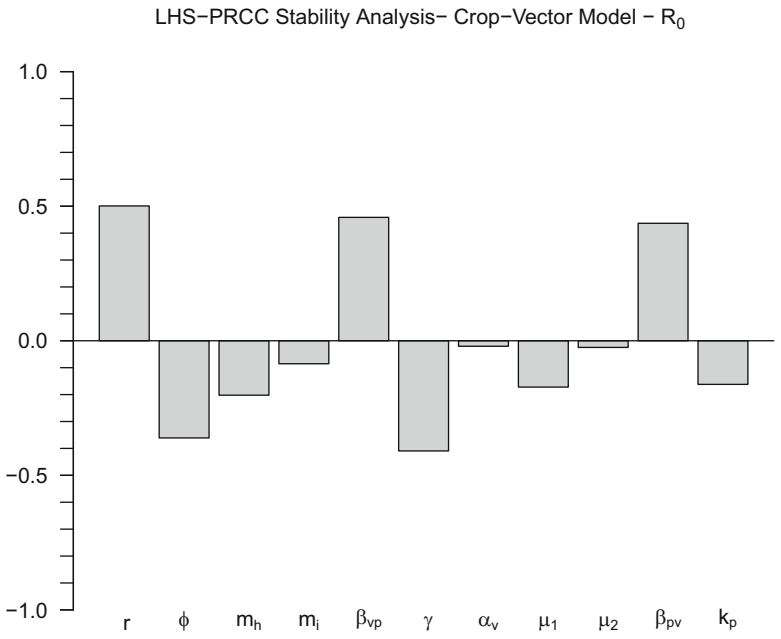


Fig. 7 LHS-PRCC sensitivity analysis of the basic reproduction numbers $\mathcal{R}_0$

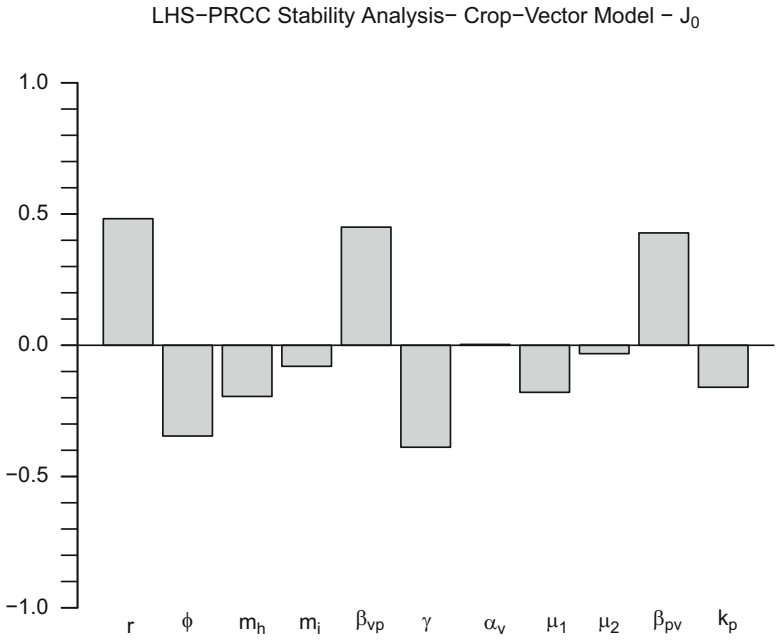
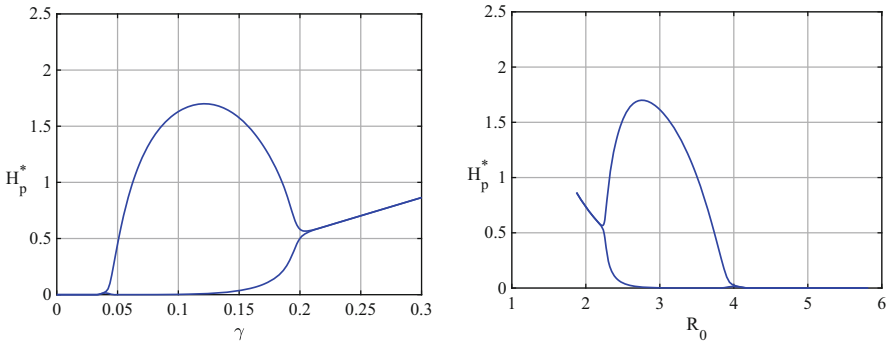
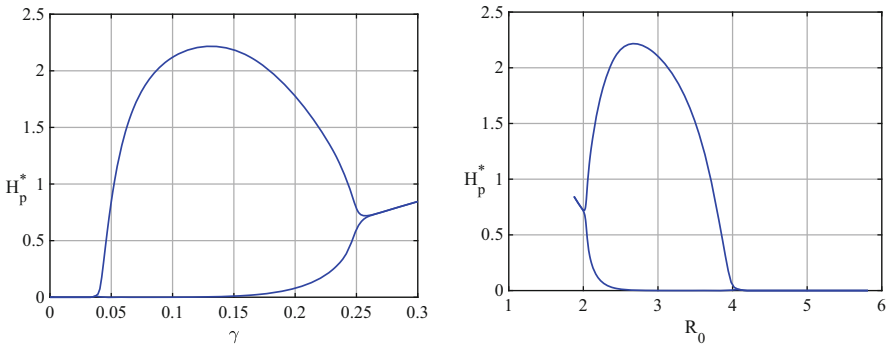


Fig. 8 LHS-PRCC sensitivity analysis of the basic reproduction number $\mathcal{J}_0$



(a) Bifurcation with respect to γ. (b) Bifurcation with respect to $\mathcal{R}_0$.

Fig. 9 Bifurcation diagram showing the minimal and maximal value of the Healthy Biomass for a given γ or $\mathcal{R}_0$ - $k_p = 0.05$ and $\delta = 0$



(a) Bifurcation with respect to γ. (b) Bifurcation with respect to $\mathcal{R}_0$.

Fig. 10 Bifurcation diagram showing the minimal and maximal value of the Healthy Biomass for a given γ or $\mathcal{R}_0$ - $k_p = 0.05$ and $\delta = 0.2$

γ, where Hopf bifurcation can take place, but also amplifies the amplitude of the oscillations. Last but not least, these numerical results confirm that our approach in the theoretical part, i.e., studying subsystem (10), p. 72, was appropriate to show that a Hopf bifurcation can occur.

So far, roguing seems to be the best way to control the Cassava Mosaic Disease. However, as our study shows, roguing can amplify a possible periodic behavior in the system. It means that, at some moment, the disease seems to almost disappear, such the decision of stopping the roguing can be taken prematurely. Similar oscillatory behavior has been observed in plant epidemiological models, see for example [24, 43]. However, it is important to notice that the model presented here is different from these works. In particular, their investigations are based on numerical simulations.

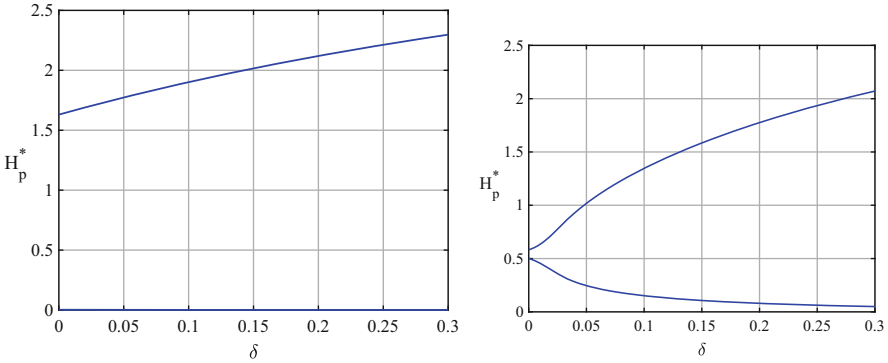


Fig. 11 Bifurcation diagram showing the maxima and minima of Healthy plant biomass and Infected vector for different values of δ with $k_p = 0.005$ and (**a**) $\gamma = 0.1$—(**b**) $\gamma = 0.2$

However, thanks to the sensitivity analysis provided in the previous subsection, some additional control strategies could be taken into account. For instance, reducing β_{vp} and β_{pv} is possible. In fact, in Kenya, tomato crops are protected against whiteflies using eco-friendly nets [37]. The results are very promising and this control approach seems to be suitable and also sustainable. However, the use of nets is only possible with young cassava plants, less than 1.5m high. The nets also may change locally the environmental parameters (temperature, humidity) of the plants, and thus may have an impact on their (photosynthetic) growth, and eventually, foster other kind of diseases, e.g., fungi. In fact, there exists several other control strategies, including the use of natural enemies, natural pesticides or pesticides extracted from plants, fungi like Beauveria bassiana, etc., and traps. Indeed, traps are widely used around the world, using pheromones or food to attract and catch insects [1], and/or for mating disruption [2].

Of course, additional bio-control strategies could be considered in our model: for example, adding resistant plants (and find, like in [3], the right percentage of resistant cassava variety to plant to reduce the impact of the disease), use entomopathogenic fungi or mycoinsecticides to kill whiteflies, etc. Last but not least, traps, using pheromones, either to catch the insects [1], and/or to disrupt mating [2]. In fact, they are many control strategies that could be tested using modeling and simulations, but also field experiments.

6 Conclusion

In this chapter, we presented a modeling example of a crop vector-borne disease with particular focus on cassava and the cassava mosaic virus disease. We showed that interesting mathematical models can be built that require various mathematical tools to be studied. Our model is relatively generic and could be applied to other

plant–vector–virus interactions, and, of course, could be improved, taking into account vector aggregation and dispersal, like in [23], or include explicitly the spatial component, like in [9].

As alluded earlier, many other control measures may be used, but the proposed model suggests roguing is the most effective way to control the disease. This involves the uprooting of infected plants from the field. However, it has also been reported, see [31], that roguing is unpopular among farmers due to the resulting reduction in plant population. Our modeling effort suggests that the induced cycles may be misleading as the farmers may prematurely stop the roguing process.

In addition to roguing, it could be interesting to consider a proportion of a new variety of cassava, less susceptible to the virus or eventually resistant (like in [3]), and then estimate theoretically the proportion needed, in a plot to reduce the epidemiological risk.

However, our interest is not only in the development and the study of new biological control models, but also in the questions highlighted by the modeling process, the analytical results and the numerical simulations. Indeed, despite a comprehensive literature review, various biological processes seem to be partially known, for instance, the transmission rates of the virus from the plant to the whitefly and from the whitefly to the plant. The plant biomass growth rates were only rigorously studied through experiments in Australia. Surprisingly, such rates were not estimated in Africa or South America despite several projects on the production of several cassava varieties under various conditions. Certainly, these rates may change, in particular, for diseased plants (i.e., estimate for m_i), and for plants infested by susceptible whiteflies (i.e., estimate for ϕ). In contrast, very complex plant models have been calibrated, but only on healthy plants. However, the reader has to be aware that field experiments can be very difficult to conduct, sometimes with uncertainty in the outputs. That is why the use of models (even theoretical ones) can be useful in designing the appropriate experiments in order to feed the models and/or to improve the knowledge on the biological system. Models are not there to replace the field expert, but to support him/her in decision-making. Last but not least, models can also help to test hypothesis about unknown processes and to focus on more specific experiments. However, whatever the models, it is always important to have in mind that they are only approximation of the reality, such that their predictability will always subject to some uncertainty.

We hope that this study on Cassava Mosaic Disease has highlighted some of the challenges to face when dealing with crop (vector-borne) diseases. That is why, to conclude, we strongly believe that crop diseases modeling can be a new field of research in Mathematical Epidemiology, and, more generally, in Applied Mathematics.

Acknowledgments The authors acknowledge the support of South African DST/NRF SARChI Chair on Mathematical Models and Methods in Bioengineering and Biosciences (M3B2) of the University of Pretoria (South Africa).

The authors thank the anonymous reviewers for their fruitful comments that greatly improve the initial manuscript.

Annexe A

Existence of a Full Endemic Equilibrium for Model (1)

To determine the endemic equilibria, we have to solve the following system:

$$\begin{cases} \dfrac{r_p H_p^\#}{1+\phi V^\#} - m_h H_p^\# P^\# - \beta_{vp} H_p^\# I_v^\# - k_p H_p^\# = 0, \\ \dfrac{r_p I_p^\#}{1+\phi V^\#} + \beta_{vp} H_p^\# I_v^\# - m_i P^\# I_p^\# - \left(\gamma + k_p\right) I_p^\# = 0, \\ \alpha_v \left(1 + \delta \dfrac{I_p^\#}{1+I_p^\#}\right) V^\# - \left(\mu_1 + \mu_2 V^\#\right) S_v^\# - \beta_{pv} S_v^\# I_p^\# = 0, \\ \beta_{pv} S_v^\# I_p^\# - \left(\mu_1 + \mu_2 V^\#\right) I_v^\# = 0, \end{cases} \tag{19}$$

where the $^\#$ denote the value at equilibrium. However we can study different cases, like $H_p^\# > 0$ and, in particular, $H_p^\# = 0$, corresponding to the Full Disease Equilibrium, FDE.

Thus, assuming $H_p^\# = 0$ leads to

$$P^\# = I_p^\# = \frac{1}{m_i}\left(\frac{r_p}{1+\phi V^\#} - \left(\gamma + k_p\right)\right),$$

with

$$V^\# = \frac{1}{\mu_2}\left(\alpha_v \left(1 + \delta \frac{I_p^\#}{1+I_p^\#}\right) - \mu_1\right). \tag{20}$$

Thus using $(19)_2$, we deduce

$$m_i I_p^\# + \left(\gamma + k_p\right) = \frac{r_p}{1+\phi V^\#} = \frac{r_p \mu_2 \left(1+I_p^\#\right)}{\mu_2\left(1+I_p^\#\right) + \phi\left(\alpha_v\left(1+I_p^\#+\delta I_p^\#\right) - \mu_1\left(1+I_p^\#\right)\right)}.$$

Reducing to the same denominator, we derive

$$\left(m_i I_p^\# + \left(\gamma + k_p\right)\right)\left(\left(1+I_p^\#\right)\left(\mu_2 + \phi\left(\alpha_v - \mu_1\right)\right) + \delta\phi\alpha_v I_p^\#\right) = r_p \mu_2 \left(1+I_p^\#\right),$$

that is

$$\left(m_i I_p^{\#} + \left(\gamma + k_p\right)\right)\left(\left(\mu_2 + \phi\left(\alpha_v - \mu_1\right)\right) + \left(\mu_2 + \phi\left(\alpha_v - \mu_1\right) + \delta\phi\alpha_v\right) I_p^{\#}\right) = r_p\mu_2\left(1 + I_p^{\#}\right).$$

Expanding all terms in the previous equality, we obtain a second order equation

$$\left(m_i\left(\mu_2 + \phi\left(\alpha_v - \mu_1\right)\right) + \left(\gamma + k_p\right)\left(\delta\phi\alpha_v + \left(\mu_2 + \phi\left(\alpha_v - \mu_1\right)\right)\right) - r_p\mu_2\right) I_p^{\#} + m_i\left(\left(\mu_2 + \phi\left(\alpha_v - \mu_1\right) + \delta\phi\alpha_v\right)\right)\left(I_p^{\#}\right)^2 + \left(\left(\gamma + k_p\right)\left(\mu_2 + \phi\alpha_v - \mu_1\right) - r_p\mu_2\right) = 0$$

that is

$$\left(B + \delta m_i\phi\alpha_v\right)\left(I_p^{\#}\right)^2 + \left(B + A + \left(\gamma + k_p\right)\phi\alpha_v\delta\right) I_p^{\#} + A = 0,$$

with

$$A = \left(\gamma + k_p\right)\left(\mu_2 + \phi\left(\alpha_v - \mu_1\right)\right) - r_p\mu_2 + \phi\left(\gamma + k_p\right)\left(\alpha_v - \mu_1\right) - \left(r_p - \left(\gamma + k_p\right)\right)\mu_2,$$

$$B = m_i\left(\mu_2 + \phi\left(\alpha_v - \mu_1\right)\right) > 0.$$

Assuming

$$\frac{\left(\gamma + k_p\right)\left(\mu_2 + \phi\left(\alpha_v - \mu_1\right)\right)}{r_p\mu_2} < 1,$$

then $A < 0$. Then, we compute

$$\Delta = \left(B + A + \left(\gamma + k_p\right)\phi\alpha_v\delta\right)^2 - 4A\left(B + \delta m_i\phi\alpha_v\right) > 0,$$

such that, we deduce the following real positive root

$$I_p^{\#} = \frac{1}{2\left(B + \delta m_i\phi\alpha_v\right)}\left(-\left(B + A + \left(\gamma + k_p\right)\phi\alpha_v\delta\right) + \sqrt{\Delta}\right) > 0. \tag{21}$$

Since we always assume that $r > \gamma + k_p$, we can consider the following two cases:

- when $\phi = 0$ (no impact of the vectors on the plant growth rate), FDE always exists. Using (21), we deduce

$$I_p^{\#} = \frac{-A}{B} = \frac{r_p - \left(\gamma + k_p\right)}{m_i} > 0.$$

  Then, the other values $V^{\#}$ (using (20)), $S_v^{\#}$ (using (19_3)), and $I_v^{\#}$ follow.

- When $\phi > 0$ and such that

$$\phi < \frac{\left(r_p - \left(\gamma + k_p\right)\right)\mu_2}{\left(\gamma + k_p\right)\left(\alpha_v - \mu_1\right)}, \tag{22}$$

then, according to (21), FDE always exists.

References

1. Anguelov, R., Dufourd, C., Dumont, Y.: Simulations and parameter estimation of a trap-insect model using a finite element approach. Math. Comput. Simul. **133**, 47–75 (2017)
2. Anguelov, R., Dufourd, C., Dumont, Y.: Mathematical model for pest-insect control using mating disruption and trapping. Appl. Math. Model. **52**, 437–457 (2017)
3. Anguelov, R., Lubuma, J.M.-S., Dumont, Y.: Mathematical analysis of vector-borne diseases on plants. In: Guo, Y., Kang, M.Z., Dumont, Y. (eds.) Plant growth modeling, simulation, visualization and applications. Proceedings PMA12: The Fourth International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications, pp. 22–29, Shanghai, China, 31 October–3 November 2012. IEEE Press, Beijing (2012)
4. Bacaër, N.: A Short History of Mathematical Population Dynamics. Springer, London (2011)
5. Borkar, S.G.: History of Plant Pathology, Agriculture Series, 248 p. Woodhead Publishing, New Delhi (2017)
6. Brauer, F.: Mathematical epidemiology: Past, present, and future. Infect. Dis. Model. **2**(2), 113–127 (2017)
7. Busenberg, S., Cooke, K.: Vertically Transmitted Disease: Models and Dynamics, vol. 23. Springer, New York (1993)
8. Castillo-Chavez, C., Feng, Z., Huang, W.: On the computation of R_0 and its role on global stability. In: Castillo-Chavez, C., Blower, S., Van der Driessche, P., Kirschner, D., Yakubu, A.A. (eds.) Mathematical Approaches for Emerging and Reemerging Infectious Diseases, An Introduction, pp. 229–250. Springer, New York (1998)
9. Chapwanya, M., Dumont, Y.: On crop vector-borne diseases. Impact of virus lifespan and contact rate on the traveling-wave speed of infective fronts. Ecol. Complex. **34**, 119–133 (2018)
10. Colvin, J., Omongo, C.A., Govindappa, M.R., Stevenson, P.C., Maruthi, M.N., Gibson, G., Seal, S.E., Muniyappa, V.: Host-Plant Viral Infection Effects on Arthropod-Vector Population Growth, Development and Behaviour: Management and Epidemiological Implications. Adv. Virus Res. **67**, 419–452 (2006). Academic Press
11. Cushing, J.M., Diekmann, O.: The many guises of $\mathcal{R}_0$ (a didactic note). J. Theor. Biol. **404**, 295–302 (2016)
12. Diekmann, O., Heesterbeek, J.A.P., Roberts, M.G.: The construction of next-generation matrices for compartmental epidemic models. J. R. Soc. Interface **7**(47), 873–885 (2009)
13. Dhooge, A., Govaerts, W., Kuznetsov, Yu.A., Meijer, H.G.E., Sautois, B.: New features of the software MatCont for bifurcation analysis of dynamical systems. MCMDS **14**(2), 147–175 (2008)
14. Fargette, D.: Epidemiologie de la mosaïque Africaine du manioc en Côte d'Ivoire, pp. 203. Ph.D. thesis, Science and Technology University, Montpellier (1985)
15. Fargette, D., Vie, K.: Modeling the temporal primary spread of African cassava mosaic virus into plantings. Phytopathology **84**(4), 378–382 (1994)
16. Fargette, D., Fauquet, C., Grenier, E., Thresh, J.M.: The spread of African cassava mosaic virus into an within cassava fields. J. Phytopathol. **130**(4), 289–302 (1990)

17. Fauquet, C., Fargette, D.: African cassava mosaic virus: etiology, epidemiology and control. Plant Dis. **74**(6), 404–11 (1990)
18. Food and Agriculture Organization of the United Nations (FAO): Save and Grow: Cassava. A Guide to Sustainable Production Intensification, Rome (2013)
19. Gabriel, L.F., Streck, N.A., Roberti, D.R., Chielle, Z.G., Uhlmann, L.O., da Silva, M.R., da Silva, S.D.: Simulating cassava growth and yield under potential conditions in Southern Brazil. Agron. J. **106**(4), 1119–1137 (2014)
20. Goodman, M.M., Galinat, W.C.: The history and evolution of Maize. Crit. Rev. Plant Sci. **7**(3), 197–220 (1988)
21. Guckenheimer, J., Holmes, P.: Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. Springer (2002)
22. Hawkes, J.G., Francisco-Ortega, J.: The early history of the potato in Europe. Euphytica **70**, 1–7 (1993)
23. Hebert, M.P., Allen, L.J.S.: Disease outbreaks in plant-vector-virus models with vector aggregation and dispersal. Discrete Contin. Dyn. Syst. B **21**(7), 2169–2191 (2016)
24. Holt, J., Jeger, M.J., Thresh, J.M., Otim-Nape, G.W.: An epidemiological model incorporating vector population dynamics applied to African cassava mosaic virus disease. J. Appl. Ecol. **34**, 793–806 (1997)
25. Isendahl, C.: The domestication and early spread of manioc (Manihot Esculenta Crantz): A brief synthesis. Lat. Am. Antiq. **22**(4), 452–468 (2011)
26. Jeger, M.J., Holt, J., van den Bosch, F., Madden, L.V.: Epidemiology of insect-transmitted plant viruses: modelling disease dynamics and control interventions. Physiol. Entomol. **29**, 291–304 (2004)
27. Jeger, M.J., Madden, L.V., van den Bosch, F.: Plant virus epidemiology: Applications and prospects for mathematical modeling and analysis to improve understanding and disease control. Plant Disease **102**(5), 837–854 (2018)
28. Keating, B.A., Evenson, J.P., Fukai, S.: Environmental effects on growth and development of cassava (Manihot esculenta Crantz.) II. Crop growth rate and biomass yield. Field Crops Res. **5**, 283–292 (1982)
29. Kuznetsov, Yu.A.: Elements of Applied Bifurcation Theory, 3rd edn. Springer (2004) (1998)
30. Lebon, A., Mailleret, L., Dumont, Y., Grognard, F.: Direct and apparent compensation in plant-herbivore interactions. Ecol. Model. **290**, 192–203 (2014)
31. Legg, J.P., Lava Kumar, P., Makeshkumar, T., Tripathi, L., Ferguson, M., Kanju, E., Ntawuruhunga, P., Cuellar, W.: Cassava virus diseases: biology, epidemiology, and management. Adv. Virus Res. **91**, 85–142 (2015)
32. Lucas, G.B., Campbell, C.L., Lucas, L.T.: History of Plant Pathology, in Introduction to Plant Diseases: Identification and Management, pp. 15–19. Springer US, Boston (MA) (1992)
33. Macfayden, S., Paull, C.A., Boykin, L.M., Barro, P.J., Maruthi, M.N., Otim, M.H., Kalyebi, A., Vassão, D.G., Sseruwagi, P., Tay, W.T., Delatte, H., Seguni, Z.S., Colvin, J., Omongo, C.: Cassava whitefly, Bemisia tabaci (Gennadius) (Hemiptera: Aleyrodidae) in East African farming landscapes: a review of the factors determining abundance. Bull. Entomol. Res. **108**(5), 565–582 (2018)
34. Marino, S., Hogue, I.B., Ray, C.J., Kirschner, D.E.: A methodology for performing global uncertainty and sensitivity analysis in systems biology. J. Theor. Biol. **254**(1), 178–196 (2008)
35. Maruthi, M.N., Colvin, J., Seal, S., Thresh, J.M.: First report of a district begomovirus infecting cassava in Zanzibar. Plant Disease **86**, 187 (2002)
36. Matthews, R.B., Hunt, L.A.: GUMCAS: a model describing the growth of cassava (Manihot esculenta L. Crantz). Field Crops Res. **36**(1), 69–84 (1994)
37. Mutisya, S., Saidi, M., Opiyo, A., Ngouajio, M., Martin, T.: Synergistic effects of agronet covers and companion cropping on reducing whitefly infestation and improving yield of open field-grown tomatoes. Agronomy (Basel) **6**(42), 14 p (2016)
38. Nelson, S.: Sooty Mold. Plant Disease PD-52. Mänoa, Honolulu, Hawaï, Cooperative Extension Service, College of Tropical Agriculture and Human Resources, University of Hawaii (2008)

39. Paine, C.E., Marthews, T.R., Vogt, D.R., Purves, D., Rees, M., Hector, A., Turnbull, L.A.: How to fit nonlinear plant growth models and calculate growth rates: an update for ecologists. Methods Ecol. Evol. **3**, 245–256 (2012)
40. Patil, B.L., Fauquet, C.M.: Cassava mosaic geminiviruses: actual knowledge and perspectives. Mol. Plant Pathol. **10**, 685–701 (2009)
41. Tironi, L.F., Streck, N.A., Gubiani, P.I., Benedetti, R.P., Freitas, C.P.O.: Simanihot: A process-based model for simulating growth, development and productivity of cassava. Engenharia Agrícola **37**(4), 471–482 (2017)
42. Tresh, J.M., Cooter, R.J.: Strategies for controlling cassava mosaic virus disease in Africa. Plant Pathol. **54**, 587–614 (2008)
43. van den Bosch, F., de Roos, A.M.: The dynamics of infectious diseases in orchards with roguing and replanting as control strategy. J. Math. Biol. **35**(2), 129–157 (1996)
44. Van den Driessche, P., Watmough, J.: Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math. Biosci. **180**(1–2), 29–48 (2002)
45. Waltman, P.: A Second Course in Elementary Differential Equations. Dover, New York (2004)
46. Zhang, X., Holt, J., Colvin, J.: A general model of plant-virus disease infection incorporating vector aggregation. Plant Pathol. **49**, 435–444 (2000)

A Multistage Mosquito-Centred Mathematical Model for Malaria Dynamics that Captures Mosquito Gonotrophic Cycle Contributions to Its Population Abundance and Malaria Transmission

Miranda I. Teboh-Ewungkem, Gideon A. Ngwa, and Mary Y. Fomboh-Nforba

1 Introduction

Many mathematical models for malaria transmission dynamics have been derived and analysed since the pioneering work of Sir Ronald Ross [26]. Some of these models are based on the assumption that the human and mosquito populations are constant, while others attempt variable human and mosquito populations [9, 11, 15–18, 20]. Other studies point to the fact that climatic factors will affect the global malaria burden problem in the future [13, 24, 33]. However, very few models exist where the demographic and reproducing life style of the malaria transmitting vector, the *Anopheles sp* mosquito, are built into the model construction process. In this paper, we consider a general mosquito–human–malaria interactive framework where the mosquito is allowed to undergo up to N *gonotrophic cycles*[1] during its entire reproductive life, where N is a positive integer greater than unity.

The idea of studying mathematical models for malaria transmission that takes into consideration the mosquito's gonotrophic cycles was used in [22]. However, given that the number of gonotrophic cycle counts was set to $N = 3$ in that paper,

[1]The cyclic path starting from the first episode of blood feeding to resting for egg maturation to oviposition and then back to blood feeding that is repeated several times during the mosquito's entire reproductive life is referred to as the *gonotrophic cycle*. The length of the gonotrophic cycle can be measured by calculating the average of the lengths of the intervals between successive batches of eggs during the mosquito's reproductive life.

M. I. Teboh-Ewungkem (✉)
Department of Mathematics, Lehigh University, Bethlehem, PA, USA
e-mail: mit703@lehigh.edu

G. A. Ngwa · M. Y. Fomboh-Nforba
Department of Mathematics, University of Buea, Buea, Cameroon

M. I. Teboh-Ewungkem, G. A. Ngwa (eds.), *Infectious Diseases and Our Planet*,
Mathematics of Planet Earth 7, https://doi.org/10.1007/978-3-030-50826-5_5

the authors had to consider a feedback mechanism whereby all vectors that were in their k-th gonotrophic cycle where $k > 3$ were re-classed into gonotrophic cycle three through a pull-back term. The main weakness of such a pull-back term meant that some mosquitoes were given infinite life spans and that affected the size of the equilibrium solutions and the threshold parameters. The main objective of the current paper is to remove the truncation point in the gonotrophic cycle count of the mosquito and then to mathematically study and assess the benefits that considering the gonotrophic cycles bring into the model. The postulated benefits include:

(i) The ability to quantify the reproductive gains that accrue to the mosquito population because of its interactions with the human. This reproductive gain is captured by requiring that a mosquito that successfully completes a gonotrophic cycle will lay eggs which will in turn contribute to the next generation of adult mosquitoes and thus eventually lead to the increase of the mosquito population through its normal developmental pathway. Thus for the mosquito, the advantage of going out to quest for and harvest a blood meal from humans outweighs the chance of death.
(ii) The ability to identify control points at different stages in the gonotrophic cycle chain where vector control measures can be applied. For example, targeting the breeding site, or the questing mosquitoes or the resting mosquitoes will reduce the number of mosquitoes available to eventually quest for blood in human populations or lay eggs for future adult mosquito populations. In fact, targeting and reducing the questing mosquito's population will be reducing chances of transmitting malaria infections.
(iii) The ability to implicitly include the extrinsic incubation period of malaria into a model that has the semblance of a susceptible-infectious model in the mosquito population. This is achieved in this paper by allowing only those mosquitoes that have completed at least two gonotrophic cycles from the time of first infection to be infectious to humans.
(iv) The ability to assess how each blood meal episode contributes to the basic offspring number of the mosquito insect as well as the reproduction number of the malaria disease.

It has been difficult, if not impossible to capture these listed benefits in previous mathematical models for malaria transmission which do not explicitly include the gonotrophic cycle. The final objective of this paper is to produce a model that yields what we may describe as an improved formula for the basic reproduction number for malaria.

The rest of the paper is organized as follows: In Sect. 2, we show a detailed derivation of the model we shall study in this paper. There, we describe the compartmentalization in the human and mosquito populations and define the state variables to be used. The flow rates in the model are explained and eventually the general mathematical model is then derived. The derived model is scaled and its properties examined. In Sect. 3, we present the mathematical analyses of both the disease-free and epidemiological models. There, it is shown that the infection-free model, which is a demographics model for mosquito populations, exhibits very rich

and diverse dynamics than the disease-free system in many mathematical models for malaria transmission. In that section, we compute the basic offspring number for the model as well as the epidemiological model's basic reproduction number. We round up the paper with discussion on the results of our paper in Sect. 4.

2 Derivation of the Model Equations

2.1 The Compartmentalization Adopted in the Human and Mosquito Populations

1. *Disease dynamics within the human population.* In the mathematical model for the dynamics of malaria transmission described in [20], the authors divide the human population into four compartments representing the disease status of the human as explained in Table 1. The human compartments are (1) the Susceptible humans S_h, (2) the Infected but not infectious humans (Exposed humans) E_h, (3) the Infected, infectious and clinically ill humans I_h and (4) the clinically recovered and partially immune but mildly infectious humans R_h. This compartmentalization allows for the possibility of an infected, infectious and clinically ill human to recover from clinical symptoms of the malaria infection but still retain some form of mild infectivity to the mosquito through the class of asymptomatic immune malaria carriers, R_h. The class of asymptomatic immune

Table 1 Types of human compartments and their description

Human type	Description
S_h	Density of susceptible humans at time t. These are humans who are not yet infected with the malaria parasite but can become infected if bitten by an infectious mosquito
E_h	Density of humans who have been infected by the malaria parasite but they are not yet infectious to mosquitoes and neither do they show any clinical symptoms of being infected. In this state, the parasite is either lying quietly in the liver or growing and multiplying in the red blood cells in the non-transmissible form. In a continuous developmental process it will take about 12 days for the parasite to develop and be in a state where it can be transmitted to a mosquito or we can see clinical symptoms of being infected in the human
I_h	Density of infected, infectious and clinically ill humans at time t. This class of humans can die of their infection if no treatment is taken. They are also highly infectious to questing mosquitoes
R_h	Density of recovered, partially immune and mildly infectious humans in the human population at time t. Members in this class are created when recovery from clinical symptoms from the I_h class is accompanied by the acquisition of partial immunity. Members from this class are reduced when they lose their immunity at a given rate to join the susceptible class. We will also refer to them as the asymptomatic immune class

carriers was identified as representing a substantial reservoir of infection in the human population. We retain this compartmentalization here as it represents a general form for an idealization of the different compartments that make up the disease dynamics within the human population. So a susceptible human who picks up the infection from an infectious bite, with a force of infection $g(S_h, \boldsymbol{I}_Q)$, from a questing female *Anopheles sp* mosquito drawn from the vector $\boldsymbol{I}_Q$ whose form will be described below, will become a human of type E_h and will go through an incubation period before he/she can become a clinically ill and infectious human of type I_h at the rate $\nu_h > 0$, where $\frac{1}{\nu_h}$ is approximately the length of the incubation period of the disease in the human. Afterwards, he/she can recover both from clinical symptoms and from the infection, at the rate r_h, to join the susceptible class or he/she can recover only from clinical symptoms with an acquisition of some form of immunity to further infection, while retaining mild infectiousness, to join the partially immune class, of type R_h at the rate σ_h. Individuals in the partially immune class lose immunity at rate. Death from natural causes occurs, at rate μ_h, in each human compartment and natural births, at rate λ_h, are also allowed to occur in each compartment. Vertical transmission in humans is not allowed so that all newborn humans are susceptible. Additional deaths due to disease can also be factored into the analysis by allowing some of the clinically ill and infectious humans the possibility of dying from their infection, at rate γ_h. The total human population at any time t, N_h, is then the sum over all the compartments; $N_h = S_h + E_h + I_h + R_h$. The description of all the compartments is shown in Table 1, while the general flow chart illustrating the flow of infection within the human population is shown in Fig. 1. With the above description, the equations that model the disease dynamics within the human population take the form

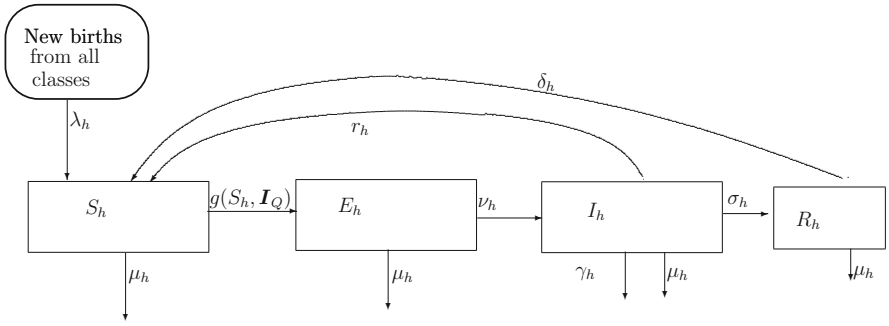


Fig. 1 Figure showing the flow of the infection within the human population. The force of infection is denoted by $g(S_h, \boldsymbol{I}_Q)$ where $\boldsymbol{I}_Q$ is a vector containing the reservoir of infected mosquito classes as explained in the text. Natural deaths occur in all classes at rate μ_h and births in all classes that enter into the susceptible class occur at the same rate of λ_h per human

$$\frac{dS_h}{dt} = \lambda_h N_h + r_h I_h + \delta_h R_h - g(S_h, \boldsymbol{I}_Q) - \mu_h S_h; \tag{1}$$

$$\frac{dE_h}{dt} = g(S_h, \boldsymbol{I}_Q) - (\nu_h + \mu_h) E_h; \tag{2}$$

$$\frac{dI_h}{dt} = \nu_h E_h - (r_h + \sigma_h + \mu_h + \gamma_h) I_h; \tag{3}$$

$$\frac{dR_h}{dt} = \sigma_h I_h - (\delta_h + \mu_h) R_h; \tag{4}$$

with appropriate initial conditions at time $t = 0$. The model represented by Eqs. (1)–(4) differs from the model in Ngwa and Shu [20] only in the form of the force of infection, $g(S_h, \boldsymbol{I}_Q)$. The nature of the vector $\boldsymbol{I}_Q$ is discussed in Sect. 2.2.

2. *The mosquito's population dynamics based on its gonotrophic cycle.* In the mosquito's population, we shall approach the disease compartmentalization of susceptible, exposed (infected) and infectious in an indirect route that passes via a physiological compartmentalization of the mosquito's population. That is, we shall base the compartmentalization on the fact that the mosquito undergoes a reproductive cycle called the *gonotrophic cycle*, so that at any one time each adult female mosquito's physiological state (well fed with a sugar meal, well nourished with a blood meal, rested after blood feeding, oviposited, nulliparous, mated/fertilized, etc.) on this cycle can be characterized. In what follows, we consider only female adult *Anopheles sp* mosquitoes since the males only survive on nectar and so, apart from the fact that they help in fertilizing the females, they do not pose any immediate threat to humans. The very nature of the gonotrophic cycle requires that a newly emerged adult female mosquito gets fertilized, searches and takes a blood meal, searches for a resting place and rests, and then searches for a breeding site where she lays eggs to complete the first gonotrophic cycle. Given that she stores the spermatozoa in a special pouch called a spermatheca, she does not need to mate again before the second and subsequent egg laying episodes, [31, 32]. Thus, it is assumed that all subsequent gonotrophic cycles starting from the second constitute only *three main steps* or locations namely: (i) being at the breeding site (oviposition site), (ii) being at human habitats site as questing mosquitoes and (iii) resting for egg maturation after blood feeding. We can use directed arrows to represent the flow of the mosquitoes as follows: from breeding site $\longrightarrow$ human habitat site for questing for blood meal $\longrightarrow$ Resting for egg development $\longrightarrow$ back to breeding site to lay eggs $\longrightarrow$ human habitat site for questing for blood $\cdots$. So every egg laying episode is preceded by blood feeding and resting episodes and the cycle is repeated until the mosquito dies of old age if she is not killed at any of the locations. The important point about these gonotrophic cycles is that every successfully completed cycle culminates with the laying of eggs that mature through the mosquito's metamorphic pathway to eventually increase the population density of the adult mosquitoes. An increase in the adult mosquito's population

density means the availability of more biting mosquitoes that can facilitate the transmission of disease where malaria disease (or any other mosquito borne disease of humans) transmission is also possible. We have captured here a natural process whereby, in the presence of malaria disease, a successful mosquito–human interaction may not only lead to transfer of the infection, but also means an increase in the number of mosquitoes to subsequently take part in the disease transmission process. This is why we have referred to the system studied in this paper as the *mosquito–human–malaria interactive framework*. To complete the characterization of the framework, we next describe how we can use the length of the gonotrophic cycle as a timer to approximate the chronological age of the mosquito and also identify mosquitoes that were infected early in their adult life and which will most likely be the candidate infectious mosquitoes to humans if they survive subsequent gonotrophic cycles.

3. *Embedding the mosquito's physiological age/disease status in its gonotrophic cycle counter.* The density of breeding site mosquitoes is denoted as type B mosquitoes, that of questing mosquitoes as type Q and that of resting mosquitoes as type R as explained in Table 2, and the general framework through which the different classes of mosquitoes relate with each other is shown in Fig. 2. In addition to the identification of the mosquitoes into the three broad types of breeding site, questing and resting mosquitoes, as described in Table 2, we subdivide each type into yet smaller classes indicating disease status, as well as into distinct age stages based on the number of gonotrophic cycles that each mosquito would have had. For example, for a given $y \in \{B, Q, R\}$, we write S_{y_k}, $k \geq 1$, to denote a susceptible mosquito of type y at reproductive stage k, and $I_{y_{k,j}}$, $k \geq j \geq 1$, to denote an infected mosquito at reproductive stage k that first picked up the infection at reproductive stage j. By extension, all the parameters of the system also have subscripted notation to capture its association with the particular reproductive stage mosquito. For example, ρ_k is the flow rate to the breeding site of susceptible rested mosquitoes at reproductive stage k, S_{R_k}, while that of $I_{R_{k,j}}$ would be $\rho_{k,j}$. We assume that mosquitoes of all types with a higher gonotrophic cycle counter are older than mosquitoes of the all types but at lower gonotrophic cycle counter. That is, at each cycle $k \geq 2$, mosquitoes of type B_k are older than mosquitoes of type R_{k-1}, while for each $k \geq 1$, we assume that mosquitoes of type B_k are younger than mosquitoes of type R_k. On

Table 2 Types of mosquitoes and their description

Mosquito type	Description
Q	Questing mosquitoes found at the human habitat site at time t
B	Breeding site mosquitoes at time t. These constitute all newly emerged adult mosquitoes together with any that have just returned to the breeding site to lay eggs
R	Resting mosquitoes at time t. These constitute all adult mosquitoes that have successfully acquired a blood meal and are now resting in view of returning to the breeding site to lay eggs after their resting period

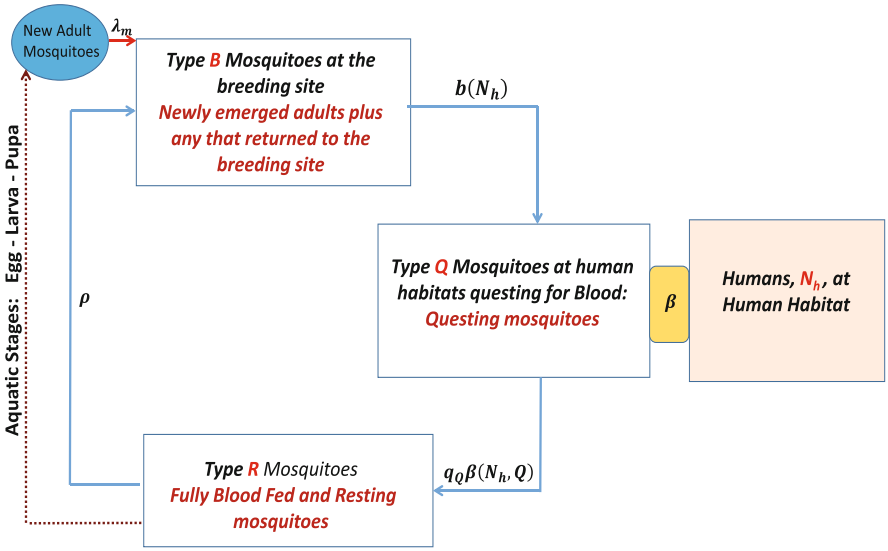


Fig. 2 Idealization of mosquito's movement at each gonotrophic cycle. Breeding site mosquitoes are attracted to human habitats at rate $b(N_h)$, where they become questing mosquitoes of type Q. Mosquitoes of type Q interact with humans N_h. Upon successful acquisition of a blood meal through the interactive exposure rate $\beta(N_h, Q)$, the mosquitoes of type Q become resting mosquitoes of type R. After the requisite resting period, the mosquitoes of type R that survive migrate again to the breeding site at rate ρ where they lay eggs that eventually contribute to the new adult mosquito population through the new adult mosquitoes' compartment. Upon successful arrival at the breeding site to lay eggs, the returned mosquitoes become breeding site mosquitoes of type B but with a higher gonotrophic cycle counter. q_Q is the probability of successfully completing a blood feeding episode to move into the resiting and egg maturation phase

the overall scale, we assume that mosquitoes of type B_1 are the youngest while mosquitoes of type R_N, $N > 1$ are the oldest. The gonotrophic cycle counter is thus used as a way to measure the adult mosquito's chronological age. This is the same compartmentalization as used in [22], but instead of ending the gonotrophic cycle count at 3, we generalize and assume that each adult female mosquito will undergo up to N reproductive cycles represented at each k, for $k = 1, 2, \cdots, N$, by the idealization on Fig. 2. Note that N here will be determined by how long a mosquito lives and how often it feeds and lays eggs, during its lifetime. In [23], it is argued that in the wild, based on a conceptualization of days and activities in the adult mosquito's life for the possible number of times the mosquito lays eggs during its reproductive life, the number N can be as large as $N = 4$ for a mosquito that lives for 24 days. The detailed definition of each of the compartments is shown in Table 3. The parameters of the system that we shall derive are shown in Table 4. In Fig. 3 we display the flow chart for a mosquito–human interactive system where each mosquito can undergo up to N complete gonotrophic cycles. That is, for a disease-free system in which we have mosquitoes interacting with humans and reproducing through the mechanisms

Table 3 The compartmentalization of the mosquito vectors as a function of reproductive stages k and j, for $k = 1, 2, 3, \cdots, N$ and $j = 1, 2, 3, \cdots, N$, as well as according to disease status: infected, I, or non-infected S. Only mosquitoes of type Q can interact with humans and so the disease can be transmitted from a human to a mosquito and vice versa only through mosquitoes of type Q. Furthermore, only infected mosquitoes of type $I_{Q_{k,j}}$ with $k - j \geq 2$ can be infectious to humans

Mosquito type	Description
S_{Q_k}	Density of Susceptible questing mosquitoes at reproductive stage k at the human habitat site at time t. This class of mosquitoes can be infected by humans after a successful blood feeding episode
$I_{Q_{k,j}}$	Density of infected questing mosquitoes at reproductive stage k that were first infected with the malaria parasite at reproductive stage j. This class of mosquitoes can be infectious to humans only if $k - j \geq 2$
S_{R_k}	Density of susceptible resting mosquitoes at reproductive stage k at time t. These were susceptible mosquitoes at reproductive stage k-1 that succeeded in their blood feeding quest without picking up the parasite and arrived at the breeding site to breed becoming susceptible breeding site mosquitoes of a higher reproductive stage
$I_{R_{k,j}}$	Density at time t of infected resting mosquitoes at reproductive stage k that were first infected with the malaria parasite at reproductive stage j. If they successfully arrive at the breeding site to breed, they become infected breeding site mosquitoes at reproductive stage $k + 1$
S_{B_k}	Density of susceptible breeding site mosquitoes at reproductive stage k at time t. These are vectors that will leave the breeding site to human habitats in search of a blood meal
$I_{B_{k,j}}$	Density of breeding site infected mosquitoes at reproductive stage k that were first infected with the malaria parasite at reproductive stage j

explained above. This is the conceptualization of the system studied in [23]. Counting the number of compartments for this disease-free system gives $3N$ mosquito compartments. In the presence of malaria, the number of compartments in the mosquito populations grows considerably and so it will be important to carefully examine how each of these compartments come about and how they contribute to the general mosquito–human–malaria interactive framework.

From the compartmentalization adopted, only mosquitoes of type Q interact with humans so that the infection can pass from humans to mosquitoes only through the interaction of type S_Q mosquitoes interacting with either infectious humans of type (I_h) or the mildly infectious recovered and partially immune humans of type (R_h). Figure 4 shows the full flow of the infection within the mosquito population for a mosquito–human–malaria interactive framework where each mosquito can undergo up to a maximum of three gonotrophic cycles. The figure clearly indicates the points where the infection can be passed from the human population into the mosquito population. The first column of compartments, comprising only the S_* variables constitute the material shown in Fig. 3. New infections can pass from mosquitoes to humans only through an interaction between a susceptible human of type (S_h) with an infected questing mosquito that is infectious. That is, mosquitoes of type $I_{Q_{k,j}}$,

Table 4 The parameters of the system and their quasi-dimension. Time in days (T), Bites (b), Vectors (V). All parameters are non-negative. Quasi-dimension of a parameter is its unit of measurement

Parameter	Description	Quasi-dimension
L_k	Carrying capacity (see Remark 2(e)) of vectors of type S_{R_k}	V
$L_{k,j}$	Carrying capacity of vectors of type $I_{R_{k,j}}$	V
λ_k	Function measuring the rate of oviposition by vectors of type S_{R_k}. Here, λ_{0_k} is the limiting rate of oviposition when the population size of resting mosquitoes is small compared with the carrying capacity of the environment L_k at reproductive stage k	T^{-1}
$\lambda_{k,j}$	Function measuring the rate of oviposition by vectors of type $I_{R_{k,j}}$. Here, $\lambda_{0_{k,j}}$ is the limiting rate of oviposition when the population size of resting mosquitoes is small compared with the carrying capacity of the environment $L_{k,j}$ at reproductive stage k	T^{-1}
ρ_k	Flow rate of reproducing and rested susceptible mosquitoes at reproductive stage k to the breeding site	T^{-1}
$\rho_{k,j}$	Flow rate of rested infected mosquitoes at reproductive stage k that where first infected at reproductive stage j	T^{-1}
a_k	Rate of flow of breeding site vectors from the breeding site. As noted in Sect. 2.2 item 1, a_k is weighted with the quantity $\frac{N_h}{(N_h+\kappa}$ to produce the parameter $b_k(N_h) = a_k \frac{N_h}{(N_h+\kappa}$ a_k is weighted to defined the rate b_k as explained in the text	T^{-1}
μ_*	Natural death rate of variable $*$. For example, μ_h is the natural death rate of the humans, and $\mu_{S_{Q_k}}$ is the natural death rate of susceptible questing mosquito at reproductive stage k and $\mu_{I_{R_{k,j}}}$ is the natural death rate of a resting infected vector at reproductive stage k that was first infected at stage j	T^{-1}
b_{Q_k}	Biting rate of questing mosquito at reproductive stage k	$Vb^{-1}T^{-1}$
q_{Q_k}	Probability of a successful blood meal upon an effective contact between questing mosquitoes at reproductive stage k and humans	1
$p_{Q_k h}$	Infectivity of mosquitoes to humans. $p_{Q_k h} \in [0, 1]$ is a measure of the chance that an infectious questing mosquito at reproductive stage k will transfer the infection to the human after a successful blood feeding encounter	bV^{-1}
p_{hQ_k}	Infectivity of infectious humans to questing mosquitoes. $p_{hQ_k} \in [0, 1]$ is a measure of the chance that a susceptible questing mosquito at reproductive stage k will pick up the infection from an infectious human after a successful blood feeding encounter	bV^{-1}
$\tilde{p}_{hQ_k}$	Infectivity of asymptomatic partially immune humans to questing mosquitoes. $\tilde{p}_{hQ_k} \in [0, 1]$ is a measure of the chance that a susceptible questing mosquito at reproductive stage k will pick up the infection from an asymptomatic partially immune human malaria carrier after a successful blood feeding encounter	bV^{-1}
λ_h	Natural birth rate for humans	T^{-1}
ν_h	Rate of onset of human infectiousness. $\frac{1}{\nu_h}$ is the incubation period	T^{-1}
δ_h	Rate of loss of acquired immunity in humans	T^{-1}
σ_h	Rate of acquisition of immunity by humans	T^{-1}
r_h	Rate of recovery in humans	T^{-1}
γ_h	Disease induced death rate in humans	T^{-1}

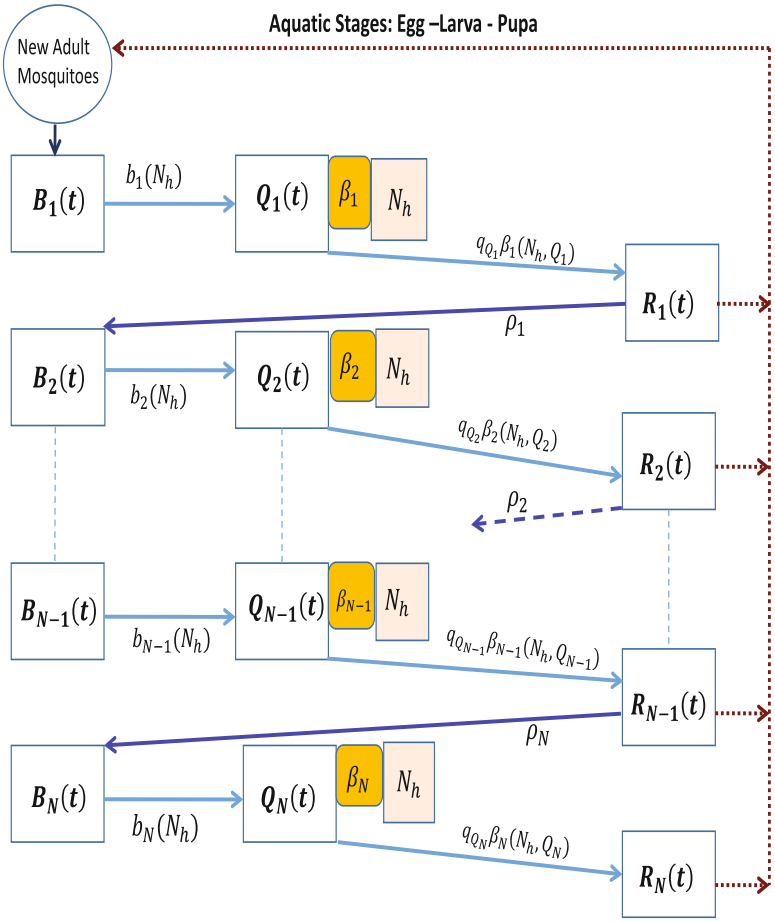


Fig. 3 Idealization of mosquito's movement according to gonotrophic cycles showing the mosquito only flows. Breeding site mosquitoes at reproductive stage k are attracted to human habitats at rate $b_k(N_h)$, where they become questing mosquitoes of type Q_k. Mosquitoes of type Q_k interact with humans N_h with exposure rate $\beta_k(N_h, Q_k)$. Upon successful acquisition of a blood meal through the interactive exposure rate $q_{Q_k}\beta_k(N_h, Q_k)$, the mosquitoes of type Q_k become resting mosquitoes of type R_k. After the requisite resting period, the mosquitoes of type R_k that survive migrate again to the breeding site at rate ρ_k where they lay eggs that eventually contribute to the new adult mosquito population through the new birth term. Upon successful arrival at the breeding site to lay eggs, the R_k mosquitoes become breeding site mosquitoes of type B_{k+1} and the cycle continues. The Mosquitoes of type R at reproductive stage N, R_N, lay their final batch of eggs and then die of old age and no longer enter the cycle

$k \geq j + 2$. The insistence that $k \geq j + 2$ ensures that we capture the minimum extrinsic incubation period; that is, the minimum time period required by an infected mosquito for the disease to mature to a level that the mosquito can now be infectious to humans. Here, we are assuming that each mosquito will spend approximately the

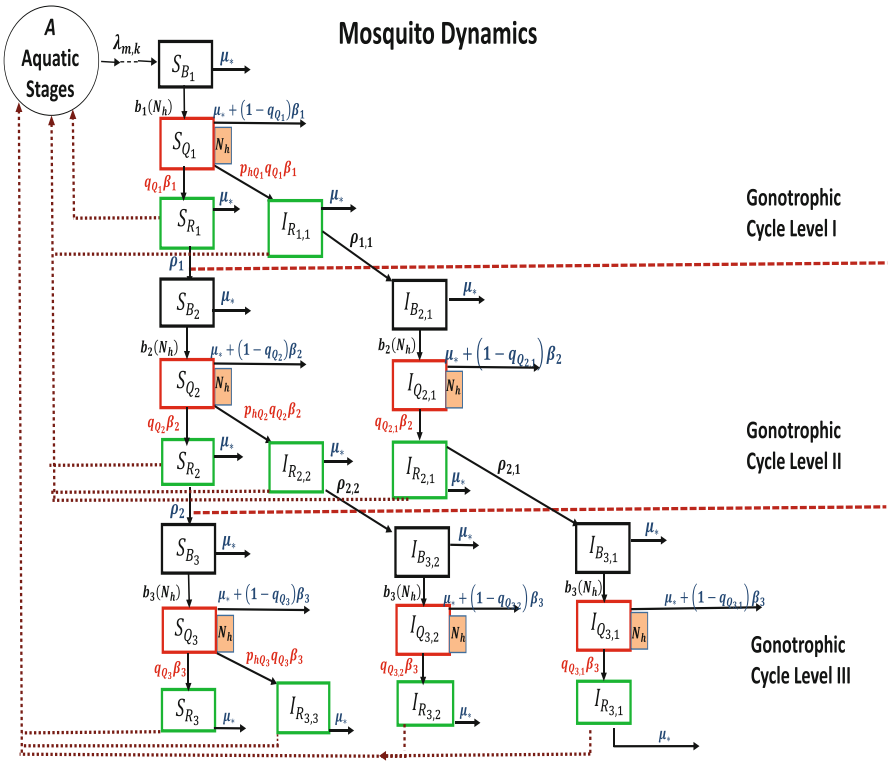


Fig. 4 The full mosquito–human–malaria interactive framework for $n = 3$ showing the flow of infection and movement of mosquitoes within the mosquito population. In the mosquito–human–malaria interactive framework used in this paper, only mosquitoes of type Q can interact with humans and infection can pass from human to mosquito only through the successful interaction between a susceptible mosquito of type Q, at reproduction stage k, S_{Q_k}, and an infectious human. This is the source of the new infected mosquito compartment $I_{R_{k,k}}$, $k = 1, 2, 3$ shown. Each new infected mosquito compartment starts a new branch of infected mosquitoes' path that eventually contribute to the source of infectious mosquitoes within the mosquito population as explained in the text. Contributions into the new adult mosquito pool after each blood meal episode by a fed and rested mosquito of type R are shown by the brown dotted lines. Additional infection of already infected mosquitoes is not considered

same length of time to complete each gonotrophic cycle. Although this can be an unrealistic assumption, given the perilous environment in which the mosquito must live and search for blood meals, we continue to use 2 as the minimum number of gonotrophic cycles whose cumulative time length is equivalent to the time lapse required for the infection to mature in the mosquito.[2]

[2]We cannot use $N = 1$, i.e. the length of one gonotrophic cycle to be equivalent to the length of the extrinsic incubation period for malaria in mosquitoes, as $N = 1$ may be too short.

The total adult mosquito population implicated in the model constitutes only those anthropophilic mosquitoes that consistently seek for blood meals within the human population. If N_v is the total mosquito population size, then $N_v = N_R + N_Q + N_B$ where N_R, N_B and N_Q are respectively the sizes of the total resting, breeding site and questing mosquitoes, given by

$$N_R = \sum_{k=1}^{N} S_{R_k} + \sum_{j=1}^{N}\sum_{k=j}^{N} I_{R_{k,j}},\ N_Q = \sum_{k=1}^{N} S_{Q_k} + \sum_{j=1}^{N-1}\sum_{k=j+1}^{N} I_{Q_{k,j}},$$
$$N_B = \sum_{k=1}^{N} S_{B_k} + \sum_{j=1}^{N-1}\sum_{k=j+1}^{N} I_{B_{k,j}}, \tag{5}$$

where S_{*_k} and $I_{*_{k,j}}$ for $* \in \{R, B, Q\}$, are as defined in Table 3. In this formulation, k tracks the reproductive stage of the adult mosquito, and hence the age, meanwhile j tracks the reproductive stage at which the adult mosquito was first infected with the malaria parasite.

Taking into consideration the full breadth of possible number of gonotrophic cycles and considering that disease dynamics adds to the complexity of the problem, we believe that it will be informative to have an idea of the size of the system. We already know that in the human population there are four compartments representing disease status with two of these compartments potentially infectious to mosquitoes. This gives us up to about a 50% chance of having an infectious human compartment in the modelling framework. We now work out the size of the system by calculating the total number of possible compartments in the mosquito population within the current framework. We state and prove the following result:

Lemma 1 (On the Size of the System) *Let there be given a mosquito–human–malaria dynamical system interactive framework whereby the mosquito can undergo up to N gonotrophic cycles during its entire reproductive life. Assume that at each gonotrophic cycle, there are three possible compartments: Q, B and R representing the three phases of breeding, questing and resting as conceived above. If in addition a mosquito in each of the compartments can be in any one of the disease states of infected (infectious) or susceptible, and if $\mathcal{M}(N)$ is the total number of mosquito compartments for the system, then $\mathcal{M}(N) = \frac{N}{2}(5 + 3N)$ where N is the maximum number of gonotrophic cycles possible for each adult mosquito.*

Proof At gonotrophic cycle one, we have the three starting susceptible compartments S_{B_1}, S_{Q_1} and S_{R_1}. At this cycle, a questing mosquito can become infected to give the additional infected resting compartment $I_{R_{1,1}}$. This gives a maximum number of 4 possible compartments at reproductive or gonotrophic cycle 1. The two possible R compartments from gonotrophic cycle 1 each give rise to three compartments at gonotrophic cycle 2 as follows: vectors from the susceptible compartment S_{R_1} migrate to the breeding site as susceptible vectors of type B at gonotrophic cycle 2, S_{B_2} which eventually follow the mosquito behavioural pattern

to produce S_{Q_2} and S_{R_2}. The infected resting mosquitoes, $I_{R_{1,1}}$ migrate to the breeding site to become infected breeding site mosquitoes at reproductive stage 2, $I_{B_{2,1}}$. These, also following the mosquito's behavioural pattern, produce $I_{Q_{2,1}}$ and $I_{R_{2,1}}$. The susceptible questing mosquitoes at reproductive stage 2 can become infected to produce infected resting mosquitoes at reproductive stage 2; $I_{R_{2,2}}$.

All mosquitoes that enter the next stage as infected mosquitoes do not add additional new infections into the mosquito sub-system. So we have a total of only 7 mosquito compartments at reproductive stage 2. Thus in general, each R compartment at the current level will produce three new compartments in the next level with a possibility of generating one additional new R compartment at that new level whenever new infections pass from the humans into the mosquito population. We thus have the sequence $(c_k)_{k\geq 1}$ where each $c_k = 3k + 1$ is the number of compartments at the k-th gonotrophic cycle. Thus, the total number of compartments after all N gonotrophic cycles is

$$\mathcal{M}(N) = \sum_{k=1}^{N} c_k = \sum_{k=1}^{N} (3k+1) = \frac{1}{2} N(5+3N); \tag{6}$$

as required. See illustration of compartment count in Fig. 5 for $N = 4$. □

2.2 The Exposure, Flow, Contact, Infectivity and Recruitment Rates

1. *The mosquito flow rate from breeding site to human habitat sites.* The flow rate from breeding site to human habitat is derived by considering the blood factor index for the mosquito and accounting for those mosquitoes that eventually choose to take a blood meal from the human population. If a_k is the rate of flow of vectors from breeding site to vertebrate habitat sites, then a_k is weighted by the quantity $\frac{N_h}{N_h+\kappa}$ where κ is a measure of the existence of an alternative blood source for the mosquito [22], so that $b_k(N_h) = a_k \frac{N_h}{N_h+\kappa}$ is a measure of the effective rate of flow of breeding site mosquitoes at reproductive stage k to the human habitats. It is understood that the fraction $1 - \frac{N_h}{N_h+\kappa}$ of a_k constitutes that fraction that searches for blood meals from non-human sources.
2. *The flow rate of resting/rested vectors to the breeding site.* After the acquisition of a blood meal, the adult female mosquito finds a suitable resting place where she rests for her eggs to mature. After a successful completion of the rest period, the rested mosquito that is ready to lay her eggs must migrate to the breeding site. Two states of vectors are considered and their flow rates differentiated accordingly: The flow rate to the breeding site of susceptible rested mosquitoes at reproductive stage k, denoted by ρ_k and the flow rate of rested infected adult mosquitoes at reproductive stage k, that were first infected (with malaria)

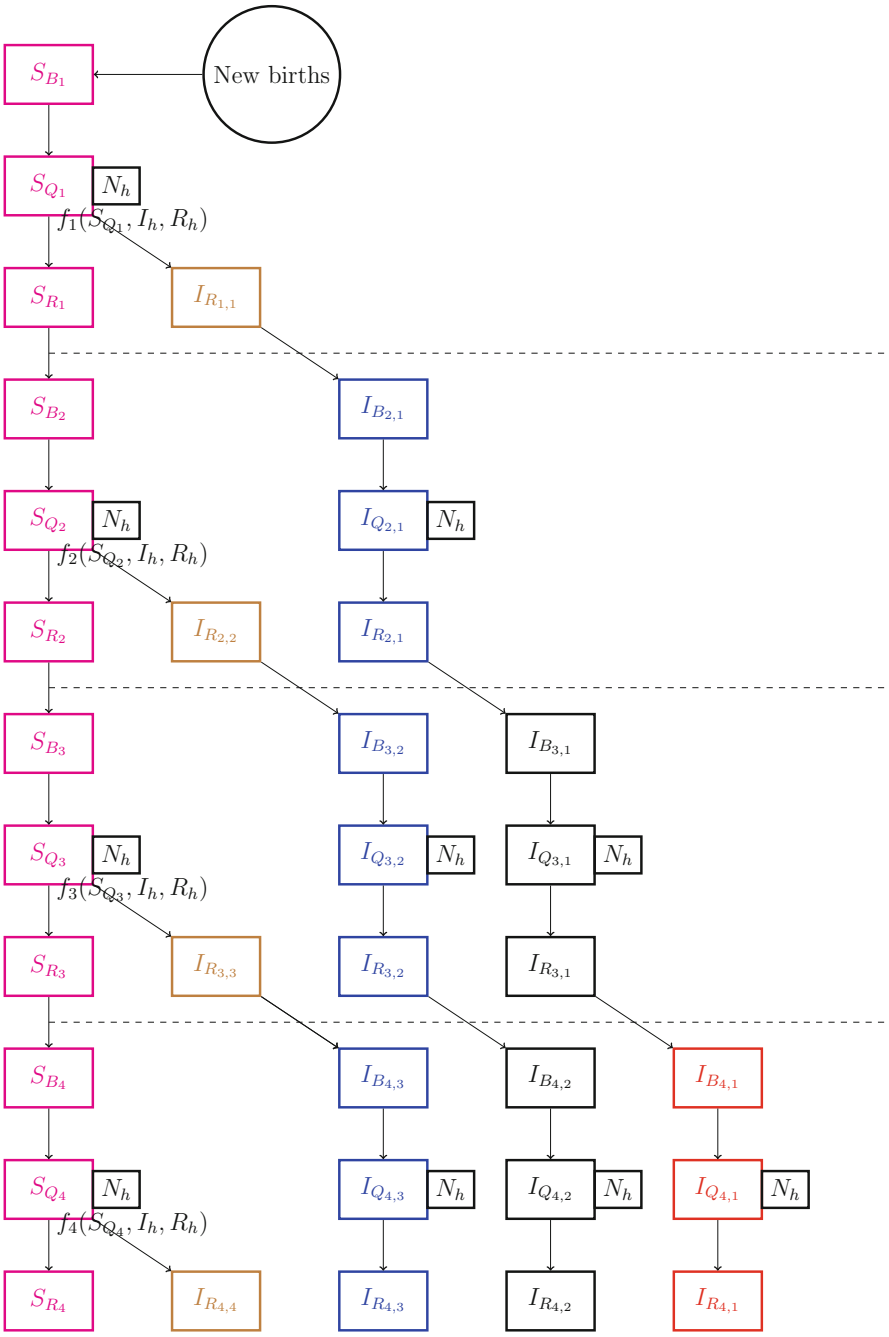


Fig. 5 Illustration of the gonotrophic cycle flow chart for $N = 4$. The cycles are demarcated with dashed lines. At gonotrophic cycles 1, 2, 3 and 4 there are respectively 4, 7, 10 and 13 possible mosquito compartments. So that at gonotrophic cycle n there will be $3n+1$ mosquito compartments

at reproductive stage j, $\rho_{k,j}$. A resting mosquito at reproductive stage k that successfully arrives at the breeding site to lay its eggs shall become a breeding site mosquito now at reproductive stage $k+1$.

3. *Exposure rate of humans to questing mosquitoes.* As in [22], the exposure rate of humans of a general type, say Y_h, $Y \in \{S, E, I, R\}$, when interacting with questing mosquitoes of type Q at reproductive stage k, represented generally as $X_Q \in \{S_{Q_k}, I_{Q_{k,j}}\}$, is denoted β_k and takes the form $\beta_k(Y_h, X_Q) = \frac{b_{Q_k} X_Q Y_h}{N_h}$ where b_{Q_k} is the human biting rate of the mosquito at reproductive stage k and N_h is the total human population available to the questing mosquitoes. To be specific, if X_Q is a susceptible questing mosquito at reproductive stage k, S_{Q_k}, then we have $\beta_k(Y_h, X_Q) = \beta_k(Y_h, S_{Q_k})$. If the questing mosquito is already infected, the infected mosquito is identified with its current reproductive stage and the reproductive stage where infection first took place through the double subscript notation so that $X_Q = I_{Q_{k,j}}$, a questing mosquito of type I at reproductive stage k, that was first infected at reproductive stage j and $\beta_k(Y_h, X_Q) = \beta_k(Y_h, I_{Q_{k,j}})$. This double subscript notation allows us to track the chronological age of the mosquito through its gonotrophic cycles counter as well as from when it first picked up the infection from the human. This way, the incubation period of the disease within the mosquito population is implicitly built into the modelling framework by assuming that the equivalent of at least two gonotrophic cycles episodes must elapse from time of first infection to the time when the questing mosquito is infectious to humans.
4. *Effective contact rates.* We consider two levels of having an effective contact between the humans and the mosquitoes: In the first instance, we consider an interactive contact that is effective in the sense that the contact leads to the transfer of infection from the human to the mosquito or from mosquito to human, and in the second instance, we consider a contact which is effective in the sense that the mosquito ingests enough blood to be able to satisfy its reproductive need. In the current formulation, we assume that when a mosquito engages in an interaction and fails to get the required blood meal, it is assumed killed in the process. In a more realistic setting, we would consider a case where a fraction of the mosquitoes that fail to get the required blood meal lives to try again as many times as it is required. We do not consider this in the present formulation, but only the following possibilities:

   (a) The questing mosquito at reproductive stage k that successfully takes a blood meal with probability q_{Q_k} shall become a resting mosquito of type R at reproductive stage k. So, it fails to take the blood meal upon trying with probability $1-q_{Q_k}$ which, in this case, is the probability of certain death. Two types of resting mosquitoes are identified in this work: susceptible resting mosquitoes at reproductive stage k whose density is denoted by S_{R_k} and infected mosquitoes at reproductive stage k that first picked up the infection at reproductive stage j, whose density is denoted by $I_{R_{k,j}}$.
   (b) A susceptible questing mosquito at reproductive stage k that successfully takes a blood meal with probability q_{Q_k} from an infectious human and

also succeeds in picking up an infection from human with probability p_{hQ_k} shall become an infected resting mosquito and the time counter to measure the length of the period of it being infected starts from the counter of the reproductive stage where it first got infected. So, for emphasis, we use the notation $I_{R_{k,j}}$ to denote the density of infected resting vectors at reproductive stage k that first picked up the malaria infection at reproductive stage j.

(c) An infectious questing mosquito at reproductive stage k that was first infected at reproductive stage j that successfully takes a blood meal from a susceptible human with probability q_{Q_k} and also transfers the infection to the human with probability p_{Q_kh} shall become a resting infected mosquito at reproductive stage k that was first infected at reproductive stage j.

(d) Where the mosquito or human is already infected, further infection is not considered. That is, we do not consider super-infection. However, every successful blood meal episode leads to oviposition by the mosquito which in turn leads to the eclosion of new adult mosquitoes that go to increase the total adult mosquito population. Where infection is present and not transferred, the probability of failure to transfer is $1 - p_{hQ_k}$ or $1 - p_{Q_kh}$, respectively.

5. *Infectivity of humans to mosquitoes.* The incubation period of the disease in humans, when caused by *Plasmodium falciparum*, has been estimated to be about 12 (9–14) days. This incubation period can be longer for other *Plasmodium* species of the parasite [27]. So, from the time of first infection with *Plasmodium falciparum* parasites, it takes about 12 days for the disease to develop in the human to the level where the human can become infectious to the mosquito. This fact is captured in the model by including a compartment of the human population wherein the humans are exposed to the infection but not yet infectious to mosquitoes. After the incubation period, the human can then become infectious where the rate of onset of infectiousness is inversely proportional to the residence time in the incubation phase. Infectious humans can recover without gain of immunity to join the susceptible class, or they can recover from clinical illness with a substantial gain of immunity to enter a partially immune class wherein members of that class are immune to clinical symptoms of malaria but are still mildly infectious to mosquitoes. This phenomenon of incomplete immunity permitting disease transmission has been known for some time now [2–4], and represents one of the main reasons why malaria eradication is difficult, among others; the reservoir of infection in the human population includes both symptomatic and asymptomatic immune carriers. To derive the expression for the infectivity of the human to the mosquito, we simply multiply the exposure rate of the mosquitoes to human, as derived above, with the probability of an effective contact between the questing susceptible mosquito at reproductive stage k, q_{Q_k} with the probability of the infectious human infecting the reproductive stage k questing mosquito, p_{hQ_k} if the human is from class I_h and $\tilde{p}_{hQ_K}$, if the human is from class R_h. Therefore, if $f_k(S_Q, I_h, R_h)$ is the force of infection for the stage k questing mosquitoes, that is the rate of change of new infections into reproductive stage k questing mosquitoes, then

$$f_k(S_{Q_k}, I_h, R_h) = p_{hQ_k} q_{Q_k} \beta_k(I_h, S_{Q_k}) + \tilde{p}_{hQ_k} q_{Q_k} \beta_k(R_h, S_{Q_k}). \quad (7)$$

These new infections will enter the $I_{R_{k,k}}$ compartment as a starting point of each new infection entering the mosquito population. Formula (7) captures the fact that asymptomatic immune malaria carriers can be infective to mosquitoes; however, we must expect that $p_{hQ_k} > \tilde{p}_{hQ_k}$ to capture the fact that infectivity of the type R_h humans is less than that of the I_h humans and that the reservoir of infection in the human population includes both symptomatic and asymptomatic immune malaria carriers.

6. *Infectivity of mosquitoes to humans.* The incubation period of the disease in mosquitoes can be as low as 10 days, [5, 27, 29] but can be made shorter in higher temperatures [5]. The length of the incubation period in mosquitoes can also be different for different species of the malaria parasite. So, from the moment the mosquito first picks up the infection we have to wait at least 10 days for the disease to mature in the mosquito before the mosquito can bring back the infection into the human population. During these 10 days, the reproductive cycle activities of blood feeding and egg laying continue as they will still continue after the 10 days incubation period. We implicitly model the incubation period of the disease in the mosquito population by requiring that an equivalent time length, measured by a cumulative time lapse equivalent to the length of at least two gonotrophic cycles, be completed by the mosquito before it can become infectious to the human. Thus, a questing mosquito at gonotrophic cycle k, that picked up the infection as a questing mosquito at gonotrophic cycle j, $I_{Q_{k,j}}$, is considered infectious to humans only if $k - j \geq 2$. Therefore only older mosquitoes that were infected much earlier in their gonotrophic cycle count, and which have undergone at least two gonotrophic cycle counts since first infection, shall be infectious to humans.

From the forgoing, we deduce that not all infected mosquito compartments contain infectious mosquitoes. So, it will be informative to work out the size of the *reservoir of infection (*ROI_V*)* in the mosquito population. Let $M < \mathcal{M}(N)$ be the number of compartments containing infectious questing mosquitoes. We seek to quantify the number of infectious questing mosquito compartments, $I_{Q_{k,j}}, 3 \leq k < N, 1 \leq j \leq N-2, k \geq j+2$. We can state and prove the following result on the actual size of $M(N)$ for a dynamical system where the total number of reproductive cycles possible is N.

Lemma 2 (On the Number of the Infectious Questing Mosquito Compartments) *Consider a system where the total number of gonotrophic cycles is N. Suppose that an infected mosquito at reproductive stage k that was first infected at reproductive stage j, $I_{Q_{k,j}}$, requires at least two gonotrophic cycles before it can become infectious to humans. If M is the number of infectious compartments in the vector population, then $M = \frac{1}{2}(N-1)(N-2)$.*

Proof If each infected mosquito requires at least two gonotrophic cycles from time of first infection to the onset of infectiousness, then a mosquito that picked up the infection in its first gonotrophic cycle will become infectious to humans

as from its third gonotrophic cycle. We thus have $I_{Q_{k,1}}$, for $k = 3, 4, 5, \cdots, N$, giving a total of $N - 2$ infectious questing compartments. Next, a mosquito that picked up the infection during its second gonotrophic cycle will become infectious to humans as from the 4th gonotrophic cycle. We thus have in this case $I_{Q_{k,2}}$, for $k = 4, 5, \cdots, N$, giving a total of $N - 3$ infectious questing compartments. Continuing in this way, we find that a mosquito that picked up the infection during its $N - 2$ gonotrophic cycle will become infectious at gonotrophic cycle N, giving only one infectious questing compartment, $I_{Q_{N,N-2}}$. Questing mosquitoes that pick up the infection either at the $N - 1$ or N gonotrophic cycles shall die before the maturation of the disease, and so this class of infected vectors will not play a part in the spread of the infection, although they will contribute to the size of the total mosquito population. Thus the total number of infectious questing compartments can be obtained by summing up the number of compartments associated with infectious questing and is

$$M(N) = \sum_{i=1}^{N-2} i = \frac{1}{2}(N-2)(N-1), \tag{8}$$

as required. □

Remark 1 (Generalization of Lemma 2) We can attempt a generalization of the result of Lemma 2 as follows: If n is the number of cycles that must elapse from time of first infection to time of onset of infectiousness of the mosquito, then from the biology of the Anopheles mosquito and the incubation period of the malaria infection in the mosquito, we deduce that $n \geq 2$, every thing being equal. Thus in this general case, M the number of infectious compartments would be given by $M = \sum_{i=1}^{N-n} i = \frac{1}{2}(N-n)(N-n+1),\ n \geq 2$. We must conclude, therefore, that M so calculated by Lemma 2 is the largest realistic integer that may be used as an indicator of the size of the number of infectious questing mosquito compartments in this framework.

The importance of Lemma 2 lies in the fact that we can combine its results with those of Lemma 1 to work out the probability of finding an infectious mosquito compartment in the entire mosquito sub-system, which result can be used to work out the chances of passing the infection from mosquito to human. In trying to understand the issue of chances of finding an infected mosquito in the system as well as the number of possible mosquito compartments of the system, we must settle two important parameters of the compartmental modelling process: (i) The number of compartments in the mosquito demographic framework and (ii) the length of the incubation period of the disease in a mosquito as captured by gonotrophic cycle count. Bearing these two facts it mind, we start by noting that the total number of mosquito compartments that eventually arise will be linked to the number of compartments originally conceived for the mosquito demographics model framework at each gonotrophic cycle. In the derivation of the result of Lemma 1, the demographic model has three compartments to capture

the breeding, questing and resting phases of the adult mosquito's reproductive life and shown in Fig. 2. It is possible to adjust the demographic model's compartmentalization to include other important steps in the adult mosquito's life such as searching for first sugar meal after adult eclosion, swarming, mating, first rest after mating, etc. This will drastically increase the number of compartments in the demographic model's compartmentalization framework. However, we do not think that the chances of finding an infected mosquito compartment in the system are equivalent to the probability of finding an infectious mosquito, unless the assumption can be made that the proportion of mosquitoes within each of the defined compartments are same or close to being the same. In the second instance, we also note that the number of compartments containing infected and infectious mosquitoes that shall be found in the system will be determined by a parameter equal to the number of gonotrophic cycles whose cumulative length of time is equivalent to the length of the incubation period of the disease in the mosquitoes. In the derivation of the results of Lemma 2, this number was set to 2, and is the number used in this manuscript, but was later generalized to $n \geq 2$ as in Remark 1. We can now state and prove the following result:

Lemma 3 (On the Probability of Finding an Infectious Mosquito Compartment) *Let there be given a mosquito–human–malaria dynamical system interactive framework in which the mosquito can undergo up to a maximum of N gonotrophic cycles during its entire reproductive life. Assume that to complete one gonotrophic cycle, the mosquito must pass through m distinct compartments where only one of these compartments represent interaction with humans and through which infection can pass into (or out of) the mosquito population. Let $\mathcal{M}_m(N)$ be the total number of compartments that this system can generate, and $M_n(N)$ the total number of infectious mosquito compartments in the system, where n is the number of cycles that must elapse from time of first infection to time of onset of infectiousness of the mosquito. Let $P_{m,n}(k)$ be the probability of finding an infectious compartment at gonotrophic cycle k. Then,*

$$P_{m,n}(k) = \begin{cases} 0, & 0 \leq k \leq n; \\ \frac{(k-n)(k-n+1)}{k(mk+m+2)}, & n \leq k \leq N, \end{cases} \quad \text{and} \quad \lim_{k\to\infty} P_{m,n}(k) = \frac{1}{m} \tag{9}$$

Proof Given that of the m compartments, infection can pass through only one of them, only one new infected mosquito compartment can be produced at each gonotrophic cycle level. Following the same argument as used in Lemma 1, we find that given m compartments at the start, at gonotrophic cycle k we have $c_k = mk + 1$ compartments and so the total for all cycles, $\mathcal{M}_m(N)$ is given by

$$\mathcal{M}_m(N) = \sum_{k=1}^{N} c_k = \sum_{k=1}^{N} (mk + 1) = \frac{1}{2} N(mN + m + 2). \tag{10}$$

From Remark 1 we have that if $M_n(k)$ is the number of infected and infectious compartments in the system at gonotrophic cycle k, where we require at least n gonotrophic cycles before infected vectors can become infectious, then

$$M_n(k) = \begin{cases} 0, & 0 \le k \le n; \\ \frac{1}{2}(k-n)(k-n+1), & n \le k \le N. \end{cases} \tag{11}$$

Therefore from standard probabilistic arguments we have that,

$$P_{m,n}(k) = \frac{M_n(k)}{\mathcal{M}_m(k)} = \begin{cases} 0, & 0 \le k \le n; \\ \frac{(k-n)(k-n+1)}{k(mk+m+2)}, & n \le k \le N, \end{cases} \quad \text{and} \quad \lim_{k\to\infty} P_{m,n}(k) = \frac{1}{m}$$

as required. □

So while the probability of finding an infected compartment increases with increasing number of gonotrophic cycles, the rate of increase is decreasing. So we should not expect to get more infectious compartments from the system just by allowing the mosquitoes to undergo more gonotrophic cycles. We conjecture that this is linked to the fact that the infection must mature in the mosquito before being available for transmission as well as to the fact that the *"bottleneck"* requiring that the infection must pass through the questing mosquito compartment limits the possibilities. In Fig. 3, we illustrate the behaviour of the probability of finding an infectious mosquito compartment in the system for different values of m and n. The shape of the graph is similar for various m and k values as described in formula (9).

By considering the limiting behaviour of the probability $P_{m,n}(k)$, we deduce that the chances of finding an infectious compartment in the human population are higher (up to 50%) than that of finding an infectious compartment in the mosquito population (less than 33% for the case $m = 3$ and 25% for the case $m = 4$). This information may be useful in determining the infectivity of the biting mosquito if the assumption is that the population sizes of mosquitoes in these compartments are about the same. Formula (9) shows that even though the length of the incubation period is important in determining when a compartment is infectious, this parameter's effect, once set, diminishes as the number of gonotrophic cycles increases. On the other hand, the parameter m is more important if we make the assumption that the population of mosquitoes within each mosquito compartments is about the same. However, if the population of mosquitoes in the infectious classes is smaller than the population size in the other compartments, then the importance of m is diminished. Perhaps we should not expect better results with more compartments in the demographic compartmentalization framework, but more on the distribution of the mosquitoes within these compartments. In what follows we continue to use the case $m = 3$ and $n = 2$ (Fig. 6).

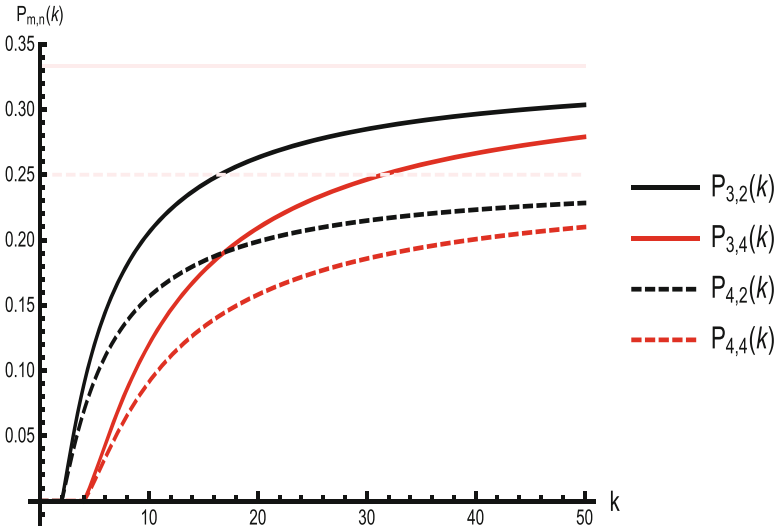


Fig. 6 The probability, $P_{m,n}(k)$, of finding an infectious mosquito compartment in the entire mosquito–human system is plotted as a function of the number of gonotrophic cycles reached for the cases $(m, n) = (3, 2)$ and $(3, 4)$, illustrated by the black and red solid curves, and also the cases $(m, n) = (4, 2)$ and $(4, 4)$, illustrated by the black and red dashed curves. The probability changes very rapidly in a narrow range of values of gonotrophic cycles. As the number of gonotrophic cycles increases further, the probability approaches $\frac{1}{3}$ for the cases when $m = 3$ and illustrated by the solid light pink line, and it approaches $\frac{1}{4}$ when $m = 4$ as illustrated by the dashed light pink line. If the number of compartments in the demographic model framework is increased, the chances of finding an infected mosquito compartment reduce. This result is only as important as the assumption that in the compartmentalization choices, the proportion of mosquitoes in the various compartments will be close or as close to being the same

The number of infectious mosquito compartments may be displayed in a vector $\boldsymbol{I}_Q \in \mathbb{R}^M$ where M is given by Eq. (8) and $\boldsymbol{I}_Q = (I_{Q_{k,j}}, 3 \leq k \leq N,\ 1 \leq j \leq N-2, k > j) \in \mathbb{R}^M$. We shall refer to $\boldsymbol{I}_Q$ as the *reservoir of infection vector* (ROI_V). We now use the entries of ROI_V to derive the force of infection in the human population. Let $g(S_h, \boldsymbol{I}_Q)$ be the force of infection in the human population. Then g is modelled by considering all those interactions between susceptible humans S_h and all the infected and infectious questing mosquitoes found in the mosquito's ROI_V. As mentioned above, employing the convention that each infected mosquito must go through at least two gonotrophic cycles before becoming infectious will implicitly build the incubation period of the disease into the mosquito population. The force of infection in the human population is therefore a sum over all those human-mosquito effective contacts, $\beta_k(S_h, I_{Q_{k,j}})$, that lead to acquisition of blood meal with probability q_{Q_k} together with the probability of transferring the infection to the human with probability $p_{Q_k h}$. Thus multiplying and summing up we have the expression

$$g(S_h, \boldsymbol{I}_Q) = \sum_{j=1}^{N-2} \sum_{k=j+2}^{N} p_{Q_k h} q_{Q_k} \beta_k(S_h, I_{Q_{k,j}}). \tag{12}$$

The force of infection so constructed takes into consideration all those infected and infectious anthropophilic mosquitoes that are participating actively in the dynamics.

7. *The rate of recruitment of new adult mosquitoes.* The actual existence of mosquitoes to continue to the next generations depends on the fact that mosquitoes of type R find suitable breeding sites to lay their eggs. It may be that a mosquito will choose a particular breeding site over another depending on several factors that could include the absence of predators, presence of other larvae at that breeding site or proximity from the resting place. Thus, the relationship between the mosquitoes of type R and the newly emerging adults cannot simply be assumed to be a linear response. This perhaps necessitates an assumption that the adult mosquito eclosion rate is density dependent. Some sources, for example [7], use a delay modelling argument, to derive a formula for the rate of emergence of new adults in a delayed differential equation framework. Others, see for example [1, 12], approach the problem of modelling the rate of new adult mosquito eclosion by including at least one (or more) state variable to represent the aquatic stages of the mosquito and then evoke the idea that the limitation of the carrying capacity of the aquatic pond will introduce competition within the aquatic stages of the mosquito's population as a source of nonlinearity and density dependence on the dynamics. Here, we simply assume that the net effect of the activities of the adult mosquitoes of type R is to contribute to the density of adult mosquito in the next generation through a birth term at a rate whose size is quantified by the birth rate function $\lambda_R : [0, \infty) \to \mathbb{R}$. The function λ_R, so described and fixed, in general, is assumed to depend in a nonlinear way on the size of the mosquitoes of type R that eventually survive the resting phase and then are in a position to lay eggs when they return to the breeding sites. Here, we assume that the form of the real valued function λ_R must satisfy desired properties, which among others, will guarantee the continued existence of a buoyant adult mosquito population so that the growth dynamics of the mosquitoes, in the absence of malaria infection, is internally stable from a mathematical and physical stand point. We write down the following definition:

Definition 1 (Recruitment Functions) For the sake of mathematical and biological realism, a function $\lambda_R : [0, \infty) \to \mathbb{R}$ is a suitable recruitment rate function if λ_R is smooth and in addition should satisfy the following:

(1) $\lambda_R(0_+) > 0$, where $\lambda_R(0_+) = \lim_{x \to 0^+} \lambda_R(x)$.
(2) $\lambda_R'(x)$ exists with $\lambda_R'(x) < 0, \quad \forall x \geq 0$.
(3) $\lim\limits_{x \to +\infty} \lambda_R(x) < \lim\limits_{x \to 0^+} \lambda_R(x)$.
(4) The function $x\lambda_R(x)$ is continuously differentiable, bounded above and unimodal so that there exists $\hat{x} > 0$ such that for $0 < x < \hat{x}$, $x\lambda_R(x)$ is

strictly monotone increasing and for $x > \hat{x}$, $x\lambda_R(x)$ is strictly monotone decreasing.

Remark 2 (Consequences of the Assumptions on λ_R) Each of the conditions considered in the above has consequences on the expectations of the behaviour of the function λ_R as follows:

(a) Condition (1) ensures that λ_R is non-negative *for small values* of its argument and represents the rate of production of new x, per x, per time so that the quantity $x\lambda_R(x)$ represents the net rate of production of new x per time.
(b) Condition (2) ensures that λ_R is a monotone decreasing function of its argument.
(c) Condition (4) ensures that $x\lambda_R(x)$ has a positive maximum value given by $\hat{x}\lambda_R(\hat{x})$, where $\hat{x} \in [0, \infty)$ satisfies the equation $\lambda_R(\hat{x}) + \hat{x}\lambda_R'(\hat{x}) = 0$.
(d) Condition (3) ensures that the equation $x'(t) = x(t)\lambda_R(x(t)) - \mu_v x(t)$, where $\mu_v > 0$ can be seen as a natural death rate parameter per x, and which represents a form of the equation for the dynamics of mosquitoes in the absence of infection, has a non-zero steady state solution x^* satisfying the equation $x^*\lambda_R(x^*) - \mu_v x^* = 0$ which is stable. Observe that such a non-zero steady state, x^*, will be found through the formula $x^* = \lambda_R^{-1}(\mu_v)$, which exists and is positive owing to the monotonicity of λ_R whenever $\lim_{x\to\infty} \lambda_R(x) < \mu_v < \lambda(0+)$.
(e) All conditions put together ensure the existence of a carrying capacity[3] L such that for $x < L$, $\frac{dx}{dt} > 0$ and thus the population $x(t)$ is increasing with time and for $x > L$, $\frac{dx}{dt} < 0$ and thus $x(t)$ is decreasing with time t. Examples of birth functions found in the biological literature that satisfy (1)-(4) may be found in Brännström and Sumpter [6].

In a general analysis, we may wish to investigate the effect and outcome of different birth functions on the dynamics of the system.

In the context of the generalized model presented in this paper, mosquitoes of type R can be in one of two states: Susceptible mosquitoes of type R at reproductive stage k; S_{R_k}, or infected mosquitoes of type R at reproductive stage k that were first infected at reproductive stage j; $I_{R_{k,j}}$. Each of these do contribute to the next generation of adult mosquitoes upon successful completion of resting period.[4] To differentiate the contributions, from each type R mosquito, to the next generation of the new adult mosquito's population, we write $\lambda_k(S_{R_k}) = \lambda_{S_{R_k}}(S_{R_k})$ and $\lambda_{k,j}(I_{R_{k,j}}) = \lambda_{I_{R_{k,j}}}(I_{R_{k,j}})$. Then, we set $\lambda_{0_k} = \lambda_k(0+) =$

[3]The carrying capacity of a biological species in an environment is the maximum population size of the species that the environment can sustain indefinitely, given the food, habitat, water, and other necessities available in the environment.

[4]Here, the subscript R in the definition of λ_R shall be replaced with either S_{R_k} or $I_{R_{k,j}}$ as the case maybe when we want to consider contributions from the different types of R mosquitoes into the function λ_R.

$\lim_{S_{R_k}\to 0}\lambda_{S_{R_k}}(S_{R_k})$ and $\lambda_{0_{k,j}} = \lambda_{k,j}(0+) = \lim_{I_{R_{k,j}}\to 0}\lambda_{k,j}(I_{R_{k,j}})$. To be specific, we use a simple density dependent form of the birth function that satisfies conditions (1)–(4) above and define, at each reproductive stage k the linear functions

$$\lambda_k(S_{R_k}) = \lambda_{0_k}\left(1 - \frac{S_{R_k}}{L_k}\right), \text{ and } \lambda_{k,j}(I_{R_{k,j}}) = \lambda_{0_{k,j}}\left(1 - \frac{I_{R_{k,j}}}{L_{k,j}}\right) \quad (13)$$

where $\lambda_{0_k} > 0$ and $\lambda_{0_{k,j}} > 0$ are constants measuring the limiting size of the oviposited egg cluster by rested mosquitoes at reproductive stage k when population numbers of those mosquitoes are small, and L_k and $L_{k,j}$ are parameters linked to the environmental carrying capacity[5] for stage k vectors of type R. The form of λ_k or $\lambda_{k,j}$ prescribed by Eq. (13) can go negative when $S_{R_k} > L_k$ or $I_{R_{k,j}} > L_{k,j}$ which will then show that for these values of S_{R_k} and $I_{R_{k,j}}$, the rate of oviposition is negative signifying declining population numbers. This behaviour actually represents a realistic mathematical idealization of the population growth process and so represents a good starting point in considering nonlinear dynamics for the mosquito's population. Other advantages in using these forms are that they are linear and could serve as the first linear approximation for any nonlinear function that satisfies conditions and assumptions of Definition 1. The main reason we continue to use the linear birth rate is because of mathematical tractability of the resulting equations based on this linear birth rate model. More nonlinear functions have been used in malaria modelling. See, for example, [17, 21].

The functional response of the resting mosquitoes in contributing to the general mosquito population size will be determined by the way in which we model inter-specific competition (if any) between the members of the rested/egg laying mosquitoes. Two formulations are possible:

(a) In the first instance we can assume that λ_k and $\lambda_{k,j}$ are a functions of the total size of the resting mosquitoes N_R where N_R is given by (5) so that we have the expression

$$\text{New adults} = \sum_{k=1}^{N}\rho_k\lambda_k(N_R)S_{R_k} + \sum_{j=1}^{N}\sum_{k=j}^{N}\rho_{k,j}\lambda_{k,j}(N_R)I_{R_{k,j}}. \quad (14)$$

[5]While λ_{0_k} and $\lambda_{0_{k,j}}$ may be different for different values of k and j because the size of the brood of eggs gets smaller with increasing number of gonotrophic cycles [27, page 68], it may be reasonable to assume that L_k and $L_{k,j}$ may be the same for all values of k and j since the carrying capacity of the environment is determined by external environmental factors such as temperature and precipitation, which once set will not change during the breeding season. However, we continue to differentiate these in what we write here for the sake of generalization adopted so far.

(b) In the second instance, we assume that each mosquito specifically lives a separate life style so that we can consider a rate of recruitment of new adult mosquitoes into the system to come from contributions from the independent classes of resting mosquitoes and write down an expression of the form

$$\text{New adults} = \sum_{k=1}^{N} \rho_k \lambda_k(S_{R_k}) S_{R_k} + \sum_{j=1}^{N} \sum_{k=j}^{N} \rho_{k,j} \lambda_{k,j}(I_{R_{k,j}}) I_{R_{k,j}}. \quad (15)$$

Note: if $L_k \to \infty, L_{k,j} \to \infty$ so that $\frac{S_{R_k}}{L_k} \to 0, \frac{I_{R_{k,j}}}{L_{k,j}} \to 0$ and $\lambda_k = \lambda_{k,j} = \lambda_{0_k} = \lambda_{0_{k,j}} = L$, the constant function, as in [22], Eqs. (14) and (15) will give the same results. However, it is reasonable to note that because of environmental variability and evolutionary stochasticity, the rate of oviposition of all the rested and egg laying mosquitoes need not be the same and use the simple linear function prescribed by (13). For the actual form of the expression for new adults we shall prefer the form (15) over the form (14) simply by advancing the argument that each adult mosquito lives an independent life and its rate of oviposition will not be determined by the availability and presence of other mosquitoes of that type; though it is fairly reasonable to assume that the rate of survival of offspring after egg laying will be determined by the environmental carrying capacity of the breeding site where the mosquitoes go to breed. So nonlinearity in the adult mosquito eclosion rate is captured by evoking the limitations imposed by the size of the breeding site through the form of formula (13). Therefore in what follows we model the rate of new adult mosquito eclosion by the expression

$$\text{New adults} = \sum_{k=1}^{N} \rho_k \lambda_{0_k} \left(1 - \frac{S_{R_k}}{L_k}\right) S_{R_k} + \sum_{j=1}^{N} \sum_{k=j}^{N} \rho_{k,j} \lambda_{0_{k,j}} \left(1 - \frac{I_{R_{k,j}}}{L_{k,j}}\right) I_{R_{k,j}}. \quad (16)$$

The terms in the right of formula (16) show the contributions from the different types of vectors: the susceptible rested vectors at reproductive stage k, S_{R_k}, and the rested vectors at reproductive stage k that were infected at reproductive stage j, $I_{R_{k,j}}$. We do not, in general, expect each of these types of vectors to contribute equally to the size of the next generation mosquitoes.

2.3 The Mathematical Equations

The form of the flow chart showing the flow in the mosquito dynamics is illustrated in Fig. 5, for $N = 4$. In that figure, the different gonotrophic cycle levels are clearly demarcated with the dashed lines and each lower level feeds into the higher

level while the points where human interactions with mosquitoes are possible are shown by the attached N_h-box. The number of mosquito compartments for each gonotrophic cycle level is determined by the number of possible susceptible and infected mosquito compartments and follows an arithmetic sequence whose first term is 4 and common difference 3. Thus, if c_k is the number of mosquito compartments at gonotrophic cycle k, then $c_k = 1+3k$. New adult mosquitoes enter the system through the S_{B_1} compartment as new births, and contributions to the new births' state come about as a result of eggs laid by type R mosquitoes according the formula shown in (14) or (15). Thus, only mosquitoes of type Q can interact with humans and only mosquitoes that have successfully interacted with humans can change status to mosquitoes of type R and eventually enter the next gonotrophic cycle. The arrows show the flows in and out of each compartment. Using standard rate of chemical reaction framework, we can write down the following equations:

$$\frac{dS_{B_1}}{dt} = \text{New adults} - \left(b_1(N_h) + \mu_{S_{B1}}\right) S_{B_1}; \tag{17}$$

$$\frac{dS_{B_k}}{dt} = \rho_{k-1} S_{R_{k-1}} - \left(b_k(N_h) + \mu_{S_{Bk}}\right) S_{B_k},\ k = 2, 3, \cdots, N; \tag{18}$$

$$\begin{aligned}\frac{dS_{Q_k}}{dt} &= b_k(N_h) S_{B_k} - \left(\beta_k(S_h, S_{Q_k}) + \beta_k(E_h, S_{Q_k}) + \beta_k(I_h, S_{Q_k}) + \beta_k(R_h, S_{Q_k})\right) \\ &\quad -\mu_{S_{Qk}} S_{Q_k},\ k = 1, 2, 3, \cdots, N;\end{aligned} \tag{19}$$

$$\begin{aligned}\frac{dS_{R_k}}{dt} &= q_{Q_k}\beta_k(S_h, S_{Q_k}) + q_{Q_k}\beta_k(E_h, S_{Q_k}) + (1 - p_{hQ_k}) q_{Q_k}\beta_k(I_h, S_{Q_k}) \\ &\quad +(1 - \tilde{p}_{hQ_k}) q_{Q_k}\beta_k(R_h, S_{Q_k}) - (\rho_k + \mu_{S_{Rk}}) S_{R_k},\ k = 1, 2, 3, \cdots, N;\end{aligned} \tag{20}$$

$$\frac{dI_{R_{k,k}}}{dt} = f_k(S_{Q_k}, I_h, R_h) - (\rho_{k,k} + \mu_{I_{R_{k,k}}}) I_{R_{k,k}},\ k = 1, 2, 3, \cdots, N; \tag{21}$$

$$\frac{dI_{B_{k,j}}}{dt} = \rho_{k-1,j} I_{R_{k-1,j}} - (b_k(N_h) + \mu_{I_{B_{k,j}}}) I_{B_{k,j}},\ k = 2, 3, \cdots, N,\ j = 1, 2, \cdots, N-1; \tag{22}$$

$$\begin{aligned}\frac{dI_{Q_{k,j}}}{dt} &= b_k(N_h) I_{B_{k,j}} - \left(\beta_k(S_h, I_{Q_{k,j}}) + \beta_k(E_h, I_{Q_{k,j}}) + \beta_k(I_h, I_{Q_{k,j}}) + \beta_k(R_h, I_{Q_{k,j}})\right) \\ &\quad -\mu_{I_{Q_{k,j}}} I_{Q_{k,j}},\ k = 2, 3, \cdots, N,\ j = 1, 2, \cdots, N-1;\end{aligned} \tag{23}$$

$$\begin{aligned}\frac{dI_{R_{k,j}}}{dt} &= \left(q_{Q_k}\beta_k(S_h, I_{Q_{k,j}}) + q_{Q_k}\beta_k(E_h, I_{Q_{k,j}}) + q_{Q_k}\beta_k(I_h, I_{Q_{k,j}}) + q_{Q_k}\beta_k(R_h, I_{Q_{k,j}})\right) \\ &\quad -\left(\rho_{k,j} + \mu_{I_{R_{k,j}}}\right) I_{R_{k,j}},\ k = 2, 3, \cdots, N,\ j = 1, 2, \cdots, N-1;\end{aligned} \tag{24}$$

$$\frac{dS_h}{dt} = \lambda_h N_h + r_h I_h + \delta_h R_h - g(S_h, \boldsymbol{I}_Q) - \mu_h S_h; \tag{25}$$

$$\frac{dE_h}{dt} = g(S_h, \boldsymbol{I}_Q) - (\nu_h + \mu_h) E_h; \tag{26}$$

$$\frac{dI_h}{dt} = \nu_h E_h - (r_h + \sigma_h + \mu_h + \gamma_h) I_h; \tag{27}$$

$$\frac{dR_h}{dt} = \sigma_h I_h - (\delta_h + \mu_h) R_h; \tag{28}$$

Where $f_k(S_{Q_k}, I_h, R_h)$ and $g(S_h, \boldsymbol{I}_Q)$ are given respectively by (7) and (12) and all the parameters are as described in Table 4. From the form of $\beta_k(X_h, Y_{Q_k}) = b_{Q_k}\frac{X_h Y_{Q_k}}{N_h}$ derived in Sect. 2.2 (item 3) earlier, the system of equations then takes the definite form

$$\frac{dS_{B_1}}{dt} = \text{New adults} - \left(b_1(N_h) + \mu_{S_{B1}}\right) S_{B_1}; \tag{29}$$

$$\frac{dS_{B_k}}{dt} = \rho_{k-1} S_{R_{k-1}} - \left(b_k(N_h) + \mu_{S_{Bk}}\right) S_{B_k},\ k = 2, 3, \cdots, N; \tag{30}$$

$$\frac{dS_{Q_k}}{dt} = b_k(N_h) S_{B_k} - \left(b_{Q_k} + \mu_{S_{Qk}}\right) S_{Q_k},\ k = 1, 2, 3, \cdots, N; \tag{31}$$

$$\frac{dS_{R_k}}{dt} = q_{Q_k} b_{Q_k} S_{Q_k} - f_k(S_{Q_k}, I_h, R_h) - (\rho_k + \mu_{S_{Rk}}) S_{R_k},\ k = 1, 2, 3, \cdots, N; \tag{32}$$

$$\frac{dI_{R_{k,k}}}{dt} = f_k(S_{Q_k}, I_h, R_h) - (\rho_{k,k} + \mu_{I_{R_{k,k}}}) I_{R_{k,k}},\ k = 1, 2, 3, \cdots, N; \tag{33}$$

$$\frac{dI_{B_{k,j}}}{dt} = \rho_{k-1,j} I_{R_{k-1,j}} - (b_k(N_h) + \mu_{I_{B_{k,j}}}) I_{B_{k,j}},\ k = 2, 3, \cdots, N,\ j = 1, 2, \cdots, N-1; \tag{34}$$

$$\frac{dI_{Q_{k,j}}}{dt} = b_k(N_h) I_{B_{k,j}} - (b_{Q_k} + \mu_{I_{Q_{k,j}}}) I_{Q_{k,j}},\ k = 2, 3, \cdots, N,\ j = 1, 2, \cdots, N-1; \tag{35}$$

$$\frac{dI_{R_{k,j}}}{dt} = q_{Q_k} b_{Q_k} I_{Q_{k,j}} - \left(\rho_{k,j} + \mu_{I_{R_{k,j}}}\right) I_{R_{k,j}},\ k = 2, 3, \cdots, N,\ j = 1, 2, \cdots, N-1; \tag{36}$$

$$\frac{dS_h}{dt} = \lambda_h N_h + r_h I_h + \delta_h R_h - g(S_h, \boldsymbol{I}_Q) - \mu_h S_h; \tag{37}$$

$$\frac{dE_h}{dt} = g(S_h, \boldsymbol{I}_Q) - (\nu_h + \mu_h) E_h; \tag{38}$$

$$\frac{dI_h}{dt} = \nu_h E_h - (r_h + \sigma_h + \mu_h + \gamma_h) I_h; \tag{39}$$

$$\frac{dR_h}{dt} = \sigma_h I_h - (\delta_h + \mu_h) R_h; \tag{40}$$

for a given set of initial conditions at time $t = 0$. An appropriate form of initial conditions would be those that start off the process with some initial density of the form

$$S_{B_k}(0) = S^0_{B_k},\ S_{R_k}(0) = S^0_{R_k},\ S_{Q_k}(0) = S^0_{Q_k} \tag{41}$$

$$I_{B_{k,j}}(0) = I^0_{B_{k,j}},\ I_{R_{k,j}}(0) = I^0_{R_{k,j}},\ I_{Q_{k,j}}(0) = I^0_{Q_{k,j}} \tag{42}$$

$$S_h(0) = S_h^0,\ E_h(0) = E_h^0,\ I_H(0) = I_H^0,\ R_h(0) = R_h^0, \tag{43}$$

where the variables with superscript 0 are typical variables at time $t = 0$ whose values will be provided as initial conditions. We should be careful to differentiate the epidemiologically realistic initial conditions with the system where there is no disease in the system. It is informative to note that the continuous dependence of the system on initial conditions means that if we start off this system with a set of initial conditions for which all the disease variables are set to zero, the system will continue to be disease-free for all subsequent time. An anticipated result in the analysis that we shall report in this paper shall be to find conditions in the full epidemiological model whereby starting the system with non-zero disease variables will lead to the eventual establishment of the infection in the population. This combined demographic and epidemiological model thus offers us a unique pathway for studying epidemiological and ecological parameters concurrently.

Though we have indicated the absence of a conservation argument to conserve the number of mosquitoes leaving the breeding site through the restriction of considering only anthropophilic mosquitoes, we can appreciate the size of the total active mosquito populations in the dynamics by adding up the relevant equations in the derived system. Studying the size of the total populations will give us an indication on the boundedness of the system under consideration, as well as provide a way of comparing the model with existing results in the literature. Recall that the total breeding site, questing and resiting mosquitoes are denoted respectively by N_B, N_Q and N_R and their size is calculated by computing the sum (5). Thus we have the following equations for the rate of change of the respective subtotals:

$$\begin{aligned}\frac{dN_B}{dt} &= \sum_{k=1}^{N} \frac{dS_{B_k}}{dt} + \sum_{j=1}^{N-1}\sum_{k=j+1}^{N} \frac{dI_{B_{k,j}}}{dt} \\ &= \text{New adults } + \sum_{k=2}^{N} \rho_{k-1} S_{R_{k-1}} + \sum_{j=1}^{N-1}\sum_{k=j+1}^{N} \rho_{k-1,j} I_{R_{k,j}} \\ &\quad - \sum_{k=1}^{N} (b_k(N_h) + \mu_{S_{Bk}}) S_{B_k} - \sum_{j=1}^{N-1}\sum_{K=j+1}^{N} (b_k(N_h) + \mu_{I_{B_{k,j}}}) I_{B_{k,j}}. \end{aligned} \tag{44}$$

Similarly, we calculate the rate of change of the total questing mosquito population as

$$\begin{aligned}\frac{dN_Q}{dt} &= \sum_{k=1}^{N} \frac{dS_{Q_k}}{dt} + \sum_{j=1}^{N-1}\sum_{k=j+1}^{N} \frac{dI_{Q_{k,j}}}{dt} \\ &= \sum_{k=1}^{N} b_k(N_h) S_{B_k} + \sum_{j=1}^{N-1}\sum_{k=j+1}^{N} b_k(N_h) I_{B_{k,j}}\end{aligned}$$

$$-\sum_{k=1}^{N}(b_{Q_k}+\mu_{S_{Q_k}})S_{Q_k}-\sum_{j=1}^{N-1}\sum_{K=j+1}^{N}(b_{Q_k}+\mu_{I_{Q_{k,j}}})I_{Q_{k,j}}. \quad (45)$$

and for the resting mosquitoes we have

$$\begin{aligned}\frac{dN_R}{dt} &= \sum_{k=1}^{N}\frac{dS_{R_k}}{dt}+\sum_{j=1}^{N}\sum_{k=j}^{N}\frac{dI_{R_{k,j}}}{dt}\\ &= \sum_{k=1}^{N}(b_{Q_k}q_{Q_k}S_{Q_k}-(\rho_k+\mu_{S_{R_k}})S_{R_k}-(\rho_{k,k}+\mu_{I_{R_{k,k}}})I_{R_{k,k}})\\ &\quad+\sum_{j=1}^{N-1}\sum_{k=j+1}^{N}(q_{Q_k}b_{Q_k}I_{Q_{k,j}}-(\rho_{k,j}+\mu_{I_{R_{k,j}}})I_{R_{k,j}}). \quad (46)\end{aligned}$$

The rate of change of the total human population, $N_h'(t)$, is found by adding up the relevant equations to have

$$\frac{dN_h}{dt}=(\lambda_h-\mu_h)N_h-\gamma_h I_h. \quad (47)$$

Equation (47) shows the dependence of the size of the population on disease related deaths. Now, if we set $\gamma_h = 0$ and $\lambda_h = \mu_h$, as in [22], the total human population will be constant. In this example, we would have a reduced system where our analysis will focus on understanding disease spread in a constant human and variable mosquito populations. Considering a constant total human population also allows us to reduce the dimension of our system by one as one of the states in the human compartment can be obtained once the other three states are known.

In general, however, from Eqs. (44)–(47), we can write down some bounds for the total population as follows:

1. *Bounds within the total human population:* For a realistic population demographics model, the functions $\lambda_h, \mu_h : [0,\infty) \to \mathbb{R}$ will have desired properties that ensure that in the absence of the disease we have a bounded non-zero human population as a basis for the modelling exercise. In [20], the natural birth rate in the human population, here λ_h, was assumed to be constant while the natural death rate in the human population, here μ_h was assumed to be a linear monotone increasing function of N_h. Here, we simply assume that λ_h is a non-zero monotone non-increasing function of its argument, while μ_h is a non-zero monotone non-decreasing function of its arguments. In fact, any nonlinear functional form for λ_h satisfying the conditions required by Assumption 1 will serve as a suitable natural birth rate function for the human population. All what we will require is that the form of the birth and death rates be such that the Eq. (47) has a bounded non-zero solution at all times. If we select the forms

$\lambda_h(N_h) = \lambda_1 - \lambda_2 N_h$ where $\lambda_1 > 0$ and $\lambda_2 \geq 0$, and $\mu_h(N_h) = \mu_1 + \mu_2 N_h$ where $\mu_1 > 0$ and $\mu_2 \geq 0$, then for a non-zero λ_2 and μ_2, Eq. (47) will experience exponential decay at all times if $\lambda_1 \leq \mu_1$, unbounded exponential growth whenever $\lambda_1 > \mu_1$ and $\mu_2 = \lambda_2 = \gamma_h = 0$, and bounded growth whenever $\max\{\lambda_2, \mu_2\} > 0$ and $\lambda_1 > \mu_1$. In any of the circumstances, we deduce that for the appropriate forms for the birth and death rates λ_h and μ_h, we will have the bound

$$\frac{dN_h}{dt} \leq (\lambda_h(N_h) - \mu_h(N_h))N_h, \tag{48}$$

as a bound for Eq. (47).

2. *Bounds within the mosquito populations:* For the variables within the mosquito population, we can write down some bounds as well by taking into considerations the definitions of the parameters of the system as shown in Table 4. Let,

$$\rho_v = \max_{1\leq k,j\leq N}\{\rho_k, \rho_{k,j}\}, \quad \mu_v = \max_{1\leq k,j\leq N}\{\mu_{S_{\varphi_k}}, \mu_{I_{\varphi_{k,j}}}\},$$
$$b_v = \max_{1\leq k\leq N}\{b_k(N_h)\}; \tag{49}$$
$$b_Q = \max_{1\leq k\leq N}\{b_{Q_k}\}, \quad p_Q = \max_{1\leq k\leq N}\{p_{Q_k}\}, \quad q_Q = \max_{1\leq k\leq N}\{q_{Q_k}\}. \tag{50}$$

Using these bounds in the equations for the total mosquito populations given by Eqs. (44)–(46), and using the definitions for N_R, N_Q and N_B defined by Eq. (5), we have the following bounds:

$$\begin{aligned}
\frac{dN_B}{dt} &\leq \text{New adults} + \rho_v N_R - (\mu_v + b_v)N_B;\\
\frac{dN_Q}{dt} &\leq b_v N_B - (\mu_v + b_Q)N_Q;\\
\frac{dN_R}{dt} &\leq b_Q q_Q N_Q - (\rho_v + \mu_v)N_R.
\end{aligned} \tag{51}$$

The solution of system (51) bounds the solutions of Eqs. (44)–(46). If $N_v = N_B + N_Q + N_R$ is the total active mosquito population, we have the bound

$$\frac{dN_v}{dt} \leq \text{New adults} - b_Q(1 - q_Q)N_Q - \mu_v N_v. \tag{52}$$

Inequality (52) shows the dependence of the total size of the active mosquito population on b_Q and q_Q and most importantly the size of the questing mosquitoes N_Q. If q_Q is close to zero, many of the questing mosquitoes die and affect the final size of the mosquito population. On the other hand, if q_Q is very near unity, many of the questing mosquitoes do not die during feeding. We have here a mechanism for controlling the total mosquito population. For acceptable forms of the mosquito

birth rate function λ_R satisfying conditions of Definition 1, the expression for new adults is bounded and we can then easily establish that the solutions of Eq. (52) are bounded. This in turn will show that the equations in system (51) are indeed bounded thus establishing the boundedness of the solutions of the derived system given by Eqs. (29)–(40). We make the following remark on the nature of the solutions of the bounding system:

Remark 3 The inequalities in system (51) are sharp in the sense that there exists a choice of parameters of the original system where we have equality. In the particular case where it is assumed that the respective death rates, the respective biting rates, the respective successful feeding probabilities across all gonotrophic cycles are equal, we will have equality in (51) and the system is equivalent to the system derived and studied in [19].

We also situate the types of solutions that are of interest to us in the following definition.

Definition 2 In line with the biological relevance, a solution of any differential equation involving a state variable of the system studied herein is called realistic if it is non-negative and bounded.

In the absence of infection the system reduces to the infection-free model whose mathematical equations are then obtained simply be setting all infected and infectious compartments to zero in system (29)–(40). The result is a demographic model for the dynamics of populations of anthropophilic mosquitoes that takes into consideration the blood feeding and reproductive cycles of the female mosquitoes. The infection-free model clearly shows the dependence of the dynamics of the populations of the mosquitoes on their ability to successfully acquire blood from humans, in this case susceptible humans. The infection-free model is given by the system

$$\frac{dS_{B_1}}{dt} = \sum_{k=1}^{N} \rho_k \lambda_{0,k} \left(1 - \frac{S_{R_k}}{L_k}\right) S_{R_k} - \left(b_1(N_h) + \mu_{S_{B1}}\right) S_{B_1}; \tag{53}$$

$$\frac{dS_{B_k}}{dt} = \rho_{k-1} S_{R_{k-1}} - \left(b_k(N_h) + \mu_{S_{Bk}}\right) S_{B_k},\ k = 2, 3, \cdots, N; \tag{54}$$

$$\frac{dS_{Q_k}}{dt} = b_k(N_h) S_{B_k} - \left(b_{Q_k} + \mu_{S_{Qk}}\right) S_{Q_k},\ k = 1, 2, 3, \cdots, N; \tag{55}$$

$$\frac{dS_{R_k}}{dt} = q_{Q_k} b_{Q_k} S_{Q_k} - (\rho_k + \mu_{S_{Rk}}) S_{R_k},\ k = 1, 2, 3, \cdots, N; \tag{56}$$

$$\frac{dS_h}{dt} = \lambda_h N_h - g(S_h, \mathbf{0}) - \mu_h S_h = (\lambda_h - \mu_h) S_h; \tag{57}$$

for a given set of initial conditions at time $t = 0$. Appropriate form for the initial conditions for the infection-free model would be those that start off the process with some initial density of the form

$$S_{B_1}(0) = S^0_{B_1}, S_{B_k}(0) = 0,\ S_{R_k}(0) = 0,\ S_{Q_k}(0) = 0, S_h(0) = S^0_h. \tag{58}$$

We note that in the infection-free model , $N_h = S_h$ leading to the simplification indicated for the S_h equation. Additionally, system (53)–(56) is a version of the reproductive stage structured model for the dynamics of malaria vector derived and studied in [23]. In [23], a mass action incidence function was used to model contacts between questing mosquitoes and humans which, if used here, would have yielded the inflow term $q_{Q_k} b_{Q_k} S_{Q_k} N_h$ in Eq. (56), as opposed to the inflow term $q_{Q_k} b_{Q_k} S_{Q_k}$ as shown, obtained as a result of the use of a standard incidence function to model contacts as in this manuscript. Though the notation is altered and the exposure rates are derived differently, the two systems are essentially identical in form and so the following results carry over.

Theorem 1 *The system* (53)–(56) *with* λ_R *given by* (13) *is well posed from a mathematical and physical stand point in the sense that a solution exists for each given set of initial conditions that is unique, non-negative and bounded.*

Proof See section 2.3 of [23]. □

Thus the system derived in this paper generalizes the systems studied earlier. We shall start the analysis by considering a scaling and non-dimensionalization.

2.4 Scaling and Non-dimensionalization

In the model derived above, the main physical dimension of the system is that of time. However, we have parameters and rates that are defined in terms of other parameters. In fact, a state variable or parameter that measures the number of individuals of certain type has a dimension-like quality (or quasi-dimensional unit) associated with it, [14]. To remove the dimension-like character on the parameters and variables, we make the following change of variables:

$$\tilde{S}_{B_k} = \frac{S_{B_k}}{S^0_{B_k}},\ \tilde{S}_{Q_k} = \frac{S_{Q_k}}{S^0_{Q_k}},\ \tilde{S}_{R_k} = \frac{S_{R_k}}{S^0_{R_k}},\ \tilde{I}_{B_{k,j}} = \frac{I_{B_{k,j}}}{I^0_{B_{k,j}}},$$
$$\tilde{I}_{Q_{k,j}} = \frac{I_{Q_{k,j}}}{I^0_{Q_{k,j}}},\ \tilde{I}_{R_{k,j}} = \frac{I_{R_{k,j}}}{I^0_{R_{k,j}}},\ \tau = \frac{t}{T^0}, \tag{59}$$

where the quantities with superscript zero are reference variables. Since we are considering a constant human population, we scale the human variables with N_h to have the system

$$\tilde{S}_h = \frac{S_h}{N_h},\ \tilde{E}_h = \frac{E_h}{N_h},\ \tilde{I}_h = \frac{I_h}{N_h},\ \tilde{R}_h = \frac{R_h}{N_h}, \tag{60}$$

so that $\tilde{S}_h + \tilde{E}_h + \tilde{I}_h + \tilde{R}_h = 1$ since $S_h + E_h + I_h + R_h = N_h$. From here, we then have $\tilde{E}_h = 1 - \tilde{S}_h - \tilde{I}_h - \tilde{R}_h$ and then set

$$
\begin{aligned}
S^0_{B_k} &= \frac{\rho_{k-1} S^0_{R_{k-1}}}{b_k(N_h)+\mu_{S_{B_k}}};\ k \geq 2,\ S^0_{Q_k} = \frac{b_k(N_h) S^0_{B_k}}{b_{Q_k}+\mu_{S_{Q_k}}};\ k \geq 1,\ S^0_{R_k} = \frac{q_{Q_k} b_{Q_k} S^0_{Q_k}}{\rho_k+\mu_{S_{R_k}}};\ k \geq 1 \\
I^0_{R_{k,k}} &= \frac{q_{Q_k} b_{Q_k} S^0_{Q_k}}{\rho_{k,k}+\mu_{I_{R_{k,k}}}};\ k \geq 1,\ I^0_{B_{k,j}} = \frac{\rho_{k-1,j} I^0_{R_{k-1,j}}}{b_k(N_h)+\mu_{I_{B_{k,j}}}};\ k \geq 2,\ j \geq 1, k > j, \qquad (61) \\
I^0_{Q_{k,j}} &= \frac{b_k(N_h) I^0_{B_{k,j}}}{b_{Q_k} + \mu_{I_{Q_{k,j}}}},\ I^0_{R_{k,j}} = \frac{q_{Q_k} b_{Q_k} I^0_{Q_{k,j}}}{\rho_{k,j} + \mu_{I_{R_{k,j}}}};\ k \geq 2,\ j \geq 1, k > j.
\end{aligned}
$$

The scaling of time will affect the time scales in the problem under consideration. In the current modelling problem under consideration, we have the following time scales:

1. *The incubation periods:* The length of the incubation period of the disease in the mosquito, also termed the *extrinsic incubation period*, under favourable conditions of the vector is dependent on ambient temperature and humidity. It has been reported that optimum conditions for sporogony are between 25°C and 30°C and it ceases below 16°C, and that, above 35°C, sporogony slows down considerably and it is also delayed by intermittent low temperatures, [27]. The actual length of the incubation period depends on the species of Plasmodium involved. The incubation period in humans is dependent on the general health and immune status of the person concerned and on the species of Plasmodium involved. The incubation periods are summarized in Table 5, which shows an average minimum incubation period of 12 days in humans and 10 days in mosquitoes. This time scale is short when compared with the life span of the human.
2. *The life span of the adult female Anopheles sp mosquito.* The average life expectancy of vectors of human malaria is 20–25 days and the average daily death rate is 4–5%, [27]. Taking into consideration the dangers that the mosquitoes go through in order to reproduce, it is normal to expect that many mosquitoes will die before completing their full life span; which for some species

Table 5 The lengths of the intrinsic and extrinsic incubation periods of malaria in humans and *Anopheles sp* mosquitoes for different species of *Plasmodium sp* parasites. Adapted from [27]

Type of plasmodium	Incubation days in humans (intrinsic mean)	Incubation days in mosquito (extrinsic mean) at 25°C
P. falciparum	9–14 (12)	10
P. malariae	18–40 (28) or longer	28
P. vivax	12–17 (15) or up to 6–12 months	10
P. ovale	16–18 (17) or longer	16

can go up to a month. However, in the model derived here where we have used the gonotrophic cycle count to measure the physical age of the mosquito at each time, we have established that mosquitoes at higher gonotrophic cycle counter are older than the ones at the start of the gonotrophic cycle counter. In [23], a procedure was developed for calculating the death rates of the mosquito at each gonotrophic cycle level. The formula captured the fact that mosquitoes of type R_N will be oldest adults in the system and thus will have the highest death rate while mosquitoes of type B_1 will be the youngest and as such will have the smallest death rate. We approximate death rate simply by calculating the reciprocal of remaining life days. We note, however, that the time frame representing the life span of any of the adult mosquitoes is short when compared with the life span of a human.

3. *The duration of each gonotrophic cycle.* The duration of the gonotrophic cycle is dependent on temperature and, in the tropics, at temperatures above 23°C, it usually lasts 2–4 days, but in the colder temperate climates it may take many days or even weeks. The time scale for this cycle is short when compared with the life span of the human. However, it has also been reported that a female *Anopheles sp* mosquito in the wild may eventually successfully complete about 4–5 gonotrophic cycles during its entire reproductive life. For the purpose of modelling, we can therefore estimate a time period of 5–6 days for each gonotrophic cycle to be completed. This length of time is also short, when compared with the life span of the human.

The above points show that there are several time scales in the problem under study and so in scaling time, we should be careful to capture the features modelled. Here we scale time with the death rate of the resting mosquitoes at reproductive stage N, R_N so as to be able to capture the gonotrophic cycles in our modelling. This will also justify our constant human population approximation in the sense that one time unit is then small enough to track the gonotrophic cycles as well as allow us to work with short enough time frames to warrant a constant human approximation. Thus we set

$$T^0 = \frac{1}{\rho_N + \mu_{S_{R_N}}}, \tag{62}$$

and then define the dimensionless parameter groupings

$$\begin{aligned}
\alpha_k &= \frac{\lambda_{0_k}\rho_k S^0_{R_k}}{S^0_{B_1}(b_1(N_h)+\mu_{S_{B_1}})},\ \alpha_{k,j}=\frac{\lambda_{0_{k,j}}\rho_{k,j} I^0_{R_{k,j}}}{S^0_{B_1}(b_1(N_h)+\mu_{S_{B_1}})},\ \tilde{L}_k=\frac{L_k}{S^0_{R_k}},\ L_{k,j}=\frac{L_k}{I^0_{R_{k,j}}}\\
\tilde{b}_k &= (b_k(N_h)+\mu_{S_{B_k}})T^0,\ \tilde{b}_{k,j}=(b_k(N_h)+\mu_{I_{B_{k,j}}})T^0,\ \tilde{\rho}_k=(\rho_k+\mu_{S_{R_k}})T^0,\\
\tilde{\rho}_{k,j} &= (\rho_{k,j}+\mu_{I_{R_{k,j}}})T^0,\ \tau_k=(b_{Q_k}+\mu_{S_{Q_k}})T^0,\ \tau_{k,j}=(b_{Q_k}+\mu_{I_{Q_{k,j}}})T^0\\
\tilde{\lambda} &= \lambda_h T^0,\ \tilde{\delta} = \delta_h T^0,\ \tilde{\sigma}=\sigma_h T^0,\ \tilde{\nu} = \nu_h T^0,
\end{aligned} \tag{63}$$

$$\tilde{r} = r_h T^0, \;\; g_{k,j} = \frac{p_{Q_k h} q_{Q_k} b_{Q_k} I^0_{Q_{k,j}} T^0}{N_h};$$

where $S^0_{B_1}$ is the initial size of the susceptible breeding site mosquito population, and the terms $S^0_{R_k}, I^0_{B_{k,j}}$ are as defined in Eq. (61). This then leads to the scaled system

$$\frac{d\tilde{S}_{B_1}}{d\tau} = \tilde{b}_1 \left(\sum_{k=1}^{N} \alpha_k \left(1 - \frac{\tilde{S}_{R_k}}{\tilde{L}_k} \right) \tilde{S}_{R_k} + \sum_{j=1}^{N} \sum_{k=j}^{N} \alpha_{k,j} \left(1 - \frac{\tilde{I}_{R_{k,j}}}{L_{k,j}} \right) \tilde{I}_{R_{k,j}} - \tilde{S}_{B_1} \right); \tag{64}$$

$$\frac{d\tilde{S}_{B_k}}{d\tau} = \tilde{b}_k \left(\tilde{S}_{R_{k-1}} - \tilde{S}_{B_k} \right), k = 2, 3, \cdots, N; \tag{65}$$

$$\frac{d\tilde{S}_{Q_k}}{d\tau} = \tilde{\tau}_k \left(\tilde{S}_{B_k} - \tilde{S}_{Q_k} \right), k = 1, 2, 3, \cdots, N; \tag{66}$$

$$\frac{d\tilde{S}_{R_k}}{d\tau} = \tilde{\rho}_k \left(\tilde{S}_{Q_k} - (p_{h Q_k} \tilde{I} + \tilde{p}_{h Q_k} \tilde{R}) \tilde{S}_{Q_k} - \tilde{S}_{R_k} \right), k = 1, 2, 3, \cdots, N; \tag{67}$$

$$\frac{d\tilde{I}_{R_{k,k}}}{d\tau} = \tilde{\rho}_{k,k} \left((p_{h Q_k} \tilde{I} + \tilde{p}_{h Q_k} \tilde{R}) \tilde{S}_{Q_k} - \tilde{I}_{R_{k,k}} \right), k = 1, 2, 3, \cdots, N; \tag{68}$$

$$\frac{d\tilde{I}_{B_{k,j}}}{d\tau} = \tilde{b}_{k,j} \left(\tilde{I}_{R_{k-1,j}} - \tilde{I}_{B_{k,j}} \right), k = 2, 3, \cdots, N-1, j = 1, 2, \cdots, N, k > j; \tag{69}$$

$$\frac{d\tilde{I}_{Q_{k,j}}}{d\tau} = \tilde{\tau}_{k,j} \left(\tilde{I}_{B_{k,j}} - \tilde{I}_{Q_{k,j}} \right), k = 2, 3, \cdots, N-1, j = 1, 2, \cdots, N, k > j; \tag{70}$$

$$\frac{d\tilde{I}_{R_{k,j}}}{d\tau} = \tilde{\rho}_{k,j} \left(\tilde{I}_{Q_{k,j}} - \tilde{I}_{R_{k,j}} \right), k = 2, 3, \cdots, N-1, j = 1, 2, \cdots, N, k > j; \tag{71}$$

$$\frac{d\tilde{S}_h}{d\tau} = \tilde{\lambda}(1 - \tilde{S}_h) + \tilde{r}\tilde{I}_h + \tilde{\delta}\tilde{R}_h - \sum_{j=1}^{N-2} \sum_{k=j+2}^{N} g_{k,j} \tilde{S}_h \tilde{I}_{Q_{k,j}} \tag{72}$$

$$\frac{d\tilde{I}_h}{d\tau} = \tilde{\nu}(1 - \tilde{S}_h - \tilde{I}_h - \tilde{R}_h) - (\tilde{\lambda} + \tilde{\sigma} + \tilde{r})\tilde{I}_h \tag{73}$$

$$\frac{d\tilde{R}_h}{d\tau} = \tilde{\sigma}\tilde{I}_h - (\tilde{\delta} + \tilde{\lambda})\tilde{R}_h. \tag{74}$$

3 Mathematical Analyses

We now analyse the model's equations and starting with the infection-free model. In what follows, for notational simplicity, we drop the tilde on each of the notations.

3.1 The Infection-Free Model: Existence and Stability of Steady States

As indicated above the infection-free model is obtained simply by setting the disease variables to zero in the scaled system (64)–(74). This gives the system

$$\frac{dS_{B_1}}{d\tau} = b_1\left(\sum_{k=1}^{N}\alpha_k\left(1-\frac{S_{R_k}}{L_k}\right)S_{R_k} - S_{B_1}\right); \tag{75}$$

$$\frac{dS_{B_k}}{d\tau} = b_k\left(S_{R_{k-1}} - S_{B_k}\right), k = 2, 3, \cdots, N; \tag{76}$$

$$\frac{dS_{Q_k}}{d\tau} = \tau_k\left(S_{B_k} - S_{Q_k}\right), k = 1, 2, 3, \cdots, N; \tag{77}$$

$$\frac{dS_{R_k}}{d\tau} = \rho_k\left(S_{Q_k} - S_{R_k}\right), k = 1, 2, 3, \cdots, N; \tag{78}$$

$$\frac{dS_h}{d\tau} = \lambda(1 - S_h), \quad 0 = \nu_h(1 - S_h). \tag{79}$$

The fact that we have a constant human population shows up very clearly in the last equation, Eq. (79).

Theorem 2 (On the Existence of Steady States for the Infection-Free Model) *The disease-free system (75)–(79) admits a trivial steady state solution* $\mathbf{0}$ *for the mosquito population with* $S_h = 1$*, which always exists for all parameter values of the system. In addition, there is a possibility for the existence of a non-trivial steady state solution, whose existence and size are determined by the size of a threshold parameter* $\mathcal{N}$ *in the sense that if* $\mathcal{N} \leq 1$*, the only steady state is the trivial steady state and when* $\mathcal{N} > 1$*, the non-trivial steady state exists and can be computed.*

Proof The steady states for the infection-free model, when they exist, are solutions of the equations when the time derivatives are set to zero. This gives the steady state solutions

$$S^*_{B_k} = S^*_{R_k} = S^*_{Q_k},\ k = 1, 2, 3, \cdots, N, \quad S^*_{B_1} = 0 \text{ or } S^*_{B_1} = \frac{\sum_{k=1}^{N}\alpha_k - 1}{\sum_{k=1}^{N}\left(\frac{\alpha_k}{L_k}\right)}, \tag{80}$$

with $S_h = 1$. Let $\boldsymbol{S_{B_k}} = (S_{B_1}, S_{B_2}, \cdots, S_{B_N})$, $\boldsymbol{S_{Q_k}} = (S_{Q_1}, S_{Q_2}, \cdots, S_{Q_N})$ and $\boldsymbol{S_{R_k}} = (S_{R_1}, S_{R_2}, \cdots, S_{R_N})$. Then, we can use this notation to write down the form of the two steady state solutions:

$$(\boldsymbol{S_{B_k}}, \boldsymbol{S_{Q_k}}, \boldsymbol{S_{R_k}}) = (\mathbf{0}, \mathbf{0}, \mathbf{0}); \quad \text{the trivial steady state,} \tag{81}$$

$$(\boldsymbol{S^{DFE}_{B_k}}, \boldsymbol{S^{DFE}_{Q_k}}, \boldsymbol{S^{DFE}_{R_k}}) = (\boldsymbol{S^{*}_{B_1}}, \boldsymbol{S^{*}_{B_1}}, \boldsymbol{S^{*}_{B_1}}); \quad \text{the persistence steady state,} \tag{82}$$

where $\boldsymbol{S^{DFE}_{B_k}} \in \mathbb{R}^N$, $\boldsymbol{S^{DFE}_{Q_k}} \in \mathbb{R}^N$ and $\boldsymbol{S^{DFE}_{R_k}} \in \mathbb{R}^N$ are N-tupples or vectors in $\mathbb{R}^N$ each of whose coordinates is exactly $S^*_{B_1}$ given by (80). Notice that in Eq. (80) the non-zero solution $S^*_{B_1}$ exists as a realistic steady state of the system in the sense of Definition 2 only if the quantity $\mathcal{N} > 1$ where

$$\mathcal{N} = \sum_{k=1}^{N} \alpha_k = \sum_{k=1}^{N} \left(\frac{\lambda_{0,k} \rho_k S^0_{R_k}}{S^0_{B_1}(b_1(N_h) + \mu_{S_{B_1}})} \right), \tag{83}$$

and the case $\mathcal{N} = 1$ reducing the non-zero solution to the zero solution. □

We note that $\mathcal{N}$ is a unique threshold parameter for the system. When $\mathcal{N} \leq 1$, the mosquito population goes extinct and only a contact thriving susceptible human population exists. The conditions for the existence of the persistence steady state solution, namely $\mathcal{N} > 1$, translate to the conditions for the establishment of a non-zero mosquito population in the environment whenever the persistence steady state is stable. We shall identify the quantity $\mathcal{N}$ so constructed with the basic offspring number.

Definition 3 (Basic Offspring Number) The basic offspring number is the number of new adult mosquitoes that arise from one adult female reproducing mosquito during its entire period of reproductivity in the absence of density dependence.

The basic offspring number is a threshold parameter analogous to the basic reproduction number in epidemiological modelling. It offers a control criterion for mosquito population as we expect that when this quantity is strictly less than unity, the trivial equilibrium solution where $S^*_{B_k} = 0$ is globally and asymptotically stable as the only steady state of the system. In fact we can state and prove the following result:

Theorem 3 (On the Global Stability of the Trivial Steady State) *The steady state solution given by* $\boldsymbol{S^*_{B_k}} = \boldsymbol{S^*_{R_k}} = \boldsymbol{S^*_{Q_k}} = \mathbf{0}$ *for all* $k = 1, 2, \cdots, N$, *with* $S_h = 1$, *which always exists for values of the parameters of the system, is globally and asymptotically stable whenever* $\mathcal{N} \leq 1$.

Proof We use the direct method of Lyapunov. We construct the Lyapunov function. Let $V : \mathbb{R}^{3N} \to \mathbb{R}$ be real valued linear function. For positive constants A_k, C_k and D_k, $k = 1, 2, 3 \cdots, N$, define V by

$$V(\boldsymbol{S_{B_k}}, \boldsymbol{S_{Q_k}}, \boldsymbol{S_{R_K}}) = \sum_{k=1}^{N} \left(A_k S_{B_k} + C_k S_{Q_k} + D_k S_{R_k}, \right) \tag{84}$$

then $V(\mathbf{0}, \mathbf{0}, \mathbf{0}) = 0$ and $V(\boldsymbol{S_{B_k}}, \boldsymbol{S_{Q_k}}, \boldsymbol{S_{R_K}}) > 0$ for non-zero values of arguments. We show that there exists a choice of constants A_k, C_k and D_k, $k = 1, 2, 3 \cdots, N$, for which the orbital derivative of V is always negative near the zero state. To do this, calculate the total derivative of V

$$\begin{aligned}
\frac{dV}{d\tau} &= \sum_{k=1}^{N} \left(A_k \frac{dS_{B_k}}{d\tau} + C_k \frac{dS_{Q_k}}{d\tau} + D_k \frac{dS_{R_k}}{d\tau} \right) \\
&= A_1 b_1 \left(\sum_{k=1}^{N} \alpha_k \left(1 - \frac{S_{R_k}}{L_k} \right) S_{R_k} - S_{B_1} \right) + \sum_{k=2}^{N} A_k b_k \left(S_{R_{k-1}} - S_{B_k} \right) \\
&\quad + \sum_{k=1}^{N} \left(C_k \tau_k \left(S_{B_k} - S_{Q_k} \right) + D_k \rho_k \left(S_{Q_k} - S_{R_k} \right) \right),
\end{aligned}$$

where the derivatives on the right hand side as given by (75)–(79) have been substituted. Rearranging the terms we have

$$\begin{aligned}
\frac{dV}{d\tau} &= \sum_{k=1}^{N-1} \left(b_1 A_1 \alpha_k + b_{k+1} A_{k+1} - \rho_k D_k \right) S_{R_k} + \left(b_1 A_1 \alpha_N - \rho_N D_N \right) S_{R_N} \\
&\quad + \sum_{k=1}^{N} \left(\left(C_k \tau_k - A_k b_k \right) S_{B_k} + \left(\rho_k D_k - C_k \tau_k \right) S_{Q_k} \right) - \sum_{k=1}^{N} \left(A_1 b_1 \frac{\alpha_k}{\tilde{L}_k} S_{R_k}^2 \right).
\end{aligned}$$

To proceed, choose the constants C_k, D_k and A_K, for $k = 1, 2, 3, \cdots, N$ so that

$$C_k \tau_k = A_k b_k = \rho_k D_k, \quad \rho_N D_N = b_1 A_1 \alpha_N. \tag{85}$$

Applying these, we have the relation

$$\frac{dV}{d\tau} = \sum_{k=1}^{N-1} \left(b_1 A_1 \alpha_k + b_{k+1} A_{k+1} - \rho_k D_k \right) S_{R_k} - \sum_{k=1}^{N} \left(A_1 b_1 \frac{\alpha_k}{\tilde{L}_k} S_{R_k}^2 \right). \tag{86}$$

Let us examine the coefficients of S_{R_k} in (86): When $k = N - 1$, we have that the coefficient of the variable $S_{R_{N-1}}$ is $b_1 A_1 \alpha_{N-1} + b_N A_N - \rho_{N-1} D_{N-1}$. Using the relation (85), we find that $b_N A_N = \rho_N D_N = b_1 A_1 \alpha_N$ so that the coefficient of $S_{R_{N-1}}$ is $b_1 A_1 \alpha_{N-1} + b_1 A_1 \alpha_N - \rho_{N-1} D_{N-1}$. So if we chose $b_1 A_1 (\alpha_{N-1} + \alpha_N) = \rho_{N-1} D_{N-1}$, this coefficient will be zero. Similarly, if $k = N - 2$, the coefficient of

$S_{R_{N-2}}$ in (86) is $b_1A_1\alpha_{N-2}+b_{N-1}A_{N-1}-\rho_{N-2}D_{N-2}$ which on using the relations in (85) and the value of $\rho_{N-1}D_{N-1}$ just computed, we have that the coefficient of $S_{R_{N-2}}$ in that expression is $b_1A_1(\alpha_{N-2}+\alpha_{N-1}+\alpha_N)-\rho_{N-2}D_{N-2}$. Again the choice $\rho_{N-2}D_{N-2}=b_1A_1(\alpha_{N-2}+\alpha_{N-1}+\alpha_N)$ will make the coefficient $S_{R_{N-2}}$ in the expression in (86), to also vanish. We continue like this and set up the recurrence relation

$$\rho_{N-i}D_{N-i}=b_1A_1\sum_{j=0}^{i}\alpha_{N-j},\quad i=0,1,2\cdots,N-2. \tag{87}$$

which in turn, knocks out the coefficients of $S_{R_k}, k=N-1, N-2, N-3, \cdots, 3, 2$. The coefficient of S_{R_2} is eventually eliminated by setting $i=N-2$ in (87) to have the final expression

$$\frac{dV}{d\tau}=\left(b_1A_1\sum_{k=1}^{N}\alpha_k-\rho_1D_1\right)S_{R_1}-\sum_{k=1}^{N}\left(A_1b_1\frac{\alpha_k}{\tilde{L}_k}S_{R_k}^2\right). \tag{88}$$

The choice of $\rho_1D_1=b_1A_1$ leads to the orbital derivative

$$\frac{dV}{d\tau}=b_1A_1(\mathcal{N}-1)S_{R_1}-\sum_{k=1}^{N}\left(A_1b_1\frac{\alpha_k}{\tilde{L}_k}S_{R_k}^2\right), \tag{89}$$

where $\mathcal{N}$ is given by (83) and A_1 is an arbitrary positive real number. It is now clear from (89) that $V'(\tau)<0, \forall\tau$ whenever $\mathcal{N}\leq 1$. When $\mathcal{N}\leq 1$, the Lyapunov–LaSalle Theorem, [10], assures us that all paths in the positive orthant $\mathbb{R}_+^{3N}\setminus\{\mathbf{0}\}$ approach the largest compact invariant subset $\{\mathbf{0}\}\subset\mathbb{R}_+^{3N}$ wherein $V'(\tau)=0$. Hence $S_{B_k}\to 0$, $S_{Q_k}\to 0$ and $S_{R_k}\to 0, \forall k$ as $\tau\to\infty$ whenever $\mathcal{N}\leq 1$, since $\{\mathbf{0}\}$ is the only omega limit point of any orbit starting in the interior of $\mathbb{R}_+^{3N}$ whenever $\mathcal{N}\leq 1$. □

Remark 4 The parameter $\mathcal{N}$ so identified in (83) is the unique threshold parameter for the disease-free system with the property that for $0\leq\mathcal{N}\leq 1$, the only steady state of the system is the trivial steady state solution which is globally and asymptotically stable and as $\mathcal{N}$ increases beyond unity, there is transcritical bifurcation at $\mathcal{N}=1$ leading to the birth of the non-trivial steady state, given by (82).

The stability of the persistent vector steady state solution derived above may be determined, starting from its linear stability, by looking for the sign of the eigenvalues of the Jacobian matrix evaluated at the steady states. So, if λ is an eigenvalue of the Jacobian matrix at the vector persistent steady state, and we denote the Jacobian matrix at the steady state by $J(\boldsymbol{S^*})$, where $\boldsymbol{S^*}=(\boldsymbol{S^*_{B_k}}, \boldsymbol{S^*_{Q_k}}, \boldsymbol{S^*_{R_k}})\in\mathbb{R}^{3N}$ is the vector of steady state values. Then we have

$$J(\boldsymbol{S}^*) = \begin{pmatrix} -\boldsymbol{B} & \boldsymbol{O} & \boldsymbol{R}(\boldsymbol{S}^*_{\boldsymbol{R}_k}) \\ \boldsymbol{T} & -\boldsymbol{T} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{P} & -\boldsymbol{P} \end{pmatrix}, \text{ and } |\lambda \boldsymbol{I}_{3N} - J(\boldsymbol{S}^*)| = 0, \tag{90}$$

where $\boldsymbol{B}$, $\boldsymbol{P}$, $\boldsymbol{T}$, $\boldsymbol{O}$ and $\boldsymbol{R}$ are $N \times N$ matrices defined by

$$\boldsymbol{B} = \text{diag}\{b_k\}, \ \boldsymbol{T} = \text{diag}\{\tau_k\}, \ \boldsymbol{P} = \text{diag}\{\rho_k\}, k = 1, 2, 3, \cdots N$$

$$\boldsymbol{R}(\boldsymbol{S}^*_{\boldsymbol{R}_k}) = \begin{pmatrix} b_1\alpha_1\left(1 - \frac{2S^*_{R_1}}{L_1}\right) & b_1\alpha_2\left(1 - \frac{2S^*_{R_2}}{L_2}\right) & \cdots & b_1\alpha_N\left(1 - \frac{2S^*_{R_N}}{L_N}\right) \\ b_2 & 0 & \cdots & 0 \\ 0 & b_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_N \end{pmatrix},$$

and for a general c_k, $\text{diag}\{c_k\}$, for $k = 1, 2, \cdots, N$ is the $N \times N$ diagonal matrix whose kth entry is c_k. Computation of the determinant shown in the second term of equation (90) yields a polynomial of degree $3N$ whose form and structure can be completely determined as has been derived in [23]. In fact the following results about the disease-free system have been studied and established in [23]:

1. The basic offspring number $\mathcal{N}$, also known as the vectorial basic reproduction number for this system, can be computed as the dominant eigenvalue of the positive linear operator by deriving the next generation operator of the system based on ideas from epidemiological models in [8, 30].
2. The computations of the $3N \times 3N$ determinant previewed by (90) can be reduced to the computation of an $N \times N$ determinant through the use of block matrices as proved in [25, 28].
3. The persistent steady state of the system is locally and asymptotically stable for a range of values of $\mathcal{N}$ and can be driven to instability with growing oscillations as $\mathcal{N}$ further passes through a critical value $\mathcal{N}_c$ whose value can be determined in a set of calculations dependent on N, the total number of gonotrophic cycles that the system would have undergone (Theorem 9 in [23]).
4. The initial period of the oscillations, at the point where the system undergoes a Hopf bifurcation, can be computed in terms of the parameters of the system (Lemma 3 in [23]).

From the above, we deduce that the disease-free model exhibits more dynamic variability than is the case with the disease-free state in mathematical models for malaria that do not include a demographic component in the population dynamics of the mosquito. This dynamic nature will then be used to study the epidemiological model in the next section. The main result of this section is the identification of the number $\mathcal{N}$ such that for $\mathcal{N} \leq 1$ the population size of the mosquitoes decays to zero and for $\mathcal{N} > 1$, this population establishes itself in the environment. $\mathcal{N} > 1$

will therefore serve as a logical starting point for the analysis of the epidemiological model derived in this paper.

3.2 The Epidemiological Model

In this subsection, we analyse the epidemiological model under the simplifying assumption that the probabilities p_{hQ_k}, $\tilde{p}_{hQ_k}$ and p_{Q_kh} are the same at each epidemiological gonotrophic cycle level. That is $p_{hQ_k} = p_{hQ}$, $\tilde{p}_{hQ_k} = \tilde{p}_{hQ}$ and $p_{Q_kh} = p_{Qh}$ for all k. This assumption reduces to the fact that all mosquitoes transmit the infection with the same efficiency, irrespective of the physiological age of the insect. It may eventually be informative to allow all these probabilities to be different to study differential infectivity of infectious questing mosquitoes based on the idea that older infectious questing mosquitoes may be more infectious than younger ones. We note, however, that from the definition in Eq. (12), we did capture the fact that only questing mosquitoes at a certain gonotrophic stage may be infectious. We start with the following theorem:

Theorem 4 (On the Existence of Steady States for the Epidemiological Model) *The system defined by Eqs.* (64)–(74) *has at least three steady states: the trivial steady state where all variables are zero, the disease-free steady state where all disease variables are zero and the endemic steady state where disease and demographic variables co-exist.*

Proof The steady state solutions are obtained by setting the right hand side of each of Eqs. (64)–(74) to zero. From the scaling done we easily deduce that all steady state solutions can be written in terms of $S^*_{Q_1}$ as follows: Set $\xi = p_{hQ}I^*_h + \tilde{p}_{hQ}R^*_h$ to have

$$S^*_{Q_k} = S^*_{B_k} = (1-\xi)^{k-1}S^*_{Q_1},\ S^*_{R_k} = (1-\xi)^k S^*_{Q_1},\ k = 1, 2, 3, \cdots, N.$$
$$I^*_{R_{k,k}} = \xi(1-\xi)^{k-1}S^*_{Q_1}, k = 1, 2, \cdots, N.$$

For any integer $m > 1$, we have

$$I^*_{Q_{j+m,j}} = I^*_{B_{j+m,j}} = I^*_{R_{j+m-1,j}} = \cdots = I^*_{R_{j,j}} = \xi(1-\xi)^{j-1}S^*_{Q_1}, j = 1, 2, \cdots, N$$

this establishes the steady state values

$$S^*_{Q_1} = S^*_{B_1},\ S^*_{Q_k} = S^*_{B_k} = S^*_{R_{k-1}},\ S^*_{R_k} = (1-\xi)^k S^*_{Q_1}, k = 2, 3, \cdots, N \quad (91)$$
$$I^*_{Q_{k,j}} = I^*_{B_{k,j}} = I^*_{R_{k-1,j}},\ I_{R_{k,j}} = \xi(1-\xi)^{j-1}S^*_{Q_1},\ k \geq j,\ j = 1, 2, 3, \cdots, N \quad (92)$$

where all the expressions are given in terms of $S^*_{Q_1}$. Setting and substituting these into the equation for S_{B_1} (Eq. (64)) when the time derivative is set to zero, we get the equation

$$\sum_{k=1}^{N} \alpha_k \left(1 - \frac{(1-\xi)^k S^*_{Q_1}}{L_k}\right)(1-\xi)^k S^*_{Q_1} + \sum_{j=1}^{N}\sum_{k=j}^{N} \alpha_{k,j} \left(1 - \frac{\xi(1-\xi)^{j-1} S^*_{Q_1}}{L_{k,j}}\right)\xi(1-\xi)^{j-1} S^*_{Q_1} - S^*_{Q_1} = 0,$$

leading to the two solutions:

$$S^*_{Q_1} = 0 \text{ and } S^*_{Q_1} = \frac{A_1(\xi)}{A_2(\xi)}, \tag{93}$$

where

$$A_1(\xi) = \left(\sum_{k=1}^{N} \alpha_k (1-\xi)^k + \sum_{j=1}^{N}\sum_{k=j}^{N} \alpha_{k,j}\xi(1-\xi)^{j-1}\right) - 1, \quad \text{and}$$

$$A_2(\xi) = \sum_{k=1}^{N}\left(\frac{\alpha_k(1-\xi)^{2k}}{L_k}\right) + \sum_{j=1}^{N}\sum_{k=j}^{N}\left(\frac{\alpha_{k,j}\xi^2(1-\xi)^{2(j-1)}}{L_{k,j}}\right).$$

When these solutions are substituted into the corresponding equations for S^*_h, I^*_h and R^*_h, we have equations from which we can work out the value of the endemic steady states. We note that the non-zero solution is given in terms of ξ, and ξ is the weighted combination of the steady state proportions I_h and R_h which are both less than 1. Therefore $0 \leq \xi < 1$ with the case $\xi = 0$ corresponding to when there is no disease in the human population. In fact, when $\xi = 0$, we recover the disease-free steady state solution which we have studied above, and when $0 < \xi < 1$ we enter the endemic steady state zone and there is a steady state solution for all variables both in the human and mosquito populations and the result of the theorem is established. □

Remark 5 The existence of the endemic steady state solution, $S^*_{Q_1} \neq 0$, as a positive steady state solution of the system depends on the sign of $A_1(\xi)$ and, in turn, on the relative size of the quantity $\left(\sum_{k=1}^{N} \alpha_k(1-\xi)^k + \sum_{j=1}^{N}\sum_{k=j}^{N} \alpha_{k,j}\xi(1-\xi)^{j-1}\right)$ when compared with unity. In this regard, we can deduce some results as follows:

1. We, at this stage, can start by accepting this quantity as a threshold parameter that determines whether or not there shall be a flow of infection within the mosquito population in the full epidemiological model, especially as its size determines

the existence of non-zero densities of infected mosquitoes in the system at equilibrium. In fact, a non-zero distribution of infected vectors at equilibrium can exists only when the quantity $\mathcal{R}$ defined by

$$\mathcal{R}(\xi) = \left(\sum_{k=1}^{N} \alpha_k (1-\xi)^k + \sum_{j=1}^{N} \sum_{k=j}^{N} \alpha_{k,j} \xi (1-\xi)^{j-1} \right) \tag{94}$$

is such that $\mathcal{R} > 1$. We then define this quantity as the number of secondary infections that will arise from one primary infectious mosquito in the full epidemiological model during the entire period of infectiousness of the primary infected mosquito when placed in a completely susceptible population. This quantity so defined depends on the size of I_h and R_h and increases with these quantities.

2. The quantity $\mathcal{R}$ so identified may be seen as being related to the epidemiological model's basic reproduction number $\mathcal{R}_0$ which we shall determine below. $\mathcal{R}$ clearly has the threshold character of the reproduction number in the vector population.
3. When $\xi = 0$, there is no disease in the human population but there is a thriving susceptible population of mosquitoes coexisting with susceptible humans, and $\mathcal{R}$ reduces to the basic offspring number which we have computed in Sect. 3.1 above and given by Eq. (83). That is, $\mathcal{R}(0) = \mathcal{N}$. In general, however, whenever $0 < \xi < 1$, we have that $\mathcal{N}(1-\xi) < \mathcal{R}(\xi) < \sum_{k=1}^{N} \alpha_k + \sum_{j=1}^{N} \sum_{k=j}^{N} \alpha_{k,j} = \mathcal{N} + \sum_{j=1}^{N} \sum_{k=j}^{N} \alpha_{k,j}$. $\mathcal{R}$ may therefore serve as a parameter that extends the demographic model's results on existence of steady states to those of the full epidemiological model.

3.3 The Epidemiological Model's Basic Reproduction Number

The full epidemiological model has been shown to possess the following equilibrium solutions: a trivial equilibrium solution $\mathbf{0}$ (81), which corresponds to the existence of a susceptible human population without disease infection and an extinct mosquito population; the disease-free equilibrium solution, which corresponds to a susceptible human population interacting with a disease-free (susceptible) mosquito population, with equilibrium solution given by Eq. (82); and the endemic equilibrium given by the second of Eq. (93), parameterized with the parameter ξ. From an epidemiological perspective, all what we want to establish is whether there does exist a parameter that determines the existence of the endemic equilibrium alluded to in Remark 5. To this effect we calculate the basic reproduction number of the system.

To compute the basic reproduction number we use the standard procedure as explained in [8, 30] wherein the basic reproduction number is calculated as the dominant eigenvector of a linear operator. In this case we can consider only the equations where the disease is in progression and by considering the equations in the sub-system

$$\frac{d\tilde{I}_{R_{k,k}}}{d\tau} = \rho_{k,k}\left((p_{hQ_k}\tilde{I} + \tilde{p}_{hQ_k}\tilde{R})S_{Q_k} - \tilde{I}_{R_{k,k}}\right), \quad k = 1, 2, 3, \cdots, N; \tag{95}$$

$$\frac{d\tilde{I}_{B_{k,j}}}{d\tau} = b_{k,j}\left(\tilde{I}_{R_{k-1,j}} - \tilde{I}_{B_{k,j}}\right), \quad k{-}2, 3, \cdots, N{-}1, \ j = 1, 2, \cdots, N, \ k > j; \tag{96}$$

$$\frac{d\tilde{I}_{Q_{k,j}}}{d\tau} = \tau_{k,j}\left(\tilde{I}_{B_{k,j}} - \tilde{I}_{Q_{k,j}}\right), \quad k{-}2, 3, \cdots, N-1, j = 1, 2, \cdots, N, k > j; \tag{97}$$

$$\frac{d\tilde{I}_{R_{k,j}}}{d\tau} = \rho_{k,j}\left(\tilde{I}_{Q_{k,j}} - \tilde{I}_{R_{k,j}}\right), \quad k{-}2, 3, \cdots, N-1, \ j = 1, 2, \cdots, N, \ k > j; \tag{98}$$

$$\frac{d\tilde{E}_h}{d\tau} = \sum_{j=1}^{N-2}\sum_{k=j+2}^{N} g_{k,j}S_h\tilde{I}_{Q_{k,j}} - (\tilde{\nu}_h{+}\tilde{\lambda}_h)\tilde{E}_h, \tag{99}$$

$$\frac{d\tilde{I}_h}{d\tau} = \tilde{\nu}_h\tilde{E}_h - (\tilde{\lambda}_h{+}\tilde{\sigma}_h + \tilde{r}_h)\tilde{I}_h, \tag{100}$$

$$\frac{d\tilde{R}_h}{d\tau} = \tilde{\sigma}_h\tilde{I}_h - (\tilde{\delta}_h + \tilde{\lambda}_h)\tilde{R}_h. \tag{101}$$

To set up the local drift matrix that determines the flow of infection within the infectious compartments, we let $\boldsymbol{x} = (\boldsymbol{I_{R_{k,j}}}, \boldsymbol{I_{B_{k,j}}}, \boldsymbol{I_{Q_{k,j}}})^T \in \mathbb{R}^{\mathcal{M}(N)-3N}$ for $k, j \in 1, 2, \cdots, N, \quad k \geq j$, and then write the reduced sub-system containing only disease progression variables in the form

$$\frac{d\boldsymbol{x}}{d\tau} = \mathcal{F}(\boldsymbol{x}) - \mathcal{V}(\boldsymbol{x}),$$

where

$$\mathcal{F}(\boldsymbol{x}) = \begin{pmatrix} \rho_{1,1}\left(p_{hQ_1}\tilde{I} + \tilde{p}_{hQ_1}\tilde{R}\right) S_{Q_1} \\ \rho_{2,2}\left(p_{hQ_2}\tilde{I} + \tilde{p}_{hQ_2}\tilde{R}\right) S_{Q_2} \\ \vdots \\ \rho_{N,N}\left(p_{hQ_N}\tilde{I} + \tilde{p}_{hQ_N}\tilde{R}\right) S_{Q_N} \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \sum_{j=1}^{N-2}\sum_{k=j+2}^{N} g_{k,j} S_h \tilde{I}_{Q_{k,j}} \\ 0 \\ 0 \end{pmatrix}, \quad \mathcal{V}(\boldsymbol{x}) = \begin{pmatrix} \rho_{1,1}\tilde{I}_{R_{1,1}} \\ \rho_{2,2}\tilde{I}_{R_{2,2}} \\ \vdots \\ \rho_{N,N}\tilde{I}_{R_{N,N}} \\ -\rho_{2,1}\left(\tilde{I}_{Q_{2,1}} - \tilde{I}_{R_{2,1}}\right) \\ -\rho_{3,2}\left(\tilde{I}_{Q_{3,2}} - \tilde{I}_{R_{3,2}}\right) \\ \vdots \\ -\rho_{N,N-1}\left(\tilde{I}_{Q_{N,N-1}} - \tilde{I}_{R_{N,N-1}}\right) \\ -b_{2,1}\left(\tilde{I}_{R_{1,1}} - \tilde{I}_{B_{2,1}}\right) \\ -b_{3,2}\left(\tilde{I}_{R_{2,2}} - \tilde{I}_{B_{3,2}}\right) \\ \vdots \\ -b_{N,N-1}\left(\tilde{I}_{R_{N-1,N-1}} - \tilde{I}_{B_{N,N-1}}\right) \\ -\tau_{2,1}\left(\tilde{I}_{B_{2,1}} - \tilde{I}_{Q_{2,1}}\right) \\ -\tau_{3,2}\left(\tilde{I}_{B_{3,2}} - \tilde{I}_{Q_{3,2}}\right) \\ \vdots \\ -\tau_{N,N-1}\left(\tilde{I}_{B_{N,N-1}} - \tilde{I}_{Q_{N,N-1}}\right) \\ (\tilde{\nu}_h + \tilde{\lambda}_h)\tilde{E}_h \\ -\tilde{\nu}_h\tilde{E}_h + (\tilde{\lambda}_h + \tilde{\sigma}_h + \tilde{r}_h)\tilde{I}_h \\ -\tilde{\sigma}_h\tilde{I}_h + (\tilde{\delta}_h + \tilde{\lambda}_h)\tilde{R}_h \end{pmatrix}.$$

We then derive the local drift matrices F and V from $\mathcal{F}$ and $\mathcal{V}$ by computing $F_{i,j} = \frac{\partial \mathcal{F}_i}{\partial x_j}(DFE)$ and $V_{i,j} = \frac{\partial \mathcal{V}_i}{\partial x_j}(DFE)$ and $R_0(N) = \max_\lambda\{|\lambda| : \lambda \text{ is an eigenvalue of } FV^{-1}\}$. To be in a position to state the formula for the basic reproduction number for an arbitrary N, we use an inductive approach on N to calculate the value of $R_0(N)$, $N = 3, 4, \cdots$, where $R_0(N)$ is the basic reproduction number when the total number of gonotrophic cycles that the mosquito can complete during its entire reproductive life is $N > 2$.

For $N = 3$ we find that

$$R_0(3) = \frac{\sqrt{\nu}\sqrt{S_h^{DFE}}\sqrt{g_{3,1}S_{Q_1}^{DFE}}\sqrt{(\delta+\lambda)p_{hQ} + \tilde{p}_{hQ}\sigma}}{\sqrt{\delta+\lambda}\sqrt{\lambda+\nu}\sqrt{\gamma+\lambda+r+\sigma}}. \tag{102}$$

Thus ending the process at $N = 3$, we notice that only susceptible questing vectors infected at gonotrophic cycle 1 would have contributed to the value of R_0. This is indeed the expected result since we allow at least two gonotrophic cycles to pass before the vector can become infectious to humans. For $N = 4$, we have the expression

$$R_0(4) = \frac{\sqrt{\nu}\sqrt{S_h^{DFE}}\sqrt{(g_{3,1}+g_{4,1})S_{Q_1}^{DFE} + g_{4,2}S_{Q_2}^{DFE}}\sqrt{(\delta+\lambda)p_{hQ} + \tilde{p}_{hQ}\sigma}}{\sqrt{\delta+\lambda}\sqrt{\lambda+\nu}\sqrt{\gamma+\lambda+r+\sigma}}. \tag{103}$$

For $N = 4$, we see the contributions to the value of the reproduction number from susceptible questing vectors first infected at gonotrophic cycles 1 and 2, with the appropriate weighting factors. At $N = 5$ we have

$$R_0(5) = \frac{\sqrt{\nu}\sqrt{S_h^{DFE}}\sqrt{(\delta+\lambda)p_{hQ}+\tilde{p}_{hQ}\sigma}\sqrt{(g_{3,1}+g_{41}+g_{5,1})S_{Q_1}^{DFE}+(g_{4,2}+g_{5,2})S_{Q_2}^{DFE}+g_{5,3}S_{Q_3}^{DFE}}}{\sqrt{\delta+\lambda}\sqrt{\lambda+\nu}\sqrt{\gamma+\lambda+r+\sigma}}, \tag{104}$$

showing contributions from susceptible questing vectors first infected at gonotrophic cycles 3, 4 and 5. Continuing with the calculation, we can then derived the general formula for the basic reproduction number by combining these together and deducing. We have thus proved the following result.

Theorem 5 (On the General Formula for the Basic Reproduction Number) *Let the assumptions of Lemma 1 continue to hold. Let $R_0(N)$ be the basic reproduction number for the epidemiological model built on the human–mosquito interactive framework. Then $R_0(N)$ is given by the formula*

$$R_0(N) = \frac{\sqrt{\nu}\sqrt{S_h^{DFE}}\sqrt{(\delta+\lambda)p_{hQ}+\tilde{p}_{hQ}\sigma}\sqrt{\sum_{j=1}^{N-2}\sum_{k=j+2}^{N}g_{k,j}S_{Q_j}^{DFE}}}{\sqrt{\delta+\lambda}\sqrt{\lambda+\nu}\sqrt{\gamma+\lambda+r+\sigma}}. \tag{105}$$

Remark 6 We note that from the scaling used herein, $S_{Q_k}^{DFE} = S_{B_1}^*$ for all $k = 1, 2, \cdots, N$ as computed in (80) and so the expression for the formula for $R_0(N)$ simplifies accordingly and we can, using the fact that for non-negative values of $R_0(N)$, $R_0(N) > 1 \Leftrightarrow R_0(N)^2 > 1$ and $0 \leq R_0(N) < 1 \Leftrightarrow R_0(N)^2 < 1$, to then consider R_0^2 as the basic reproduction number and use $\mathscr{R}_0 = R_0(N)^2$ as the basic reproduction number for our system, where

$$\mathscr{R}_0 = \frac{\nu\left((\delta+\lambda)p_{hQ}+\tilde{p}_{hQ}\sigma\right)\left(\sum_{j=1}^{N-2}\sum_{k=j+2}^{N}g_{k,j}\right)S_h^{DFE}S_{Q_1}^{DFE}}{(\delta+\lambda)(\lambda+\nu)(\gamma+\lambda+r+\sigma)}. \tag{106}$$

This expression clearly shows the form and dependence of the system and reproduction number on the disease parameters.

To appreciate the contributions into the two threshold parameters $\mathcal{N}$ and $\mathscr{R}_0$ we have computed and displayed in formulas (83) and (106), we return to the original variables, passing through the scaled variables and parameter groupings (61), (62) and (63). We define the quantities

$$w_k = \frac{b_{Q_k}}{b_{Q_k}+\mu_{S_{Q_k}}}\frac{b_k(N_h)}{b_k(N_h)+\mu_{S_{B_k}}}\frac{\rho_k}{\rho_k+\mu_{S_{R_k}}}q_{Q_k}, k = 1, 2, \cdots, N \tag{107}$$

$$z_k = \frac{b_{Q_{k-1}}}{b_{Q_k}+\mu_{S_{Q_k}}}\frac{b_k(N_h)}{b_k(N_h)+\mu_{S_{B_k}}}\frac{\rho_{k-1}}{\rho_{k-1}+\mu_{S_{R_{k-1}}}}q_{Q_{k-1}}, k=2, \cdots, N \tag{108}$$

$$w_{k,j} = \frac{b_{Q_k}}{b_{Q_k} + \mu_{I_{Q_{k,j}}}} \frac{b_k(N_h)}{b_k(N_h) + \mu_{I_{B_{k,j}}}} \frac{\rho_{k,j}}{\rho_{k,j} + \mu_{I_{R_{k,j}}}} q_{Q_k}, k > j \tag{109}$$

$$z_{k,j} = \frac{b_{Q_{k-1}}}{b_{Q_k} + \mu_{I_{Q_{k,j}}}} \frac{b_k(N_h)}{b_k(N_h) + \mu_{I_{B_{k,j}}}} \frac{\rho_{k-1,j}}{\rho_{k-1,j} + \mu_{I_{R_{k-1,j}}}} q_{Q_{k-1}}, k > j, \tag{110}$$

then, for any $p \geq 1$, we have

$$I^0_{Q_{j+p,j}} = \prod_{i=1}^{p} z_{j+i,j} S^0_{Q_j}, \; S^0_{Q_j} = \prod_{i=2}^{j} z_j S^0_{Q_1}, j = 2, 3, \cdots, N, \text{ and } S^0_{Q_1} = \frac{b_1(N_h)}{b_{Q_1} + \mu_{S_{Q_1}}} S^0_{B_1}. \tag{111}$$

It is straightforward to see that $\frac{b_1(N_h)}{b_{Q_1}+\mu_{S_{Q_1}}} \prod_{k=2}^{N} z_k = \prod_{k=1}^{N} w_k$ after rearrangement of terms. Similar rearrangements can be made to write down the products involving $z_{k,j}$ in terms of the products $w_{k,j}$. With these characterizations, we have

$$\alpha_k = \lambda_{0,k} \prod_{i=1}^{k} w_i \Rightarrow \mathcal{N} = \sum_{k=1}^{N} \left(\lambda_{0,k} \prod_{i=1}^{k} w_i \right), \tag{112}$$

written in terms of the original parameters. Thus for any integer $p \geq 1$, we have the expression

$$g_{j+p,j} = \frac{p_{Q_{j+p}h} q_{Q_{j+p}} b_{Q_{j+p}}}{N_h(\rho_N + \mu_{S_{R_N}})} I^0_{Q_{j+p,j}}, \tag{113}$$

using the expression for $I^0_{Q_{j+p,j}}$ computed in (111). The expressions in (112) and (113) completely demonstrate how we can write down the basic offspring number $\mathcal{N}$ and the epidemiological model's basic reproduction number $\mathscr{R}_0$ given in (83) and (106) respectively in terms of the original parameters of the system. We note that the terms in the sum for (112) and (113) are ever decreasing in size since each of the weights $w_i < 1$ and $w_{k,j} < 1$. This captures the fact that contributions to the new mosquito population from each egg laying is decreasing with increasing number of gonotrophic cycle. These results confirm the results earlier reported in [22] for a limited form on this model where the gonotrophic cycle count was truncated at $N = 3$. This result is biologically consistent with the biology of the *Anopheles sp* mosquito, as it is known that depending on the species, and on the quality and size of the blood meal, a female *Anopheles* lays 50–200 eggs during a single oviposition, usually at night. *Successive egg batches tend to decrease in size* and there may be seasonal variations in the number of eggs laid per batch [27].

4 Results and Discussions

We set out to derive and study a model that takes into account the demography of mosquito populations, its reproductive life and gonotrophic cycles' count. The model was derived based on a restricted form of homogeneous mixing contingent on the idea that the mosquito has a human biting habit. Though consideration was made on the possibility that the mosquito can seek for blood meals from alternative blood sources, this zoophilic characteristic of the mosquito's life traits was not fully integrated into the model equations. We ended up with a system of equations governing the dynamics of a human–mosquito–human interactive framework where the mosquito can undergo up to N gonotrophic cycles: the cycle that starts at egg laying, followed by blood meal questing within human populations, resting for maturation of eggs and again ending with egg laying. In the model we used the gonotrophic cycle counter as a measure of physical age of the mosquito, so that a mosquito that has undergone the largest number of gonotrophic cycles in the system is considered to be among the oldest mosquitoes, while those that are still in the first gonotrophic cycle are considered to be among the young ones. In the presence of malaria disease in the model, the length of the gonotrophic cycle was fixed and used as a means to measure and capture the incubation period of the disease within the mosquito population. This was set by requiring that a mosquito that picks up the malaria infection after a successful interaction with humans will have to wait for at least two gonotrophic cycles before bringing back the infection to the human population. All mosquitoes in the system are assumed to complete the requisite number of gonotrophic cycles unless they dies during the interaction with the humans.

The size of the entire system was computed by explicitly counting the total number of mosquito compartments in the system. The probability of finding an infected/infectious compartment in the mosquito system was computed by comparing the total number of infectious mosquito compartments with the total number of mosquito compartments by explicitly calculating the disease reservoir of infection in the mosquito population, ROI_V. The results show that this probability is determined by the number of susceptible mosquito compartments that the disease-free system can have at each gonotrophic cycle. In any event it was seen that the probability can be as large as 50% in the case where we have only one susceptible compartment and one infected/infectious compartment.

We do not believe the results presented above is equivalent to the result of finding an infectious mosquito, unless there is a one-to-one correspondence between the number of compartments considered and the distribution of the mosquitoes in the various compartments. For example, if there are M compartments at each gonotrophic cycle k, and a total of $N_{m_k} > M$ mosquitoes at each corresponding gonotrophic cycle k, then the results so obtained could be used as a proxy for the probability of finding an infectious mosquito if the M proportions $\frac{M}{N_{m_k}}$ is the same and a constant and sums up to 1. However, if this is not the case, then increasing the number of compartments does not necessarily produce any more

interesting results. However, the flows from the susceptible compartments to the infectious compartments are important as they will play a part in determining the size of the questing mosquito populations, susceptible or infectious. However, from a biological and ecological standpoint, it is worth asking the question: what is the maximum number of compartments required that would capture the human–mosquito interaction system and approximates as close as possible the number of mosquitoes (in our case questing mosquitoes) interacting with humans?

The disease-free model was analysed and the results show that there exists a threshold parameter, $\mathcal{N}$, which we identified as the basic offspring number, with the property that for $\mathcal{N} \leq 1$ the system has only the zero or trivial steady state solution which is globally and asymptotically stable. We note that this is the state in which the mosquito population is extinct; however, there is a thriving population of only susceptible humans. Our analysis also shows that when $\mathcal{N} > 1$, the trivial steady state which always exists for all forms of the parameters, loses stability, giving rise to a non-trivial steady state via a transcritical bifurcation. These results have been observed before [19, 22, 23]. They are again echoed here in this general model where the number of gonotrophic cycles is arbitrary.

A study of the epidemiological model also reveals the existence of a basic reproduction number whose nature and form has been completely determined in terms of the original parameters of the system. These results were obtained by assuming that the human population is constant and that disease induced death rate is zero. These simplifications facilitated the parameterization of the steady states in the epidemiological model in terms of the infectious proportions in the human population I_h and R_h. It was satisfying to notice that when there is no infection in the system, the system collapses to the disease-free model whose form has been studied in earlier works. The novelty in the work done in this paper lies in its complete characterization of the contributions from each mosquito from each gonotrophic cycle count and the ability to build in the incubation period of the disease within the mosquito population using the gonotrophic cycle counter.

The length of the extrinsic incubation period was measured here by allowing two gonotrophic cycles to elapse before the infected mosquito becomes infectious. This choice was deemed realistic by rationalizing that for a mosquito to pass the infection from one human to another, it must bite two humans at two different and distinct periods of time and the assumption is that the second bite will already be infectious. Since the choice of 2 may not be true in general, especially in regions with extended dry or wet period where this value could be higher (see for example [27]), a better option to be considered in future research, would be to set the number, n, of gonotrophic cycles whose cumulative time length is equivalent to the length of the extrinsic incubation period as one of the parameters of the system, and even allow this parameter to vary with temperature and rainfall. However, since it is clear that such a parameter must be at least 2, we decided to use $n = 2$ in the analyses. In this paper, we discussed how this arbitrary parameter can affect the size of the number of infectious compartments of the system.

From the point of view of control, three control points are clearly possible from the modelling framework used here. (i) On the one hand, we can block human–

mosquito contact by preventing the questing mosquitoes from taking blood meals. This will certainly prevent the passage of the infection from mosquito to humans and vice versa and at the same time reduce the population sizes of the anthropophilic mosquitoes in the system. (ii) One could identify resting places of the mosquitoes and spray these with insecticide. This will kill the resting mosquitoes and cut the link between current adult and next generation adults, as the resting blood fed mosquitoes will be prevented from laying eggs. (iii) One could use larvicides or other control measures to kill the breeding site mosquitoes and aquatic forms. Each of the enumerated control strategies is mechanical and has the net effect of reducing the size of the epidemiological model's basic reproduction number which will in turn control the spread of the malaria disease.

We believe that we have achieved the primary objective of the current paper which was to derive a malaria transmission model for a human–mosquito interactive framework where the mosquito is allowed to undergo up to N gonotrophic cycles during its entire reproductive life, and then calculate the different threshold parameters of the derived system; namely, the basic offspring number in the vector demographic model and the basic reproduction number in the epidemiological model. The analysis of the model derived here is not complete as we need to study the mathematical form and structure of the solutions of the entire system of equations as well as the explicit forms of the steady state solutions. The nature of the rate of production of new adult mosquitoes, given here as λ_R, needs to be studied from a general nonlinear function analytic point of view. The full model represented here offers us a unique opportunity to conduct a global sensitivity analysis of the system on its parameters. These and other aspects of the model shall be subject of subsequent work.

Acknowledgments MIT-E acknowledges the support of Lehigh University and GAN acknowledges the grants and support of the Cameroon Ministry of Higher Education through the initiative for the modernization of research in Cameroon's Higher Education. All authors also acknowledge the sponsorship of the Commission for Developing Countries (CDC) in conjunction with the International Mathematics Union (IMU) through the CDC-ADMP (African Diaspora Mathematicians Program) grant that made it possible for interactive collaborative work during the 2018 and 2019 grant sponsored visits to the University of Buea.

References

1. Anguelov, R., Dumont, Y., Lubumaa, J.: Mathematical modeling of sterile insect technology for control of Anopheles mosquito. Comput. Math. Appl. **64**, 374–389 (2012)
2. Aron, J.L.: Dynamics of acquired immunity boosted by exposure to infection. Math. Biosci. **64**, 249–253 (1983)
3. Aron, J.L.: Acquired immunity dependent upon exposure in an SIRS epidemic model. Math. Biosci. **88**, 37–47 (1988)
4. Aron, J.L.: Mathematical modeling of immunity to malaria. Math. Biosci. **90**, 385–396 (1988)
5. Baton, L.A., Ranford-Cartwright, L.C.: Spreading the seeds of million-murdering death: metamorphoses of malaria in the mosquito. Trends Parasitol. **21**(12), 573–580 (2005)

6. Brännström, Å., Sumpter, D.J.T.: The role of competition and clustering in population dynamics. Proc. R. Soc. B **272**, 2065–2072 (2005)
7. Cook, K., van den Driessche, P., Zou, X.: Interaction of maturation delay and nonlinear birth in population and epidemic models. J. Math. Biol. **39**, 332–352 (1999)
8. Diekmann, O., Heesterbeek, J.A.P., Metz, J.A.J.: On the definition and the computation of the basic reproduction ratio R_0 in rodels for infectious diseases in heterogeneous populations. J. Math. Biol. **28**, 365–382 (1990)
9. Gurarie, D., Karl, S., Zimmerman, P.A., King, C.H., Pierre, T.G.St., Davis, T.M.E.: Mathematical modeling of malaria infection with innate and adaptive immunity in individuals and agent-based communities. PLoS One **7**(3), e34040 (2012)
10. Hale, H.K.: Ordinary Differential Equations. Wiley, New York (1969)
11. Labadin, J., Kon, C.M.L., Juan, S.F.S.: Deterministic malaria transmission model with acquired immunity. In: Proceedings of the World Congress on Engineering and Computer Science, vol. 2, pp. 779–784 (2009)
12. Lutambi, A.M., Penny, M.A., Chitnis, N., Smith, T.: Mathematical modelling of mosquito dispersal in a heterogeneous patchy environment. Math. Biosci. **241**(2), 198–216 (2013)
13. Martens, P., Kovats, R.S., Nijhof, S., de Vries, P., Levermore, M.J.T., Bradley, D.J., Cox, J., McMichael, A.J.: Climate change and future populations at risk of malaria. Global Environ. Change **9**, S89–S107 (1999)
14. Nåsell, I.: *Hybrid Models of Tropical Infections*. Lecture Notes in Biomathematics. Springer, Berlin, Heidelberg, New York, Tokyo (1985)
15. Ngonghala, C.N., Ngwa, G.A., Teboh-Ewungkem, M.I.: Periodic oscillations and backward bifurcation in a model for the dynamics of malaria transmission. Math. Biosci. **240**(1), 45–62 (2012)
16. Ngonghala, C.N., Teboh-Ewungkem, M.I., Ngwa, G.A.: Persistent oscillations and backward bifurcation in a malaria model with varying human and mosquito populations: implications for control. J. Math. Biol. **70**(7), 1581–1622 (2015)
17. Ngonghala, C.N., Teboh-Ewungkem, M.I., Ngwa, G.A.: Observance of period-doubling bifurcation and chaos in an autonomous ODE model for malaria with vector demography. Theoretical Ecology **9**(3), 337–351 (2016)
18. Ngwa, G.A.: Modelling the dynamics of endemic malaria in growing populations. DCD Ser. B **4**(4), 1173–1202 (2004)
19. Ngwa, G.A.: On the population dynamics of the malaria vector. Bull. Math. Biol. **68**(8), 2161–2189 (2006)
20. Ngwa, G.A., Shu, W.S.: A mathematical model for endemic malaria with variable human and mosquito populations. Math. Comput. Modell. **32**(7), 747–763 (2000)
21. Ngwa, G.A., Niger, A.M., Gumel, A.B.: Mathematical assessment of the role of non-linear birth and maturation delay in the population dynamics of the malaria vector. Appl. Math. Comput. **217**, 3286–3313 (2010)
22. Ngwa, G.A., Teboh-Ewungkem, M.I., Dumont, Y., Ouifki, R., Banasiak, J.: On a three-stage structured model for the dynamics of malaria transmission with human treatment, adult vector demographics and one aquatic stage. J. Theor. Biol. **481**, 202–222 (2019)
23. Ngwa, G.A., Wankah, T.T., Fomboh-Nforba, M.Y., Ngonghala, C.N., Teboh-Ewungkem, M.I.: On a reproductive stage-structured model for the population dynamics of the malaria vector. Bull. Math. Biol. **76**, 2476–2516 (2014)
24. Patz, J.A., Olson, S.H.: Malaria risk and temperature: Influences from global climate change and local land use practices. Proc. Natl. Acad. Sci. **103**(15), 5635–5636 (2006)
25. Powel, P.D.: Calculating determinants of block matrices. arXiv:1112.4379v1 [math.RA] (2011)
26. Ross, R.: The Prevention of Malaria. John Murray, London (1911)
27. Service, M.W., Townson, H.: The Anopheles vector. In: Warrell, D.A., Gilles, H.M. (eds.), Essential Malariology, chapter 4, pp. 59–84. CRC Press, Taylor & Francis Group, (2002)
28. Silvester, J.R.: Determinants of block matrices. Mathematical Gazzet **85**(501), 460–467 (2000)

29. Teboh-Ewungkem, M.I., Yuster, T.: A within-vector mathematical model of Plasmodium falciparum and implications of incomplete fertilization on optimal gametocyte sex ratio. J. Theor. Biol. **264**(2), 273–286 (2010)
30. van den Driessche, P., Watmough, J.: Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math. Biosci. **180**, 29–48 (2002)
31. Villarreal, C., Fuentes-Maldonado, G., Rodriguez, M.H., Yuval, B.: Low rates of multiple fertilization in parous Anopheles albimanus. J. Am. Mosq. Control Assoc. **10**(1), 67–69 (1994)
32. Yuval, B., Fritz, G.N.: Multiple mating in female mosquitoes evidence from a field population of Anopheles freeborn (Diptera: Culicidae). Bull. Entomol. Res. **84**(01), 137–139 (1994)
33. Zhou, G., Minakawa, N., Githeko, A.K., Yan, G.: Association between climate variability and malaria epidemics in the East African highlands. Proc. Natl. Acad. Sci. **101**, 2375–2380 (2004)

Charles Darwin Meets Ronald Ross: A Population-Genetic Framework for the Evolutionary Dynamics of Malaria

Kristan A. Schneider

1 Introduction

Aside tuberculosis and HIV/AIDS, malaria ranks among the most important infectious diseases. It continues to be a major burden to reach global development, rendering its control and elimination a public-health priority. Half of the world's population lives at risk of malaria, and, although a substantial decline in cases and deaths have been achieved since 2010, it still causes annually over 200 million infections and half a million deaths [36]. After years of achievements in control and prevention the trend of reducing malaria burden reversed in 2018 [36].

Malaria is a vector-borne infectious disease, pathogenic to humans and animals, caused by haploid, unicellular, eukaryotic parasites belonging to the group of *Plasmodium*. Several species of *Anopheles* mosquitoes serve as disease vectors—a discovery for which Sir Ronald Ross was awarded the Nobel Prize in 1902. *Plasmodium* species are well-adapted to their hosts, five of which are pathogenic to humans with *P. falciparum* being the most important in terms of morbidity and mortality, followed by *P. vivax*. Other human malaria species include *P. ovale*, *P. malariae*, and *P. knowlesi*. Although progress has been made over the last years in vaccine development, with the first vaccine currently undergoing pilot trials, the levels of immunization are still relatively low [9] and vaccine development is hampered by the parasite's ability to hide from the immune system.

Malaria control relies heavily on drug treatments for radical cure and prophylaxis, as well as on vector control measures, such as indoor-residual insecticide spraying and the use of insecticide-treated bednets. Control efforts are challenged by the spread of insecticide and drug resistance as well as HRP2/3 gene deletions in *P.*

K. A. Schneider (✉)
Department of Applied Computer- and Biosciences, University of Applied Sciences Mittweida, Mittweida, Germany
e-mail: kristan.schneider@hs-mittweida.de; schneid2@hs-mittweida.de

M. I. Teboh-Ewungkem, G. A. Ngwa (eds.), *Infectious Diseases and Our Planet*, Mathematics of Planet Earth 7, https://doi.org/10.1007/978-3-030-50826-5_6

falciparum causing false negative rapid diagnostic test (RDT) results—all particular instances of Darwinian evolution.

Resistance against chloroquine (CQ), sulfadoxine-pyrimethamine (SP), mefloquine (MQ) are widespread in *P. falciparum*, rendering these drugs inefficient [2, 31]. Currently the spread of artemisinin resistance is a source of concern. In fact, 13 non-synonymous mutations in the k13 gene of *P. falciparum* have been associated with slow parasite clearance [12, 37]. While CQ resistance evolved decades ago in *P. falciparum*, CQ remains the drug of choice to treat *P. vivax* malaria—however, resistance is becoming more widespread [23]. The emergence of insecticide resistance is an evolutionary process occurring in the vector species. On the contrary, evolution of drug resistance, HRP2/3 deletions, and adaptations to hide from immune recognition are evolutionary processes happening in the parasite population. The topic to which this chapter is devoted to.

A general evolutionary-genetic framework for malaria is introduced, applicable to all (human) malaria species. The complex transmission cycle (Fig. 1) shapes the evolutionary dynamics of malaria. The framework is an abstraction of the complex malaria transmission cycle based on biological facts rather than assumptions, while capturing the organism's characteristics and removing unnecessary complexity. The framework generalizes the models of [29, 30] and is capable to model the spread of drug resistance, HRP2/3 deletions or other evolutionary processes. The model is

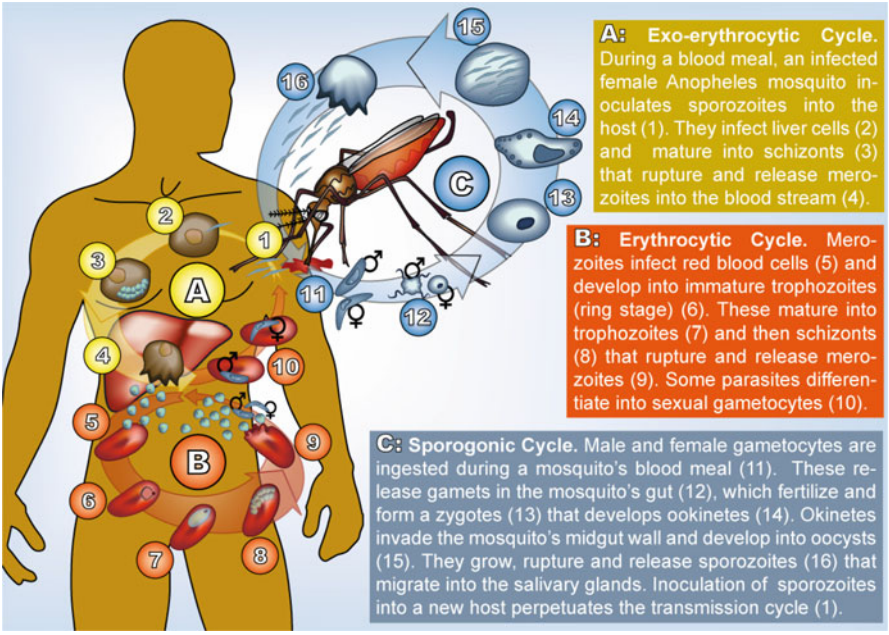


Fig. 1 Transmission cycle of human malaria. All species have the same cycle, but parasites life stages have different morphology (illustrated here for *P. falciparum*). In *P. vivax* and *P. ovale* dormant hypnozoites remain in the liver

applicable to all (human) malaria species, however, it does not model hypnozoites as they occur in *P. vivax* and *P. ovale*.

The presence of multiple genetically distinct haplotypes of the pathogen infecting the same individual host is properly addressed (as it is subsumed by MOI), which is an important feature of malaria correlating with transmission intensities. Particularly, this allows formalizing the differences between haplotype frequencies and prevalences—the former being more relevant in an evolutionary, the latter in an epidemiological or clinical context. These terms are often used synonymously in the empirical literature leading to wrong inferences.

In Sect. 3, the focus shifts to the evolution of drug resistance, assuming a single locus at which multiple mutations act synergistically to confer increasing levels of drug resistance, as it occurs, for instance, at codons 51, 59, 108, 164 in the *Pfdhfr* gene or 436, 437, 613 in the *Pfdhps* gene in *P. falciparum* to confer SP resistance. These results generalize those of [29]. The evolutionary dynamics of drug-resistant haplotype frequencies—unlike haplotype prevalences—are shown to be independent of transmissions intensities (if intra-host competition of parasites is neglected). Furthermore, the effect of genetic hitchhiking at neutral genetic markers flanking the target of drug resistance is studied, generalizing the results of [30]. The hitchhiking pattern is similar as in standard population-genetic theory [13], but the processes of selection and recombination cannot be decoupled due to the characteristics of malaria transmission. Although, the spread of resistance is independent of transmission intensities, the effect of genetic hitchhiking is not. Importantly, the reduction in genetic variation surrounding the target of selection can be genomewide if transmission intensities are low (due to increased inbreeding immanent in the malaria transmission cycle), as explained by Eq. (21) in Sect. 3.3.

Section 4 is devoted to employ the framework for estimating transmission intensities, more precisely MOI, allele frequencies, and selection parameters from molecular data. The methods are illustrated by a hypothetical example. Notably, they have been previously applied to empirical data [16, 27].

An explanation why drug-resistance evolution is faster in *P. falciparum* than in *P. vivax* is given in Sect. 5. Namely, a formal argument how differences in the life histories of the two species, particularly the onset of gametogenesis and longevity of gametocytes, translate into more efficient selection in *P. falciparum*.

The topics covered in this chapter are by far not comprehensive. Rather they compile work that has been done previously, while generalizing the results and deriving new ones. While the framework is mathematically not trivial, the presentation tries to reach out to a less mathematically sophisticated audience, by motivating and verbally describing the model. Because the beauty of the mathematical structure might not be obvious to all readers, proofs are shifted to the appendices.

2 Evolutionary Dynamics

Malaria is a complex disease involving the pathogen, the vector, and the human or animal host linked together by a complex transmission cycle (cf. Fig. 1).

The evolutionary dynamics of malaria are tightly linked to the malaria transmission cycle. During this cycle *Plasmodium* undergoes several steps of multiple fission in the human host and mosquito vector including merogony, gametogony, and sporogony. It starts with sporozoites in a female mosquito's salivary glands. A handful of sporozoites is inoculated into the human blood stream during a mosquito's blood meal (notably such an infectious contact is not necessarily infective), starting the exo-erythrocytic cycle. Sporozoites travel through the blood vessels into the liver, where they infect hepatocytes (liver cells) and mature into schizonts. When these rupture they release merozoites into the blood stream initiating the erythrocytic cycle. Merozoites invade erythrocytes starting a series of asexual replications cycles. First, merozoites form ring stage trophozoites within the red blood cells that mature into schizonts that rupture and release new merozoites. Some of the parasites differentiate into sexual gametocytes, which no longer reproduce in the human host. However, ingestion of male and female gametocytes by a mosquito starts the sporogonic cycle. In the mosquito's gut, gametocytes release gametes that fertilize to form a zygote. These develop into ookinetes that burrow into the mosquito's midgut wall to develop into oocysts. Growth and division of each oocyst produces thousands of sporozoites that move into the salivary glands.

This chapter is concerned with the evolutionary dynamics of the pathogen itself, i.e., with the temporal change in the distribution of parasite haplotypes (characterized by alleles at loci under consideration) under the influence of mutation, selection, and recombination. Notably these evolutionary forces occur at different stages in the life history of the parasite. While selection, e.g., for drug-resistance-conferring haplotypes or for strategies to hide from human immune recognition, occurs in the human host, recombination happens during sporogony inside the mosquito vector, and mutations might occur at all life stages inside the host or vector. To tailor a mathematical model to the malaria transmission cycle, it is crucial to clearly define at which life-stage the haplotype distribution is considered.

A convenient census point is the end of the sporogonic cycle (cf. Fig. 2), i.e., the distribution of haplotypes in the population of sporozoites in the mosquitoes' salivary glands, ready to be injected into human hosts, is considered. In fact, sporozoites of the whole mosquito population are pooled together to form the parasite population. As this population is large, it is sufficient to describe it by their relative haplotype frequencies. This setup follows to follow the parasite population in steps of sexual reproduction. Namely, with a full transmission cycle only one step of sexual reproduction occurs, namely inside the mosquito vector, while several steps of asexual reproduction occur inside the human host.

Clearly, the sporozoite population is changing constantly as mosquitoes die and new ones become infected. However, after a mosquito's blood meal it would typically rest to breed and not participate in transmission for a while. Moreover,

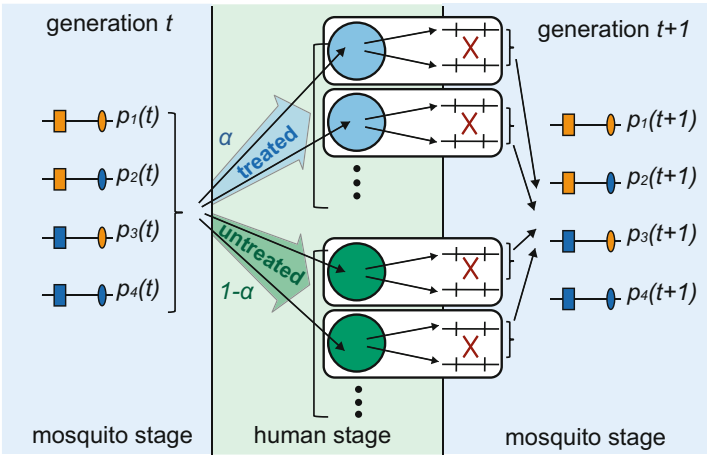


Fig. 2 Idealized transmission cycle: illustration of the idealized transmission cycle for the evolutionary model. Four haplotypes (two biallelic) loci and two strata (treated and untreated hosts) are illustrated. Each host is infected by randomly drawing haplotypes from generation t. With probability α a host receives treatment. Selection occurs differently within treated and untreated hosts. Recombination occurs exclusively between haplotypes exiting the same host. After recombination, haplotypes in the mosquitoes are pooled together to derive their distribution in generation $t + 1$

Anopheles is night active and the mosquito population size undergoes seasonal changes. Finally, there is a time delay between mosquitoes becoming infected and becoming infectious. For these reasons modeling the sporozoites' haplotype distribution on a continuous time scale—although mathematically tempting—would be inaccurate. The important changes in the haplotype distribution occur because mosquitos infect new hosts and new mosquitoes become infected. As a step of idealization, all mosquitoes are assumed to take their blood meals at the same time. It is therefore appropriate to follow the haplotype distribution in discrete time units corresponding to full transmission cycles (cf. Fig. 2). The number of transmission cycles per year depends on transmission intensities. Furthermore, these are not uniformly distributed over the year, especially in regions with seasonal malaria.

2.1 Idealizing the Transmission Cycle

Haplotypes are characterized by their allelic expressions at one or several loci. Let H be the number of possible haplotypes and denote their frequencies in generation t by $P_1(t), \ldots, P_H(t)$ or collectively by the vector $\boldsymbol{P}(t)$. To calculate the change in haplotype frequencies from one generation to the next, the transmission cycle is followed in the steps described below, corresponding to the idealization depicted in

Fig. 2. Throughout this chapter the terms “haplotypes” and “clones” will be used synonymously.

2.1.1 Vector-Host Transmission

As an idealization, a large (infinite) number of hosts is assumed. Clearly, hosts exhibit heterogeneity in exposure to vectors, levels of natural and acquired immunity, access to treatment, and malaria treatment itself (e.g., different treatment policies apply to children, pregnant women, severe cases, different treatment compliance). This is idealized as hosts falling into discrete strata (classes). In generation t, a given host falls into class k ($k = 1, 2, \ldots$) with probability $\alpha_k(t)$. This distribution is modeled time-dependently and thus may subsume changes in treatment policies or other interventions. In the following, if not necessary, dependence on time t is omitted to simplify notation.

Hosts might be infected by one or several mosquitoes (roughly) at the same time (super-infection), i.e., within a few days. At each infectious contact the mosquito transmits exactly one parasite haplotype to the host (cf. Figs. 2 and 4). Hence, transmission of several parasite haplotypes from one mosquito (co-infection) is ignored. In generation t, a host belonging to class k is (super-) infected exactly m times with probability $\kappa_m^{(k)}(t)$, $\left(\sum_{m=1}^{\infty} \kappa_m^{(k)}(t) = 1\right)$. The distribution $(\kappa_m^{(k)})$ describes multiplicity of infection (MOI; cf. Figs. 3 and 4), depending on time and the stratum to which the host belongs. Typical choices for $(\kappa_m^{(k)})$ are discussed below. Note that—unlike here—MOI is often referred to as the number of different clones (haplotypes) within an infection. The former is rather a matter of convenience, as it describes observable information, while neglecting the possibility of multiple infections with the same haplotypes, which is unobservable. The problem with this definition is that the number of different clones does not just depend on transmission intensities but on the haplotype distribution. Further note that ignoring co-infections effectively implies that parasites transmitted during co-infections are randomly mixed (independent) and can be hence subsumed as super-infections. Properly modeling co-infections requires knowledge of the distribution of different haplotype combinations in the mosquitoes salivary glands, besides that it is largely unknown how many different co-infection haplotypes successfully arrive in the liver and lead to merozoite offspring at detectable levels—however, it is believed to be just a handful.

If a host is super-infected by m haplotypes, these are drawn with replacement from the mosquito pool, i.e., according to a multinomial distribution with parameters m and $P_1, \ldots, P_H$.

Given that a host is super-infected m times, the probability of transmitting haplotype i exactly m_i times ($i = 1, \ldots, H$) is

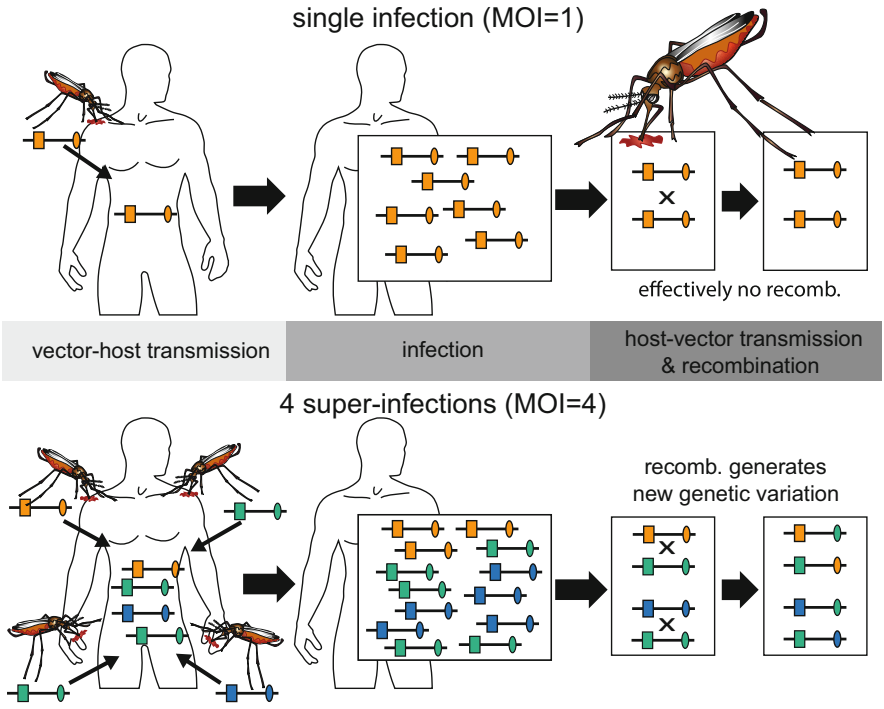


Fig. 3 Illustration of relationship between inbreeding and MOI. Top: A single-clone infection (MOI = 1) results in effectively no recombination. Bottom: different haplotypes exiting a multi-clone infection can generate new genetic variation by recombination. The illustration shows MOI = 4, with three different haplotypes. Note that in principal the same haplotype could infect four times—also effectively resulting in no recombination. However, this happens only with small probability

$$\binom{m}{\boldsymbol{m}} \boldsymbol{P}^{\boldsymbol{m}}, \tag{1}$$

where the vector $\boldsymbol{m} = (m_1, \ldots, m_H)$ indicates how many times the host was super-infected with the respective haplotypes, satisfying $|\boldsymbol{m}| := \sum_{i=1}^{H} m_i = m$, $\binom{m}{\boldsymbol{m}} = \frac{m!}{\prod_{i=1}^{H} m_i!}$ is a multinomial coefficient, and $\boldsymbol{P}^{\boldsymbol{m}} := \prod_{i=1}^{H} P_i^{m_i}$. All infecting sporozoite haplotypes are assumed to successfully migrate into the human liver. This is no loss of generality, as sporozoites not migrating to the liver can be subsumed by MOI m.

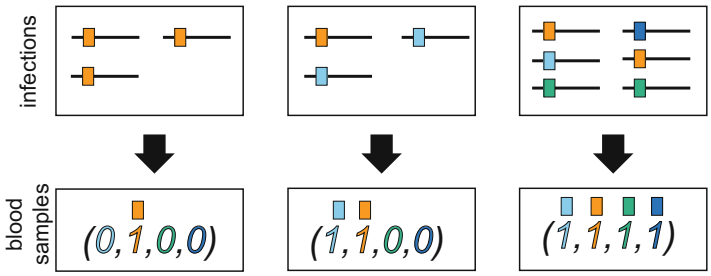


Fig. 4 Illustration of infections (top tow) and information assayed from blood samples. Top row shows three infections with MOI 3, 3, and 6, respectively, infecting with one, two, and four haplotypes (characterized by alleles at one locus). This information gets lost when looking at parasites in patients' blood samples. Only the absence and presence of haplotypes (alleles) can be detected, coded by 0 and 1

2.1.2 Intra-Host Dynamics: Exo-Erythrocytic and Erythrocytic Cycle

Arrived in the liver, sporozoites form schizonts that rupture and finally release merozoites into the blood stream. These invade red blood cells, replicate clonally inside them until they rupture and release new merozoites, determined to further infect red blood cells. A relatively small fraction of merozoites develop gametocytes. These are sexual forms that do not further reproduce inside the host but are responsible for further disease transmission (see Fig. 1). Importantly, no step of sexual reproduction or recombination occurs inside the human host. In the blood stream, anti-malarial drugs clear out merozoites. Most drugs do not affect gametocytes. The reproduction and clearance rates of merozoites, the latter due to anti-malarial drugs or human immune response, are haplotype specific. Drug-tolerant or drug-resistant parasites have a slower clearance rate than drug-sensitive haplotypes but are usually associated with metabolic costs, manifesting in a lower reproduction rate. This results in different selection pressures for sensitive and resistant parasites.

Because all steps of parasite replication inside the human host are asexual, selection acting on parasites inside the host can be modeled as a single step (cf. Fig. 2). Selection pressures naturally vary between hosts as these depend on drug treatment, levels of natural and acquired immunity, etc., which is modeled as hosts belonging to different strata, i.e., a host falls into class k with probability α_k. The "absolute" frequency of haplotype i within a host belonging to class k with MOI m, which was super-infected m_j times with haplotype j ($j = 1, \ldots, H$), is $\frac{m_i}{m} W^{(k)}_{\boldsymbol{m},i}$. The fitness $W^{(k)}_{\boldsymbol{m},i}$ is interpreted as the expected number of gametocyte descendants of one copy of an infecting sporozoite haplotype i in a host within class k at the time of a mosquito bite (see Sect. 5 for a comprehensive discussion).

Notably, mathematically less-skilled readers are encouraged to read through this section but shall feel free to skip the formulae or proceed directly to Sect. 3.

2.1.3 Host-Vector Transmission

At its blood meal, a mosquito ingests a fraction γ of female and male gametocytes, proportional to the haplotype frequencies within the host. (An equal sex ratio is assumed in the description, but this assumption is not necessary.) Hence, a mosquito picks up $\gamma \frac{m_i}{m} W^{(k)}_{\boldsymbol{m},i}$ male and female haplotype-i gametocytes from a host in class k, which was super-infected by m parasites containing exactly m_i i-haplotypes.

2.1.4 Sporogonic Cycle

In the gut of the mosquito, gametocytes release gametes, between which recombination occurs immediately after the blood meal (see Fig. 1). Particularly, only parasites descending from the same host might recombine (see Fig. 3). If the mosquito took its blood meal from a host in stratum k originally infected with m_i haplotypes i ($i = 1, \ldots, H$, $|\boldsymbol{m}| = m$), the probability that a female j-gamete fertilizes a male l-gamete is

$$\frac{\gamma \frac{m_j}{m} W^{(k)}_{\boldsymbol{m},j}}{\gamma W^{(k)}_{\boldsymbol{m}}} \cdot \frac{\gamma \frac{m_l}{m} W^{(k)}_{\boldsymbol{m},l}}{\gamma W^{(k)}_{\boldsymbol{m}}} = \frac{m_j W^{(k)}_{\boldsymbol{m},j} m_l W^{(k)}_{\boldsymbol{m},l}}{m^2 {W^{(k)}_{\boldsymbol{m}}}^2},$$

where

$$\gamma W^{(k)}_{\boldsymbol{m}} := \gamma \sum_{i=1}^{H} \frac{m_i}{m} W^{(k)}_{\boldsymbol{m},i}$$

is the frequency of parasite haploids in the gut of the mosquito. The above probability is the relative frequency of j-haplotypes times that of l-haplotypes within the mosquito's gut. Multiplying the above with $\gamma W^{(k)}_{\boldsymbol{m}}$ gives the absolute number of such fertilizations, i.e.,

$$\gamma \frac{m_j W^{(k)}_{\boldsymbol{m},j} m_l W^{(k)}_{\boldsymbol{m},l}}{m^2 W^{(k)}_{\boldsymbol{m}}}.$$

Therefore, the absolute frequency of sporozoite haplotype i in the mosquito population descending from gametocytes from hosts belonging to stratum k that were infected with m parasites is

$$P_i^{*(k,m)} = \sum_{|\boldsymbol{m}|=m} \binom{m}{\boldsymbol{m}} \boldsymbol{P}^{\boldsymbol{m}} \sum_{j,l=1}^{H} \frac{\gamma m_j W^{(k)}_{\boldsymbol{m},j} m_l W^{(k)}_{\boldsymbol{m},l}}{m^2 W^{(k)}_{\boldsymbol{m}}} R(jl \to i),$$

where $R(jl \rightarrow i)$ is the probability that fusion of male and female gametes of haplotypes j and l, respectively, results in an offspring sporozoite of haplotype i.

2.1.5 Sporozoite Population

Averaged over the MOI distributions and all strata, the absolute frequency of haplotype i in the next generation's sporozoite population is

$$P_i^* = \gamma \sum_{k=1}^{\infty} \alpha_k \sum_{m=1}^{\infty} \kappa_m^{(k)} \sum_{j,l=1}^{H} f_{j,l}^{(m,k)} R(jl \rightarrow i),$$

where

$$f_{j,l}^{(m,k)} = \sum_{|\boldsymbol{m}|=m} \binom{m}{\boldsymbol{m}} \boldsymbol{P}^{\boldsymbol{m}} \frac{m_j m_l W_{\boldsymbol{m},j}^{(k)} W_{\boldsymbol{m},l}^{(k)}}{m^2 W_{\boldsymbol{m}}^{(k)}}.$$

The relative frequency of haplotype i in the next generation's sporozoite population is thus

$$P_i' = \frac{P_i^*}{\sum_{j=1}^{H} P_j^*}. \tag{2}$$

Formulating the dynamics in terms of generation time t yields

$$P_i(t+1) = \frac{P_i^*(t)}{\sum_{j=1}^{H} P_j^*(t)} \tag{3a}$$

with

$$P_i^*(t) = \gamma \sum_{k=1}^{\infty} \alpha_k^{(t)} \sum_{m=1}^{\infty} \kappa_m^{(k,t)} \sum_{j,l=1}^{H} f_{j,l}^{(m,k,t)} R(jl \rightarrow i) \tag{3b}$$

and

$$f_{j,l}^{(m,k,t)} = \sum_{|\boldsymbol{m}|=m} \binom{m}{\boldsymbol{m}} \boldsymbol{P}(t)^{\boldsymbol{m}} \frac{m_j m_l W_{\boldsymbol{m},j}^{(k)} W_{\boldsymbol{m},l}^{(k)}}{m^2 W_{\boldsymbol{m}}^{(k)}}, \tag{3c}$$

where the potential time dependence of the parameters is indicated by the superscript t.

Note that it is not necessary to model fitness time-dependently, reflecting, e.g., changes in treatment policies, as this is subsumed by the change in probability of belonging to the different strata over time.

If the distribution of MOI remains constant over time, vector competence and transmission intensities are implicitly assumed to remain constant in each stratum. The distribution of MOI, $(\kappa_m^{(k)})$, varies across strata. By assuming the same distribution of MOI in all strata, i.e., by replacing $(\kappa_m^{(k)})$ by (κ_m), it is implicitly assumed that the pattern of infection is the same across all strata, i.e., in statistical terms the distributions (κ_m) and (α_k) are independent.

2.2 MOI Distributions

As mentioned earlier, typical examples for the distribution of MOI are the conditional Poisson or conditional negative-binomial distributions.

First, consider only one stratum. MOI following a Poisson distribution can be regarded as a null model. It emerges from assuming infective mosquito bites to be rare and independent events and hosts having the same (small) probability of becoming infected. Hence, within a given time window, the number of superinfections follows a Poisson distribution. As only infected hosts are considered, one needs to condition on at least one infection, resulting in the conditional Poisson distribution

$$\kappa_m = \frac{1}{\mathrm{e}^{\lambda}-1}\frac{\lambda^m}{m!}, \text{ where } \lambda > 0 \text{ and } m = 1, 2, \ldots. \tag{4}$$

Its probability generating function, to be used later, is given by

$$U(t) := \mathrm{E}(t^m) = \sum_{m=1}^{\infty} \kappa_m t^m = \frac{1}{\mathrm{e}^{\lambda}-1}\sum_{m=1}^{\infty}\frac{(t\lambda)^m}{m!} = \frac{\mathrm{e}^{t\lambda}-1}{\mathrm{e}^{\lambda}-1}. \tag{5}$$

The function $U(t)$ can be used to calculate the mean and variance of MOI, since $U(\mathrm{e}^t)$ is the moment generating function. Namely, by differentiation at $t = 0$ one obtains

$$\mathrm{E}\, m = \frac{\lambda\, \mathrm{e}^{\lambda}}{\mathrm{e}^{\lambda}-1} \tag{6a}$$

and

$$\mathrm{Var}\, m = \frac{\mathrm{e}^{\lambda}\left(\mathrm{e}^{\lambda}-\lambda-1\right)\lambda}{\left(\mathrm{e}^{\lambda}-1\right)^2}. \tag{6b}$$

Hosts being infected with the same probability (heterogeneous hosts) might not always be justified. Groups of hosts might experience different biting rates, etc. Considering several strata, with MOI following conditional Poisson distributions in each of them, MOI follows a mixture of Poisson distributions, which is no longer a Poisson distribution. The mean MOI is

$$\mathrm{E}\, m = \sum_{k=1}^{\infty} \alpha_k \frac{\lambda_k \, \mathrm{e}^{\lambda_k}}{\mathrm{e}^{\lambda_k} - 1} \tag{7a}$$

while its variance becomes

$$\mathrm{Var}\, m = \sum_{k=1}^{\infty} \alpha_k \frac{\lambda_k (\lambda_k + 1) \, \mathrm{e}^{\lambda_k}}{\mathrm{e}^{\lambda_k} - 1} - \left(\sum_{k=1}^{\infty} \alpha_k \frac{\lambda_k \, \mathrm{e}^{\lambda_k}}{\mathrm{e}^{\lambda_k} - 1} \right)^2 . \tag{7b}$$

Clearly, the model is hopelessly over-parameterized if the number of strata is large and MOI follows a different Poisson distribution in every stratum.

An alternative to modeling discrete strata is to assume a continuum of strata. In the limit of infinitely many groups, with the Poisson parameters being Gamma-distributed across groups, the negative-binomial distribution emerges. (Note that the negative-binomial distribution has alternative definitions that are less general than the one as a Gamma–Poisson mixture.) The conditional negative-binomial distribution is defined as

$$\kappa_m = \frac{\Gamma(\eta + m)}{(1 - \nu^\eta) m! \Gamma(\eta)} (1 - \nu)^m \nu^\eta, \text{ where} \nu \in (0, 1) \text{and} \eta \in \mathbb{R}^+. \tag{8}$$

Its probability generating function (cf. Appendix 1 for a derivation) is given by

$$U(t) = \left(\left(1 - t(1 - \nu) \right)^{-\eta} - 1 \right) \frac{\nu^\eta}{1 - \nu^\eta}. \tag{9}$$

Mean and variance of MOI are given by

$$\mathrm{E}\, m = \frac{\eta(1 - \nu)}{\nu \, (1 - \nu^\eta)} \tag{10a}$$

and

$$\mathrm{Var}\, m = \frac{\eta(\nu - 1) \left((\eta + 1)\nu^\eta - \eta \nu^{\eta+1} - 1 \right)}{\nu^2 \, (\nu^\eta - 1)^2}. \tag{10b}$$

Assuming a negative-binomial distribution allows reducing the number of MOI parameters as compared to a mixture of Poisson distributions. However, the negative-binomial assumption does not remedy the need to model different strata, as also fitness differs among groups. Anyhow, fewer strata need to be modeled.

2.3 Haplotype Frequencies and Prevalences

The dynamics of malaria derived above can be adapted for evolutionary-genetic analyses as discussed in the following sections. Clearly, the changes in haplotype frequencies are the quantities of interest. However, in malaria many considerations revolve around clinical and epidemiological aspects. For instance, in the case of drug resistance the frequency change of resistance-conferring haplotype frequencies over time is irrelevant for the medical sector. Of interest is the probability of infections with haplotypes causing delayed treatment success or even failure. Similarly, epidemiologists are more interested in the current exposure to potential resistance-conferring haplotypes on a spatial scale. Obviously, the evolutionary dynamics are intertwined with epidemiological and clinical observations, and it is crucial to understand the connection between quantities of interest in different fields. This leads to the definition of "haplotype prevalence."

Haplotype i's prevalence is the probability of "observing" this haplotype in an infection, more precisely, the probability of an infection carrying this haplotype. If every infection was infected by exactly one haplotype, frequency and prevalence would coincide. However, super-infections occur, distorting this argument.

Therefore, the prevalence of haplotype j, i.e., the probability of being infected with j, is derived in Appendix 2 to be

$$q_j := 1 - \sum_{k=1}^{\infty} \alpha_k U_k (1 - P_j). \tag{11}$$

This implies that haplotype j's prevalence always exceeds its frequency (if super-infections occur with positive probability; cf. Appendix 2). The above is a rather general expression, which simplifies for particular choices of MOI distributions.

The simplest assumption is MOI following the same conditional Poisson distribution in all strata. In this case, prevalences become

$$q_j = \frac{1 - \mathrm{e}^{-\lambda P_j}}{1 - \mathrm{e}^{-\lambda}}. \tag{12}$$

Assuming the same negative-binomial distribution across strata yields

$$q_j = \frac{1 - \nu^{\eta}(\nu - \nu P_j + P_j)^{-\eta}}{1 - \nu^{\eta}}. \tag{13}$$

The relationship between frequency and prevalence is illustrated in Fig. 5. Prevalence exceeding frequency is obvious, with the amount of excess being mediated by the mean MOI (cf. Eqs. (6a) and (10a)). Whereas frequency and prevalence are similarly in low transmission areas (mean MOI $\sim$1), their discrepancy is striking in high-transmission areas (mean MOI $>$2), where a frequency of 0.8 leads to a prevalence of almost 0.95.

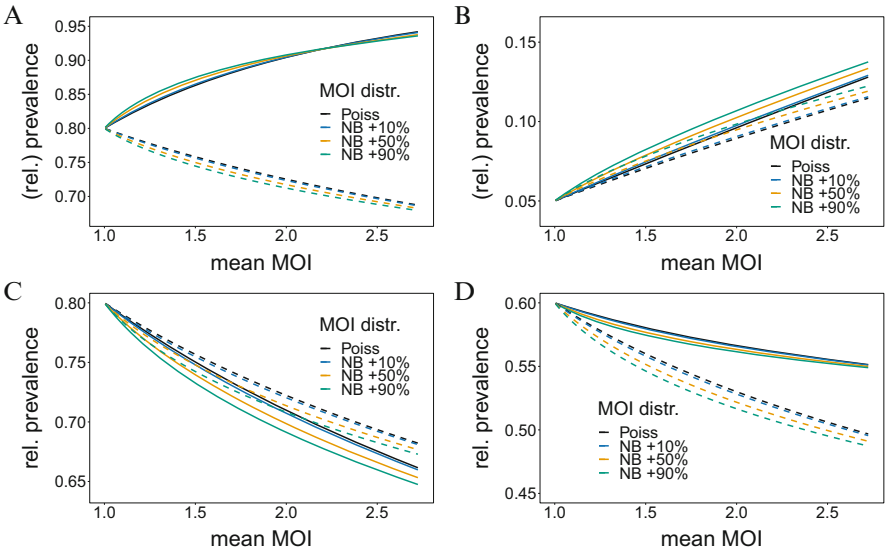


Fig. 5 (**a**, **b**) Prevalence vs. frequency: shown are the prevalence (solid) and relative prevalence (dashed) of haplotype 1 corresponding to frequency $P_1 = 0.8$ (**a**) and $P_1 = 0.05$ (**b**) as a function of mean MOI following either (6a) or (10a). Prevalence depends only on the frequency P_1 and the MOI distribution. To derive relative prevalences, which depends additionally on the whole haplotype distribution, a second haplotype with frequencies $P_2 = 1 - P_1$ was assumed. Colors correspond to different MOI distributions. The conditional Poisson and negative-binomial distributions were used, such that the latter has 10, 50, and 90% higher variance than the Poisson distribution given by (6b) or (10b). (**c**, **d**) Similar as in panels (**a**, **b**) but shown are only relative prevalences corresponding to P_1. (**c**) Solid lines for frequency distribution $\boldsymbol{P} = (0.8, 0.19, 0.01)$, dashed lines for $\boldsymbol{P} = (0.8, 0.1, 0.1)$. (**d**) Solid lines for distribution $\boldsymbol{P} = (0.6, 0.4)$, dashed lines for $\boldsymbol{P} = (0.6, 0.2, 0.2)$

This exposes the problems of confusing frequency with prevalence, particularly in the context of drug resistance, in which the chances to contain resistance might be underestimated when the two quantities are treated synonymously. Notably, comparison between low- and high-transmission areas should always be done with respect to frequency rather than prevalence.

Sometimes prevalences are easier to estimate than frequencies. In this case frequencies are sometimes estimated as relative prevalences, i.e., as

$$x_j := \frac{q_j}{\sum_{l=1}^{H} q_l}. \tag{14}$$

These estimates are far from optimal, because they lack a clear interpretation, and depend not only on the actual haplotype frequency and MOI distributions, but also on the prevalences of all other haplotypes. Figure 5a, b shows that although

prevalences always exceed frequency, relative prevalences might exceed or fall below frequency. While the relation between frequency and prevalence can be deduced quite generally from (11), the relationship between relative prevalences and frequency is unclear because it depends on the haplotype distribution as a whole as illustrated in Fig. 5c, d. Relative prevalences are particular unfortunate frequency estimates in the cases of seasonal malaria. Namely, assuming that haplotype frequencies (P_j) remain constant, changes in MOI due to dry and wet seasons result in oscillations in x_j (cf. also Sects. 3 and 4). Such changes were reported, e.g., in [1], at several mutations at the drug-resistance-associated genes *Pfmdr1*, *Pfcrt*, *Pfdhfr*, *Pfdhps* in *P. falciparum*. Some of the seasonal fluctuations were reported to be significant, with a trend of higher relative prevalence in the dry season, but a reversed trend for a low-frequency mutation at *Pfdhps*. These observations can be conveniently explained by Fig. 5 and attributed to changes in MOI without the need to hypothesize about the underlying evolutionary process.

Finally, it should be mentioned that haplotype prevalence must not be confused with disease prevalence. Disease prevalence is the relative frequency of infected individuals within the host population, hence diseases prevalence is a feature of the human population. Haplotypes' prevalences are attributed only to the population of currently infected individuals, whereas haplotype frequencies are properties of the parasite population.

3 Evolutionary Dynamics: Selection for Drug Resistance

The framework of the last section can be adapted to model selection for drug resistance as in [29, 30]. The following derivations assume absence of intra-host competition. Consequently, the fitness of haplotype i in stratum k is independent of the distribution of parasites super-infecting the host, i.e., $W^{(k)}_{\boldsymbol{m},i} = W^{(k)}_i$.

3.1 Selection at a Single Locus

Further assume that alleles conferring drug resistance are segregating only at the first locus, while all other loci are neutral. Assume that n alleles are segregating at the selected locus, where each allele codes for a different level of drug resistance, as it occurs, for instance, at the *Pfdhfr* or *Pfdhps* loci (where point mutations at, respectively, codons 51, 59, 108, and 164, or 436, 437, 540, 581, and 613 act synergistically to increasing levels of resistance [cf. 5, 8, 16, 21, 33]. Hence, the number of possible haplotypes is a multiple of n, i.e., $H = nL$. By appropriate labeling, haplotypes $(u-1)L+1, \ldots, uL$ carry the same allele at the resistance-conferring locus ($u = 1, \ldots, L$). The fitnesses of these haplotypes are denoted by $w^{(k)}_u$. The frequency of the u-th allele at the resistance-conferring locus is

$$p_u = \sum_{i=(u-1)L+1}^{uL} P_i .$$

As shown in Appendix 3 the evolutionary dynamics are given by

$$p_u(t) = \frac{w_u^t p_u(0)}{\sum_{v=1}^{n} w_v^t p_v(0)}, \quad \text{where} w_v = \sum_{k=1}^{\infty} \alpha_k w_v^{(k)} \tag{15}$$

is the average fitness of allele v across all strata.

This frequency dynamics depend on average fitnesses and initial frequencies, but they are independent of MOI and hence transmission intensities (in a deterministic model without intra-host competition). Importantly, generation time t is counted in transmission cycles not in calendar years. In high-transmission areas many more cycles occur per year. Thus, if, e.g., 20 transmission cycles (holo-endemic) occur instead of 4 (seasonal malaria), the same selection pressures lead to a five times faster real-time spread of resistance. In addition, stochastic events, e.g., random drift and bottlenecks, affect the evolutionary dynamics differently in low *vs.* high-transmission areas. Namely, the probability of initial occurrence and spread of resistance depend on parasite population size, which correlates with transmission intensities and hence also MOI. Furthermore, fitnesses vary across transmission areas, as levels of host-acquired and natural immunity (modeled by the distribution α_k) as well as treatment policies (modeled by $w_j^{(k)}$) change.

The dynamics (15) are illustrated in Fig. 6 for $n = 4$ alleles, corresponding to a drug-sensitive wildtype and accumulated point mutations as it occurs in the *Pfdhfr* or *Pfdhps* genes conferring SP resistance. Resistance is associated with metabolic costs, c, while the probability to clear parasites, d, is depleted by a factor $1 - s$ for each point mutation. Both costs and delayed parasite clearance act multiplicatively. Hence, fitnesses are given by

$$w_u^{(1)} = (1-c)^{u-1} \quad \text{and} \quad w_u^{(2)} = \big((1-d(1-s)^{u-1}\big)(1-c)^{u-1}$$

resulting in

$$w_u = \alpha_1(1-c)^{u-1} + (1-\alpha_1)\big(1-d(1-s)^{u-1}\big)(1-c)^{u-1}. \tag{16}$$

Note that in reality mutations occur at different time points in the evolutionary process, and initially only one copy of the mutation occurs. In the dynamics (15) this can be approximated by choosing the initial frequencies sufficiently small, so they reach the desired frequency at the appropriate time. The limiting behavior of (15) is straightforward: the mutation with the largest selective advantage will ultimately become fixed. However, in drug-resistance evolution, one is typically concerned with ongoing evolutionary processes rather than those in the past, which emphasizes

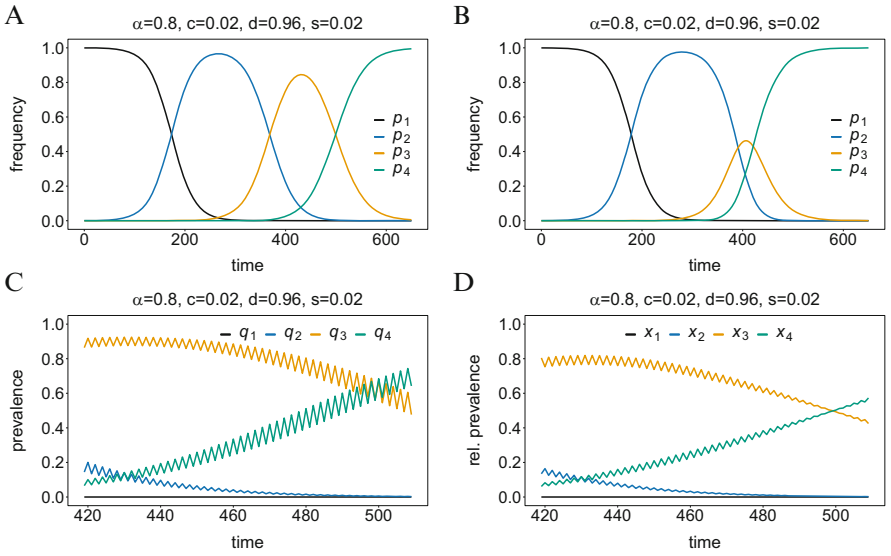


Fig. 6 (**a**, **b**) Evolutionary dynamics (15) for one selected locus with one sensitive and three resistance-conferring alleles as described in the text. Parameters are at the top of the plot panels. The initial frequencies of the mutant alleles were $p_2(0) = 5 \times 10^{-4}$, $p_3(0) = 3 \times 10^{-10}$ and $p_4(0) = 1 \times 10^{-17}$ (A) and $p_2(0) = 4 \times 10^{-4}$, $p_3(0) = 10^{-10}$ and $p_4(0) = 5 \times 10^{-17}$ (**b**). (**c**) Prevalences, according to (12), corresponding to the dynamics in panel (**a**) for a narrower time window, assuming MOI to follow a Poisson distributions with seasonally changing parameters, corresponding to dry ($\lambda = 0.5$) and wet ($\lambda = 1.5$) seasons. (**d**) Like (**c**) but for relative prevalences according to (14)

both qualitative and quantitative behavior of (15). They depend not only on selection coefficients, which will differ across endemic regions, but also on the random process underlying the occurrence of mutations. Fig. 6a, b illustrate two realizations of (15) with the same selection coefficients but different initial frequencies, resulting in quite different dynamics. From Fig. 6 it becomes clear that some mutations will never reach high frequencies, if mutations conferring stronger resistance already spread. The observation depends on the selective pressures, the time points at which the mutations occur and the time point of the cross-sectional survey. In reality stochastic effects will further confound the dynamics. Therefore, the observation of differences in frequency distributions from cross-sectional samples from separate endemic areas must not per se be attributed to different selective environments.

Regarding *Pfdhfr* or *Pfdhps*, in the empirical literature different mutant frequency distributions across endemic areas are mentioned, e.g., McCollum et al. [16]. These reflect cross-sectional snapshots at different times in the evolutionary processes in various locations. Additionally, the evolutionary processes exhibit quantitative differences because mutations occur randomly at different time points, resulting in changes as seen when comparing Fig. 6a with b.

3.2 Frequency vs. Prevalence in a Time-Dynamical Context

The importance of distinguishing between frequency and prevalence was already mentioned. In a time-dynamical rather than a cross-sectional context this becomes even more important. Malaria-endemic areas are often characterized by dry and wet seasons, with epidemic peaks during the latter. This obviously depends on the increase in mosquito population size and hence transmission during the rainy seasons. The change in transmission intensities leads to periodic changes in the MOI distribution. While this is not affecting the evolutionary dynamics of the spread of drug resistance, it affects the prevalence dynamics (cf. Fig. 6).

As an example, assume an endemic region with alternating rainy and dry malaria seasons. Suppose further, MOI follows a conditional Poisson distribution among all strata with parameters λ_r and λ_d in the rainy and dry seasons, satisfying $\lambda_d < \lambda_r$. The evolutionary dynamics and corresponding prevalence dynamics (for a shorter time range) are shown in Fig. 6. While the frequency dynamics in Fig. 6a, b are relatively straightforward, the prevalence dynamics of Fig. 6c show oscillating behavior, emphasizing that frequencies are the correct quantities to consider on an evolutionary scale. Importantly, the amount of fluctuations depends on the allele frequencies. Fig. 6 also shows corresponding relative prevalences, x_j given by (14). These quantities show fluctuations, which are less pronounced (compare Figs. 6c with d). The relative prevalences underestimate the dominating allele frequency. While the prevalence, i.e., the clinical occurrence, exceeds 80%, relative prevalence stays below 80%. Also the oscillatory trend with higher values in the rainy and lower in the dry season is reversed for the dominating alleles (this changes around generation 5000), while the trend is identical for alleles at lower frequency. Such oscillations have been observed in the empirical literature (e.g., [1, 3, 7, 20]). Particularly, the observations in [1] discussed in Sect. 2.3 coincide with the patterns of Fig. 6c, d. Sometimes even statistical test were employed in the literature to support seasonal fluctuations. This is unfortunate, because the excess of oscillations depend on the underlying frequencies, and might even vanish (see Fig. 6d around generation 500), besides that the problem of multiple testing is encountered. The correct way to test for seasonal fluctuations would be to estimate the MOI parameter across seasons and test for significant differences as described in [27].

3.3 Genetic Hitchhiking

To study the effect of genetic hitchhiking, besides the resistance-associated locus, it suffices to consider one additional neutral locus. As a simplifying assumption only two alleles are segregating at the selected locus, one sensitive and one resistance-conferring, denoted by A_1 and A_2, whereas L alleles $(B_1, \ldots, B_L)$ are segregating at the neutral locus. (In the case of drug resistance these are interpreted as a drug-sensitive wildtype and a resistance-conferring mutant allele.) This setup yields

$2L$ haplotypes, $A_1B_1, \ldots A_1B_L, A_2B_1, \ldots, A_2B_L$, labeled by $1, \ldots, 2L$. Hence, haplotypes $1, \ldots, L$ are called wildtype haplotypes, while $L+1, \ldots, 2L$ are referred to as mutant haplotypes. Only the selected locus determines fitness. Since the absence of intra-host competition is assumed, i.e., $W^{(k)}_{\boldsymbol{m},i} = W^{(k)}_i$, it follows that haplotypes $1, \ldots, L$ (wildtypes) and $L+1, \ldots, 2L$ (mutants) have the same fitness, denoted by $w^{(k)}_1$ and $w^{(k)}_2$, respectively, i.e., $W^{(k)}_i = w^{(k)}_1$ for $i = 1, \ldots, L$ and $W^{(k)}_i = w^{(k)}_2$ for $i = L+1, \ldots, 2L$.

The frequencies of the neutral allele B_i conditioned on a wildtype or mutant background are, respectively,

$$R_i = \frac{P_i}{\sum_{j=1}^{L} P_j} \quad \text{and} \quad Q_i = \frac{P_{L+i}}{\sum_{j=1}^{L} P_{L+j}}. \tag{17}$$

The dynamics of $Q_i(t)$ are derived in Appendix 4 (cf. Eq. (A.4)). In the limit $t \to \infty$, $Q_i(t)$ approaches its limit $\hat{Q}_i$ (given by Eq. (A.5) in Appendix 4), which is approximated by

$$\hat{Q}_i \approx R_i(0) - \big(R_i(0) - Q_i(0)\big) p_0^{\frac{r\tilde{\vartheta}}{w_2(\log w_2 - \log w_1)}} \tag{18}$$

(cf. Appendix 6).

From the equilibrium allele frequencies $\hat{Q}_i$ the expected relative equilibrium heterozygosity, i.e., the ratio of the expected equilibrium heterozygosity and the initial heterozygosity, is calculated in Appendix 5 (cf. Eq. (A.6)), which is approximated by

$$\tilde{\mathscr{H}}(r) = 1 - p_0^{\frac{2r\tilde{\vartheta}}{w_2(\log w_2 - \log w_1)}}, \tag{19}$$

where

$$\tilde{\vartheta} := \sum_{k=1}^{\infty} \alpha_k \sum_{m=1}^{\infty} \kappa^{(k)}_m \tilde{\vartheta}^{(k,m)} \tag{20}$$

is an inbreeding adjustment (cf. Appendix 6). In the above equation the denominator in the exponent corresponds to the strength of selection. By normalization, fitnesses can be rewritten as $w_1 = 1$ and $w_2 = 1 + s$, where s is the selective advantage of the resistance-conferring allele. Typically s is small ($s \in [0, 0.1]$), such that $w_2(\log w_2 - \log w_1) \approx s$. Hence,

$$\tilde{\mathscr{H}}(r) \approx 1 - p_0^{2r\tilde{\vartheta}/s}. \tag{21}$$

This has exactly the same form as the heterozygosity in a standard hitchhiking model, except for the inbreeding adjustment $\tilde{\vartheta}$. Since, only parasites can recombine

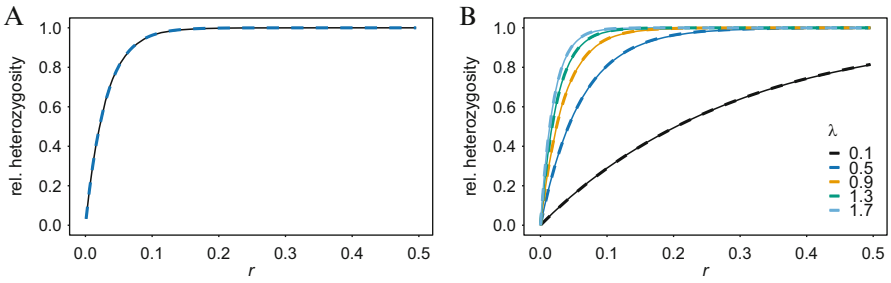


Fig. 7 Expected relative heterozygosities as functions of the recombination rate. Solid lines show exact solution, dashed lines approximations. (**a**) Three strata are assumed with fitness parameters $W_{11} = 1$, $W_{21} = 0.99$, $W_{12} = 0.5$, $W_{22} = 0.51$, $W_{13} = 0.5$, $W_{23} = 0.65$, $\alpha_1 = 0.15$, $\alpha_2 = 0.35$, $\alpha_1 = 0.5$, $\lambda_1 = 0.7$, $\lambda_2 = 1.7$, $\lambda_3 = 1.2$, and $p_0 = 10^{-4}$. (**b**) Same parameters as in (**a**), but with $\lambda = \lambda_1 = \lambda_2 = \lambda_3$, as shown in the legend

that exited the same host, mating between parasites (in the total parasite population) is not random (it is random within each mosquito). Single-clone infections ($m_k = m$ for some k, mainly due to MOI $= 1$) lead effectively to no recombination (recombining identical clones). MOI is therefore mediating the amount of recombination, with higher average MOI corresponding to more recombination (cf. Fig. 3). The factor $\tilde{\vartheta}$ is difficult to interpret as it involves fitness parameters and MOI. Particularly, this factor shows that the process of selection and recombination cannot be separated in malaria, a fact inherited from the transmission cycle.

Figure 7 illustrates the hitchhiking effect for different MOI distributions. Strikingly, for low MOI the effect of reduce equilibrium heterozygosity can be genomewide. This observation coincides with more pronounced hitchhiking patterns in areas of low and intermediate transmission, compared with high transmission [cf. 14–16]. Interpreting more pronounced hitchhiking patterns in low/intermediate transmission areas as stronger selection could be incorrect, as this is more likely attributed to different levels of transmission. Hence, estimates for transmission intensities in terms of MOI are also important in correctly interpreting population-genetic statistics.

4 Parameter Estimation

The population-genetic framework introduced here is parameterized in a way that allows to estimate model parameters from molecular data and detect potential genes under selection. Molecular data is obtained by taking blood samples from N infected malaria patients, and assaying alleles at resistance-associated and neutral loci of the parasites found in the blood samples. Typically several SNPs at resistance-associated (selected) loci are assayed (e.g., at codons 51, 59, 108 at *Pfdhfr* or 436, 437, 613 at *Pfdhps*), as well as microsatellites (STR) at positions that will and

will not be affected by genetic hitchhiking, i.e., at positions flanking resistance-associated loci and at different chromosomes.

4.1 Distribution of Classes

The distribution of the different classes α_k can be estimated from census data and data on the implementation of healthcare policies and transmission interventions. First, the classes need to be specified, e.g., age classes, particular population or risk groups (e.g., pregnant women), treatment types, rural vs. urban populations, etc. These estimates can be refined by household surveys asking patients about transmission interventions, e.g., the use of insecticide-threaded bednets, the frequency of indoor-residual spraying, compliance with chemotherapy and travel records.

4.2 Estimating MOI and Frequency Spectra

The distribution of MOI, $\kappa_m^{(k)}$, can be estimated for every class k. This requires molecular data from the respective sub-populations. As the estimation procedure will be the same for each stratum, it suffices to consider a single class and the superscript k can be waived.

For rare and independent infections, the null model for MOI is the conditional Poisson distribution (4) (cf. Sect. 2.2), characterized by a single parameter λ. The aim is to estimate the Poisson parameter λ. Notably, if two sub-populations are mixed up, with MOI following Poisson distributions with different parameters, MOI in the joint population does not follow a Poisson distribution any longer (cf. Eq. (7) in Sect. 2.2)—a fact to be considered in the sampling design.

Schneider and Escalante [27] and Schneider [25] refined the maximum-likelihood approach by Hill and Babiker [10] to estimate λ from a single molecular marker. Assume n alleles (SNPs, haplotypes in a short non-recombining region, STR repeats) at the locus with frequencies $p_1, \ldots, p_n$, collectively denoted by $\boldsymbol{p}$.

In each blood sample, the absence and presence of the n alleles can be assayed (Fig. 4). Clearly, several alleles might be present in a sample due to super-infections. Note that the presence of only a single allele does not imply MOI $= 1$, as the patient might have been super-infected several times with the same allele (Figs. 3 and 4). Clearly, at least one allele must be present in each sample. In practice, however, the molecular assay might fail resulting in missing data. Here, missing at random is assumed, and N is the number of proper samples (no missing data). Let N_k denote the number of blood samples that contain allele k. (Note that $\sum_{k=1}^{n} N_k \geq N$). Here, N_k/N is the empirical prevalence of allele N_k. According to Schneider and Escalante [27], $(N_1, \ldots, N_n)$ are a sufficient statistic for the molecular data (absence and presence of alleles per sample). The log-likelihood function is given by

$$L = L(\lambda, \boldsymbol{p}) = L(\lambda, \boldsymbol{p}|N_1, \ldots, N_n) = -N \log(\mathrm{e}^{\lambda} - 1) + \sum_{k=1}^{n} N_k \log(\mathrm{e}^{\lambda p_k} - 1)\,. \tag{22}$$

Except in the cases $\sum_{k=1}^{n} N_k = N$ (i.e., no sign of super-infections) or $N_k = N$ for some k (one allele is present in all samples) it was shown in Schneider and Escalante [27] that the maximum-likelihood estimate exists and is unique. There is no explicit formula but it can be calculated by a fast-converging iteration. Namely, except in irregular situation, the maximum-likelihood estimate (MLE) $\hat{\boldsymbol{\theta}} = (\hat{\lambda}, \hat{\boldsymbol{p}})$ is given by

$$\hat{p}_k = -\frac{1}{\hat{\lambda}} \log\left(1 - \frac{N_k}{N}(1 - \mathrm{e}^{-\hat{\lambda}})\right), \tag{23a}$$

where $\hat{\lambda}$ is found by iterating

$$\lambda_{t+1} = \lambda_t - \frac{\lambda_t + \sum_{k=1}^{n} \log\left(1 - \frac{N_k}{N}(1 - \mathrm{e}^{-\lambda_t})\right)}{1 - \sum_{k=1}^{n} \frac{N_k}{N\,\mathrm{e}^{\lambda_t} - N_k(\mathrm{e}^{\lambda_t} - 1)}}, \tag{23b}$$

which converges monotonically and at quadratic rate from any initial value $\lambda_1 \geq \hat{\lambda}$. Hence, it is guaranteed to find the MLE as long as the initial value λ_1 is chosen to be sufficiently large.

As shown in Schneider [25], the observed haplotype prevalences, N_k/N, are the expected prevalences given the MLE, and the estimator has the typical desirable asymptotic properties, i.e., it is asymptotically unbiased, strongly consistent and efficient. Moreover, also the finite sample properties are good. Namely, frequency estimates are almost unbiased, $\hat{\lambda}$ has very low bias for $N \geq 150$, and the Cramér–Rao lower-bounds (i.e., the minimal possible variance among unbiased estimators; cf. [25] for an explicit formula) is a good approximation for the estimator's variance. This bound can be employed for sample size determination to achieve given precision goals. Furthermore, it is relatively robust against model violations. For a discussion on profile-likelihood and asymptotic confidence intervals as well as statistical tests, the reader is referred to Schneider and Escalante [28]. A convenient R script to calculate MOI alongside a comprehensive description is available in the supplemental materials of Schneider [25].

As an application, random samples of size $N = 150$ were generated according to the Poisson model with parameters $\lambda = 1$, and six frequency distributions taken from the dynamics in Fig. 6a, at generations $t = 400$, 410, 420, 430, 440, and 450. The true frequency distributions, the prevalence counts N_k, and the MLE are summarized in Table 1.

While the frequency estimates are sufficiently accurate, the true MOI parameter tends to be overestimated, except in two cases. This is due to the relatively skewed frequency distributions (cf. [25]).

Table 1 Frequency distributions $p_1, \ldots, p_4$ of the dynamics of Fig. 6a at generations $t = 400 + \tau_l$ ($l = 1, \ldots, 6$). Shown are prevalence counts for a sample of size 150, randomly generated from the conditional Poisson model, the MLE, i.e., $\hat{p}_k$ and $\hat{\lambda}$. The true MOI parameter is $\lambda = 1$

l	τ_l	p_1	p_2	p_3	p_4	N_1	N_2	N_3	N_4	$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$	$\hat{p}_4$	$\hat{\lambda}$
1	0	2.05×10^{-5}	0.333	0.453	0.214	0	62	86	47	0	0.312	0.461	0.228	0.911
2	10	9.14×10^{-6}	0.230	0.461	0.309	0	58	94	50	0	0.274	0.496	0.231	1.073
3	20	3.79×10^{-6}	0.148	0.438	0.414	0	32	95	72	0	0.141	0.504	0.355	1.063
4	30	1.48×10^{-6}	0.089	0.390	0.520	0	22	78	95	0	0.096	0.396	0.508	1.026
5	40	5.44×10^{-7}	0.051	0.329	0.620	0	6	64	107	0	0.029	0.341	0.631	0.742
6	50	1.92×10^{-7}	0.028	0.266	0.706	0	9	62	121	0	0.035	0.279	0.686	1.251

Next, it is shown how such estimates can be employed to infer selection coefficients.

4.3 Estimating Selection Coefficients

The evolutionary dynamics (15) yield the frequency ratio of alleles k and l at time t to be

$$\frac{p_k(t)}{p_j(t)} = \frac{w_k^t p_k(0)}{w_j^t p_j(0)},$$

or

$$\log \frac{p_k(t)}{p_j(t)} = t \log \frac{w_k}{w_j} + \log \frac{p_k(0)}{p_j(0)}.$$

Particularly, if fitness is scaled such that $w_1 = 1$, the above reduces to

$$\log \frac{p_k(t)}{p_1(t)} = a_k + t b_k \tag{24a}$$

with

$$a_k := \log \frac{p_k(0)}{p_1(0)} \tag{24b}$$

and

$$b_k := \log \frac{w_k}{w_1} = \log w_k, \tag{24c}$$

which reveals the linear relationship of the log-ratio of allele frequencies and time. Hence, the coefficients a_k and b_k are transformations of the frequencies at the reference time point $t = 0$ and the fitnesses, respectively. Clearly, the parameters a_k and b_k can be estimated by a linear regression. This idea was already pursued by Nair et al. [19] and McCollum et al. [16] without formalizing the argument.

Assume allele frequency estimates from T cross-sectional samples at time points $\tau_1, \ldots, \tau_T$ are available. The frequency estimates are obtained as described in Sect. 4.2. Let

$$y_{k,l} := \log \frac{\hat{p}_k(\tau_l)}{\hat{p}_1(\tau_l)} \tag{24d}$$

denote the log-ratio of the frequencies of alleles k and 1 ($k = 1, \ldots, n$) at the l-th time point, collectively denoted by the $T \times (n-1)$-matrix $\boldsymbol{Y}$. Let

$$\boldsymbol{B} = \begin{pmatrix} a_2 & \ldots & a_n \\ b_2 & \ldots & b_n \end{pmatrix} \quad \text{and} \quad \boldsymbol{X} = \begin{pmatrix} 1 & \tau_1 \\ \vdots & \vdots \\ 1 & \tau_T \end{pmatrix}.$$

Then, the linear model

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{\varepsilon} \tag{25}$$

emerges. A least-square fit yields the solution

$$\hat{\boldsymbol{B}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}. \tag{26}$$

By denoting the column vectors of $\boldsymbol{X}$ by $\mathbf{1}$ and $\boldsymbol{\tau}$, this is more concretely rewritten as

$$\hat{a}_j = \frac{\|\boldsymbol{\tau}\|_2^2\langle \mathbf{1}|\boldsymbol{y}_j\rangle - \langle \boldsymbol{\tau}|\mathbf{1}\rangle\langle \boldsymbol{\tau}|\boldsymbol{y}_j\rangle}{T\|\boldsymbol{\tau}\|_2^2 - \langle \boldsymbol{\tau}|\mathbf{1}\rangle^2} = \frac{\sum_{l=1}^{T}\tau_l^2\sum_{l=1}^{T}y_{l,j} - \sum_{l=1}^{T}\tau_l\sum_{l=1}^{T}\tau_l y_{l,j}}{T\sum_{l=1}^{T}\tau_l^2 - \left(\sum_{l=1}^{T}\tau_l\right)^2} \tag{27a}$$

and

$$\hat{b}_j = \frac{T\langle \boldsymbol{\tau}|\boldsymbol{y}_j\rangle - \|\boldsymbol{\tau}\|_2^2\langle \mathbf{1}|\boldsymbol{y}_j\rangle}{T\|\boldsymbol{\tau}\|_2^2 - \langle \boldsymbol{\tau}|\mathbf{1}\rangle^2} = \frac{T\sum_{l=1}^{T}\tau_l y_{l,j} - \sum_{l=1}^{T}\tau_l^2\sum_{l=1}^{T}y_{l,j}}{T\sum_{l=1}^{T}\tau_l^2 - \left(\sum_{l=1}^{T}\tau_l\right)^2}. \tag{27b}$$

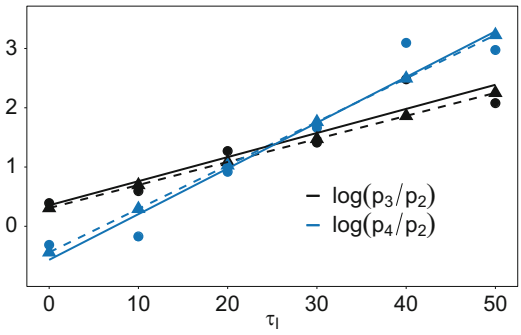


Fig. 8 Linear regression to estimate fitness parameters from six cross-sectional samples according to Table 1. Dots show estimates, triangle exacts values. Point $\tau_1 = 0$ corresponds to generation $t = 400$ of the dynamics of Fig. 6a

The fitnesses are then estimated as $\hat{w}_j = \operatorname{Exp} \hat{b}_j$ with $w_1 = 1$. Such point estimates do not properly address uncertainty in the data. Confidence intervals can be readily constructed using bootstrap techniques. Clearly, more advanced approaches might be chosen to estimate fitness.

Estimating fitness parameters for each stratum, i.e., $w_j^{(k)}$ will be difficult. By appropriate sample design, investigations can be restricted to homogeneous population strata.

As an illustration take the frequency estimates from Sect. 4.2 summarized in Table 1. Clearly, the first allele (sensitive allele) was not observed in the data, and its frequency is hence estimated as 0. This is a common observation in real applications. Thus, the dependent variable is defined as $y_{k,l} := \log \frac{\hat{p}_{k+1}(\tau_l)}{\hat{p}_2(\tau_l)}$, and the ratios $\frac{w_3}{w_2} = 1.0396$ and $\frac{w_4}{w_2} = 1.0761$ are estimated. One obtains the estimates $\frac{\hat{w}_3}{\hat{w}_2} = 1.0415$ and $\frac{\hat{w}_4}{\hat{w}_2} = 1.0801$, being very close to the true values (see Fig. 8).

5 Fitness

So far fitnesses were unspecified model parameters. To ultimately control the evolutionary dynamics of drug resistance, it is necessary to influence fitness parameters. To do so, more involved considerations regarding fitness are indispensable. For simplicity, a single locus under selection with a sensitive and a resistance allele is considered. The average fitnesses w_1 and w_2 are the likelihoods of sensitive and resistant sporozoites, respectively, to cause an infection, which leads to gametocytes offspring picked up by a new mosquitoes, in which sporozoite offspring is produced. In fact fitnesses are a convolution of events during the malaria life-cycle, called fitness components (cf. [26]). In particular fitness depends on the average merozoite and gametocyte dynamics, referred to as parasitemia and gametocytemia.

Following Schneider and Escalante [26], it is convenient to contrast *P. falciparum* and *P. vivax* while assuming (qualitatively) the same (hypothetical) parasitemia. While these two most relevant human malaria pathogens are more or less similar

they also have important differences in their life history. In particular, gametogenesis is initiated with the erythrocytic cycle's onset in *P. vivax*, while it occurs in *P. falciparum* with a delay of at least 6.5 days [4, 11, 17]. Moreover longevity of gametocytes differs among species with approximate life spans of 2 days in *P. vivax* as compared with 14–48 days in *P. falciparum*. Furthermore, *P. vivax* infections involve liver-stage hypnozoites that can cause malaria relapses. In the end, these differences in fitness components directly affect fitness and the evolutionary dynamics, and explain the slow spread of *P. vivax* compared with *P. falciparum*. (A detailed model description is found in Schneider and Escalante [26], here only a brief summary is presented.)

Suppose average merozoite dynamics are given by the Holding Type III function convoluted with an exponential decay function modeling drug treatment,

$$P_*^{(*)}(\tau) = \frac{P_{\max}^{(*)}}{1 + \xi \rho_*^{-\tau}} \, \mathrm{e}^{-\delta_* \max\{0, \tau - \tau_0\}}, \tag{28}$$

where τ is the time, measured in days from the erythrocytic cycle's onset, with τ_0 being the onset of drug treatment. Furthermore, the asterisk in the superscript is a placeholder for "f" or "v" referring to *P. falciparum* or *P. vivax*, respectively, while that in the subscript is a placeholder for 1 or 2 referring to the sensitive or resistant allele, $P_{\max}^{(*)}$ is the threshold parasitemia, ρ_* the growth rate of parasitemia, ξ a common scaling factor, and δ_* is the parasite clearance rate imposed by drug treatment. These dynamics are illustrated in Fig. 9a.

These dynamics are purely hypothetical and model a population average. It can be replaced by more realistic dynamics. Importantly, *P. falciparum* and *P. vivax* differ only quantitatively but not qualitatively, namely in the maximum parasite counts, $P_{\max}^{(f)}$ and $P_{\max}^{(v)}$, respectively. However, the dynamics of sensitive and resistant parasites differ. Typically, resistant is associated with metabolic costs and an associated slower replication of merozoites, reflected here by $\rho_2 < \rho_1$. But, sensitive parasites are cleared faster than resistant ones, resulting in $\delta_2 < \delta_1$. In the illustration in Fig. 9a, the differences between average intra-host parasitemia dynamics of sensitive and resistant haplotypes are rather subtle.

Parasitemia dynamics determine the clinical episode of the diseases, but the derived gametocytes dynamics are responsible for transmission. The latter are derived from the former. Taking longevity of gametocytes, denoted by $\beta^{(*)}$, and the onset of gametogenesis, $\tau_1^{(*)}$, into account, the number gametocytes are derived from accumulating (integrating) gametocytes being produced proportionally to parasitemia, by a constant $\varphi^{(*)}$, namely by

$$G_*^{(*)}(\tau) = \int_{\max\{0, \tau - \beta^{(*)}\}}^{\tau} \varphi^{(*)} P_*^{(*)}(z) g(z - \tau_1^{(*)}) dz, \tag{29}$$

where $g(z)$ is a function to smooth out the discontinuity in the number of gametocytes being produced. It is defined by

$$g(z) := \begin{cases} 0 & \text{for} \quad z < 0, \\ 64z^3(1-z)^3 & \text{for} \quad 0 \le z < 0.5, \\ 1 & \text{for} \quad z \ge 0.5. \end{cases}$$

Gametocytemia for resistant and sensitive *P. falciparum* and *P. vivax* haplotypes derived from their respective parasitemia dynamics are illustrated in Fig. 9b. While the latter are qualitatively identical, their derived gametocytemia dynamics are very different across species for two reasons. First, the onset of gametogenesis is earlier in *P. vivax*, and there are more gametocytes derived from sensitive parasites than from resistant ones—a trend that reverses after drug treatment starts. In *P. falciparum* gametogenesis starts approximately at the same time as drug treatment, thus—in contrast to *P. vivax*—only the resistant parasites' advantage is captured by gametocyte dynamics. Second, longevity of gametocytes is different among malaria species. This amplifies the effect of later gametogenesis in *P. falciparum*, namely the time span in which resistant *P. vivax* parasites are in excess of sensitive ones is much shorter than in *P. falciparum*.

The probability of transmitting malaria depends on the abundance of gametocytes, where the number of gametocytes need to exceed a threshold for successful transmission. Hence, the fitness ratio w_2/w_1 corresponds to the ratio of the areas under the curve of resistant and sensitive gametocyte dynamics, i.e.,

$$\frac{\int_{a_2^{(*)}}^{b_2^{(*)}} G_2^{(*)}(\tau)d\tau}{\int_{a_1^{(*)}}^{b_1^{(*)}} G_1^{(*)}(\tau)d\tau}, \tag{30}$$

where $a_*^{(*)}$ and $b_*^{(*)}$ are the time intervals in which gametocyte count surpasses the transmission threshold (chosen to be 10 parasites per μL in the numerical example). This ratio is different for *P. falciparum* and *P. vivax*. As clearly visible from Fig. 9b that it is higher for *P. falciparum*. Numerical integration of the dynamics from Fig. 9b yields 1.0671 for the fitness ratio in *P. falciparum* and 1.00332 in *P. vivax*, i.e., a twenty times stronger selective advantage for resistance-conferring mutations in *P. falciparum*. (Note that these numbers are irrelevant as the model is just mechanistically.) Plugging these ratios into the evolutionary dynamics (15) clearly leads to a the much faster spread of resistance in *P. falciparum*. In fact, the spread is ten times faster (cf. Fig. 9c). While this argument reflects the worldwide observation of drug-resistant *P. vivax* parasites being absent or rare [18, 22, 32, 34, 35], the concrete dynamics of Fig. 9c are just an illustration. Not

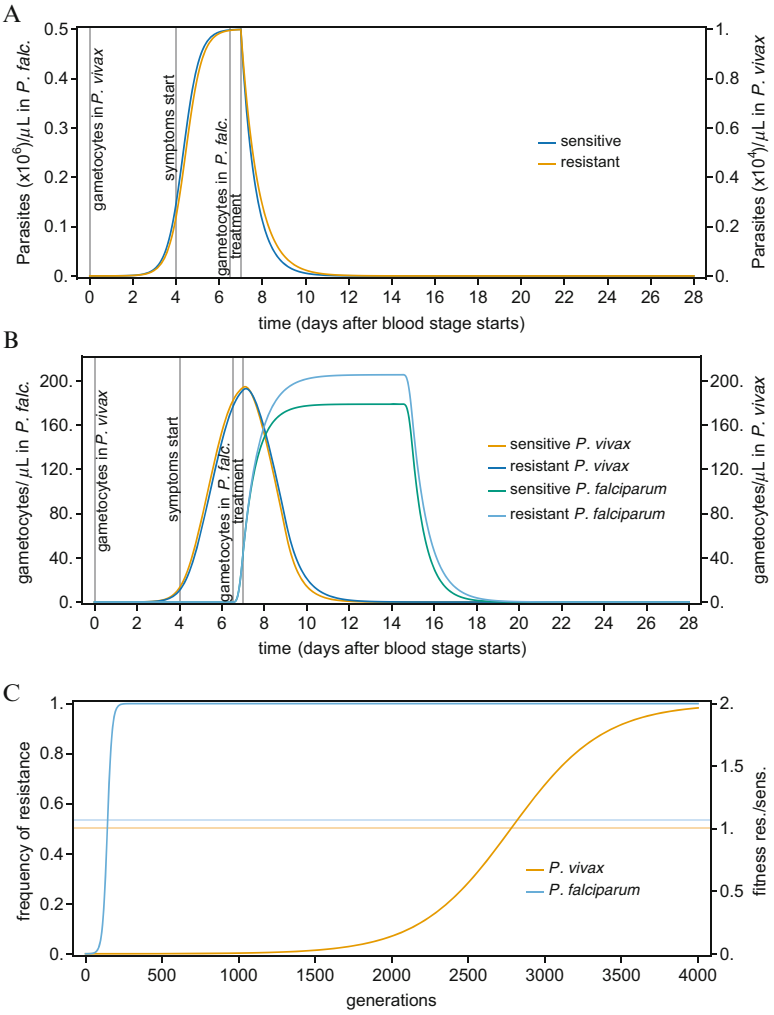


Fig. 9 Illustration of parasitemia and gametocytemia for *P. falciparum* and *P. vivax* according to (28) (**a**), (29) (**b**). Parameters are $P_{\max}^{(v)} = 10^4/\mu L$, $P_{\max}^{(f)} = 5 \times 10^5/\mu L$, $\xi = 10^5$, $\rho_1 = 14.2857$, $\rho_2 = 13.3333$, $\delta_1 = 1.5$, $\delta_2 = 1.25$, $\tau_0 = 7$, $\tau_1^{(v)} = 0$, $\tau_1^{(f)} = 6.5$, $\varphi^{(v)} = 0.01$, $\varphi^{(f)} = 0.0004$, $\beta^{(v)} = 2$, and $\beta^{(f)} = 8$. (**c**) The resulting fitness rations were used in the model (15) with $p_2(0) = 10^{-4}$. The vertical lines indicate fitness values

only are the dynamics and hence the fitnesses not derived from empirical evidence, also the initial frequencies are somewhat arbitrary. In fact, the malaria species will have very different population sizes, hence initial frequencies of the mutants, being the reciprocal of the population sizes, will be different. Anyhow the mechanistic model is derived from common knowledge and species-specific differences in the

parasite life history, and explains the puzzling phenomenon of different speeds of drug-resistance evolution in a simple manner.

6 Discussion

Since Darwin's influential book "On the Origin of Species" [6] evolutionary theory and the modern evolutionary synthesis in the 1920s, evolutionary genetics evolved into a flourishing discipline. The advances in malaria research since Ronalds Ross's first mathematical model [24]—which most of his contemporaries did not understand—are not less striking. However, the combination of both, evolutionary genetics and mathematical modeling of malaria, has not yet reached its full potential. While mathematical models of malaria are widespread, tailored population-genetic models and their application to empirical data remain the exception. The level of sophistication in research on standard population-genetic organisms like *Drosophila*, to which theoretical knowledge and statistical methods are tailored to, has not been reached the field of malaria. Namely, rather than tailoring population-genetic theory to the characteristics of *Plasmodium*, standard text-book methods are incorrectly applied.

Here, a general population-genetic framework to study the evolutionary genetics of malaria was introduced, generalizing the work of Schneider and Kim [29, 30]. The model itself is applicable to all malaria species, and considers heterogeneity in the host population, in transmission intensities, and selective regimes, notably, in a time-dependent way. However, the full generality of the model was not exploited fully here. Particularly, there was a focus on drug-resistance evolution, assuming selection at a single locus, a restriction, which is per se not necessary. Other evolutionary processes such as the spread of HRP2/3 deletions can be readily modeled. However, hypnozoites are not explicitly incorporated, which might render applications to *P. vivax* and *P. ovale* incorrect. Anyhow, the effect of relapses occurring from hypnozoites can be approximated by host strata with different MOI distributions.

The restriction to drug resistance being mediated by a single locus is both illustrative and informative. It serves as a null model to understand this evolutionary process. It was shown—however, for a more general setting than in Schneider and Kim [29], who restricted themselves to two alleles—that the dynamics at the selected locus (15) are independent of MOI and hence transmission intensities. However, the real-time spread depends on transmission intensities, because the number of generations per year scales with transmission. Indeed the dynamics (15) are equivalent to the standard haploid selection model. These dynamics motivated the procedure to estimate selection coefficients by a linear regression (27).

The independence of the dynamics from transmission intensities vanishes if a second selected locus or even a neutral one is considered. Here, the effect of genetic hitchhiking was studied in a more general way than in Schneider and Kim [30]. The hitchhiking effect, i.e., the reduction in expected heterozygosity, was

recursively derived (A.6) and an accurate approximation (21) established. The latter has a similar form as in the standard hitchhiking model of Maynard Smith and Haigh [13]. However, recombination is reduced by an inbreeding adjustment (20), which is a complex convolute of MOI and selection coefficients. Particularly, the process of selection and recombination cannot be decoupled in malaria. Importantly, recombination is restricted in malaria, because only haplotypes exiting the same host can recombine. Single-clone infections lead to effectively no recombination. If selection is sufficiently strong and transmission is low, so that multi-clone infections are the exception, the effect of genetic hitchhiking can be genomewide (cf. Fig. 7). However, if transmission is not extremely low, the initial levels of genetic variation are restored at unlinked loci, justifying the study of selection at a single locus for drug resistance (because resistance-conferring loci are typically located on different chromosomes).

The absence of intra-host competition is crucial for the independence of the evolutionary dynamics at the selected locus and MOI. However, the importance of intra-host competition might be overestimated. The observation of fluctuating frequencies of resistance-conferring mutations in areas of seasonal malaria, with alternating wet and dry seasons, led to the hypothesis of intra-host parasite competition playing an important role (e.g., [1, 3, 7, 20]). Namely, the counterintuitive observation of lower frequencies/prevalences of resistance-conferring mutations during the rainy seasons, during which malaria prevalence was clearly and correctly shown to increase, led to the rational of higher intra-host competition between resistance and sensitive parasites in patient with high MOI. However, these studies used relative prevalences as frequency estimates (notably not all studies explain how they derived their estimates). As shown here, the observations are just pathologies arising from relative prevalences, which are a highly problematic statistic, particularly because they might severely underestimate the probability of infections carrying resistant alleles (cf. Fig. 6). Hence, relative prevalences, which lack a clear interpretation, are problematic in an epidemiological or clinical context as well as in an evolutionary-genetic one. Prevalences would have a more intuitive behavior, with an increase in the rainy and a decrease in the dry seasons (cf. Fig. 6). However, in an evolutionary context, haplotype frequencies are the relevant quantities. These considerations put the relevance of intra-host competition into perspective. Notably, further explanations for temporally changing "frequencies" include different drug pressures due to seasonality. Also this hypothesis is questionable, given the simple explanation that improper statistics were used. Moreover, the importance of intra-host competition could be overestimated from molecular assays, which might behave differently in the presence of multiple haplotypes in an infection—with parasites competing in the molecular assays rather than the host.

As a further advantage the framework allows to estimate MOI and allele (haplotype) frequencies. This was exemplified by reporting on a maximum-likelihood (ML) approach. However, it can be replaced by moment estimations or Bayesian methods. Anyhow, the ML method has the usual desirable asymptotic properties and also excellent finite sample properties as shown in Schneider [25]. These results can also be used in study design to determine sample size. The MLE (assuming a Pois-

son model) has an intuitive interpretation: the observed allele/haplotype prevalences coincide with the true prevalences given the MLE was the true parameter vector. Statistical tests for differences in MOI between seasons can be found in Schneider and Escalante [27]. Testing for differences in the MOI parameter is more appropriate than testing for differences in prevalences or even relative prevalences between seasons. The reason is that the problem of multiple testing is not encountered.

Finally, the puzzling observation of the much faster spread of drug resistance in *P. falciparum* compared with *P. vivax* was explained. Namely, even if both species experience the same hypothetical drug pressure, differences in parasites' life histories, especially the onset of gametogenesis and longevity of gametocytes, lead to a more efficient translation of drug pressures in *P. falciparum*. Namely, merozoite clearance rates in infections must not be mistaken with fitness. They are just one of many fitness components that intertwine to shape fitness. The mechanistic explanation is tempting as it is solely derived from obvious differences in the two species. However, other factors not considered here presumably also affect the evolutionary observations. For instance, the presence of hypnozoites in *P. vivax* act like a seed-bank, which will slow down the evolutionary process. Furthermore, *P. vivax* is more prevalent in areas of low transmission, in which *P. falciparum* cannot be sustained. Therefore, *P. vivax* likely has fewer generation per year than *P. falciparum*, which renders any evolutionary process slower in real time.

Because fitnesses enter the framework only as parameters, any intra-host model or convolution of fitness components can be decoupled from the evolutionary process. If necessary, they can be deduced from separate models and used as plug-ins in the evolutionary framework.

In summary, the framework introduced here provides a formal foundation for evolutionary-genetic research on malaria. Some of the pitfalls often encountered in the literature were discussed—particularly pathologies, which are iconified as relevant features, whose unfortunate mechanistic explanations resemble the geocentric model of the universe. In the Olymp of malaria research the time has finally arrived for Charles Darwin to meet Ronald Ross.

Acknowledgments This work was funded by the German Academic Exchange (Project-ID 57417782) and the SMWK-SAB project "Innovationsvorhaben zur Profilschärfung an Hochschulen für angewandte Wissenschaften" (Project number 100257255). As this chapter is a compilation of research of the last years, the author is thankful for the many fruitful discussion with friends and colleagues that made this work possible. Particularly, the mistakes found by Aliou Bouba, Sulyman Iyanda, Pierre Ngougoue Ngougoue, and Christian Tsoungui are gratefully acknowledged as well as the positive and helpful comments of two anonymous reviewers on an earlier draft.

Appendix 1

Generating Function of the Negative-Binomial Distribution

The probability generating function is defined as

$$U(t) = \frac{\nu^\eta}{(1-\nu^\eta)\Gamma(\eta)} \sum_{m=1}^{\infty} \frac{\Gamma(\eta+m)}{m!}(t-t\nu)^m.$$

Its derivative is calculated from differentiation by component, which is justified by absolute convergence, as

$$\begin{aligned}
U(t)' &= \frac{\nu^\eta(1-\nu)}{(1-\nu^\eta)\Gamma(\eta)} \sum_{m=1}^{\infty} \frac{\Gamma(\eta+m)}{m!} m(t-t\nu)^{m-1} \\
&= \frac{\nu^\eta(1-\nu)}{(1-\nu^\eta)\Gamma(\eta)} \sum_{m=0}^{\infty} \frac{(\eta+m)\Gamma(\eta+m)}{m!}(t-t\nu)^m \\
&= \frac{\nu^\eta\eta(1-\nu)}{(1-\nu^\eta)\Gamma(\eta)} \sum_{m=0}^{\infty} \frac{\Gamma(\eta+m)}{m!}(t-t\nu)^m \\
&\quad + t(1-\nu)\frac{\nu^\eta(1-\nu)}{(1-\nu^\eta)\Gamma(\eta)} \sum_{m=1}^{\infty} \frac{\Gamma(\eta+m)}{m!} m(t-t\nu)^{m-1} \\
&= \eta(1-\nu)\left(U(t) + \frac{\nu^\eta}{1-\nu^\eta}\right) + t(1-\nu)U'(t).
\end{aligned}$$

Consequently, $U(t)$ satisfies the differential equation

$$U(t)' = \frac{\eta(1-\nu)}{\big(1-t(1-t\nu)\big)}\left(U(t) + \frac{\nu^\eta}{1-\nu^\eta}\right).$$

The solution of this differential equation satisfying $U(0) = 0$ is the generating function (9).

Appendix 2

Prevalence

The probability that a host belonging to class k is infected with haplotype H is

$$q_H^{(k)} := \sum_{m=1}^{\infty} \kappa_m^{(k)} \sum_{\substack{|\boldsymbol{m}|=m \\ m_H>0}} \binom{m}{\boldsymbol{m}} \boldsymbol{P}^{\boldsymbol{m}}.$$

By writing $\tilde{\boldsymbol{m}} = (m_1, \ldots, m_{H-1})$, $\tilde{\boldsymbol{P}} = (P_1, \ldots, P_{H-1})$, $\boldsymbol{m} = (\tilde{\boldsymbol{m}}, m_H)$ and using $|\tilde{\boldsymbol{P}}| = 1 - P_H$ this becomes

$$\begin{aligned}
q_H^{(k)} &= \sum_{m=1}^{\infty} \kappa_m^{(k)} \sum_{m_H=1}^{m} \sum_{|\tilde{\boldsymbol{m}}|=m-m_H} \binom{m}{(\tilde{\boldsymbol{m}}, m_H)} \boldsymbol{P}^{(\tilde{\boldsymbol{m}}, m_H)} \\
&= \sum_{m=1}^{\infty} \kappa_m^{(k)} \sum_{m_H=1}^{m} \sum_{|\tilde{\boldsymbol{m}}|=m-m_H} \binom{m}{m_H} \binom{m-m_H}{\tilde{\boldsymbol{m}}} \tilde{\boldsymbol{P}}^{\tilde{\boldsymbol{m}}} P_H^{m_H} \\
&= \sum_{m=1}^{\infty} \kappa_m^{(k)} \sum_{m_H=1}^{m} \binom{m}{m_H} P_H^{m_H} \sum_{|\tilde{\boldsymbol{m}}|=m-m_H} \binom{m-m_H}{\tilde{\boldsymbol{m}}} \tilde{\boldsymbol{P}}^{\tilde{\boldsymbol{m}}} \\
&= \sum_{m=1}^{\infty} \kappa_m^{(k)} \sum_{m_H=1}^{m} \binom{m}{m_H} P_H^{m_H} (1-P_H)^{m-m_H} \\
&= \sum_{m=1}^{\infty} \kappa_m^{(k)} \left(1 - (1-P_H)^m\right) \\
&= 1 - U_k(1-P_H),
\end{aligned}$$

where $U_k(x)$ denotes the probability generating function of $(\kappa_m^{(k)})$. Clearly, by relabeling, the probability of a host, belonging to class k, being infected with haplotype j is

$$q_j^{(k)} = 1 - U_k(1-P_j).$$

Therefore, the prevalence of haplotype j, i.e., the probability of being infected with j, is

$$q_j := 1 - \sum_{k=1}^{\infty} \alpha_k U_k(1-P_j).$$

Appendix 3

Selection at a Single Locus

We calculate the frequency of the first allele at the resistance-conferring locus in the next generation. It is calculated by marginalization of (2), i.e.,

$$p_1' = \frac{\sum_{i=1}^{L} P_i^*}{\sum_{i=1}^{nL} P_i^*}.$$

Clearly, one obtains

$$\sum_{i=1}^{L} P_i^* = \sum_{k=1}^{\infty} \alpha_k \sum_{m=1}^{\infty} \kappa_m^{(k)} \sum_{i=1}^{L} P_i^{*(k,m)}.$$

Particularly, the last sum is

$$\sum_{i=1}^{L} P_i^{*(k,m)} = \sum_{i=1}^{L} \sum_{j,l=1}^{nL} f_{j,l}^{(m,k)} R(j,l \to i) = \sum_{j,l=1}^{nL} f_{j,l}^{(m,k)} \sum_{i=1}^{L} R(j,l \to i).$$

Because $R(j,l \to i) = 0$ for $i \le L$ and $j,l > L$, and $\sum_{i=1}^{L} R(j,l \to i) = 1$ for $i, j \le L$ or $i,l \le L$, by using the definition of $W_{\boldsymbol{m}}^{(k)}$, the above can be rewritten as

$$\begin{aligned}\sum_{i=1}^{L} P_i^{*(k,m)} &= \sum_{j=1}^{L} \sum_{l=1}^{nL} f_{j,l}^{(m,k)} = \sum_{|\boldsymbol{m}|=m} \binom{m}{\boldsymbol{m}} \boldsymbol{P}^{\boldsymbol{m}} \frac{1}{m^2 W_{\boldsymbol{m}}^{(k)}} \sum_{j=1}^{L} m_j W_j^{(k)} \sum_{l=1}^{nL} m_l W_l^{(k)} \\ &= \sum_{|\boldsymbol{m}|=m} \binom{m}{\boldsymbol{m}} \boldsymbol{P}^{\boldsymbol{m}} \frac{1}{m} \sum_{j=1}^{L} m_j W_j^{(k)} = \frac{w_1^{(k)}}{m} \sum_{j=1}^{L} \sum_{|\boldsymbol{m}|=m} m_j \binom{m}{\boldsymbol{m}} \boldsymbol{P}^{\boldsymbol{m}}.\end{aligned}$$

Using the property that m_j is binomially distributed with parameters m and P_j, this reduced to

$$\sum_{i=1}^{L} P_i^{*(k,m)} = \frac{w_1^{(k)}}{m} \sum_{j=1}^{L} \mathrm{E}(m_j) = \frac{w_1^{(k)}}{m} \sum_{j=1}^{L} m P_j = w_1^{(k)} \sum_{j=1}^{L} P_j = w_1^{(k)} p_1. \tag{A.1}$$

This yields

$$\sum_{i=1}^{L} P_i^* = \sum_{k=1}^{\infty} \alpha_k \sum_{m=1}^{\infty} \kappa_m w_1^{(k)} p_1$$
$$= \sum_{k=1}^{\infty} \alpha_k w_1^{(k)} p_1$$
$$= w_1 p_1,$$

where w_1 is the fitness of the first allele at the selected locus averaged over all strata.

The frequencies of the other alleles are calculated similarly, so one obtains

$$p'_u = \frac{w_u p_u}{\sum_{v=1}^{n} w_v p_v}, \tag{A.2a}$$

where

$$w_v = \sum_{k=1}^{\infty} \alpha_k w_v^{(k)} \tag{A.2b}$$

is the average fitness of allele v across all strata.

Thus, the frequency change from one generation to the next is given by

$$p_u(t+1) = \frac{w_u p_u(t)}{\sum_{v=1}^{n} w_v p_v(t)} \tag{A.3a}$$

implying the explicit solution is given by (15).

Appendix 4

Hitchhiking

If r denotes the recombination rate between the two loci, the absolute frequency of haplotype i ($i = 1, \ldots, L$) descending from hosts in stratum k, which were superinfected m times is

$$P_i^{*(k,m)} = \sum_{j,l=1}^{2L} f_{j,l}^{(m,k)} R(jl \to i)$$

$$
\begin{aligned}
&= \sum_{l=1}^{L} f_{i,l}^{(m,k)} + r \sum_{l=1}^{L} f_{i+L,l}^{(m,k)} + (1-r) \sum_{l=1}^{L} f_{i,l+L}^{(m,k)} \\
&= \sum_{l=1}^{2L} f_{i,l}^{(m,k)} + r \Bigg(\sum_{l=1}^{L} f_{i+L,l}^{(m,k)} - \sum_{l=1}^{L} f_{i,l+L}^{(m,k)} \Bigg) \\
&= \gamma \sum_{|\boldsymbol{m}|=m} \binom{m}{\boldsymbol{m}} \boldsymbol{P}^{\boldsymbol{m}} \frac{1}{m^2 W_{\boldsymbol{m}}^{(k)}} \Bigg[m_i w_1^{(k)} \sum_{l=1}^{2L} m_l W_l^{(k)} \\
&\quad + r w_1^{(k)} w_2^{(k)} \Bigg(m_{i+L} \sum_{l=1}^{L} m_l - m_i \sum_{l=1}^{L} m_{l+L} \Bigg) \Bigg] \\
&= \gamma w_1^{(k)} \sum_{|\boldsymbol{m}|=m} \binom{m}{\boldsymbol{m}} \boldsymbol{P}^{\boldsymbol{m}} \frac{m_i}{m} \\
&\quad + \gamma r w_1^{(k)} w_2^{(k)} \sum_{|\boldsymbol{m}|=m} \binom{m}{\boldsymbol{m}} \boldsymbol{P}^{\boldsymbol{m}} \frac{1}{m^2 W_{\boldsymbol{m}}^{(k)}} \Bigg(m_{i+L} \sum_{l=1}^{L} m_l - m_i \sum_{l=1}^{L} m_{l+L} \Bigg).
\end{aligned}
$$

The first term simplifies to $\gamma w_1^{(k)} P_i$, because m_i is multinomial distributed with parameters m and $\boldsymbol{P}$.

Let $\boldsymbol{P}_1 = (P_1, \ldots, P_L)$, $\boldsymbol{P}_2 = (P_{L+1}, \ldots, P_{2L})$, $\boldsymbol{m}_1 = (m_1, \ldots, m_L)$, and $\boldsymbol{m}_2 = (m_{L+1}, \ldots, m_{2L})$. Then, the above simplifies to

$$
\begin{aligned}
P_i^{*(k,m)} &= \gamma w_1^{(k)} P_i + \gamma r \frac{w_1^{(k)} w_2^{(k)}}{m} \sum_{u=0}^{m} \binom{m}{u} \frac{1}{w_1^{(k)} u + w_2^{(k)} (m-u)} \\
&\quad \times \sum_{|\boldsymbol{m}_1|=u} \binom{u}{\boldsymbol{m}_1} \boldsymbol{P}_1^{\boldsymbol{m}_1} \sum_{|\boldsymbol{m}_2|=m-u} \binom{m-u}{\boldsymbol{m}_2} \boldsymbol{P}_2^{\boldsymbol{m}_2} \big(m_{i+L} u - m_i (m-u) \big) \\
&= \gamma w_1^{(k)} P_i + \gamma r \frac{w_1^{(k)} w_2^{(k)}}{m} \sum_{u=1}^{m-1} \binom{m}{u} \frac{1}{w_1^{(k)} u + w_2^{(k)} (m-u)} \\
&\quad \times \sum_{|\boldsymbol{m}_1|=u} \binom{u}{\boldsymbol{m}_1} \boldsymbol{P}_1^{\boldsymbol{m}_1} \sum_{|\boldsymbol{m}_2|=m-u} \binom{m-u}{\boldsymbol{m}_2} \boldsymbol{P}_2^{\boldsymbol{m}_2} \big(m_{i+L} u - m_i (m-u) \big) \\
&= \gamma w_1^{(k)} P_i + \gamma r \frac{w_1^{(k)} w_2^{(k)}}{m} \sum_{u=1}^{m-1} \binom{m}{u} \frac{1}{w_1^{(k)} u + w_2^{(k)} (m-u)} \\
&\quad \times \Bigg(\sum_{|\boldsymbol{m}_1|=u} \binom{u}{\boldsymbol{m}_1} \boldsymbol{P}_1^{\boldsymbol{m}_1} u \sum_{|\boldsymbol{m}_2|=m-u} \binom{m-u}{\boldsymbol{m}_2} \boldsymbol{P}_2^{\boldsymbol{m}_2} m_{i+L}
\end{aligned}
$$

$$
\begin{aligned}
&- \sum_{|\boldsymbol{m}_2|=m-u} \binom{m-u}{\boldsymbol{m}_2} \boldsymbol{P}_2^{\boldsymbol{m}_2} (m-u) \sum_{|\boldsymbol{m}_1|=u} \binom{u}{\boldsymbol{m}_1} \boldsymbol{P}_1^{\boldsymbol{m}_1} m_i \Bigg) \\
&= \gamma w_1^{(k)} P_i + \gamma r \frac{w_1^{(k)} w_2^{(k)}}{m} \sum_{u=1}^{m-1} \binom{m}{u} \frac{1}{w_1^{(k)} u + w_2^{(k)} (m-u)} \\
&\quad \times \Bigg(\sum_{|\boldsymbol{m}_1|=u} \binom{u}{\boldsymbol{m}_1} \boldsymbol{P}_1^{\boldsymbol{m}_1} u(m-u) \frac{P_{i+L}}{p_2} p_2^{m-u} \\
&\quad - \sum_{|\boldsymbol{m}_2|=u} \binom{m-u}{\boldsymbol{m}_2} \boldsymbol{P}_2^{\boldsymbol{m}_2} (m-u) u \frac{P_i}{p_1} p_1^u \Bigg) \\
&= \gamma w_1^{(k)} P_i + \gamma r \frac{w_1^{(k)} w_2^{(k)}}{m} \sum_{u=1}^{m-1} \binom{m}{u} \frac{u(m-u) p_1^u p_2^{m-u}}{w_1^{(k)} u + w_2^{(k)} (m-u)} \left(\frac{P_{i+L}}{p_2} - \frac{P_i}{p_1} \right) \\
&= \gamma w_1^{(k)} P_i - \gamma r w_1^{(k)} w_2^{(k)} \sum_{u=1}^{m-1} \binom{m-2}{u-1} \frac{p_1^u p_2^{m-u} (m-1)}{w_1^{(k)} u + w_2^{(k)} (m-u)} (R_i - Q_i) \\
&= \gamma w_1^{(k)} P_i - \gamma r p_1 p_2 (R_i - Q_i) \vartheta_{\boldsymbol{p},m}^{(k)},
\end{aligned}
$$

where

$$
\vartheta_{\boldsymbol{p},m}^{(k)} := \sum_{u=0}^{m-2} \binom{m-2}{u} p_1^u p_2^{m-2-u} \frac{(m-1) w_1^{(k)} w_2^{(k)}}{w_1^{(k)} (u+1) + w_2^{(k)} (m-1-u)}.
$$

Thus,

$$
P_i^* = \gamma w_1 P_i - \gamma r p_1 p_2 (R_i - Q_i) \vartheta_{\boldsymbol{p}},
$$

where

$$
w_1 = \sum_{k=1}^{\infty} \alpha_k \sum_{m=1}^{\infty} \kappa_m^{(k)} w_1^{(k)} = \sum_{k=1}^{\infty} \alpha_k w_1^{(k)}
$$

is the average fitness of allele A_1 and

$$
\vartheta_{\boldsymbol{p}} = \sum_{k=1}^{\infty} \alpha_k \sum_{m=1}^{\infty} \kappa_m^{(k)} \vartheta_{\boldsymbol{p},m}^{(k)}.
$$

By using $\sum_{i=1}^{L} R_i = \sum_{i=1}^{L} Q_i = 1$, the above simplifies to

$$\sum_{i=1}^{L} P_i^* = \gamma w_1 \sum_{i=1}^{L} P_i - \gamma r p_1 p_2 \vartheta_{\boldsymbol{p}} \sum_{i=1}^{L} (R_i - Q_i) = \gamma w_1 p_1,$$

implying

$$R_i' = R_i - r p_2 \frac{\vartheta_{\boldsymbol{p}}}{w_1} (R_i - Q_i).$$

Similarly, it is shown that

$$Q_i' = Q_i - r p_1 \frac{\vartheta_{\boldsymbol{p}}}{w_2} (Q_i - R_i).$$

Substraction gives

$$R_i' - Q_i' = (R_i - Q_i)\Big(1 - r\Big(\frac{p_1}{w_1} + \frac{p_2}{w_2}\Big)\vartheta_{\boldsymbol{p}}\Big).$$

Hence,

$$\begin{aligned} R_i(t) - Q_i(t) &= (R_i(0) - Q_i(0)) \prod_{\tau=0}^{t-1} \left(1 - r\Big(\frac{p_1(\tau)}{w_1} + \frac{p_2(\tau)}{w_2}\Big)\vartheta_{\boldsymbol{p}(\tau)}\right) \\ &= (R_i(0) - Q_i(0)) \prod_{\tau=0}^{t-1} \left(1 - r \frac{p_1(0) w_1^{\tau+1} + p_2(0) w_2^{\tau+1}}{p_1(0) w_1^{\tau} + p_2(0) w_2^{\tau}} \frac{\vartheta_{\boldsymbol{p}(\tau)}}{w_1 w_2}\right). \end{aligned}$$

Let $q := \frac{p_2(0)}{p_1(0)}$ and $\mu := \frac{w_2}{w_1}$. Since

$$p_1(\tau)^{m-2-u} p_2(\tau)^u = \frac{w_1^{\tau(m-2-u)} p_1(0)^{m-2-u} w_2^{\tau u} p_2(0)^u}{(w_1^{\tau} p_1(0) + w_2^{\tau} p_2(0))^{m-2}} = \frac{(q\mu^{\tau})^u}{(1 + q\mu^{\tau})^{m-2}}$$

one arrives at

$$\vartheta_{\boldsymbol{p}(\tau),m}^{(k)} = \sum_{u=0}^{m-2} \binom{m-2}{u} \frac{(q\mu^{\tau})^u}{(1+q\mu^{\tau})^{m-2}} \frac{(m-1) w_1^{(k)} w_2^{(k)}}{w_1^{(k)}(u+1) + w_2^{(k)}(m-1-u)}.$$

Therefore,

$$R_i(t) - Q_i(t) = (R_i(0) - Q_i(0)) \prod_{\tau=0}^{t-1} \Lambda_{\tau}$$

with

$$\Lambda_\tau := 1 - \frac{r}{w_2}\frac{1+q\mu^{\tau+1}}{1+q\mu^\tau}\vartheta_{\boldsymbol{p}(\tau)}$$

and

$$\vartheta_{\boldsymbol{p}(\tau)} := \sum_{k=1}^{\infty}\alpha_k \sum_{m=1}^{\infty}\kappa_m^{(k)}\vartheta_{\boldsymbol{p}(\tau),m}^{(k)}.$$

Finally, one obtains

$$Q_i(t+1) = Q_i(t) - rp_1(t)\frac{\vartheta_{\boldsymbol{p}(t)}}{w_2}(R_i(t) - Q_i(t)) \tag{A.4a}$$

$$= Q_i(0) + r\frac{Q_i(0) - R_i(0)}{w_2}\sum_{\tau=0}^{t}\frac{1+q\mu^{\tau+1}}{1+q\mu^\tau}\vartheta_{\boldsymbol{p}(\tau)}\prod_{l=0}^{\tau-1}\Lambda_l. \tag{A.4b}$$

In the limit $t \to \infty$ this becomes

$$\hat{Q}_i = Q_i(0) - \big(Q_i(0) - R_i(0)\big)A(r), \tag{A.5a}$$

where

$$A(r) = \frac{r}{w_2}\sum_{\tau=0}^{\infty}\frac{1+q\mu^{\tau+1}}{1+q\mu^\tau}\vartheta_{\boldsymbol{p}(\tau)}\prod_{l=0}^{\tau-1}\Lambda_l. \tag{A.5b}$$

Appendix 5

Equilibrium Heterozygosity

The equilibrium heterozygosity is derived to be

$$\hat{H}_e = 1 - \sum_{i=1}^{L}\hat{Q}_i^2 = 1 - \sum_{i=1}^{L}\Big(Q_i(0) - \big(Q_i(0) - R_i(0)\big)A(r)\Big)^2.$$

Assuming a single mutation at generation 0, $Q_i(0) = 1$ with probability $R_i(0)$ and $Q_i(0) = 0$ with probability $1 - R_i(0)$. Taking the expectation yields

$$\mathrm{E}(\hat{H}_e) = \mathrm{E}\left(\mathrm{E}(\hat{H}_e|R_1(0),\ldots,R_L(0))\right)$$

$$= \mathrm{E}\left(1 - \sum_{i=1}^{L}\Big(R_i(0)\big(1 - \big(1-R_i(0)\big)A(r)\big)^2 + \big(1-R_i(0)\big)\big(R_i(0)A(r)\big)^2\Big)\right)$$

$$= \mathrm{E}\left(1-\sum_{i=1}^{L}\left(R_i(0)-2R_i(0)\big(1-R_i(0)\big)A(r)+\big(1-R_i(0)\big)R_i(0)A(r)^2\right)\right)$$

$$= \mathrm{E}\left(2A(r)-A(r)^2-2A(r)\sum_{i=1}^{L}R_i(0)^2-A(r)^2\sum_{i=1}^{L}R_i(0)^2\right)$$

$$= \big(2A(r)-A(r)^2\big)\,\mathrm{E}\left(1-\sum_{i=1}^{L}R_i(0)^2\right).$$

Hence,

$$\mathscr{H}(r) = 2A(r) - A(r)^2. \tag{A.6}$$

Appendix 6

Approximating the Hitchhiking Effect

If $p_2(0) \approx 0$, then $q = p_2(0)/p_1(0) \approx 0$, and

$$\vartheta^{(k)}_{\boldsymbol{p}(\tau),m} = \sum_{u=0}^{m-2}\binom{m-2}{u}\frac{(q\mu^\tau)^u}{(1+q\mu^\tau)^{m-2}}\frac{(m-1)w_1^{(k)}w_2^{(k)}}{w_1^{(k)}(u+1)+w_2^{(k)}(m-1-u)}$$

$$\approx \frac{(m-1)w_1^{(k)}w_2^{(k)}}{w_1^{(k)}+w_2^{(k)}(m-1)} =: \tilde{\vartheta}^{(k,m)}.$$

Furthermore, let

$$\tilde{\vartheta} := \sum_{k=1}^{\infty}\alpha_k\sum_{m=1}^{\infty}\kappa_m^{(k)}\tilde{\vartheta}^{(k,m)}.$$

Hence,

$$\Lambda_{\tau,m} \approx 1-\frac{r}{w_2}\frac{1+q\mu}{1+q\mu^\tau}\tilde{\vartheta} \approx 1-\frac{r}{w_2}\tilde{\vartheta} =: \tilde{\Lambda}.$$

Thus,

$$A(r) \approx \frac{r}{w_2}\sum_{\tau=0}^{\infty}\frac{\tilde{\vartheta}}{1+q\mu^\tau}\prod_{l=0}^{\tau-1}\tilde{\Lambda} = \frac{r\tilde{\vartheta}}{w_2}\sum_{\tau=0}^{\infty}\frac{\tilde{\Lambda}^\tau}{1+q\mu^\tau} =: \tilde{A}(r).$$

This yields the approximate equilibrium frequencies

$$\hat{Q}_i \approx Q_i(0) - (Q_i(0) - R_i(0))\tilde{A}(r), \tag{A.7}$$

which have the same form as eq. 60 in [30]. Following the calculations in section 4.2. of [30],

$$\tilde{A}(r) \approx p_0^{\frac{r\tilde{\vartheta}}{w_2(\log w_2 - \log w_1)}} \tag{A.8}$$

the above equilibrium frequencies are approximated by

$$\hat{Q}_i \approx \tilde{Q}_i := R_i(0) - (R_i(0) - Q_i(0))p_0^{\frac{r\tilde{\vartheta}}{w_2(\log w_2 - \log w_1)}} \tag{A.9}$$

and the expected relative heterozygosity at equilibrium is approximately

$$\mathscr{H}(r) \approx 2\tilde{A}(r) - \tilde{A}(r)^2 \approx 1 - p_0^{\frac{2r\tilde{\vartheta}}{w_2(\log w_2 - \log w_1)}} =: \tilde{\mathscr{H}}(r). \tag{A.10}$$

References

1. Abdel-Muhsin, A.M.A., Mackinnon, M.J., Ali, E., Nassir, E.K.A., Suleiman, S., Ahmed, S., Walliker, D., Babiker, H.A.: Evolution of drug-resistance genes in *Plasmodium falciparum* in an area of seasonal malaria transmission in Eastern Sudan. J. Infect. Dis. **189**(7), 1239–1244 (2004)
2. Antony, H.A., Parija, S.C.: Antimalarial drug resistance: an overview. Trop. Parasitol. **6**(1), 30–41 (2016)
3. Babiker, H.A., Satti, G., Ferguson, H., Bayoumi, R., Walliker, D.: Drug resistant *Plasmodium falciparum* in an area of seasonal transmission. Acta Trop. **94**(3), 260–268 (2005). https://doi.org/10.1016/j.actatropica.2005.04.007
4. Coatney, G,, Collins, W., Warren, M., Contacos, P.: The Primate Malarias. U.S. National Institute of Allergy and Infectious Diseases, Bethesda (1971)
5. Cortese, J.F., Plowe, C.V.: Antifolate resistance due to new and known *Plasmodium falciparum* dihydrofolate reductase mutations expressed in yeast. Mol. Biochem. Parasitol. **94**(2), 205–214 (1998)
6. Darwin, C.: On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. John Murray, London (1859)
7. Geiger, C., Compaore, G., Coulibaly, B., Sie, A., Dittmer, M., Sanchez, C., Lanzer, M., Jänisch, T.: Substantial increase in mutations in the genes *pfdhfr* and *pfdhps* puts sulphadoxine–pyrimethamine-based intermittent preventive treatment for malaria at risk in Burkina Faso. Trop. Med. Int. Health **19**(6), 690–697 (2014). https://doi.org/10.1111/tmi.12305
8. Gregson, A., Plowe, C.V.: Mechanisms of resistance of malaria parasites to antifolates. Pharmacol. Rev. 57(1), 117–145 (2005). https://doi.org/10.1124/pr.57.1.4
9. Han, L., Hudgens, M.G., Emch, M.E., Juliano, J.J., Keeler, C., Martinson, F., Kamthunzi, P., Tegha, G., Lievens, M., Hoffman, I.F.: RTS,S/AS01 malaria vaccine efficacy is not modified by seasonal precipitation: results from a phase 3 randomized controlled trial in Malawi. Sci. Rep. **7**(1), 7200 (2017)

10. Hill, W.G., Babiker, H.A.: Estimation of numbers of malaria clones in blood samples. Proc. R. Soc. Lond. Ser. B **262**(1365), 249–257 (1995). https://doi.org/10.1098/rspb.1995.0203
11. Howes, R.E., Battle, K.E., Mendis, K.N., Smith, D.L., Cibulskis, R.E., Baird, J.K., Hay, S.I.: Global epidemiology of *Plasmodium vivax*. Am. J. Trop. Med. Hyg. **95**(6 Suppl), 15–34 (2016)
12. Imwong, M., Suwannasin, K., Kunasol, C., Sutawong, K., Mayxay, M., Rekol, H., Smithuis, F.M., Hlaing, T.M., Tun, K.M., van der Pluijm, R.W., Tripura, R.: The spread of artemisinin-resistant *Plasmodium falciparum* in the Greater Mekong subregion: a molecular epidemiology observational study. Lancet Infect. Dis. **17**(5), 491–497 (2017). https://doi.org/10.1016/S1473-3099(17)30048-8
13. Maynard Smith, J., Haigh, J.: The hitch-hiking effect of a favourable gene. Genet. Res. **23**(1), 23–35 (1974). https://doi.org/10.1017/S0016672300014634
14. McCollum, A.M., Mueller, K., Villegas, L., Udhayakumar, V., Escalante, A.A.: Common origin and fixation of *Plasmodium falciparum dhfr* and *dhps* mutations associated with sulfadoxine–pyrimethamine resistance in a low-transmission area in South America. Antimicrob. Agents Chemother. **51**(6), 2085–2091 (2007). https://doi.org/10.1128/AAC.01228-06
15. McCollum, A.M., Basco, L.K., Tahar, R., Udhayakumar, V., Escalante, A.A.: Hitchhiking and selective sweeps of *Plasmodium falciparum* sulfadoxine and pyrimethamine resistance alleles in a population from Central Africa. Antimicrob. Agents Chemother. **52**(11), 4089–4097 (2008). https://doi.org/10.1128/AAC.00623-08
16. McCollum, A.M., Schneider, K.A., Griffing, S.M., Zhou, Z., Kariuki, S., Ter-Kuile, F., Shi, Y.P., Slutsker, L., Lal, A.A., Udhayakumar, V., Escalante, A.A.: Differences in selective pressure on *dhps* and *dhfr* drug resistant mutations in western Kenya. Malar. J. **11**, 77 (2012). https://doi.org/10.1186/1475-2875-11-77
17. McKenzie, F.E., Jeffery, G.M., Collins, W.E.: Gametocytemia and fever in human malaria infections. J. Parasitol. 93(3), 627–633 (2007)
18. Mint Lekweiry, K., Ould Mohamed Salem Boukhary, A., Gaillard, T., Wurtz, N., Bogreau, H., Hafid, J.E., Trape, J.F., Bouchiba, H., Ould Ahmedou Salem, M.S., Pradines, B., Rogier, C: Molecular surveillance of drug-resistant *Plasmodium vivax* using *pvdhfr*, *pvdhps* and *pvmdr1* markers in Nouakchott, Mauritania. J. Antimicrob. Chemother. **67**(2), 367–374 (2011). https://doi.org/10.1093/jac/dkr464
19. Nair, S., Williams, J.T., Brockman, A., Paiphun, L., Mayxay, M., Newton, P.N., Guthmann, J.P., Smithuis, F.M., Hien, T.T., White, N.J., Nosten, F., Anderson, T.J.C.: A selective sweep driven by pyrimethamine treatment in Southeast Asian malaria parasites. Mol. Biol. Evol. **20**(9), 1526–1536 (2003). https://doi.org/10.1093/molbev/msg162
20. Nassir, E.K.A., Suleiman, S., Muhsin, A.M.A.A., Ahmed, S., Ali, E., Nassir, E.K.A., Walliker, D., Mackinnon, M.J., Babiker, H.A.: Evolution of drug-resistance genes in *Plasmodium falciparum* in an area of seasonal malaria transmission in Eastern Sudan. J. Infect. Dis. **189**(7), 1239–1244 (2004). https://doi.org/10.1086/382509
21. Plowe, C.V., Kublin, J.G., Doumbo, O.K.: *P. falciparum* dihydrofolate reductase and dihydropteroate synthase mutations: epidemiology and role in clinical resistance to antifolates. Drug Resist. Updat. **1**(6), 389–396 (1998). https://doi.org/https://doi.org/10.1016/S1368-7646(98)80014-9
22. Popovici, J., Kim, S., Bin, S., Run, V., Menard, D., Lek, D., Pierce-Friedrich, L., Serre, D., Cannon, M.V., Hee, K.H.D., Lee Soon-U, L.: Recrudescence, reinfection, or relapse? A more rigorous framework to assess chloroquine efficacy for *Plasmodium vivax* malaria. J. Infect. Dis. **219**(2), 315–322 (2018). https://doi.org/10.1093/infdis/jiy484
23. Price, R.N., von Seidlein, L., Valecha, N., Nosten, F., Baird. J.K., White, N.J.: Global extent of chloroquine-resistant *Plasmodium vivax*: a systematic review and meta-analysis. Lancet Infect. Dis. **14**(10), 982–991 (2014)
24. Ross, R.: Report on the Prevention of Malaria in Mauritius, p. 202. E. P. Dutton, New York (1908)
25. Schneider, K.A.: Large and finite sample properties of a maximum-likelihood estimator for multiplicity of infection. PLoS One **13**(4), 1–21 (2018). https://doi.org/10.1371/journal.pone.0194148

26. Schneider, K.A., Escalante, A.A.: Fitness components and natural selection: why are there different patterns on the emergence of drug resistance in *Plasmodium falciparum* and *Plasmodium vivax*? Malar. J. **12**, 15 (2013)
27. Schneider, K.A., Escalante, A.A.: A likelihood approach to estimate the number of co-infections. PLoS One **9**(7), e97899 (2014). https://doi.org/10.1371/journal.pone.0097899
28. Schneider, K.A., Escalante, A.A.: Correction: a likelihood approach to estimate the number of co-infections. PLoS One **13**(2), 1–3 (2018). https://doi.org/10.1371/journal.pone.0192877
29. Schneider, K.A., Kim, Y.: An analytical model for genetic hitchhiking in the evolution of antimalarial drug resistance. Theor. Popul. Biol. **78**(2), 93–108 (2010)
30. Schneider, K.A., Kim, Y.: Approximations for the hitchhiking effect caused by the evolution of antimalarial-drug resistance. J. Math. Biol. **62**, 789–832 (2011). https://doi.org/10.1007/s00285-010-0353-9
31. Sidhu, A.B.S., Verdier-Pinard, D., Fidock, D.A.: Chloroquine resistance in *Plasmodium falciparum* malaria parasites conferred by *pfcrt* mutations. Science (New York) **298**(5591), 210–213 (2002)
32. Tantiamornkul, K., Pumpaibool, T., Piriyapongsa, J., Culleton, R., Lek-Uthai, U.: The prevalence of molecular markers of drug resistance in *Plasmodium vivax* from the border regions of Thailand in 2008 and 2014. Int. J. Parasitol. Drugs Drug Resist. **8**(2), 229–237 (2018)
33. Triglia, T., Wang, P., Sims, P.F., Hyde, J.E., Cowman, A.F.: Allelic exchange at the endogenous genomic locus in *Plasmodium falciparum* proves the role of dihydropteroate synthase in sulfadoxine-resistant malaria. EMBO J. **17**(14), 3807–3815 (1998)
34. Vathsala, P., Pramanik, A., Dhanasekaran, S., Usha Devi, C., Pillai, C., Subbarao, S., Ghosh, S., Tiwari, S., Sathyanarayan, T., Deshpande, P., Mishra, G.: Widespread occurrence of the *Plasmodium falciparum* chloroquine resistance transporter (PFCRT) gene haplotype SVMNT in *P. falciparum* malaria in India. Am. J. Trop. Med. Hyg. **70**, 256–259 (2004). https://doi.org/10.4269/ajtmh.2004.70.256
35. Wellems, T.E., Plowe, C.V.: Chloroquine-resistant malaria. J. Infect. Dis. **184**(6), 770–776 (2001). https://doi.org/10.1086/322858
36. WHO: World Malaria Report 2018. World Health Organization, Geneva (2018). License: CC BY-NC-SA 3.0 IGO
37. Zaw, M.T., Emran, N.A., Lin, Z.: Updates on k13 mutant alleles for artemisinin resistance in *Plasmodium falciparum*. J. Microbiol. Immunol. Infect. **51**(2), 159–165 (2018). https://doi.org/10.1016/j.jmii.2017.06.009

Identifying the Dominant Transmission Pathway in a Multi-stage Infection Model of the Emerging Fungal Pathogen *Batrachochytrium Salamandrivorans* on the Eastern Newt

Md Rafiul Islam, Matthew J. Gray, and Angela Peace

1 Introduction

Emerging infectious diseases are a threat to biodiversity and fungal pathogens have caused rapid declines in amphibian populations around the globe [18]. Gray et al. [6] identify *Batrachochytrium salamandrivorans* (Bsal) as an emerging fungal pathogen that caused rapid die-offs of naïve salamanders in Europe and predicts North America will soon experience similar devastation if no policy actions are taken and the pathogen emerges. Due to the fact that Bsal is such a recently emerging pathogen, we currently lack epidemiological data on how it may spread temporally and spatially across North America. Recent efforts have focused on building mathematical models to gain insight on pathogen spread and identify control strategies. Using Bsal as a case study, [8] employ Spatial Distribution Models to highlight the difficulty in validating model predictions when available data is limited, as well as the importance of appropriate model selection. Schmidt et al. [23] present a compartmental population model incorporating direct transmission and spatial diffusion that identified preventing emergence as the best strategy.

In order to better understand Bsal pathogen dynamics, we develop Susceptible-Infected-Recovered-Susceptible (SIRS) type disease models for a population of Eastern Newt adults. This particular species is widely distributed across eastern North America and has been shown to be highly susceptible to Bsal [17, 21]. In some cases, Bsal can lead to death in susceptible species within 2 to 3 weeks

M. R. Islam (✉) · A. Peace
Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX, USA

Department of Mathematics, Iowa State University, Ames, IA, USA
e-mail: rafiul.islam@ttu.edu; rafiul@iastate.edu

M. J. Gray
Center for Wildlife Health, University of Tennessee Institute of Agriculture, Knoxville, TN, USA

M. I. Teboh-Ewungkem, G. A. Ngwa (eds.), *Infectious Diseases and Our Planet*,
Mathematics of Planet Earth 7, https://doi.org/10.1007/978-3-030-50826-5_7

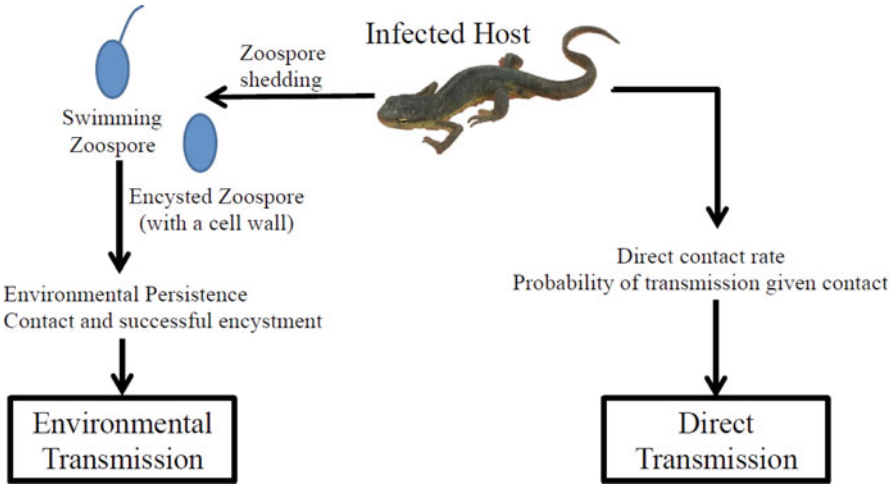


Fig. 1 Multiple transmission pathways

after exposure, but it has also been observed that some individuals can recover and clear the infection [17]. Compared to the duration of infection, adult Eastern Newts have a long lifespan potentially persisting >10 years [21, 22]. Unlike previous models [23], here we incorporate two routes of pathogen transmission: direct transmission via contact between infected and susceptible individuals and environmental transmission via shed zoospores in the water (Fig. 1). Bsal produces two types of zoospores that lead to environmental transmission, motile zoospores with flagellum and encysted zoospores with cell walls. Motile zoospores actively swim towards hosts, whereas encysted zoospores typically float on the surface [24].

We first present a base model that divides the host into four subpopulations depending on disease status (susceptible, latently infected, infectious, and recovered) as well as tracks the environmental loads of the two zoospores types. Epidemic compartmental models are commonly used to characterize the epidemiology of host–pathogen systems by providing means of estimating the invasion potential of a pathogen and surviving host population [9]. SIRS type models commonly assume that the duration of host infectiousness follows an exponential distribution [9], however, the duration of host infectiousness has been shown to be realistically closer related to a gamma-distributed [12, 25]. Following [25], we expand our base SIRS model to a full model that includes multiple stages of infection, each exponentially distributed so that the sum of the sequence of these independent exponentially distributed random variables approaches a gamma-distribution.

Since Bsal has not invaded North America yet, several parameter estimates remain unknown for eastern newts. For our simulations, we used a combination of Bsal data from eastern newts and European fire salamanders (Salamandra salamandra). We also used estimates of zoospore shedding from a closely related

chytrid species (B. dendrobatidis). We used sensitivity analyses to identify the most parameters driving transmission. In addition, by investigating the invasion probability (i.e., the basic reproduction number) we found that direct transmission is likely to be the dominant driver of pathogen dynamics for low density populations, whereas environmental transmission will dominate in high density populations.

2 Model Development

2.1 Base Model Development

We begin with a base model of ordinary differential equations where individuals are divided into four subpopulations, susceptible $S(t)$, latently infected but not infectious $L(t)$, infectious $I(t)$, and Recovered $R(t)$. Total population is denoted as $N(t) = S(t) + L(t) + I(t) + R(t)$. These state variables represent the density of individuals in an aquatic environment (e.g., pond), with units of number of individual per volume. Infected individuals shed two types of Bsal zoospores, $Z_m(t)$ and $Z_e(t)$ into the environment. These state variables have units of zoospore per volume. The schematic of the base model is shown in Fig. 2 and the equations take the following form:

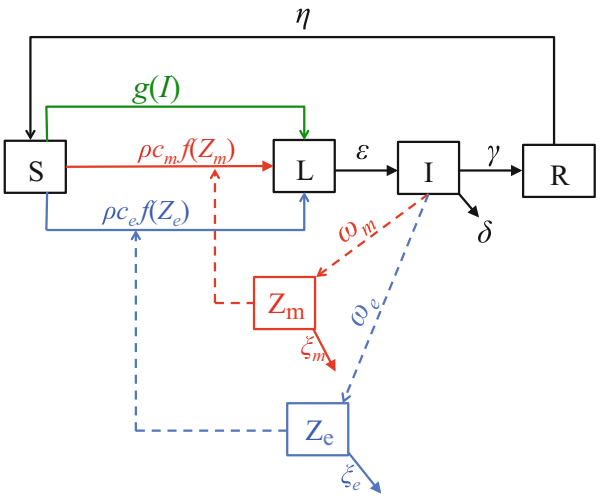


Fig. 2 Structure of the base model (1). The model tracks individuals divided into four subpopulations, susceptible $S(t)$, latently infected $L(t)$, infectious $I(t)$, and recovered $R(t)$, as well as the zoospores in the environment $Z_m(t)$ and $Z_e(t)$. The different routes of transmission are depicted using different colors, Direct transmission route is green, environmental transmission from Z_e is in blue, and from Z_m is in red. Solids lines depict the movement of individuals between compartments and dashed lines show the role of environmental zoospores

$$\frac{dS}{dt} = \underbrace{-g(I)S}_{\text{direct transmission}} \underbrace{-\rho(c_m f(Z_m,\kappa_m) + c_e f(Z_e,\kappa_e))S}_{\text{waterborne transmission}} + \underbrace{\eta R}_{\text{loss of immunity}} \tag{1a}$$

$$\frac{dL}{dt} = \underbrace{g(I)\,S}_{\text{direct transmission}} \underbrace{+\rho(c_m f(Z_m,\kappa_m) + c_e f(Z_e,\kappa_e))S}_{\text{waterborne transmission}} - \underbrace{\epsilon L}_{\text{latent becomes infectious}} \tag{1b}$$

$$\frac{dI}{dt} = \underbrace{\epsilon L}_{\text{latent becomes infectious}} - \underbrace{\delta I}_{\text{Bsal induced death}} - \underbrace{\gamma I}_{\text{Bsal recovery}} \tag{1c}$$

$$\frac{dR}{dt} = \underbrace{\gamma I}_{\text{recovery rate}} - \underbrace{\eta R}_{\text{loss of immunity}} \tag{1d}$$

$$\frac{dZ_m}{dt} = \underbrace{\omega_m I}_{\text{shed zoospores type m}} - \underbrace{\xi_m Z_m}_{\text{Bsal degradation for zoospores type m}} \tag{1e}$$

$$\frac{dZ_e}{dt} = \underbrace{\omega_e I}_{\text{shed zoospores type e}} - \underbrace{\xi_e Z_e}_{\text{Bsal degradation for zoospores type e}}. \tag{1f}$$

Susceptible individuals can become latently infected after contact with an infected individual following direct transmission rate given by function $g(I)$. Here, we consider both frequency-dependent and density-dependent direct transmission rates:

$$g(I) = \beta \frac{I}{N} \quad \text{and} \quad g(I) = \hat{\beta} I, \tag{2}$$

where β is the frequency-dependent direct transmission rate and $\hat{\beta} = \frac{\beta}{S(0)}$ is the density-dependent direct transmission rate. Susceptible individuals can also become infected after contact with environmental zoospores by the environmental transmission functions $\rho c_m f(Z_m,\kappa_m)$ and $\rho c_e f(Z_e,\kappa_e)$ where c_m and c_e are the contact rate coefficients between individuals and environmental zoospores of type m and e, respectively, ρ is the percentage of these spores that successfully encyst, and

$$f(Z_i,\kappa_i) = \frac{Z_i}{Z_i + \kappa_i}; \quad i = m, e. \tag{3}$$

Here, κ_m and κ_e are the Bsal ID50s (the doses of each zoospore type needed to infect 50% of a population). Latently infected individuals $L(t)$ have an incubation duration of $1/\epsilon$, after which they become infectious $I(t)$. Infectious individuals have a disease induced mortality rate of δ and recover at rate γ. Once recovered, we assume that an individual becomes susceptible again at the loss of immunity rate η. Infectious individuals shed zoospore type Z_m at rate ω_m and zoospore type Z_e at rate ω_e. These environmental zoospores naturally degrade at rates ξ_m and ξ_e, respectively.

2.1.1 Basic Analysis of the Base Model

We assume that the initial solution of system (1) are non-negative, i.e.,

$$(S(0), L(0), I(0), R(0), Z_m(0), Z_e(0)) \geq (0, 0, 0, 0, 0). \tag{4}$$

The model is of the form $X' = F(X)$, $X(t_0) = X_0$ where $X_0 \in \mathbb{R}^n$ and $F : \mathbb{R}^n \to \mathbb{R}^n$ is C^1. Thus by the theorem 4.1 in [1] the solution exists and is unique.

The following lemmas show that the Base Model (1) with the assumed initial conditions (4) is biologically meaningful, as solutions are positive and bounded. The proofs are in Appendix 1.

Lemma 2.1 *The solutions* $(S(t), L(t), I(t), R(t), Z_m(t), Z_e(t))$ *of system (1) are nonnegative for all* $t \geq 0$ *with the nonnegative initial conditions (4) in* $\left(\mathbb{R}_0^+\right)^5$.

Lemma 2.2 *Let*

$$\Sigma_H = \left\{(S, L, I, R) \in \left(\mathbb{R}_0^+\right)^4 \mid 0 \leq S(t) + L(t) + I(t) + R(t) \leq N(0)\right\}$$

and

$$\Sigma_Z = \left\{(Z_m, Z_e) \in \left(\mathbb{R}_0^+\right)^2 \mid 0 \leq Z_m(t) + Z_e(t) \leq \frac{(\omega_m + \omega_e)N(0)}{\xi_m + \xi_e}\right\}.$$

Define

$$\Sigma = \Sigma_H \times \Sigma_Z.$$

If $Z(0) \leq \frac{(\omega_m+\omega_e)N(0)}{\xi_m+\xi_e}$, *then the region* Σ *is bounded for the model (1).*

2.2 Full Model Development

Infectious individuals in the base model (1) recover at constant rate γ. However, the probability of recovering from the infection should increase as an individual

progresses through the disease. Here, we update the base model so the probability of recovery increases the longer the individual resides in the infected compartment. Following [25], we divide the infection compartment into n subcompartments, where the rate of recovery to advance through each subcompartment is $n\gamma$. Here, the full model breaks up the infectious stage into a sequence of n subcompartments, each exponentially distributed with mean $1/(n\gamma)$. The number of infectious stages influences the distribution for the overall duration of the infectious period ranging from and exponential distribution when $n = 1$ to resembling gamma distributions when $n > 1$ [10]. This technique also allows for incorporating different parameters for transmission and zoospore shedding rates throughout the infectious period. The full model takes the following form:

$$\frac{dS}{dt} = -\underbrace{\sum_{i=1}^{n} g(I_i)\, S}_{\substack{\text{direct}\\ \text{transmission}}} - \underbrace{\rho(c_m f(Z_m, \kappa_m) + c_e f(Z_e, \kappa_e))S}_{\substack{\text{waterborne}\\ \text{transmission}}} + \underbrace{\eta R}_{\substack{\text{loss of}\\ \text{immunity}}} \tag{5a}$$

$$\frac{dL}{dt} = \underbrace{\sum_{i=1}^{n} g(I_i)\, S}_{\substack{\text{direct}\\ \text{transmission}}} + \underbrace{\rho(c_m f(Z_m, \kappa_m) + c_e f(Z_e, \kappa_e))S}_{\substack{\text{waterborne}\\ \text{transmission}}} - \underbrace{\epsilon L}_{\substack{\text{latent becomes}\\ \text{infectious}}} \tag{5b}$$

$$\frac{dI_1}{dt} = \underbrace{\epsilon L}_{\substack{\text{latent becomes}\\ \text{infectious}}} - \underbrace{\delta I_1}_{\substack{\text{disease induced}\\ \text{mortality}}} - \underbrace{n\gamma I_1}_{\substack{\text{infection}\\ \text{advances}}} \tag{5c}$$

$$\frac{dI_2}{dt} = \underbrace{n\gamma I_1}_{\substack{\text{infection}\\ \text{advances}}} - \underbrace{\delta I_2}_{\substack{\text{disease induced}\\ \text{mortality}}} - \underbrace{n\gamma I_2}_{\substack{\text{infection}\\ \text{advances}}} \tag{5d}$$

$$\vdots$$

$$\frac{dI_n}{dt} = \underbrace{n\gamma I_{n-1}}_{\substack{\text{infection}\\ \text{advances}}} - \underbrace{\delta I_n}_{\substack{\text{disease induced}\\ \text{mortality}}} - \underbrace{n\gamma I_n}_{\substack{\text{infection}\\ \text{advances}}} \tag{5e}$$

$$\frac{dR}{dt} = \underbrace{n\gamma I_n}_{\substack{\text{recovery}\\ \text{rate}}} - \underbrace{\eta R}_{\substack{\text{loss of}\\ \text{immunity}}} \tag{5f}$$

$$\frac{dZ_m}{dt} = \underbrace{\sum_{i=1}^{n} \omega_{mi} I_i}_{\substack{\text{shed}\\ \text{zoospores m}}} - \underbrace{\xi_m Z_m}_{\substack{\text{Bsal degradation}\\ \text{for zoospores m}}} \tag{5g}$$

$$\frac{dZ_e}{dt} = \underbrace{\sum_{i=1}^{n} \omega_{ei} I_i}_{\substack{\text{shed} \\ \text{zoospores e}}} - \underbrace{\xi_e Z_e}_{\substack{\text{Bsal degradation} \\ \text{for zoospores e}}}, \tag{5h}$$

where we incorporate multiple stages of infection (I_i) with varying direct transmission rates (β_i) and varying zoospores shedding rates (ω_{mi}, ω_{ei}) for $i = 1, ..., n$ stages of infection. Similar to the base model, direct transmission is either frequency dependent or density dependent, where the ith stage direct transmission function follows:

$$g(I_i) = \begin{cases} \beta_i \frac{I_i}{N} & \text{for frequency-dependent transmission} \\ \beta_i \frac{I_i}{S_0} = \hat{\beta}_i I_i & \text{for density-dependent transmission.} \end{cases} \tag{6}$$

The number of infectious stages, n, can play an important role in the model predictions; however, this also depends on the parameterization of $\beta_1, ..., \beta_n$, $\omega_{e1}, ..., \omega_{en}$, and $\omega_{m1}, ..., \omega_{mn}$. In order to compare how varying n influences the predictions, we normalized these parameters between the different cases while assuming that the transmission and zoospores shedding rates tend to increase as time post exposure increases. Here we use the following for setting the parameter values:

$$\beta_i = \beta_{\text{base}} \frac{2i}{n(n+1)}, \qquad \omega_{ei} = \omega_{e\text{base}} \frac{2i}{n(n+1)}, \quad \text{and} \quad \omega_{mi} = \omega_{m\text{base}} \frac{2i}{n(n+1)} \tag{7}$$

for $i = 1, 2, ...n$, so that $\sum_{i=1}^{n} \beta_i = \beta_{\text{base}}$, $\sum_{i=1}^{n} \omega_{ei} = \omega_{e\text{base}}$, and $\sum_{i=1}^{n} \omega_{mi} = \omega_{m\text{base}}$.

2.3 Parameterization

While adequately parameterizing models remains a major challenge in epidemic and ecology modeling, parameter sensitivity analyses help shed light on the relative importance of the parameters. This requires baseline values, as well as ranges for the parameter space. For our simulations, we used values from the eastern newt (when available), European fire salamander, and some results from the related chytrid fungus, B. dendrobatidis. Given the high transmission rates that have been reported [24], we assume a baseline direct transmission rate of $\beta_{\text{base}} = 2$ per day. Martel et al. [16] observed that infected salamanders died within 1 week after showing severe symptoms, therefore we set the disease induced mortality rate to $\delta = 0.14$ per day. Martel et al. [17] investigated the susceptibility of 34 amphibian species to Bsal and found in many species infection resulted in mortality of all infected animals and

other infected species had the possibility of recovery. Here, we allow the possibility of recovery and investigate a range of recovery rates $\gamma \in (0.05 - 0.9)$ per day with a waning immunity at rate $\eta \in (0.05, 1)$ per day. Stegen et al. [24] observed encysted zoospores persisted in the environmental for 1 month and were more resistant to predation that motile zoospores. Therefore, we assumed the baseline zoospore degradation rates of $\xi_e = 0.03$ per day and $\xi_m = 0.05$ per day. Zoospore shedding rates are unknown for Bsal in eastern newts, however, shedding rates have been measured for a similar fungal pathogen, Batrachochytrium dendrobatidis (Bd) in frogs [14]. Using these measured ranges of Bd zoospore shedding rate averaged across the frog species, we assumed $\omega_{mbase} \in (8.6 - 345)$ thousand zoospores per day. Additionally, we assumed that ω_{ebase} is half the value of ω_{mbase}. In some cases, parameter values are unknown, for example, contact rate coefficients with each type of environmental zoospores (c_m and c_e) and the percentage of contacted spores that successfully encyst (ρ). A summary of parameters and their assumed values is given in Table 1. Given the uncertainty of several parameters for the eastern newt systems, we focus our analyses on parameter sensitivity.

3 Full Model Analysis

Analysis of the above models includes parameter sensitivity analyses, numerical simulations, and calculation of the basic reproductive number. These results assume frequency-dependent transmission functions (Eqs. 2 and 6), however, density-dependent led to similar qualitative predictions and figures are not shown.

3.1 Parameter Sensitivity Analysis

Here, we use Latin hypercube sampling (LHS), developed by McKay et al. [19] with the statistical partial rank correlation coefficient (PRCC) technique in order to perform a sensitivity analysis of the parameter space of the full model (5). The LHS/PRCC sensitivity analysis method globally explores the multi-dimensional parameter space. LHS is a stratified Monte Carlo sampling without replacement technique that gives unbiased estimates of modeling output measures subject to combinations of varying parameters. The PRCC can be used to classify how the output measures are influences by changes in a specific parameter value, while linearly discounting the effects of the other parameters [15]. The PRCC is appropriate since each parameter has a monotonic relationship with the output measures, details given in the Appendix 3. Here, a positive PRCC has a positive relationship with the output measure, whereas a negative PRCC value has an inverse relationship with the output measure. Larger PRCC values do not necessarily indicate more important parameters, however, we used a z test on transformed PRCC values to rank model parameters in terms of relative sensitivity [15]. The number of

Table 1 Model parameters

	Parameter	Unit	Base value	Range	Source
β_{base}	Base direct transmission rate	1/day	2	(0.1, 3)	[24][a]
c_m	Environmental contact rate coefficient with Z_m	1/day	0.02	(0.01, 0.05)	Assumed
c_e	Environmental contact rate coefficient with Z_e	1/day	0.01	(0.005, 0.03)	Assumed
η	Loss of immunity rate	1/day	0.1	(0.05, 1)	[13, 17][b]
$1/\epsilon$	Latency period	Days	10	(7, 14)	[13, 17][b]
δ	Disease induced mortality rate	1/day	0.14	(0.01, 0.5)	[13, 16, 17][b]
γ	Bsal recovery rate	1/day	0.1	(0.05, 0.9)	[13, 17][b]
n	Number of infected stages	–	5	NA	Assumed
ρ	% of contacted spores that encyst	–	0.75	(0.5, 1)	Assumed
$\omega_{m\text{base}}$	Base shedding rate of Z_m	$\frac{1000Zoospores}{day.individual}$	176	(8.6–345)	[14][c]
$\omega_{e\text{base}}$	Base shedding rate of Z_e	$\frac{1000Zoospores}{day.individual}$	88	(4.3–172)	Assumed $\omega_{e\text{base}} = \frac{1}{2}\omega_{m\text{base}}$
ξ_m	Degradation rate of Z_m	1/day	0.05	(0.02, 2)	[24][a]
ξ_e	Degradation rate of Z_e	1/day	0.03	(0.01, 1)	[24][a]
κ_m	ID-50	$\frac{Zoospore}{Vol}$	245K	(240K–250K)	MJG, unpublished data[b]
κ_e	ID-50	$\frac{Zoospore}{Vol}$	145K	(140K–150K)	Assumed $\kappa_e < \kappa_m$

[a]Values from *Salamandra salamandra* data
[b]Values from *Notophthalmus viridescens* data
[c]Values obtained from Bd (*Batrachochytrium dendrobatidis*) data

model simulations (or runs $= R$) need to be $\frac{4}{3}$ times greater than the number of uncertain parameters k, i.e., $R > \frac{4}{3}k$ [2, 5]. Our parameter space consist of 26 parameters and we choose 5,000 runs for our simulations. We used two output measures to classify the sensitivity of the parameters which has the monotone relation with our input variables;

1. The size of the Newt population after 150 days,
2. The maximum load of environmental zoospores at any time of the simulation 150 days.

The choice of using 150 days for the length of the simulations was to ensure the environmental zoospores concentrations achieved their maximum during the simulations. This duration also is within the expected duration to observe population collapse due to Bsal invasion in highly susceptible species [24]. The PRCC values for the Newt population density are shown in Fig. 3 and the PRCC values for the maximum zoospore load are shown in Fig. 4.

Figure 3 illustrates the most sensitive parameters affecting the size of the newt population. The Newt population increased significantly for increases in recovery rate γ. The Newt population decreased for increases in the direct transmission rates (β_i), the loss of immunity (η), the rate an infected individual becomes infectious (ϵ, inverse of the incubation period), as well as the disease induced mortality rate (δ).

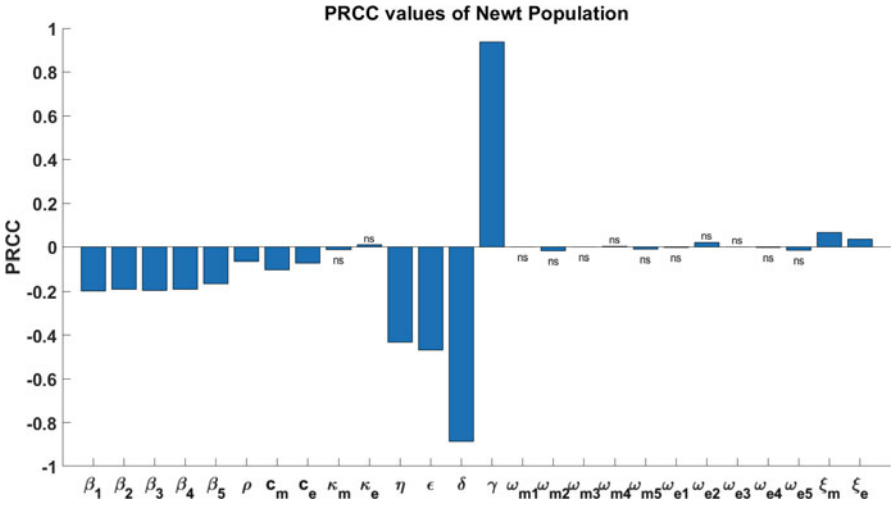


Fig. 3 Partial rank correlation coefficient (PRCC) values for each parameter in the Latin hypercube sampling (LHS) using the surviving Newt population density after 150 days as the output measure. Values marked ns are non-significant ($P > 0.05$)

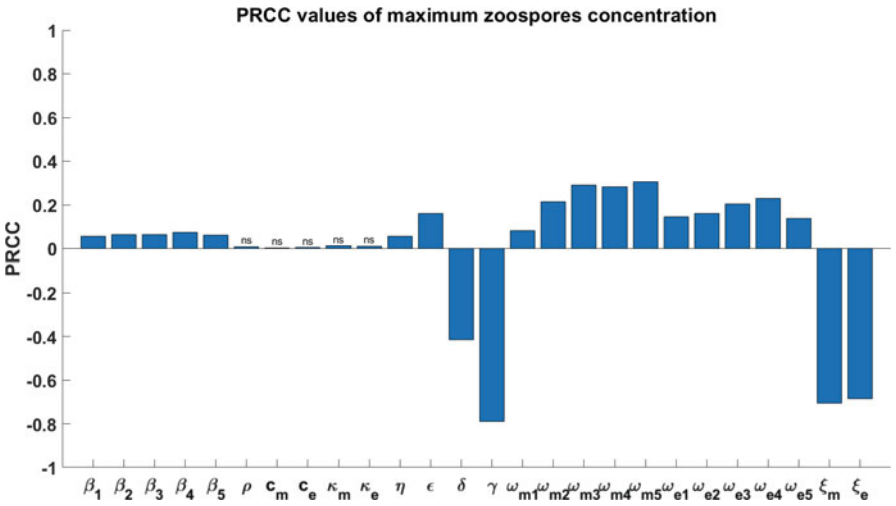


Fig. 4 Partial rank correlation coefficient (PRCC) values for each parameter in the Latin hypercube sampling (LHS) using maximum environmental zoospore load after 150 days as the output measure. Values marked ns are non-significant ($P > 0.05$)

For the maximum environmental zoospore load (Fig. 4), the zoospores concentration increased significantly for shorter incubation periods $1/\epsilon$ (i.e., larger the rate at which an infected individual becomes infectious, ϵ) and high zoospore shedding rates (ω_{mi} and ω_{ei} for $i = 1, ..., 5$). The zoospore concentration decreased significantly with larger disease induced mortality and recovery rates (δ, γ) as well as higher rates of environmental zoospore degradation (ξ_m and ξ_e).

3.2 Numerical Simulations

We varied the recovery rate (γ) to study the effect on the total Newt population, the final epidemic size, and the total maximum zoospores concentrations. The final epidemic size was calculated as the total number of cases over the duration of the simulation (150 days). As γ increased, a larger portion of the population survived the outbreak (Fig. 5a) and total maximum zoospores concentration decreased (Fig. 6). While an increase in γ increased survival of the population (Fig. 5a) it also increased the final epidemic size (Fig. 5b).

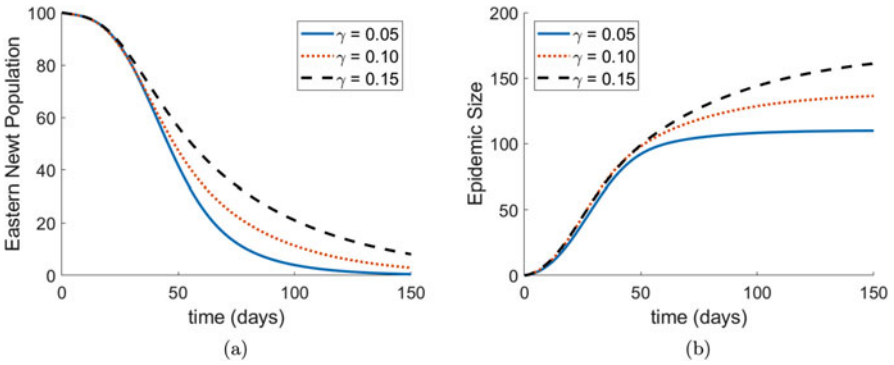


Fig. 5 (**a**) Eastern newt population and (**b**) endemic size for varying the recovery rate, γ and all other parameters are set to the baseline values from Table 1

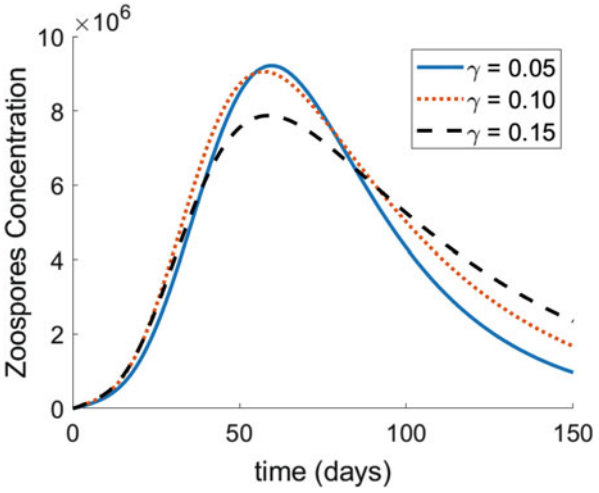


Fig. 6 Total maximum zoopores concentration varying the recovery rate, γ and all other parameters are set to the baseline values from Table 1

Varying the incubation period $(1/\epsilon)$, the model predicted that the newt population decreased only slightly for shorter period, however, the total maximum zoospores concentration increased substantially (Figs. 7 and 8).

We also investigated the influence of disease induced mortality rates (δ). While increases in δ decreased the newt population and the epidemic size (Fig. 9), it also caused a decrease in the zoospore concentration (Fig. 10).

Variations in the degradation rates of zoopores (ξ_e, ξ_m) did not substantially affect the newt population size (Fig. 11), however, it did play an important role in environmental zoospore loads (Fig. 12).

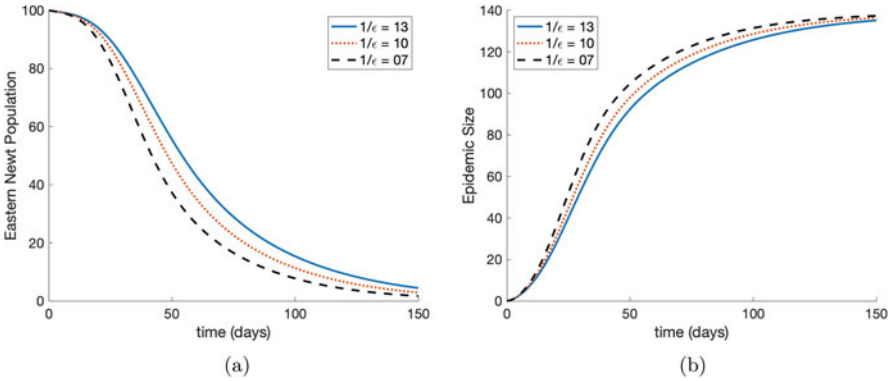


Fig. 7 (**a**) Eastern newt population and (**b**) epidemic size for varying the latency period, $1/\epsilon$ and all other parameter are set to the baseline values from Table 1

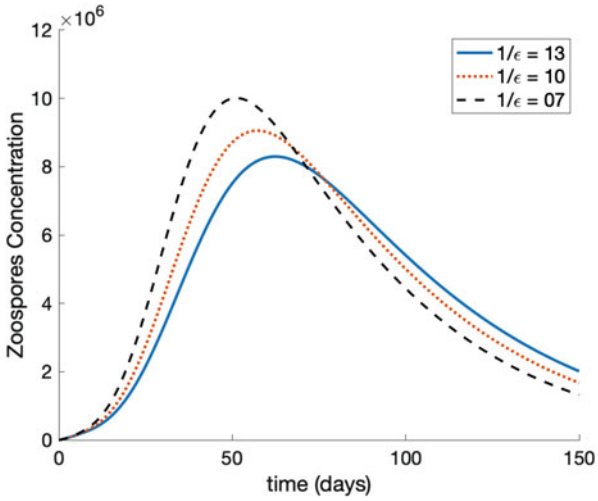


Fig. 8 Total maximum zoospores concentration varying the latency period, $1/\epsilon$ and all other parameters are set to the baseline values from Table 1

Figure 13 showed how varying the number of infectious stages n influenced model predictions. For smaller n the outbreak occurred earlier as the peak number of infected cases occurred sooner and was higher in magnitude (Fig. 13a, b). Additionally, for higher n the maximum zoospore concentration was delayed and was lower in magnitude (Fig. 13c).

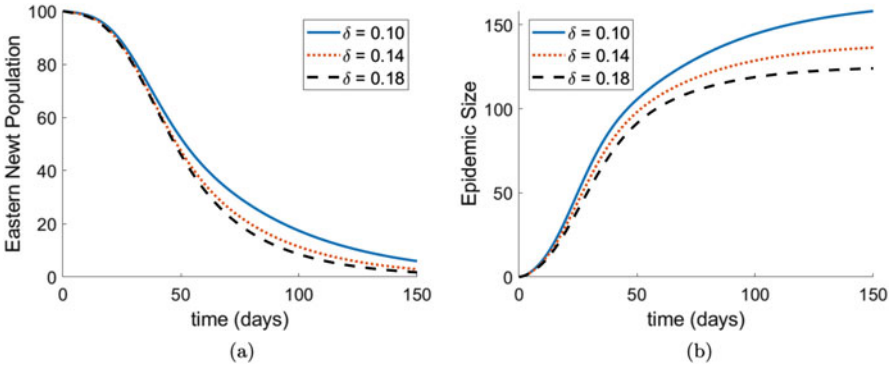
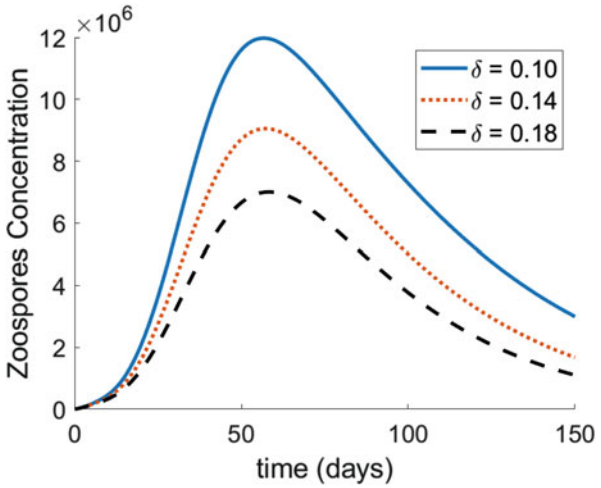


Fig. 9 (**a**) Eastern newt population and (**b**) epidemic size for varying the disease induced mortality rate δ and all other parameter are set to the baseline values from Table 1

Fig. 10 Total maximum zoospores concentration varying the disease induced mortality rate, $1/\delta$ and all other parameters are set to the baseline values from Table 1



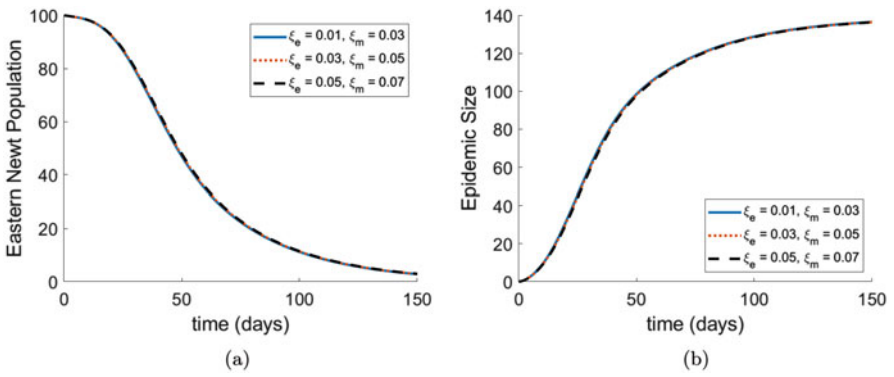


Fig. 11 (**a**) Eastern newt population and (**b**) epidemic size for varying the degradation rates of zoopores (ξ_e, ξ_m) and all other parameter are set to the baseline values from Table 1

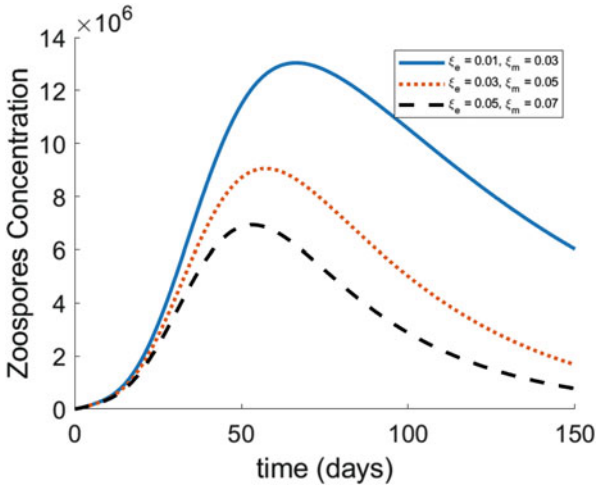


Fig. 12 Total maximum zoospores concentration varying the degradation rates of zoopores (ξ_e, ξ_m) and all other parameters are set to the baseline values from Table 1

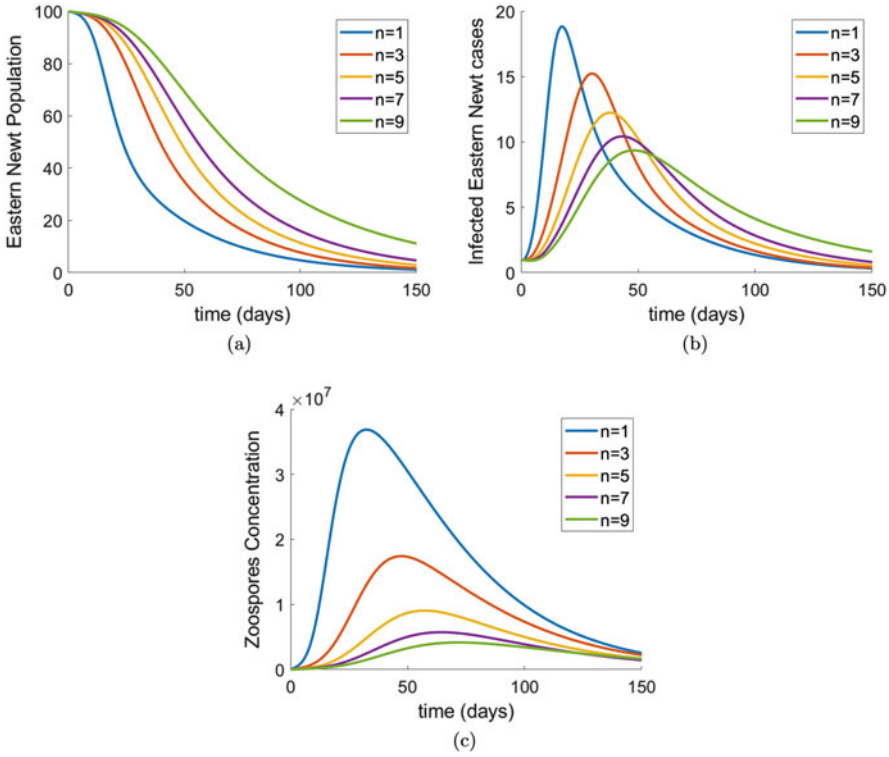


Fig. 13 (**a**) Eastern newt population (**b**) total infected eastern newt population, and (**c**) zoospores concentration for varying number of infectious stages, n using equation (7) for β_i, ω_{ei}, and ω_{mi} with all other parameters set to the baseline values in Table 1

3.3 Basic Reproductive Number

The basic reproduction number $\mathcal{R}_0$ is the expected number of secondary infections in a completely susceptible population produced by a single infectious individual during its infectious period [3]. To derive the basic reduction number, $\mathcal{R}_0$ of Bsal, we use the next-generation matrix approach for the system (5) following [4]. The analytic expression of the basic reproduction number $\mathcal{R}_0$ is as follows and the details are shown in the Appendix 2.

$$\mathcal{R}_0 = \underbrace{\sum_{i=1}^{n} \beta_i \overbrace{\left(\frac{\gamma n}{\delta+\gamma n}\right)^{i-1}}^{\text{fraction survives to the } i^{th} \text{ stage}} \overbrace{\frac{1}{\delta+\gamma n}}^{\text{duration in the } i^{th} \text{ stage}}}_{\text{direct}} + \underbrace{\frac{\rho c_m S_0 \sum_{i=1}^{n} \omega_{mi} (\gamma n)^{i-1} (\delta+\gamma n)^{n-i}}{(\delta+\gamma n)^n \xi_m \kappa_m}}_{\text{environmental for zoospores m}}$$

$$+ \underbrace{\frac{\rho c_e S_0 \sum_{i=1}^{n} \omega_{ei} (\gamma n)^{i-1} (\delta+\gamma n)^{n-i}}{(\delta+\gamma n)^n \xi_e \kappa_e}}_{\text{environmental for zoospores e}}, \tag{8}$$

where S_0 is the initial number of susceptible individuals. All transmission pathways contribute to the analytical expression of $\mathcal{R}_0$ which is a measure of potential outbreak. The expression in Eq. (8) describes the relative contribution of each pathway. We investigate the relative contributions of each pathway to $\mathcal{R}_0$ in order to identify the dominant pathway.

The dominant pathway depends on the population size (Table 2, Fig. 14). Direct transmission is the dominant pathway for small populations sizes (population size <5). As the population size increases (population size >5), environmental transmission for Z_m becomes the dominant pathway, while the environmental pathway for Z_e remains subordinate. As populations sized continue to increase (>7) then the directed transmission pathway becomes subordinate.

Table 2 Numerical values of $\mathcal{R}_0$ using parameter values from Table 1

	$S_0 = 1$	$S_0 = 5$	$S_0 = 7$	$S_0 = 30$
Total $\mathcal{R}_0$	1.15	2.35	2.95	9.84
Direct contact transmission proportion of $\mathcal{R}_0$	0.74	0.36	0.29	0.09
Environmental transmission proportion of $\mathcal{R}_0$ (Z_m)	0.16	0.39	0.43	0.55
Environmental transmission proportion of $\mathcal{R}_0$ (Z_e)	0.10	0.25	0.28	0.36

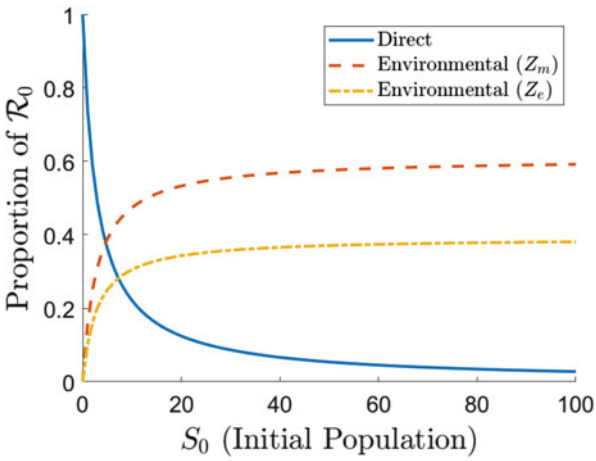


Fig. 14 Proportion of $\mathcal{R}_0$ for each transmission pathway given in Eq. (8) using parameter values from Table 1

4 Discussion

Infection disease models offer a powerful tool to better understand and predict epidemics in human and ecological systems. Bsal is an emerging infectious pathogen with potentially devastating biodiversity crisis in North America [26]. To better understand the dynamics, we developed and analyzed the first model of the highly virulent emerging pathogen Bsal on a host population of Eastern Newts that incorporates multiple transmission pathways with multiple stages of infection. The model predicts significant declines in host population shown in Fig. 5. These predictions are qualitatively consistent with empirical data captured by Stegen et al. [24], which show a collapse of 90% of a population of salamander introduced to Bsal within 6 months.

While varying key parameters can slow down population declines, the pathogen will persist as the basic reproduction number is always greater than one (Table 2). While previous models have concluded that mitigation during outbreaks is likely to fail and control efforts should focus on preventing disease emergence [23], once Bsal is established in an area, slowing the spread may help prevent nearby populations from being exposed.

Model predictions are sensitive to the number of infectious stages incorporated in the model structure (Fig. 13). Fitting the model to empirical data can help better parameterize the number of stages. For example, [20] fit a similar model to empirical data of Ranavirus cases in wood frogs Peace et al. [20].

The analytical expression of the basic reproductive number $\mathcal{R}_0$ given in Eq. (8) was used to identify the dominant transmission pathway. Our model and parameter set predicts that direct transmission is the dominant pathway for small population densities, however, environmental transmission is the dominant pathway for large population densities. These results can help guide intervention strategies. For example, in small density scenarios where direct transmission is the dominant pathway, we suggest intervention strategies focus efforts on reducing the contact rate between individuals. This can be done by increasing habitat complexity among them such as aquatic plants. For larger population densities, the model suggests that more efficient diseased control strategies would be to reduce environmental transmission. Here, the parameters that have the largest influence the maximum environmental zoospores concentrations (high PRCC values in Fig. 4) should be the main focus of intervention strategies. For example, the degradation rate of the zoospores in the environment have large influence on epidemic dynamics, which suggests the use of intervention strategies to increase ξ_m and ξ_e such as increasing UV penetration or number of zooplankton in the water [24] might reduce zoospore persistence.

The developed models consider a homogeneous population of Eastern Newts, however, disease dynamics for each individual is like to depend on their life stage (larvae, juvenile, adult). A future iteration of the model should consider a heterogeneous population and include stage structure of the host population. While our parameter sensitivity analysis highlights the important role that recovery rates can have, many species may have very low or zero probabilities of recovering from Bsal [17]. Given the hyper-susceptibility of eastern newts and fire salamanders to Bsal chytridiomycosis [17], it is possible that our simulations represent plausible scenarios in North America if Bsal in introduced. It is also possible that seasonal variations can influence disease and population dynamics, hence future iterations of the model should consider how temperature influences host contact rates, zoospore persistence, and importance of transmission pathways.

Appendix 1: Proofs of Base Model Lemmas

Proof of Lemma 2.1:

Proof We have

$$\left.\frac{dS}{dt}\right|_{\alpha(S)} = \eta R > 0$$

$$\left.\frac{dL}{dt}\right|_{\alpha(L)} = g(I)\,S + \rho(c_m f(Z_m, \kappa_m) + c_e f(Z_e, \kappa_e))S > 0$$

$$\left.\frac{dI}{dt}\right|_{\alpha(I)} = \epsilon L > 0$$

$$\left.\frac{dR}{dt}\right|_{\alpha(R)} = \gamma I > 0$$

$$\left.\frac{dZ_m}{dt}\right|_{\alpha(Z_m)} = \omega_m I > 0$$

$$\left.\frac{dZ_e}{dt}\right|_{\alpha(Z_e)} = \omega_e I > 0,$$

where $\alpha(x) = \{x(t) = 0$ and $S, L, I, R, Z_m, Z_e \in C(\mathbb{R}_0^+, \mathbb{R}_0^+)\}$ and $x \in \{S, L, I, R, Z_m, Z_e\}$. Therefore, due to the Lemma (2) in [7], any solutions $(S(t), L(t), I(t), R(t), Z_m(t), Z_e(t))$ of system (1) are nonnegative for all $t \geq 0$ with the nonnegative initial conditions (4) in $\left(\mathbb{R}_0^+\right)^5$.

Proof of Lemma 2.2:

Proof We can divide our system (1) into two parts: the host population, $N'(t) = S'(t) + L'(t) + I'(t) + R'(t)$ and the zoospores population, $Z(t) = Z_m(t) + Z_e(t)$. Adding the first four equations of system (1) yields

$$N'(t) = S'(t) + L'(t) + I'(t) + R'(t) = -\delta I \leq 0$$

which is a decreasing function of time. Therefore, $N(t) \leq N(0)$ and adding the last two equations of system (1) yields

$$Z'(t) = Z_m'(t) + Z_e'(t) = (\omega_m + \omega_e)I(t) - (\xi_m + \xi_e)Z(t)$$

$$\leq (\omega_m + \omega_e)N(0) - (\xi_m + \xi_e)Z(t).$$

A standard comparison theorem in [11] can be used to show that

$$Z(t) \leq Z(0)e^{-(\xi_m+\xi_e)t} + \frac{(\omega_m + \omega_e)N(0)}{(\xi_m + \xi_e)}(1 - e^{-(\xi_m+\xi_e)t})$$

$$= \frac{(\omega_m + \omega_e)N(0)}{(\xi_m + \xi_e)} + \left(Z(0) - \frac{(\omega_m + \omega_e)N(0)}{(\xi_m + \xi_e)}\right)e^{-(\xi_m+\xi_e)t}.$$

Therefore, if $Z(0) \leq \frac{(\omega_m+\omega_e)N(0)}{(\xi_m+\xi_e)}$, then $Z(t) \leq \frac{(\omega_m+\omega_e)N(0)}{(\xi_m+\xi_e)}$. Thus the region Σ is bounded.

Appendix 2: Basic Reproduction Number

The Jacobian matrix J of the system (5) is obtained from linearizing the system. Model (5) has a disease-free equilibria at $\mathbf{X}_0 = (S_0, \ldots, 0, 0)$, where S_0 is the initial population of susceptible individuals. Evaluating the Jacobian matrix at $\mathbf{X}_0$ yields

$$\mathbf{J} = \begin{pmatrix} 0 & 0 & -\beta_1 & -\beta_2 & \ldots & -\beta_{n-1} & -\beta_n & \eta & -\frac{\rho c_m S_0}{\kappa_m} & -\frac{\rho c_e S_0}{\kappa_e} \\ 0 & -\epsilon & \beta_1 & \beta_2 & \ldots & \beta_{n-1} & \beta_n & 0 & \frac{\rho c_m S_0}{\kappa_m} & \frac{\rho c_e S_0}{\kappa_e} \\ 0 & \epsilon & -n\gamma - \delta & 0 & \ldots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & n\gamma & -n\gamma - \delta & \ldots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & n\gamma & \ldots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ldots & -n\gamma - \delta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ldots & n\gamma & -n\gamma - \delta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ldots & 0 & n\gamma & -\eta & 0 & 0 \\ 0 & 0 & \omega_{m1} & \omega_{m2} & \ldots & \omega_{m(n-1)} & \omega_{mn} & 0 & -\xi_m & 0 \\ 0 & 0 & \omega_{e1} & \omega_{e2} & \ldots & \omega_{e(n-1)} & \omega_{en} & 0 & 0 & -\xi_e. \end{pmatrix}$$

Near the $\mathbf{X}_0$, for small perturbations $\mathbf{z} = (L, I_1, I_2, \ldots, I_n, R, Z_m, Z_e)$ the linearized infected subsystem of (5) evolves according to the following system of equations:

$$\frac{\mathrm{d}\mathbf{z}}{\mathrm{dt}} = \mathbf{M}\,\mathbf{z},$$

where

$$\mathbf{M} = \begin{pmatrix} -\epsilon & \beta_1 & \beta_2 & \ldots & \beta_{n-1} & \beta_n & 0 & \frac{\rho c_m S_0}{\kappa_m} & \frac{\rho c_e S_0}{\kappa_e} \\ \epsilon & -n\gamma - \delta & 0 & \ldots & 0 & 0 & 0 & 0 & 0 \\ 0 & n\gamma & -n\gamma - \delta & \ldots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & n\gamma & \ldots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \ldots & -n\gamma - \delta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \ldots & n\gamma & -n\gamma - \delta & 0 & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & n\gamma & -\eta & 0 & 0 \\ 0 & \omega_{m1} & \omega_{m2} & \ldots & \omega_{m(n-1)} & \omega_{mn} & 0 & -\xi_m & 0 \\ 0 & \omega_{e1} & \omega_{e2} & \ldots & \omega_{e(n-1)} & \omega_{en} & 0 & 0 & -\xi_e. \end{pmatrix}$$

We decompose the matrix $\mathbf{M}$ into transmission ($\mathbf{T}$) and transition ($\mathbf{\Sigma}$) matrices, respectively, obtaining

$$\frac{\mathrm{d}\mathbf{z}}{\mathrm{dt}} = (\mathbf{T} + \mathbf{\Sigma})\,\mathbf{z},$$

where

$$\mathbf{T} = \begin{pmatrix} 0 & \beta_1 & \beta_2 & \dots & \beta_{n-1} & \beta_n & 0 & \frac{\rho c_m S_0}{\kappa_m} & \frac{\rho c_e S_0}{\kappa_e} \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and

$$\mathbf{\Sigma} = \begin{pmatrix} -\epsilon & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ \epsilon & -n\gamma-\delta & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ 0 & n\gamma & -n\gamma-\delta & \dots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & n\gamma & \dots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & -n\gamma-\delta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & n\gamma & -n\gamma-\delta & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & n\gamma & -\eta & 0 & 0 \\ 0 & \omega_{m1} & \omega_{m2} & \dots & \omega_{m(n-1)} & \omega_{mn} & 0 & -\xi_m & 0 \\ 0 & \omega_{e1} & \omega_{e2} & \dots & \omega_{e(n-1)} & \omega_{en} & 0 & 0 & -\xi_e. \end{pmatrix}$$

The next-generation matrix with large domain is $\mathbf{K} = -\mathbf{T}\mathbf{\Sigma}^{-1}$. Since $\mathbf{T}$ has rank 1, the NGM $\mathbf{K}$ also has rank 1. Therefore, only the first row of $\mathbf{K}$ contains non-zero entries. Consequently, the spectral radius of $\mathbf{K}$ is the first entry on the diagonal, i.e., $\mathbf{K}_{1,1}$ which is equivalent to $\mathcal{R}_0$, and it is the same for both frequency and density-dependent direct transmission. Therefore,

$$\mathcal{R}_0 = \underbrace{\sum_{i=1}^{n} \beta_i \overbrace{\left(\frac{\gamma n}{\delta+\gamma n}\right)^{i-1}}^{\text{fraction survives to the } i^{th} \text{ stage}} \overbrace{\frac{1}{\delta+\gamma n}}^{\text{duration in the } i^{th} \text{ stage}}}_{\text{direct}} + \underbrace{\frac{\rho c_m S_0 \sum_{i=1}^{n} \omega_{mi} (\gamma n)^{i-1} (\delta+\gamma n)^{n-i}}{(\delta+\gamma n)^n \xi_m \kappa_m}}_{\text{environmental for zoospores m}}$$

$$+ \underbrace{\frac{\rho c_e S_0 \sum_{i=1}^{n} \omega_{ei} (\gamma n)^{i-1} (\delta+\gamma n)^{n-i}}{(\delta+\gamma n)^n \xi_e \kappa_e}}_{\text{environmental for zoospores e}},$$

where S_0 is the initial number of susceptible individuals.

Appendix 3: Monotonicity Test

Each subplot of the following Figs. 15 and 16 shows the monotonic relation between the model parameters with the output variable Eastern Newt population and maximum Zoospores concentrations, respectively, which is a required condition for the PRCC-LHS test.

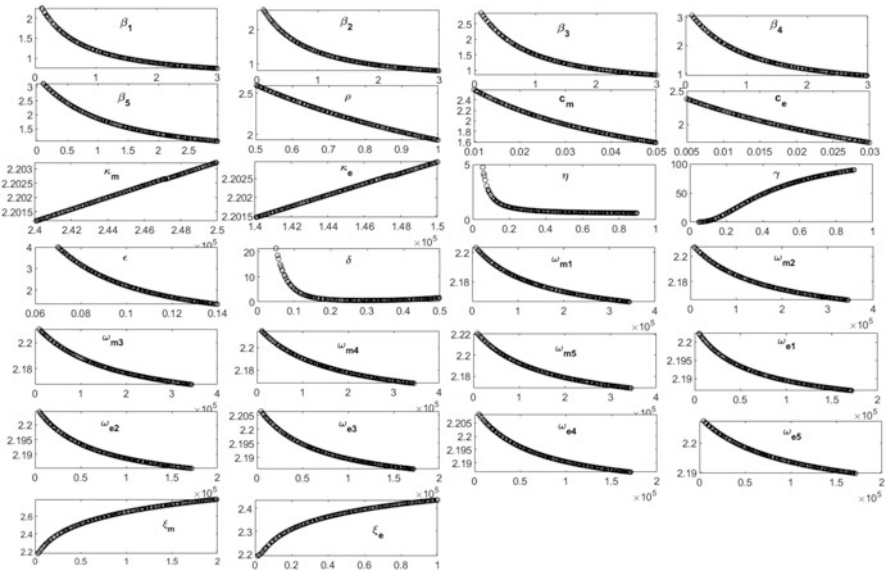


Fig. 15 Monotonicity plot of the model parameters vs. the output measure eastern newt population

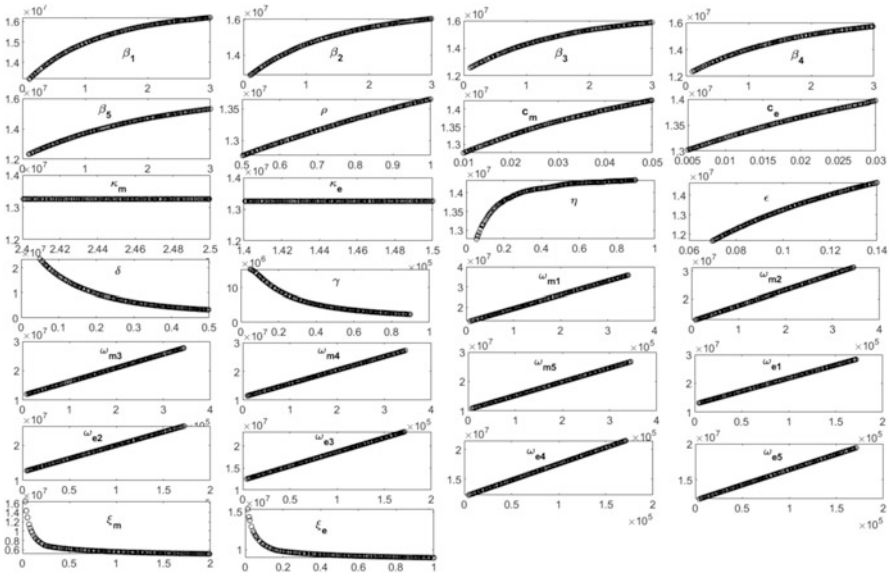


Fig. 16 Monotonicity plot of the model parameters vs. the output measure total maximum zoospores concentration

References

1. Allen, L.J.: Introduction to Mathematical Biology. Pearson, London/Prentice Hall, Upper Saddle River (2007)
2. Blower, S.M., Dowlatabadi, H.: Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. Int. Stat. Rev. **62**, 229–243 (1994)
3. Diekmann, O., Heesterbeek, J.A.P., Metz, J.A.: On the definition and the computation of the basic reproduction ratio r 0 in models for infectious diseases in heterogeneous populations. J. Math. Bio. **28**(4), 365–382 (1990)
4. Diekmann, O., Heesterbeek, J.A.P., Roberts, M.G.: The construction of next-generation matrices for compartmental epidemic models. J. R. Soc. Interface **7**(47), 873–885 (2010)
5. Gomero, B.: Latin hypercube sampling and partial rank correlation coefficient analysis applied to an optimal control problem. Master's Thesis, University of Tennessee (2012)
6. Gray, M.J., Lewis, J.P., Nanjappa, P., Klocke, B., Pasmans, F., Martel, A., Stephen, C., Olea G.P., Smith, S.A., Sacerdote-Velat, A., et al.: Batrachochytrium salamandrivorans: the North American response and a call for action. PLoS Path. **11**(12), e1005251 (2015)
7. Huo, H.F.: Permanence and global attractivity of delay diffusive prey-predator systems with the michaelis-menten functional response. Comput. Math. Appl. **49**(2-3), 407–416 (2005)
8. Katz, T.S., Zellmer, A.J.: Comparison of model selection technique performance in predicting the spread of newly invasive species: a case study with batrachochytrium salamandrivorans. Biol. Invasions **20**(8), 2107–2119 (2018)
9. Keeling, M.J., Rohani, P.: Modeling Infectious Diseases in Humans and Animals. Princeton University Press, Princeton (2011)
10. Krylova, O., Earn, D.J.: Effects of the infectious period distribution on predicted transitions in childhood disease dynamics. J. R. Soc. Interface **10**(84), 20130098 (2013)

11. Lakshmikantham, V., Leela, S., Martynyuk, A.A.: Stability Analysis of Nonlinear Systems. Springer, Berlin (1989)
12. Lloyd, A.L.: Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods. Proc. R. Soc. Lond. B Biol. Sci. **268**(1470), 985–993 (2001)
13. Longo, A.V., Fleischer, R.C., Lips, K.R.: Double trouble: co-infections of chytrid fungi will severely impact widely distributed newts. Biol. Invasions **21**(6), 2233–2245 (2019)
14. Maguire, C., DiRenzo, G.V., Tunstall, T.S., Muletz, C.R., Zamudio, K.R., Lips, K.R.: Dead or alive? Viability of chytrid zoospores shed from live amphibian hosts. Dis. Aquat. Organ. **119**(3), 179–187 (2016)
15. Marino, S., Hogue, I.B., Ray, C.J., Kirschner, D.E.: A methodology for performing global uncertainty and sensitivity analysis in systems biology. J. Theor. Bio. **254**(1), 178–196 (2008)
16. Martel, A., Spitzen-van der Sluijs, A., Blooi, M., Bert, W., Ducatelle, R., Fisher, M.C., Woeltjes, A., Bosman, W., Chiers, K., Bossuyt, F., et al: Batrachochytrium salamandrivorans sp. nov. causes lethal chytridiomycosis in amphibians. Proc. Natl. Acad. Sci. **110**(38), 15325–15329 (2013)
17. Martel, A., Blooi, M., Adriaensen, C., Van Rooij, P., Beukema, W., Fisher, M.C., Farrer, R.A., Schmidt, B.R., Tobler, U., Goka, K., et al.: Recent introduction of a chytrid fungus endangers western Palearctic salamanders. Science **346**(6209), 630–631 (2014)
18. McCraw, S., Gurr, S.: Emerging fungal threats to animal, plant and ecosystem health. Nature **484**(7393), 186194 (2012)
19. McKay, M.D., Beckman, R.J., Conover, W.J.: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics **42**(1), 55–61 (2000)
20. Peace, A., O'Regan, S.M., Spatz, J.A., Reilly, P.N., Hill, R.D., Carter, E.D., Wilkes, R.P., Waltzek, T.B., Miller, D.L., Gray, M.J.: A highly invasive chimeric ranavirus can decimate tadpole populations rapidly through multiple transmission pathways. Ecol. Model. **410**, 108777 (2019)
21. Petranka, J.W.: Salamanders of the United States and Canada. Smithsonian Institution Press, Washington (1998)
22. Roe, A.W., Grayson, K.L.: Terrestrial movements and habitat use of juvenile and emigrating adult eastern red-spotted newts, notophthalmus viridescens. J. Herpetol. **42**(1), 22–31 (2008)
23. Schmidt, B.R., Bozzuto, C., Lötters, S., Steinfartz, S.: Dynamics of host populations affected by the emerging fungal pathogen batrachochytrium salamandrivorans. R. Soc. Open Sci. **4**(3), 160801 (2017)
24. Stegen, G., Pasmans, F., Schmidt, B.R., Rouffaer, L.O., Van Praet, S., Schaub, M., Canessa, S., Laudelout, A., Kinet, T., Adriaensen, C., et al.: Drivers of salamander extirpation mediated by batrachochytrium salamandrivorans. Nature **544**(7650), 353–356 (2017)
25. Wearing, H.J., Rohani, P., Keeling, M.J.: Appropriate models for the management of infectious diseases. PLoS Med. **2**(7), 621 (2005)
26. Yap, T.A., Koo, M.S., Ambrose, R.F., Wake, D.B., Vredenburg, V.T.: Averting a North American biodiversity crisis. Science **349**(6247), 481–482 (2015)

Reducing the Global HIV Burden: The Importance of Uneven Exposure to the Results of HIV Prevention Trials

Mia Moore, Marie-Claude Boily, Deborah J. Donnell, and Dobromir Dimitrov

Mathematics Subject Classification 92D30, 62P10

1 Introduction

Infectious diseases continue to have a substantial contribution to the loss of life due to premature mortality and years lived with disability worldwide, especially in low- and middle-income countries [1]. More than 30 years after HIV was identified as the cause of AIDS, the HIV epidemic remains a significant public health burden with an estimated 0.95 million deaths attributed to AIDS in 2017 and an estimated 36.8 million people living with HIV, up from 33.3 million 5 years prior [2]. The increasing number of HIV infected worldwide is not only a result of persistent HIV transmission but also of more people living significantly longer with HIV while on antiretroviral treatment (ART). Although HIV testing capacity has increased over time, enabling more people to learn their HIV status, about 3 in 10 of people with HIV are still unaware they are infected. The stable number of new infections among adults over the recent years emphasizes the importance of HIV prevention.

This work was supported by the HIV Prevention Trials Network (HPTN), sponsored by the National Institute of Allergy and Infectious Diseases (NIAID), the National Institute on Drug Abuse (NIDA), the National Institute of Mental Health (NIMH), and the Office of AIDS Research of the National Institutes of Health (NIH), Department of Health and Human Services (DHHS) [grant number UM1AI068617].

M. Moore · D. J. Donnell · D. Dimitrov (✉)
Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA
e-mail: jrmoore@fredhutch.org; deborah@scharp.org; ddimitro@fredhutch.org

M.-C. Boily
Department of Infectious Disease Epidemiology, Imperial College, London, UK
e-mail: mc.boily@imperial.ac.uk

M. I. Teboh-Ewungkem, G. A. Ngwa (eds.), *Infectious Diseases and Our Planet*, Mathematics of Planet Earth 7, https://doi.org/10.1007/978-3-030-50826-5_8

An ambitious 95-95-95 strategy was announced by UNAIDS in 2014, aiming to end the AIDS epidemic by 2030 by achieving 95% diagnosed among all people living with HIV, 95% on ART among diagnosed, and 95% virally suppressed among treated [3]. An intermediate goal of 90-90-90 was set for 2020. Although this strategy is based on treatment, all major stakeholders understand that decisive reduction in HIV incidence cannot be achieved without investment in HIV prevention. The new national plan to end the HIV epidemic in the United States emphasizes the importance of reducing new HIV cases through evidence-based HIV prevention [4].

A series of randomized controlled trials (RCTs) has demonstrated that different interventions, such as prevention of mother-to-child transmission, male circumcision, ART as prevention, and pre-exposure prophylaxis (PrEP) reduce HIV acquisition risk [5–13]. New biomedical products are under development or currently being tested in RCTs including long-acting injectables, infusions of broadly neutralizing antibodies, and multi-dose vaccines [14–18]. RCTs of HIV biomedical prevention interventions are often designed assuming common effectiveness at all levels of HIV risk. However, in a previous modeling analysis we have demonstrated that the majority of participants in HIV prevention trials likely remain unexposed to HIV even if the RCT takes place in a high-prevalence settings [19].

Mathematics has a long history of contributions to understanding disease epidemics in humans and animals [20–24]. Mathematical models are frequently used to investigate epidemic trends and evaluate biomedical interventions aiming to curb or eliminate diseases transmission[25–29]. Models specific to HIV have been employed to simulate epidemic dynamics among the general population and high-risk groups in different geographic settings, to estimate the impact of various combinations of prevention and treatment tools (medical male circumcision, pre-exposure prophylaxis, treatment as prevention), and to inform the design of future clinical trials testing new HIV prevention products [30–35]. Two recent studies compared multiple HIV models used to evaluate the impact of ART scale-up and HIV vaccination strategies [36, 37]. These studies show that: (i) deterministic compartmental models are used more often (11 out of 20) than individual-based microsimulation models (4 out of 12) when intervention effectiveness is an outcome; (ii) Markov models (4 out of 20) and decision-tree models (1 out of 20) are prevalent in cost-effectiveness analyses; and (iii) models are more often used to simulate epidemic conditions in Africa (12 out of 20) where the majority of new HIV infections occur, compared to settings in Asia (6 out of 20) and North America (1 out of 20). In rare cases models are calibrated to multiple settings around the world to assess the potential of an intervention in different conditions [38, 39].

A few modeling studies have already investigated problems related to heterogeneity in HIV risk and its importance for the HIV epidemic. Unequal exposure to infection and differential susceptibility were shown to play a role in estimated efficacy of HIV vaccines [40]. Desai et al. also demonstrated that traditional efficacy measures, based on the hazard rate ratio or cumulative incidence ratio are sensitive to population heterogeneities modifying HIV risks [41]. Wilson et al. used a simple model to show that effectiveness of interventions depends on the number

of exposures that each trial participant has [42]. Monte-Carlo simulations of clinical trials showed that heterogeneity in risk can lead to underestimation of the efficacy of HIV prevention and suggested that the crossover design may guarantee better efficacy estimation and statistical power [43].

In this study we use compartmental mathematical models to analyze the uneven HIV exposure as a driver of heterogeneity in risk and to explore its effects on the efficacy observed in RCTs. More specifically, we define three characteristics of the exposure to HIV (magnitude, concentration, and replacement) and investigate how they create an imbalance between trial arms over time leading to attenuation of the efficacy estimates of imperfect biological interventions. Our work expands on previous modeling studies focusing on one exposure metric (magnitude) only [42] which assigned individuals constant acquisition risk over time [43]. Although partial protection can be conceptualized in different ways, we assume equal imperfect efficacy for all people in every HIV exposure. For simplicity, we refer to this protection as the "true efficacy per act" in this paper. In contrast, the "IRR effectiveness" is calculated from the results of the simulated RCTs as one minus the incidence rate ratio (IRR) of acquiring HIV, also commonly called just effectiveness. In general, it is preferable to quantify the "true efficacy" as a measure that is more likely to be time invariant and independent of the sexual behavior of each individual.

Regarded as a gold standard for effectiveness research, individually randomized trials are possible only where there is ethical equipoise between the two arms, typically when effectiveness of the new intervention is unproven. In this study we also investigate the impact of uneven HIV exposure on the outcomes of trials with step-wedge design which are preferred in situations when products with known efficacy are tested under new conditions and therefore the assignment of participants to placebo is unethical.

2 Model Description and Analysis

2.1 Metrics of HIV Exposure

We quantify the amount and distribution of HIV exposure in a population cohort with three metrics (see Table 1). The proportion of participants who have sexual contacts with infected partners over unit time determines *the concentration of exposure*. If HIV exposure is distributed among a small proportion of individuals, the exposure is highly concentrated. Conversely, if HIV exposure is distributed among large proportion of individuals, then the exposure is less concentrated. The average number of exposures of each exposed individual per unit time defines *the magnitude of exposure*. If the annual HIV incidence in two cohorts is equal, then the magnitude of exposure in the cohort with higher concentration must be larger. Allowing to a fraction of the cohort ("replacement fraction") to switch their

Table 1 Metrics of HIV exposure and efficacy

Term	Symbol	Description
Concentration of exposure	f_E	The fraction of the population at the start of the trial exposed to HIV through sexual contacts with HIV-positive partners
Magnitude of exposure	r	The rate of exposure among those exposed depending on the frequency of sexual contacts with HIV-positive partners
Replacement of exposure	σ	The rate at which people transition into (and out of) the exposed class
Per-act (true) efficacy	ϵ	The percentage by which the intervention decreases HIV infection risk for a given exposure. For example, if the risk of HIV infection due to unprotected sex with an HIV-positive individual is reduced from 1% to 0.1% with PrEP, the true efficacy is 90%.
Incidence rate ratio (IRR)		The ratio of incidence in the active arm to incidence in the control arm.
IRR effectiveness	E	The percentage reduction of HIV incidence, defined with (2), among individuals administered the intervention compared to control population not using the intervention. It is $1 - \text{IRR}$.
Hazard ratio	θ	The ratio of the risk of infection due to the introduction of the treatment, as estimated from a linear model.
GLM effectiveness	E_{GLM}	Measure of effectiveness in complex trial designs, such as the step-wedge. It is computed from a generalized linear model and related to the hazard rate via $E_{\text{GLM}} = e^{\theta}$.

exposure status (from exposed to unexposed and vice versa) each time period defines *the replacement of exposure*.

2.2 Modeling the Traditional RCT Design Description

In the traditional RCT design each participant is randomly assigned to the treatment (active) or placebo (control) group and treatment effectiveness is estimated by the difference in the observed HIV incidence between these 2 groups (arms). We model RCT using a compartmental model governed by ordinary differential equations, which are common in epidemiology. We assume that the trial cohort only represents a small subset of the general population and so any infections that occur during the trial are due to contacts with individuals outside of the trial. We divide each arm of the trial cohort into exposed and unexposed populations. Let $X_i(t)$ and $U_i(t)$ be the number of participants at time t, where the index $i \in \{c, a\}$ indicates the control or active arm. Those in the exposed population have on average r exposures per

period with a per-exposure risk of ρ_c in the control arm and $\rho_a = (1-\epsilon)\rho_c$ in the active arm, where ϵ is the per-act treatment efficacy. Trial participants are removed from follow-up upon infection. Finally, individuals move from the exposed to the unexposed population at per-capita rate κ and from the unexposed to the exposed population at rate γ. The traditional RCT model is

$$\begin{aligned} \text{Exposed} \quad & \frac{dX_i}{dt} = -\rho_i r X_i - \kappa X_i + \gamma U_i \\ \text{Unexposed} \quad & \frac{dU_i}{dt} = \kappa X_i - \gamma U_i \\ U_i(0) = (1-f_E)N \quad & X_i(0) = f_E N \end{aligned} \tag{1}$$

where f_E is the fraction of the population that is exposed to HIV and N is the number of participants in each arm at enrollment. We assume that individuals update their exposure status at a rate σ (for example, by changing sexual partners) with a fraction f_E of those individuals who change status entering (or re-entering) the exposed class. As a result, exposure replacement rates are $\kappa = \sigma(1-f_E)$ and $\gamma = \sigma f_E$. This formulation allows us to adjust the concentration, magnitude, and replacement of exposure independently via the parameters, f_E, r, and σ. Note that here we assume that the risk of infection is the same for all exposed individuals and does not change over time. We use this simple set of assumptions in order to isolate the effect of concentrating exposures in a fraction of the population.

The HIV incidence in each arm of the traditional RCT is estimated as the recorded number of new infections among the arm cohort over the duration of the trial τ, divided by the cumulative time that arm participants spent at risk for HIV infection. It is given by

$$\text{Trial Incidence} \quad I_i = \frac{\int_0^\tau \rho_i r X_i(t)\,dt}{\int_0^\tau (U_i(t) + X_i(t))\,dt} \tag{2}$$

The IRR effectiveness of HIV prevention tested in RCT is measured as the reduction in HIV incidence in the active arm compared to the control arm: $1 - I_a/I_c$.

The solution to (1) is given by

$$\begin{aligned} X_i(t) &= X_{1i}e^{\lambda_{1i}t} + X_{2i}e^{\lambda_{2i}t} \\ U_i(t) &= U_{1i}e^{\lambda_{1i}t} + U_{2i}e^{\lambda_{2i}t} \\ \lambda_{1,2} &= -(\rho_i r + \sigma) \pm B_i \\ B_i &= \sqrt{(\rho_i r - \sigma)^2 + 4\rho_i r(1-f_E)\sigma} \\ X_{1,2i} &= f_E N/2 \mp \big(\sigma(1-2f_E) - (r\rho_i + \sigma)\big)N/(2B_i) \\ U_{1,2i} &= (1-f_E)N/2 \pm \big(\sigma(1-3f_E) + \rho_i r(1-f_E)\big)N/(4B_i) \end{aligned} \tag{3}$$

which in the general case is cumbersome to study analytically. We will, therefore, focus on two special cases: (1) when all individuals are equally exposed to HIV and (2) when there is no switching between exposed and unexposed classes. We will also study the asymptotic behavior for lengthy trials.

2.2.1 Analysis of Special Case: Even Exposure

If all trial participants are exposed to HIV ($f_E = 1$), then model (1) simplifies to

$$\begin{aligned} \frac{dX_a}{dt} &= -\rho_c(1-\epsilon)rX_a \\ \frac{dX_c}{dt} &= -\rho_c rX_c \\ X_a(0) &= X_c(0) = N \end{aligned} \tag{4}$$

with solution $X_a(t) = Ne^{-\rho_c(1-\epsilon)rt}$ and $X_c(t) = Ne^{-\rho_c rt}$. The HIV incidences by arm are $I_a = \rho_c r(1-\epsilon)$ and $I_c = \rho_c(1-\epsilon)$ so the effectiveness simplifies to ϵ, the true efficacy. Therefore, in case of perfect adherence and homogeneously distributed HIV risk, the true efficacy and the IRR effectiveness will match exactly.

2.2.2 Analysis of Special Case: Uneven Exposure with Stable Partners

If a highly efficacious prevention product is tested ($1 > \epsilon \gg 0$), the risk of infection is significantly reduced (but not eliminated) in the active arm. As a result, highly exposed individuals become infected and removed from follow-up more quickly in the placebo than in the active arm, leading to a faster depletion of the "exposed" participants in the placebo arm. This creates a growing imbalance with respect to HIV exposure between trial arms over time, even if the arms were perfectly balanced at randomization.

To illustrate the effect of uneven exposure, we consider the simple case where all participants have stable partners. In such a cohort, each participant is repeatedly exposed to HIV or never exposed to HIV during follow-up, depending on the HIV status of his/her partners. The governing equations are obtained by substituting $\kappa = \gamma = 0$ into model (1).

$$\begin{aligned} \text{Exposed} \quad & \frac{dX_i}{dt} = -\rho_i rX_i \\ \text{Unexposed} \quad & \frac{dU_i}{dt} = 0 \end{aligned} \tag{5}$$

which has solution $X_i(t) = Nf_E e^{-\rho_i rt}$ and $U_i(t) = N(1-f_E)$.

The trial incidence in each arm is

$$I_i = \frac{f_E \rho_i r (1 - e^{-\rho_i r \tau})}{\tau \rho_i r (1 - f_E) + f_E (1 - e^{-\rho_i r \tau})} \tag{6}$$

which decreases as the length of the trial, τ, increases, provided that $f_E < 1$. This is not due to any change in risk magnitude over time, as r, ρ_a, and ρ_c are assumed to be constant, but instead to a selective depletion of exposed individuals.

If the fraction of exposed individuals is quite low, then this depletion can happen surprisingly rapidly. Suppose that we wish to perform an RCT in a population with an annual incidence of 5%. If $f_E = 100\%$, then $r\rho_c = 5\%$ and the observed trial incidence will be 5% regardless of the trial duration (see Sect. 2.2.1). However, if $f_E = 20\%$, then $r\rho_c = 25\%$ to ensure 5% overall incidence at enrollment and the trial incidence calculated via (6) will be only 4.5% for a 1-year RCT and 4.1% for a 2-year RCT.

This risk depletion, the selective removal of exposed individuals from the trial population due to infection, happens more rapidly in the control than in the treatment arm and creates growing imbalance over time. This is reflected in the IRR effectiveness

$$\text{IRR Effectiveness} = 1 - \frac{\rho_a (1 - e^{-\rho_a r \tau})}{\rho_c (1 - e^{-\rho_c r \tau})} \left(\frac{\tau \rho_c r (1 - f_E) + f_E (1 - e^{-\rho_c r \tau})}{\tau \rho_a r (1 - f_E) + f_E (1 - e^{-\rho_a r \tau})} \right) \tag{7}$$

When the trial is short relative to risk (very small τ), the expression (7) reduces to $\epsilon = 1 - \rho_a / \rho_c$, the true efficacy of the intervention in the exposed population. However when the trial is long ($\tau \to \infty$), then the effectiveness drops to zero unless all participants remain unexposed to HIV ($f_E = 0$).

2.2.3 Asymptotic Behavior

To understand the long-term behavior of model (1) it is helpful to make the following change of variables. Let $N_i(t) = X_i(t) + U_i(t)$ be the cohort size of trial arm i at time t, and let $x_i(t) = X_i(t)/N_i(t)$ be the proportion of the participants in arm i who are in the exposed class at time t.

$$\begin{aligned} N_i' &= -\rho_i r x_i N_i \\ x_i' &= -\rho_i r x_i - \kappa x_i + \gamma(1 - x_i) + \rho_i r x_i^2 \\ &= -\rho_i r x_i (1 - x_i) + \sigma(f_E - x_i) \\ N_i(0) &= N \qquad x_i(0) = f_E \end{aligned} \tag{8}$$

If a trial lasts sufficiently long, the fraction $x_i(t)$ will approach its asymptotically stable steady state

$$x_i^* = \frac{\sigma + \rho_i r}{2\rho_i r}\left(1 - \sqrt{1 - \frac{4\rho_i r \sigma f_E}{(\sigma + \rho_i r)^2}}\right) \tag{9}$$

even though the total uninfected population size may still be far from the equilibrium $N_i^* = 0$. We define *instantaneous incidence* $I_{i,\text{inst}}(t) = r\rho_i p_i$ to be the trial incidence between time t and $t + dt$, and the instantaneous effectiveness to be $1 - I_{a,\text{inst}}/I_{c,\text{inst}}$. The trial incidence is the mean of the instantaneous incidence throughout the entire duration of the trial, weighted by the population size $N_a(t) + N_c(t)$.

Initially, the instantaneous effectiveness is ϵ, the true efficacy. However for trials that are lengthy compared to the timescale of exposure, i.e. an HIV trial of several years during which individuals may change partners many times, the IRR effectiveness will be dominated by the instantaneous effectiveness at steady state $E^* = 1 - (1 - \epsilon)x_a^*/x_c^*$. Using the expression (9) we find that:

$$\begin{aligned} \lim_{f_E \to 0} E^* &= 1 - (1 - \epsilon)\frac{\sigma + r}{\sigma + r(1 - \epsilon)} \\ &= \epsilon \frac{\sigma/r}{\sigma/r + (1 - \epsilon)} \end{aligned} \tag{10}$$

From the above expression, we should expect that the estimated efficacy is reduced from the true efficacy when the HIV exposure is highly concentrated (small f_E). This loss of efficacy is mitigated if the magnitude of risk r is small relative to the replacement of risk σ or if the true efficacy is very close to 100%.

2.3 Modeling Stepped-Wedge Design

The stepped-wedge trial is a group-randomized design that is particularly well suited to testing the efficacy of an intervention during a roll out of a new prevention strategy without completely withholding it from any of the participants. Groups of participants are randomly selected to switch from standard of care to treatment at regular intervals (steps) until all groups are treated during the last period. Stepped-wedge trials are becoming increasingly popular in HIV research with more combination strategies of indisputably efficacious components being tested under new conditions [44–46].

For the stepped-wedge trials, simulated in this analysis, we divide our population into n_W cohorts. Let X_{ij} and U_{ij} be the number exposed and unexposed individuals in cohort $i \in \{1 \ldots n_W\}$ during time period $j \in \{0 \ldots n_W\}$ where each period is of length $\hat{\tau} = \tau/(1 + n_W)$. Then the time-evolution equations are

$$\begin{aligned}
\text{Exposed} \quad & \frac{dX_{ij}}{dt} = -\rho_c(1-\epsilon T_{ij})rX_{ij} - \kappa X_i + \gamma U_i \\
\text{Unexposed} \quad & \frac{dU_{ij}}{dt} = \kappa X_{ij} - \gamma U_{ij} \\
U_{ij}(0) = U_{i,j-1}\left(\hat{\tau}\right) \quad & X_{ij}(0) = X_{i,j-1}\left(\hat{\tau}\right) \\
U_{i0}(0) = (1-f_E)N \quad & X_{i0}(0) = f_E N
\end{aligned} \tag{11}$$

where exposure replacement rates are defined as in the traditional RCT while T_{ij} indicates whether the intervention was offered to cohort i during time period j:

$$T_{ij} = \begin{cases} 1 & j \geq i \\ 0 & j < i \end{cases} \tag{12}$$

In the stepped-wedge trial, the HIV incidence of the control and active arms is given by

$$\begin{aligned}
I_c &= \frac{\sum_{ij}(1-T_{ij})\int_0^{\hat{\tau}} \rho_c r X_{ij}(t)\,dt}{\sum_{ij}(1-T_{ij})\int_0^{\hat{\tau}} (X_{ij}(t)+U_{ij}(t))\,dt} \\
I_a &= \frac{\sum_{ij} T_{ij}\int_0^{\hat{\tau}} \rho_c r X_{ij}(t)\,dt}{\sum_{ij} T_{ij}\int_0^{\hat{\tau}} (X_{ij}(t)+U_{ij}(t))\,dt}
\end{aligned} \tag{13}$$

2.3.1 Estimating Effectiveness in Stepped-Wedge Trials

Due to the nature of stepped-wedge designs, effectiveness cannot be reliably computed with the incidence rate ratio. For example, suppose that the epidemic is becoming more intense over time. This means that when individuals start to receive the intervention, they will be facing a higher likelihood of infection, which will affect the estimate of effectiveness. Similarly, as randomization occurs at the population level rather than the individual-level, it is likely that there are differences in exposure rates between populations. If a high-risk population is among the first to receive the intervention, this may artificially lower the effectiveness as measured by IRR.

To account for these potential confounding effects, effectiveness is typically calculated by fitting a generalized linear model with time and cohort-level effects [47]. This formulation is equivalent to fitting the following differential equation to the data.

$$\begin{aligned}
\frac{dX_{ij}}{dt} &= h_{ij}X \\
\ln h_{ij} &= \alpha_i + \beta_j + \theta T_{ij}
\end{aligned} \tag{14}$$

where h_{ij} is the hazard rate or the risk of infection in cohort i during time interval j. Note that in this formulation, all individuals in the population are considered equally susceptible. The parameter θ is the log-odds-ratio of infection and can be used to estimate effectiveness via $E_{GLM} = e^{\theta}$. We will refer to this estimate of efficacy as the GLM-derived effectiveness.

The parameters α and β are used to represent the force of infection in different cohorts and at different times, respectively, and are necessary to adjust for potential confounding in the step-wedge design. Many authors treat these parameters as random effects, meaning that they are assumed to be sampled from a random distribution, typically normal. This reduces the overall number of parameters that need to be estimated—a mean and a variance for the α's rather than each one individually—and prevents overfitting in cases where there are small numbers of observations. However we will assume that our trial is large enough that all α's and β's can be reliably estimated as individual parameters, i.e. they are fixed effects. We take this approach to show that the distortions due to exposure heterogeneity are unrelated to small sample size.

The generalized linear model (14) is a descriptive or phenomenological model, rather than an explanatory model. Its practical use is to quantify the difference in infection risk of the simulated population in the presence or absence of the treatment. In this section we outline some of the errors that can occur when this model is fit to data generated via model (11). However, our intent is not to show that model (14) is somehow inferior to model (11), as in a practical setting the true underlying model is not known and model (14) is designed to be quite flexible to account for this uncertainty. Instead our goal is to isolate the effect that exposure heterogeneity can have on effectiveness estimates.

2.3.2 Increasing Exposure During a Stepped-Wedge Trial

Although the solutions to (11) are rather cumbersome to analyze, we can gain some insight into the potential issues with the stepped-wedge design using the same change of variables as in 2.2.3. Let x_i be the proportion of individuals exposed in wedge i, it's time evolution is defined by

$$\begin{aligned} x_i' &= -\rho_c(1-\epsilon T_i) r x_i + \sigma(f_E - x_i) + \rho_c(1-T_i) r x_i^2 \\ T_i &= \begin{cases} 1 & t \geq i\tau/(1+n_w) \\ 0 & t < i\tau/(1+n_w) \end{cases} \\ x_i(0) &= f_E \end{aligned} \tag{15}$$

As shown in Sect. 2.2.3, in absence of treatment the proportion exposed x_i will initially decline towards the steady state, $x_c^* \approx \sigma f_E/(\sigma + r\rho_c)$.

At the moment the treatment is introduced at $t_1 = i\tau/(n_W+1)$, the instantaneous incidence will drop from $\rho_c r x_i(t_1)$ to $(1-\epsilon)\rho_c r x_i(t_1)$. As a result, the fraction of

exposed individuals x_i will start to approach a new equilibrium $x_a^* \approx \sigma f_E/(\sigma + r\rho_c(1-\epsilon)) > x_c^*$. If $x_i(t_1) < x_a^*$, then the fraction of exposed individuals will start to increase and with it the HIV incidence in the active arm. This might give the false impression that either the force of infection is increasing or the effectiveness of the intervention is declining. In reality, individuals who are exposed can remain uninfected longer and therefore their proportion in the population increases.

2.4 Model Parameterization

We have informed the transmission and behavioral parameters in the models from literature sources representative for South Africa [48, 49]. Simulated levels of HIV incidence among the trial participants are comparable with the results observed in the control arms of recently completed trials (see Table 2). Variables controlling the level of heterogeneity are experimental and therefore explored over a wide range. The exposed fraction of RCT participants in trials conducted among women in sub-Saharan Africa was previously estimated at 20–30% in a published modeling study [19].

Table 2 HIV incidence in prevention trials in sub-Saharan Africa. The list includes randomized controlled studies conducted among uninfected people in the region. Some of the tested products are based on Tenofovir (TDF) and Emtricitabine (FTC)

Study	Population (Period)	Product	HIV incidence in control arm
HPTN 035 [50]	Women, South Africa, Tanzania, Uganda, and Zambia (2005–2008)	Vaginal Gel PRO2000	4.3%
CAPRISA 004 [51]	Women, South Africa (2007–2010)	Vaginal Gel TDF	9.1%
TDF2 [11]	Men and women, Botswana (2007–2010)	Daily oral TDF–FTC	3.1%
FEM PreP [52]	Women, Kenya, South Africa, and Tanzania (2009–2011)	Daily oral TDF–FTC	5.0%
VOICE [53]	Women, South Africa, Uganda, and Zimbabwe (2009–2012)	Oral/Topical TDF–FTC	6.0%
ASPIRE [54]	Women, Malawi, South Africa, Uganda, and Zimbabwe (2012–2015)	Dapivirine vaginal ring	4.5%

3 Main Results

3.1 Heterogeneous Exposure Leads to Low Efficacy Estimates in the Traditional RCT

As stated above, IRR effectiveness should theoretically match true efficacy in case of perfect adherence. For example, consider a population in which all individuals have an expected 40 exposures to HIV per year. If each exposure has a risk of infection of 0.3%, then the annual incidence will be 12%. If we introduce a treatment with 50% efficacy, then the incidence will drop to 6%, yielding an effectiveness of also 50%.

Now suppose that only 50% of the population is exposed to HIV, but the population as a whole has the same total number of exposures. Using the expression (2), the incidence in this population is 11.2%, and drops to 5.6% with a 50% efficacy treatment, an effectiveness of only 48%. This loss of effectiveness gets worse as the exposures become concentrated in a smaller fraction of the population (Fig. 1a), as the overall number of exposures increases (Fig. 1b), and the trial gets longer (Fig. 1c).

An individual's rate of exposure may change in time, especially, for example, if the HIV status of their main partner changes. If we assume that trial participants change from the exposed to unexposed class frequently, then the effectiveness observed in the trial will more closely match the treatment efficacy (Fig. 1d).

3.2 Exposures Rates Become Imbalanced Between Trial Arms over Time

The loss of effectiveness demonstrated in the previous section is due to a growing imbalance in the rate of exposure between control and active arms in a clinical trial. As the trial progresses, exposed individuals become infected, removed from follow-up, and referred to ART. The remaining study population is, therefore, enriched in unexposed individuals and the average rate of exposures per-capita decreases (Fig. 2a,c) when only fraction of the participants are exposed to HIV.

In a traditional RCT trial, the exposed individuals in the control arm are removed faster and therefore the exposure rate becomes increasingly lower than in the active arm. Due to the growing imbalance in exposure rates between the two arms, the incidence in the active arm may approach and eventually surpass the control arm if the trial lasts long enough (Fig. 2b).

In a stepped-wedge trial, the fraction of exposed individuals may start to rebound in the treated cohort. Figure 2c,d shows the exposure and incidence dynamics of a single cohort within a 10 year-long trial assuming different levels of concentration of exposure. The intervention is introduced into this cohort 5 years into the 10-year

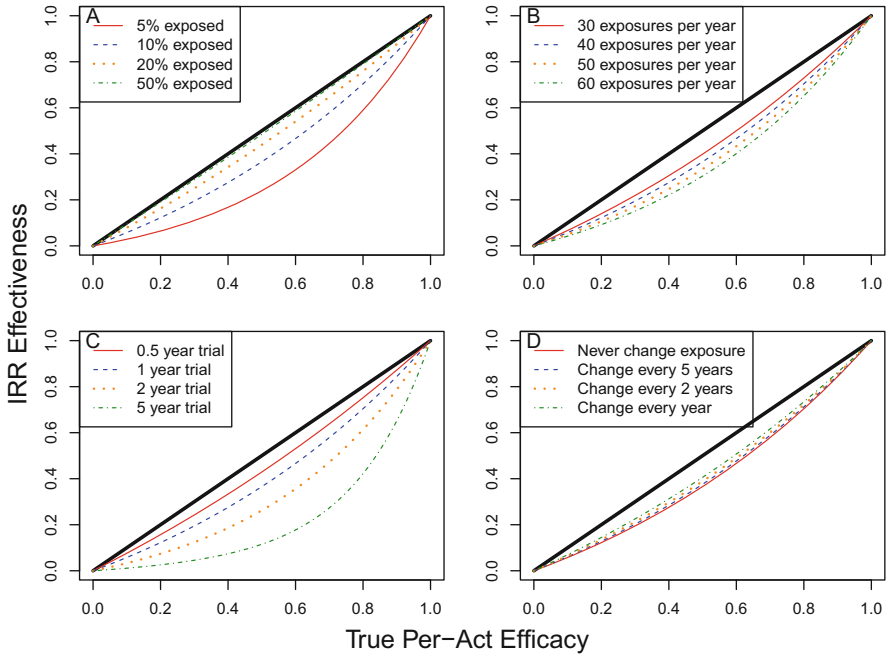


Fig. 1 Decrease in effectiveness in RCT due to concentrated exposure to HIV. Assuming 100% adherence, effectiveness should match efficacy if risk is evenly distributed among all participants (solid black line). Exposure heterogeneity leads to a decrease in effectiveness. This effect is stronger if HIV exposure is more concentrated, i.e., the exposed fraction is smaller (panel **a**), the magnitude (number) of exposures is higher (**b**), the trial is longer (**c**), and if the replacement of exposed individuals is lower (**d**). Base parameters: exposures per year: $r = 40$, risk per infection: $\rho_c = 0.3\%$, exposed fraction: $f_E = 0.1$, length of trial: $\tau = 10$ years, annual exposure update rate: A,B: $\sigma = 0$, C/D: $\sigma = 0.5$

period. These results demonstrate that an increase in HIV incidence may be expected after the initial steep decrease after the intervention is introduced (Fig. 2d).

3.3 Stepped-Wedge Trial Design Can Both Over- and Under-Estimate Efficacy

Stepped-wedge designs can lead to both over and underestimation of the effectiveness, depending upon whether effectiveness is calculated from the incidence rate ratio (IRR effectiveness) or, as is more typically done, deriving a hazard ratio from a generalized linear model (GLM effectiveness).

When HIV exposure becomes concentrated in a smaller fraction of the population, a stepped-wedge design can inflate IRR effectiveness (Fig. 3a). As with the traditional RCT, the effect becomes stronger with increasing magnitude of exposure

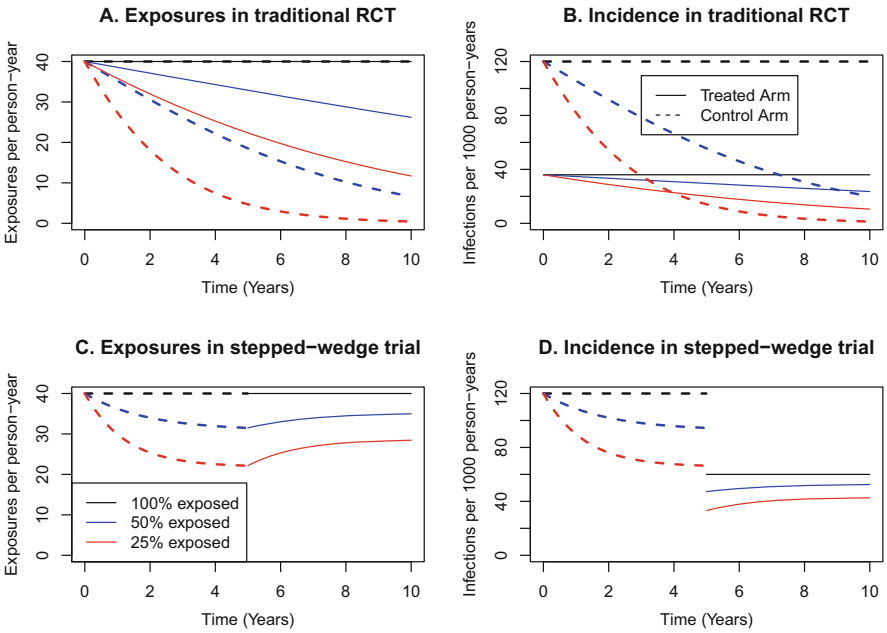


Fig. 2 Imbalances between trial arms due to concentrated exposure to HIV. When exposures are concentrated in a small fraction of the population, the treatment and control arms become imbalanced as the trial progresses. Panels **a–b** depict a traditional RCT with separate and simultaneous control and treatment arms, panels **c–d** depict a single population cohort within a stepped-wedge RCT, which receives the intervention at the same time. Replacement of risk, $\sigma = 0$ for panels **a–b** and $\sigma = 0.50$ for **c–d**. **a/c**: The number of exposures per person year. **b/d**: HIV instantaneous incidence per 1000 person years. Black = 100% exposed ($f_E = 1$), Blue = 50% exposed ($f_E = 0.5$), Red = 25% Exposed ($f_E = 0.25$)

and trial duration but it is partially mitigated when individuals change their exposure status more frequently (Fig. 3b–d). As the intervention is delayed for an initiation period, many of the exposed individuals may be removed from the population cohort before even gaining access to treatment. Therefore the population that receives the intervention will have fewer exposures than the untreated population (Fig. 5).

3.3.1 Calculating Treatment Effect Using a Generalized Linear Model Is Also Error Prone

In a stepped-wedge design, treatment is gradually introduced overtime, so any changes in the underlying risk of infection (for example, improvements in linkage to care for HIV-positive individuals) will confound the effectiveness of the treatment. Therefore, researchers typically attempt to account for these effects, as well as population effects, through a generalized linear model (14).

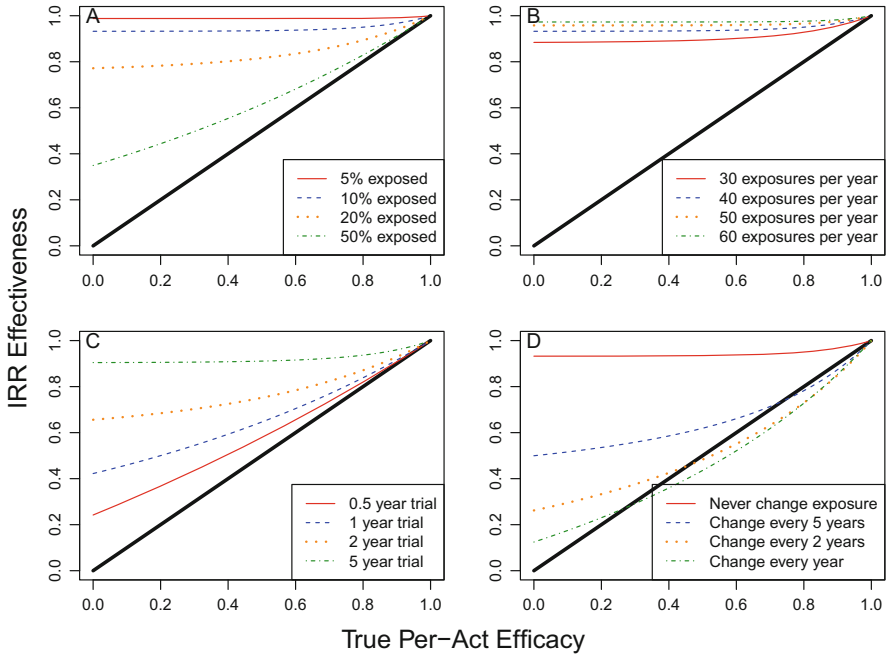


Fig. 3 Increase in effectiveness in stepped-wedge trial due to concentrated HIV exposure. Assuming 100% adherence, effectiveness should match efficacy if risk is evenly distributed among all participants (solid black line). Heterogeneity in HIV exposure leads to both, an increase or a decrease in trial effectiveness. This effect is stronger if the exposure is more concentrated (panel **a**), the magnitude of exposure is higher (**b**), the trial is longer (**c**), and if the replacement of exposed individuals is lower (**d**). Base parameters: exposures per year: $r = 40$, risk per infection: $\rho_c = 0.3\%$, exposed fraction: $f_E = 0.1$, length of trial: $\tau = 6$ years, annual exposure update rate: $\sigma = 0$, number of populations $N_W = 5$

Applying a generalized linear model to a theoretical study in which the treatment is given to five populations over a 6-year period shows an attenuation of the GLM effectiveness similar to the ERR effectiveness of the traditional RCT (Fig. 4).

Although we assume that the risk of infection is unchanged over time, the incidence goes down during each period as infected individuals are removed from the exposed group creating an illusion of decreasing risk (Fig. 5). This decreasing risk over-corrects the GLM effectiveness resulting in an underestimation rather than an overestimation of the intervention effect.

When individuals move between the exposed and unexposed class, another spurious pattern emerges in a stepped-wedge design. After the introduction of treatment, incidence initially drops, but thereafter rises as the number of people in the exposed class starts to increase again. This spurious increase in incidence could be falsely interpreted as a decreased adherence, drug resistance, or risk compensation within the population even if none of these factors are occurring (Fig. 5, lower panel).

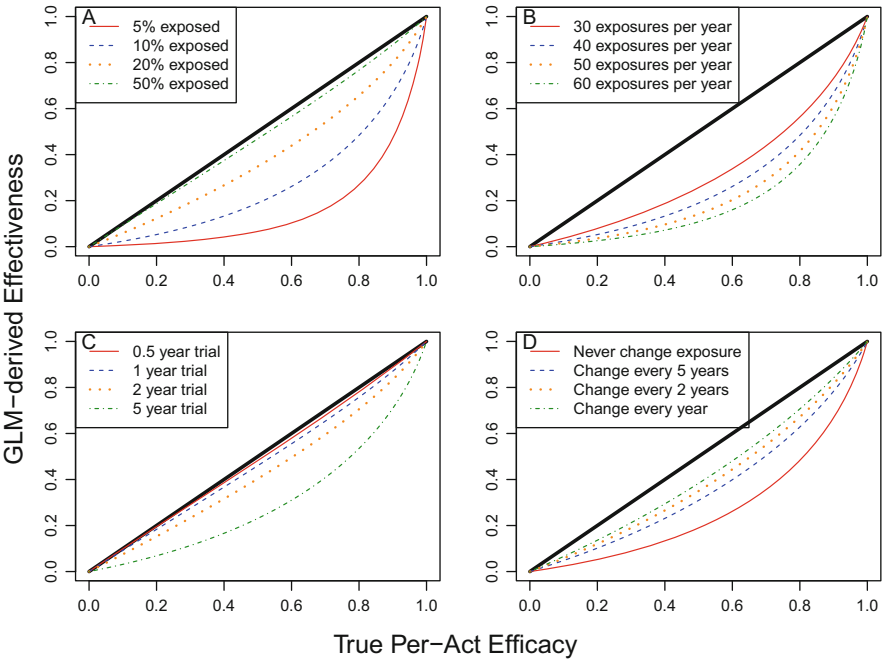


Fig. 4 The use of a linear model in stepped-wedge design restores the decrease in effectiveness. As time and treatment effects are confounded in a stepped-wedge design, a linear model is necessary to match the effectiveness observed in the traditional RCT design which is lower than true efficacy. This effect is stronger if the exposed fraction is smaller (panel **a**), the population incidence is higher (**b**), the trial is longer (**c**), and if the replacement of high-risk individuals is lower (**d**). Parameters are the same as in Fig. 3

4 Discussion

HIV prevention is a critical component in any successful strategy to reduce the global burden of AIDS. In this paper we use simple mathematical models to demonstrate that the effectiveness observed in clinical trials testing novel HIV prevention methods, depends on how evenly the exposure to HIV is distributed among trial participants.

The RCT design is the gold standard for testing prevention effectiveness of novel drugs. The intent-to-treat analysis yields an estimate of the average relative risk of infection between trial arms in the enrolled cohort, typically in the context of a very high standard of intervention delivery. Step-wedge designs were conceived in the context of evaluating efficacy within an implementation framework, typically in a situation where it is not ethical to withhold the treatment (e.g., implementation following proof of effectiveness in an individually randomized trial). Initiation of the treatment is planned for all groups/participants, however, by strategically taking advantage of a plan to roll out implementation through a set of clinics or

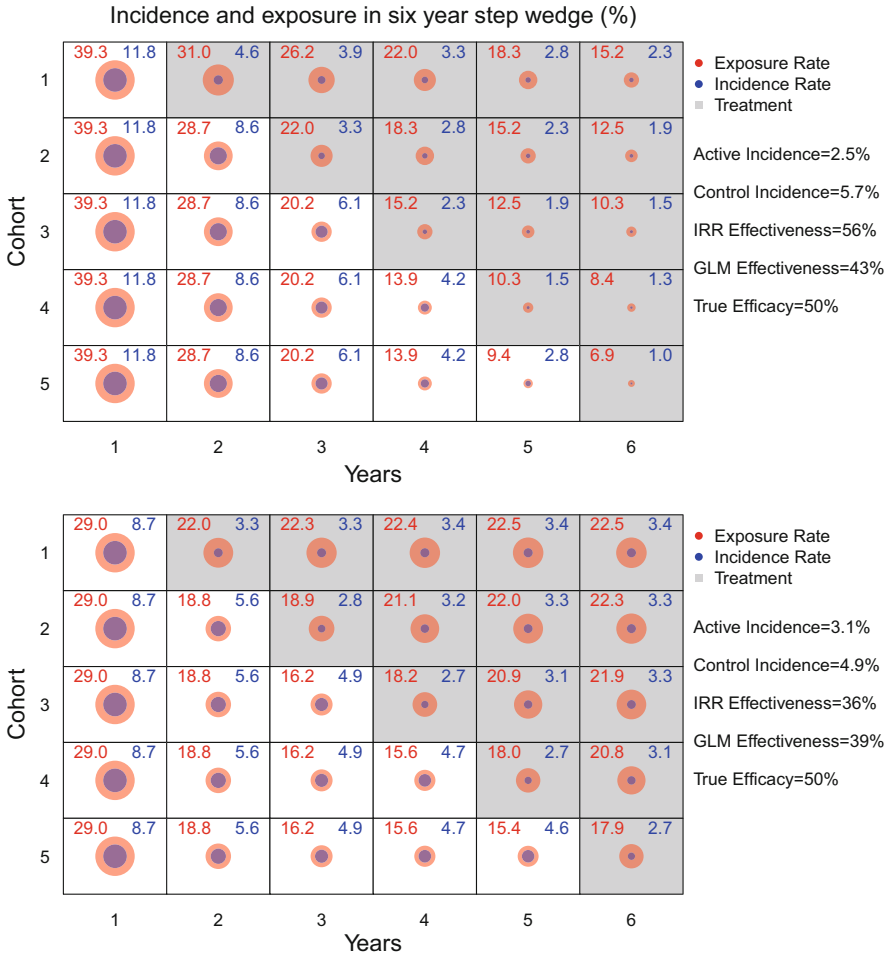


Fig. 5 Examples of stepped-wedge trial effectiveness distortion. A 6-year stepped-wedge trial was simulated in a population with 40 exposures per person per year. The true efficacy was 50%. Exposed fraction is 33% in the upper panel and 15% in the lower panel. Individuals never change exposure level in the upper panel and change at a rate $\sigma = 0.4$ in the lower panel

communities, the step-wedge design makes it possible to estimate the effectiveness of the intervention. As with all group-randomized trials, the total sample size is larger than an individually randomized trials because of intracluster correlation between groups.

Our analysis suggests that the traditional RCT design generally underestimates the true efficacy of the tested HIV prevention with the effect being stronger for more concentrated exposure (smaller fraction of participants having sex with infected partners), for longer trials and for larger overall number of exposures in the trial cohort. In comparison, the step-wedge design may either over- or underestimate the

true efficacy depending on whether the incidence rate ratio or the linear regression model is used to estimate effectiveness.

Our analysis aims to initiate discussion on at least two important questions. First, how real is the problem with concentration of exposure and what is the estimated proportion of participants in RCTs exposed to HIV? Multiple studies based on data from different epidemic settings suggest that the majority of participants in completed trials remain unexposed to HIV during their follow-up. Individual-based mathematical model which simulates detailed sexual activity of the participants in RCT among high-risk women in South Africa estimates that fewer than 30% of the participants have been exposed to HIV at least once over 1-year follow-up with up to 7% annual incidence projected [19]. Detection of HIV DNA in vaginal swabs was proposed as biomarker of HIV exposure in women. However, when tested in clinical studies it was not able to indicate exposure to HIV in uninfected women, even if they acquired HIV later in the study [55, 56]. Second, what are the possible ways to address the issue with uneven HIV exposure? Crossover designs in which participants switch their arm assignment (active vs. placebo) in the middle of the follow-up period have been suggested to alleviate the problem with heterogeneity in risk [43]. However, these designs are only applicable when the tested product is used repeatedly in the forms of pills, injections, topical gels, etc., but not for one-time or few-dose interventions such as male circumcision or HIV vaccines. Limiting enrollment to participants who are regularly exposed to HIV, such as serodiscordant couples, may reduce the effect of uneven exposure but poses questions with transferability of the observed results to populations without clear knowledge of their HIV risk [57].

Our results show that heterogeneity in exposure can decrease the effectiveness observed in RCTs. This downward bias means that larger and more expensive clinical trials may be needed in order to achieve the same power. In a recent modeling study we demonstrated that enrolling different proportions of individuals with uniformly high HIV exposure, such as female sex workers, will help reduce trial duration and improve efficacy estimates [58]. The analysis here provides additional rationale for these results and warrants further investigation of feasible recruitment strategies limiting the effects of the concentration of HIV exposure.

Acknowledgments The authors are grateful to Megan Rabone, Executive Assistant at the Vaccine and Infectious Disease Division of Fred Hutchinson Cancer Research Center for proofreading and editing the final version of the manuscript.

The authors acknowledge the NIH support to the HIV Prevention Trials Network (UM1 AI068617 and 5UM1AI068619-11).

References

1. Stanaway, J.D., Afshin, A., Gakidou, E., et al.: Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. Lancet **392**(10159), 1923–1994 (2019)

2. Frank, T.D., Carter, A., Jahagirdar, D., et al.: Global, regional, and national incidence, prevalence, and mortality of HIV, 1980–2017, and forecasts to 2030, for 195 countries and territories: a systematic analysis for the Global Burden of Diseases, Injuries, and Risk Factors Study 2017. Lancet HIV **6**(12), e831–e859 (2019)
3. Geneva Joint United Nations Programme on HIV/AIDS. Understanding fast-track; accelerating action to end the AIDS epidemic by 2030 (2015). http://www.unaids.org/sites/default/files/media_asset/201506_JC2743_Understanding_FastTrack_en.pdf. Cited Mar 2019
4. Fauci, A.S., Redfield, R.R., Sigounas, G., Weahkee, M.D., Giroir, B.P.: Ending the HIV epidemic: a plan for the United States. JAMA **321**(9), 844–845 (2019)
5. Taha, T.E., Kumwenda, N.I., Hoover, D.R., Fiscus, S.A., Kafulafula, G., Nkhoma, C., Nour, S., Chen, S., Liomba, G., Miotti, P.G., Broadhead, R.L.: Nevirapine and zidovudine at birth to reduce perinatal transmission of HIV in an African setting – a randomized controlled trial. JAMA **292**(2), 202–209 (2004)
6. Auvert, B., Taljaard, D., Lagarde, E., Sobngwi-Tambekou, J., Sitta, R., Puren, A.: Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: The ANRS 1265 trial. PLOS Med. **2**(11), e298 (2005)
7. Bailey, R.C., Moses, S., Parker, C.B., Agot, K., Maclean, I., Krieger, J.N., Williams, C.F., Campbell, R.T., Ndinya-Achola, J.O.: Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial. Lancet **369**(9562), 643–56 (2007)
8. Gray, R.H., Kigozi, G., Serwadda, D., Makumbi, F., Watya, S., Nalugoda, F., Kiwanuka, N., Moulton, L.H., Chaudhary, M.A., Chen, M.Z., Sewankambo, N.K., Wabwire-Mangen, F., Bacon, M.C., Williams, C.F., Opendi, P., Reynolds, S.J., Laeyendecker, O., Quinn, T.C., Wawer, M.J.: Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial. Lancet **369**(9562), 657–66 (2007)
9. Grant, R.M., Lama, J.R., Anderson, P.L., McMahan, V., Liu, A.Y., Vargas, L., Goicochea, P., Casapia, M., Guanira-Carranza, J.V., Ramirez-Cardich, M.E., Montoya-Herrera, O., Fernandez, T., Veloso, V.G., Buchbinder, S.P., Chariyalertsak, S., Schechter, M., Bekker, L.G., Mayer, K.H., Kallas, E.G., Amico, K.R., Mulligan, K., Bushman, L.R., Hance, R.J., Ganoza, C., Defechereux, P., Postle, B., Wang, F.R., McConnell, J.J., Zheng, J.H., Lee, J., Rooney, J.F., Jaffe, H.S., Martinez, A.I., Burns, D.N., Glidden, D.V.: Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N. Engl. J. Med. **363**(27), 2587–2599 (2010)
10. Cohen, M.S., Chen, Y.Q., McCauley, M., Gamble, T., Hosseinipour, M.C., Kumarasamy, N., Hakim, J.G., Kumwenda, J., Grinsztejn, B., Pilotto, J.H.S., Godbole, S.V., Mehendale, S., Chariyalertsak, S., Santos, B.R., Mayer, K.H., Hoffman, I.F., Eshleman, S.H., Piwowar-Manning, E., Wang, L., Makhema, J., Mills, L.A., de Bruyn, G., Sanne, I., Eron, J., Gallant, J., Havlir, D., Swindells, S., Ribaudo, H., Elharrar, V., Burns, D., Taha, E.T., Nielsen-Saines, K., Celentano, D., Essex, M., Fleming, T.R.: Prevention of HIV-1 infection with early antiretroviral therapy. N. Engl. J. Med. **365**(6), 493–505 (2011)
11. Thigpen, M.C., Kebaabetswe, P.M., Paxton, L.A., Smith, D.K., Rose, C.E., Segolodi, T.M., Henderson, F.L., Pathak, S.R., Soud, F.A., Chillag, K.L., Mutanhaurwa, R., Chirwa, L.I., Kasonde, M., Abebe, D., Buliva, E., Gvetadze, R.J., Johnson, S., Sukalac, T., Thomas, V.T., Hart, C., Johnson, J.A., Malotte, C.K., Hendrix, C.W., Brooks, J.T.: Antiretroviral preexposure prophylaxis for heterosexual HIV transmission in Botswana. N. Engl. J. Med. **367**(5), 423–434 (2012)
12. Baeten, J.M., Donnell, D., Ndase, P., Mugo, N.R., Campbell, J.D., Wangisi, J., Tappero, J.W., Bukusi, E.A., Cohen, C.R., Katabira, E., Ronald, A., Tumwesigye, E., Were, E., Fife, K.H., Kiarie, J., Farquhar, C., John-Stewart, G., Kakia, A., Odoyo, J., Mucunguzi, A., Nakku-Joloba, E., Twesigye, R., Ngure, K., Apaka, C., Tamooh, H., Gabona, F., Mujugira, A., Panteleeff, D., Thomas, K.K., Kidoguchi, L., Krows, M., Revall, J., Morrison, S., Haugen, H., Emmanuel-Ogier, M., Ondrejcek, L., Coombs, R.W., Frenkel, L., Hendrix, C., Bumpus, N.N., Bangsberg, D., Haberer, J.E., Stevens, W.S., Lingappa, J.R., Celum, C.: Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. N. Engl. J. Med. **367**(5), 399–410 (2012)
13. Molina, J.-M., Capitant, C., Spire, B., Pialoux, G., Cotte, L., Charreau, I., Tremblay, C., Le Gall, J.-M., Cua, E., Pasquet, A., Raffi, F., Pintado, C., Chidiac, C., Chas, J., Charbonneau,

P., Delaugerre, C., Suzan-Monti, M., Loze, B., Fonsart, J., Peytavin, G., Cheret, A., Timsit, J., Girard, G., Lorente, N., Preau, M., Rooney, J.F., Wainberg, M.A., Thompson, D., Rozenbaum, W., Dore, V., Marchand, L., Simon, M.-C., Etien, N., Aboulker, J.-P., Meyer, L., Delfraissy, J.-F.: On-demand preexposure prophylaxis in men at high risk for HIV-1 infection. N. Engl. J. Med. **373**(23), 2237–2246 (2015)

14. HIV Prevention Trials Network. HVTN 703/HPTN 081: a phase 2b study to evaluate the safety and efficacy of VRC01 broadly neutralizing monoclonal antibody in reducing acquisition of HIV-1 infection in women in sub-Saharan Africa. https://www.hptn.org/research/studies/703_081. Cited Mar 2019
15. HIV Prevention Trials Network. HPTN 083: A phase 2b/3 Double Blind Safety and Efficacy Study of Injectable Cabotegravir Compared to Daily Oral Tenofovir Disoproxil Fumarate/Emtricitabine (TDF/FTC), for Pre-Exposure Prophylaxis in HIV-Uninfected Cisgender Men and Transgender Women who have Sex with Men. https://www.hptn.org/research/studies/hptn083. Cited Mar 2019
16. HIV Prevention Trials Network. HPTN 084: A Phase 3 Double Blind Safety and Efficacy Study of Long-Acting Injectable Cabotegravir Compared to Daily Oral TDF/FTC for Pre-Exposure Prophylaxis in HIV-Uninfected Women. https://www.hptn.org/research/studies/hptn084. Cited Mar 2019
17. HIV Prevention Trials Network. HVTN 704/HPTN 085: A phase 2b study to evaluate the safety and efficacy of VRC01 broadly neutralizing monoclonal antibody in reducing acquisition of HIV-1 infection among men and transgender persons who have sex with men. https://www.hptn.org/research/studies/704_085. Cited Mar 2019
18. U.S. National Library of Medicine NIH. Pivotal Phase 2b/3 ALVAC/Bivalent gp120/MF59 HIV Vaccine Prevention Safety and Efficacy Study in South Africa (HVTN702). https://clinicaltrials.gov/ct2/show/NCT02968849 (2016). Visited March 2019
19. Dimitrov, D., Donnell, D., Brown, E.R.: High incidence is not high exposure: what proportion of prevention trial participants are exposed to HIV? PLoS ONE **10**(1), e0115528 (2015)
20. Kermack, W.O., McKendrick, A.G.: Contributions to the mathematical theory of epidemics–I. Bull. Math. Biol. **53**(1), 33–55 (1991)
21. Castillo-Chavez, P.C., Blower, S., Driessche, P.van d., Kirschner, D., Yakubu, A.-A. (eds.): Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods, and Theory. Springer, New York (2002)
22. Brauer, F., Castillo-Chavez, C.: Mathematical Models in Population Biology and Epidemiology. Texts in Applied Mathematics. Springer, New York (2012)
23. Dimitrov, D.T., Hallam, T.G., Rupprecht, C.E., Turmelle, A.S., McCracken, G.F.: Integrative models of bat rabies immunology, epizootiology and disease demography. J. Theor. Biol. **245**(3), 498–509 (2007)
24. Gaff, H.D., Gross, L.J.: Modeling tick-borne disease: a metapopulation model. Bull. Math. Biol. **69**(1), 265–288 (2007)
25. Feng, Z.L., Castillo-Chavez, C., Capurro, A.F.: A model for tuberculosis with exogenous reinfection. Theor. Popul. Biol. **57**(3), 235–247 (2000)
26. Reluga, T.C., Medlock, J., Perelson, A.S.: Backward bifurcations and multiple equilibria in epidemic models with structured immunity. J. Theor. Biol. **252**(1), 155–165 (2008)
27. Dimitrov, D.T., Troeger, C., Halloran, M.E., Longini, I.M., Chao, D.L.: Comparative effectiveness of different strategies of oral cholera vaccination in Bangladesh: a modeling study. PLoS Negl. Trop. Dis. **8**(12), e3343 (2014)
28. King, A.A., de Celles, M.D., Magpantay, F.M.G., Rohani, P.: Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to Ebola. Proc. Biol. Sci. **282**(1806) (2015)
29. Moore, J., Ahmed, H., Jia, J., Akondy, R., Ahmed, R., Antia, R.: What controls the acute viral infection following yellow fever vaccination? Bull. Math. Biol. **80**(1), 46–63 (2018)
30. Dimitrov, D., Kublin, J.G., Ramsey, S., Corey, L.: Are clade specific HIV vaccines a necessity? An analysis based on mathematical models. EBioMed. **2**(12), 2062–2069 (2015). PMCID: PMC4703729

31. de Montigny, S., Adamson, B.J., Mâsse, Jr. B.R., Garrison, L.P., Kublin, J.G., Gilbert, P.B., Dimitrov, D.T.: Projected effectiveness and added value of HIV vaccination campaigns in South Africa: a modeling study. Sci. Rep. **8**(1), 6066 (2018). PMCID: PMC5904131
32. Goodreau, S.M., Carnegie, N.B., Vittinghoff, E., Lama, J.R., Sanchez, J., Grinsztejn, B., Koblin, B.A., Mayer, K.H., Buchbinder, S.P.: What drives the US and Peruvian HIV epidemics in men who have sex with men (MSM)? PLoS One **7**(11), e50522 (2012). PMCID: PMC3510067
33. Vassall, A., Pickles, M., Chandrashekar, S., Boily, M.-C., Shetty, G., Guinness, L., Lowndes, C.M., Bradley, J., Moses, S., Alary, M., Vickerman, P.: Cost-effectiveness of HIV prevention for high-risk groups at scale: an economic evaluation of the Avahan programme in south India. Lancet Glob. Health **2**(9), e531–e540 (2014)
34. Mitchell, K.M., Dimitrov, D., Hughes, J.P., Xia, F., Donnell, D., Amico, K.R., Bokoch, K., Chitwarakorn, A., Bekker, L.G., Holtz, T.H., Mannheimer, S., Grant, R.M., Boily, M.C.: In what circumstances could nondaily preexposure prophylaxis for HIV substantially reduce program costs? AIDS **32**(6), 809–818 (2018)
35. Dimitrov, D.T., Boily, M. C., Baggaley, R.F., Masse, B.: Modeling the gender-specific impact of vaginal microbicides on HIV transmission. J. Theor. Biol. **288**, 9–20 (2011)
36. Eaton, J.W., Menzies, N.A., Stover, J., Cambiano, V., Chindelevitch, L., Cori, A., Hontelez, J.A.C., Humair, S., Kerr, C.C., Klein, D.J., Mishra, S., Mitchell, K.M., Nichols, B.E., Vickerman, P., Bakker, R., Bärnighausen, T., Bershteyn, A., Bloom, D.E., Boily, M.-C., Chang, S.T., Cohen, T., Dodd, P.J., Fraser, C., Gopalappa, C., Lundgren, J., Martin, N.K., Mikkelsen, E., Mountain, E., Pham, Q.D., Pickles, M., Phillips, A., Platt, L., Pretorius, C., Prudden, H.J., Salomon, J.A., van de Vijver, D.A.M.C., de Vlas, S.J., Wagner, B.G., White, R.G., Wilson, D.P., Zhang, L., Blandford, J., Meyer-Rath, G., Remme, M., Revill, P., Sangrujee, N., Terris-Prestholt, F., Doherty, M., Shaffer, N., Easterbrook, P.J., Hirnschall, G., Hallett, T.B.: Health benefits, costs, and cost-effectiveness of earlier eligibility for adult antiretroviral therapy and expanded treatment coverage: a combined analysis of 12 mathematical models. Lancet Glob. Health **2**(1), e23–e34 (2014)
37. Adamson, B., Dimitrov, D., Devine, B., Barnabas, R.: The potential cost-effectiveness of HIV vaccines: a systematic review. PharmacoEconomics **1**(1), 1–12 (2017)
38. Stover, J., Hallett, T.B., Wu, Z., Warren, M., Gopalappa, C., Pretorius, C., Ghys, P.D., Montaner, J., Schwartlander, B.: How can we get close to zero? the potential contribution of biomedical prevention and the investment framework towards an effective response to HIV. Plos One **9**(11), e111956 (2014)
39. Medlock, J., Pandey, A., Parpia, A.S., Tang, A., Skrip, L.A., Galvani, A.P.: Effectiveness of UNAIDS targets and HIV vaccination across 127 countries. Proc. Natl. Acad. Sci. 114(15), 4017–4022 (2017)
40. Halloran, M.E., Longini, I.M., Haber, M.J., Struchiner, C.J., Brunet, R.C.: Exposure efficacy and change in contact rates in evaluating prophylactic HIV-1 vaccines in the field. Stat. Med. **13**(4), 357–377 (1994)
41. Desai, K.N., Boily, M.C., Masse, B.R., Alary, M., Anderson, R.M.: Simulation studies of phase III clinical trials to test the efficacy of a candidate HIV-1 vaccine. Epidemiol. Infect. **123**(1), 65–88 (1999)
42. Wilson, D.P.: Interpreting sexually transmissible infection prevention trials by adjusting for the magnitude of exposure. Clin. Trial. **7**(1), 36–43 (2010)
43. Auvert, B., Sitta, R., Zarca, K., Mahiane, S.G., Pretorius, C., Lissouba, P.: The effect of heterogeneity on HIV prevention trials. Clin. Trial. **8**(2), 144–154 (2011)
44. Brown, C.A., Lilford, R.J.: The stepped wedge trial design: a systematic review. BMC Med. Res. Methodol. **6**(1), 54 (2006)
45. Padian, N.S., McCoy, S.I., Manian, S., Wilson, D., Schwartländer, B., Bertozzi, S.M.: Evaluation of large-scale combination HIV prevention programs: essential issues. J. Acquir. Immune Defic. Syndr. **58**(2), e23 (2011)

46. Mdege, N.D., Man, M.-S., Taylor, C.A., Torgerson, D.J.: Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. J. Clin. Epidemiol. **64**(9), 936–948 (2011)
47. Hussey, M.A., Hughes, J.P.: Design and analysis of stepped wedge cluster randomized trials. Contemp. Clin. Trial. **28**(2), 182–191 (2007)
48. Boily, M.C., Baggaley, R.F., Wang, L., Masse, B., White, R.G., Hayes, R.J., Alary, M.: Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. Lancet Infect. Dis. **9**(2), 118–129 (2009)
49. Kalichman, S.C., Simbayi, L.C., Cain, D., Jooste, S.: Heterosexual anal intercourse among community and clinical settings in Cape Town, South Africa. Sex. Transm. Infect. **85**(6), 411–415 (2009)
50. McCormack, S., Ramjee, G., Kamali, A., Rees, H., Crook, A.M., Gafos, M., Jentsch, U., Pool, R., Chisembele, M., Kapiga, S., Mutemwa, R., Vallely, A., Palanee, T., Sookrajh, Y., Lacey, C.J., Darbyshire, J., Grosskurth, H., Profy, A., Nunn, A., Hayes, R., Weber, J.: Pro2000 vaginal gel for prevention of HIV-1 infection (microbicides development programme 301): a phase 3, randomised, double-blind, parallel-group trial. Lancet **376**(9749), 1329–1337 (2010)
51. Abdool Karim, Q., Abdool Karim, S.S., Frohlich, J.A., Grobler, A.C., Baxter, C., Mansoor, L.E., Kharsany, A.B., Sibeko, S., Mlisana, K.P., Omar, Z., Gengiah, T.N., Maarschalk, S., Arulappan, N., Mlotshwa, M., Morris, L., Taylor, D., CAPRISA 004 Trial Group.: Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. Science **329**(5996), 1168–1174 (2010)
52. Van Damme, L., Corneli, A., Ahmed, K., Agot, K., Lombaard, J., Kapiga, S., Malahleha, M., Owino, F., Manongi, R., Onyango, J., Temu, L., Monedi, M.C., Mak'Oketch, P., Makanda, M., Reblin, I., Makatu, S.E., Saylor, L., Kiernan, H., Kirkendale, S., Wong, C., Grant, R., Kashuba, A., Nanda, K., Mandala, J., Fransen, K., Deese, J., Crucitti, T., Mastro, T.D., Taylor, D.: Preexposure prophylaxis for HIV infection among African women. N. Engl. J. Med. **367**(5), 411–422 (2012)
53. Marrazzo, J.M., Ramjee, G., Richardson, B.A., Gomez, K., Mgodi, N., Nair, G., Palanee, T., Nakabiito, C., van der Straten, A., Noguchi, L., Hendrix, C.W., Dai, J.Y., Ganesh, S., Mkhize, B., Taljaard, M., Parikh, U.M., Piper, J., Msse, B., Grossman, C., Rooney, J., Schwartz, J.L., Watts, H., Marzinke, M.A., Hillier, S.L., McGowan, I.M., Mike Chirenje, Z.: Tenofovir-based preexposure prophylaxis for HIV infection among African women. N. Engl. J. Med. **372**(6), 509–518 (2015)
54. Baeten, J.M., Palanee-Phillips, T., Brown, E.R., Schwartz, K., Soto-Torres, L.E., Govender, V., Mgodi, N.M., Matovu Kiweewa, F., Nair, G., Mhlanga, F., Siva, S., Bekker, L.-G., Jeenarain, N., Gaffoor, Z., Martinson, F., Makanani, B., Pather, A., Naidoo, L., Husnik, M., Richardson, B.A., Parikh, U.M., Mellors, J.W., Marzinke, M.A., Hendrix, C.W., van der Straten, A., Ramjee, G., Chirenje, Z.M., Nakabiito, C., Taha, T.E., Jones, J., Mayo, A., Scheckter, R., Berthiaume, J., Livant, E., Jacobson, C., Ndase, P., White, R., Patterson, K., Germuga, D., Galaska, B., Bunge, K., Singh, D., Szydlo, D.W., Montgomery, E.T., Mensch, B.S., Torjesen, K., Grossman, C.I., Chakhtoura, N., Nel, A., Rosenberg, Z., McGowan, I., Hillier, S.: Use of a vaginal ring containing dapivirine for HIV-1 prevention in women. N. Engl. J. Med. **375**(22), 2121–2132 (2016)
55. Penrose, K.J., Richardson, B.A., Besson, G., Dezzutti, C.S., Herold, B.C., Abdool Karim, S.S., Mellors, J.W., Parikh, U.M., M.T.N. Biomedical Sciences Working Group, and the HPTN 035 Protocol Team: Y chromosome and HIV DNA detection in vaginal swabs as biomarkers of semen and HIV exposure in women. Sex. Transm. Infect. **41**(25299415), 674–679 (2014)
56. Lemos, M.P., Lazarus, E., Isaacs, A., Dietrich, J., Morgan, C., Huang, Y., Grove, D., Andrasik, M., Laher, F., Hural, J., Chung, E., Dragavon, J., Puren, A., Gulati, R.K., Coombs, R., McElrath, M.J., Gray, G., Kublin, J.G.: Daily vaginal swabs and mobile phone sex report for assessing HIV virion exposure prospectively among a cohort of young sexually active women in South Africa (HVTN 915). J. Acquir. Immune Defic. Syndr. **81**(2) (2019)

57. Pyra, M., Brown, E.R., Haberer, J.E., Heffron, R., Celum, C., Bukusi, E.A., Asiimwe, S., Katabira, E., Mugo, N.R., Baeten, J.M., Baeten, J., Celum, C., Heffron, R., Donnell, D., Barnabas, R., Haberer, J., Haugen, H., Hendrix, C., Kidoguchi, L., Marzinke, M., Morrison, S., Morton, J., Ware, N., Wyatt, M.: Patterns of oral prep adherence and HIV risk among eastern African women in HIV serodiscordant partnerships. AIDS Behav. **22**(11), 3718–3725 (2018)
58. Wood, D., Lancaster, K.E., Boily, M.-C., Powers, K.A., Donnell, D., Cohen, M.S., Dimitrov, D.T.: Recruitment of female sex workers in HIV prevention trials: can efficacy endpoints be reached more efficiently? J. Acquir. Immune Defic. Syndr. **77**(4), 350–357 (2018)

Dynamic Regulation of T Cell Activation by Coupled Feedforward Loops

Gershom Buri, Girma Mesfin Zelleke, and Wilfred Ndifon

1 Introduction

The immune system serves to defend the body against attacks by both foreign and internal threats. Lack of knowledge about the mechanisms of the immune system has sometimes proved costly in terms of human morbidity and death, particularly, during the pre-vaccination era. The discovery of vaccination motivated the quest for a deeper understanding of the immune system in order to harness it to control various life-threatening infections. The results were more effective disease treatments and the successful prevention by vaccines of formerly fatal and disabling diseases like polio, measles, whopping cough and many others. Most of these vaccines were, however, based on the humoral component of the immune system, which has proven less effective against diseases like malaria, HIV/AIDS and TB, protection against which requires cellular immunity as well. This has prompted the search for T cell-based vaccines to supplement existing humoral vaccines. To succeed at this, however, an adequate understanding of T cell activation dynamics is necessary.

T cells are the main orchestrators of the adaptive immune response and, if not well regulated, can have perilous consequences. Both a prolonged response and a response to a wrong target can lead to destruction of body tissue and have been

G. Buri
African Institute for Mathematical Sciences, Cape Town, South Africa

SACEMA, Stellenbosch University, Cape Town, South Africa

G. M. Zelleke
African Institute for Mathematical Sciences, Limbe, Cameroon

W. Ndifon (✉)
Research Department, African Institute for Mathematical Sciences Global Secretariat, Kigali, Rwanda
e-mail: wndifon@aims.ac.za

M. I. Teboh-Ewungkem, G. A. Ngwa (eds.), *Infectious Diseases and Our Planet*, Mathematics of Planet Earth 7, https://doi.org/10.1007/978-3-030-50826-5_9

linked to pregnancy complications, stroke, heart attacks and blood clots [1]. In addition, they can lead to wastage of resources such as energy used in the production of cytokines [2, 3]. It is, therefore, imperative that T cells get activated only by the right kind of signals and for the correct length of time. How is this achieved?

T cell activation is triggered by the binding of a T cell receptor (TCR) to its cognate antigen. First reported in Yanagi et al. [4], the TCR is a specialised molecule found on the surface of a T cell that confers the ability to recognise antigenic peptides. However, in the absence of costimulation, TCR-mediated recognition of antigen is generally not sufficient to produce the necessary T cell proliferation and differentiation. Instead, it has been reported to cause a state of nonresponsiveness called anergy [5]. The two-signal theory posits that T cell activation requires two signals: one in the form of antigen recognition by the TCR and another in the form of a secondary stimulus, now known as costimulation. The requirement of two signals for lymphocyte activation was first proposed as a mechanism for distinguishing the self from the non-self in B cells [6]. This model was later extended to T cell activation by Lafferty and Cunningham [7], and supported by the identification of the CD28 receptor, with the B7-1 molecule as its ligand [8, 9]. Studies in CD28 deficient mice revealed a diminished response to a variety of pathogens [10] even though other signals could compensate for this absence to a limited extent [11].

Following the discovery of CD28, additional costimulatory molecules such as ICOS, OX40 and 4-1BB were identified [12, 13]. The complexity of this activation regulatory system, however, became more apparent with the discovery of coinhibitory receptors such as CTLA-4, PD-1, LAG-3 and BTLA. These receptors are expressed on the surfaces of activated T cells and downregulate T cell activation using different mechanisms and at distinct times [14–16]. For example, CTLA-4 binds to B7 ligands and thus inhibits T cell activation through competition with CD28 in addition to other intracellular signalling mechanisms [17]. These molecules provide an important avenue to implement the regulatory mechanisms of T cells, and their absence in mice has been associated with autoimmunity [18, 19].

Amongst the key cellular players of the immune system, with antagonistic roles, are effector and regulatory T cells. Effector T cells serve to coordinate immune responses and clear health threats, whilst regulatory T cells downregulate immune responses effected by effector T cells. In doing so, regulatory T cells play a critical role in maintaining peripheral tolerance by suppressing activation of self-reacting effector T cells [20, 21]. With the uncertainty about the existence of regulatory T cells in the beginning [22, 23], costimulation was initially defined solely with respect to effector T cells [24, 25]. Moreover, based on in vitro data, regulatory T cells were initially considered anergic [26, 27]. However, following a renewed interest in regulatory T cells, the discovery that costimulation is necessary for their activation played a critical role in our understanding of the underlying mechanisms [28–30]. It was established that some costimulatory molecules, for example, CD28 and IL-2, had dual (opposing) effects on immune responses. This had a consequence of further confounding the mechanisms influencing the net balance of regulatory vs. effector T cell responses, a significant determinant of overall immune response outcomes [28, 31, 32]. In addition, it made the prediction of therapeutic outcomes

aimed at blocking any of these signalling pathways difficult [24, 33, 34]. Attempts have since been made to elucidate this conundrum. Vogel et al. [35] report a lower dependence on costimulation by regulatory vs. effector T cells, whilst Hombach et al. [36] report a higher costimulation threshold for regulator T cells compared with effector T cells, in an infection of healthy human subjects.

Here, we use the concepts and mathematical machinery of network science to shed further light on these complex regulatory interactions. Despite the discovery of various other cosignalling molecules, signal transduction via CD28 receptors and the B7 molecules remains the best defined initiator of T cell activation [37, 38]. We restrict our attention to this classical version of the two-signal theory, which espouses necessity of costimulation via CD28-B7 signalling, even though some quantitative descriptions regard costimulation as a facilitator rather than a necessity [10, 39]. We save investigation of the dynamic interplay between costimulation and coinhibition for the future. In addition, we apply the two-signal requirement to both effector and regulatory T cells. We show that the two-signal requirement induces a novel type-1 coherent feedforward loop. Feedforward loops (FFL) are amongst the most significant transcriptional regulatory motifs found in diverse gene systems [40–42], and in other biological networks [43, 44]. Notably, this motif is capable of signal processing as a sign-sensitive delay element. We argue that this is a desirable property of any regulatory system, more so, of T cell activation.

The mathematical instantiation of our hypothesis combines models for gene transcription regulation [41] with a canonical model for virus growth, widely used to investigate viral and immune system dynamics [45–48]. A comprehensive literature review reveals that no previous study has taken such a mechanistic approach to model immune system dynamics based on the two-signal theory of T cell activation. Initial attempts to model the downstream signalling pathways of the TCR and costimulatory receptors include a Boolean network model by Saez-Rodriguez et al. [49], which was later extended by Beyer et al. [50]. These models rely on logical formalism, are mainly computational and consider more detailed costimulation mechanisms.

Next, we analyse the implications of our novel hypothesis on the interaction between effector and regulatory T cells. The result is a coupled coherent-incoherent feedforward loop. A number of mathematical models have previously been proposed to capture the dynamics of this interaction in different pathological conditions [51–55]. Still, these are mainly empirical models that assume already activated T cell populations. For example, a related model by Sontag [56] analyses this interaction by combining negative feedback and type 1 incoherent feed forward loops to model immune responses to antigen presentation.

2 Results

Briefly, a typical immune response to an infectious pathogen commences with the engulfing and subsequent processing of the pathogens by antigen-presenting cells

(APCs) of the innate immune system. Whilst bound to MHC class II molecules, pathogen-derived peptides are displayed for recognition by CD4 T cells [57]. Upon successful recognition of such peptide-MHC complexes (pMHCs), the T cells are activated, and go on to orchestrate the rest of the immune response against the infection via various avenues including enhancement of CD8 T cell and B cell responses. For convenience, we will collectively refer to these CD4 and CD8 T cells as effector T cells. In situations of autoimmunity or an excessive immune response, regulatory T cells intervene to suppress or regulate effector cell activity [58]. In the following, we will study the dynamics of T cell activation and the interaction between effector and regulatory T cells using mathematical models inspired by recent advances in network science.

2.1 *T cell Activation is Underpinned by a Type-1 Coherent FFL (C1-FFL) with AND Logic*

As mentioned above, activation of T cells requires at least two signals: one signal in the form of TCR-pMHC binding, and the second mediated by costimulatory proteins found on the surfaces of T cells and APCs. The B7 molecules are only expressed on professional APCs, and their production peaks after the APCs are activated by microbial products [58]. The combined action of the two signals activates T cells to proliferate and begin to differentiate into effector T cells [58, 59]. In this way, the two-signal requirement induces a FFL in which the antigen signal (A) regulates the costimulatory signal (C), and both jointly regulate T cell (T_e) activation (Fig. 1). This particular case of the FFL is "coherent" because the sign of the direct regulation path (from A to T_e) is the same as the overall sign of the indirect path through C. This FFL motif is of type-1 because all the paths are positive (activators). Also, since both signals are required to activate the T cell, the input function of T_e activation follows an AND logic gate. In contrast, an OR-gate would have been appropriate if either of the two signals was sufficient to activate the T cell.

2.2 *The C1-FFL Induces a Sign-Sensitive Delay in T cell Activation and Prevents Responses to Transient Antigen Signals*

To permit assessment of the dynamical effects of the C1-FFL on T cell activation, we instantiated the C1-FFL in a mathematical model (see Sect. 4). Numerical simulations based on the mathematical model show that the C1-FFL can serve as a sign-sensitive delay element in T cell activation (Fig. 2). Specifically, the C1-FFL responds with a delay to a step-like antigen stimulus in the ON direction, and rapidly to the step in the opposite direction. The delay (T_{ON}) depends on the costimulation

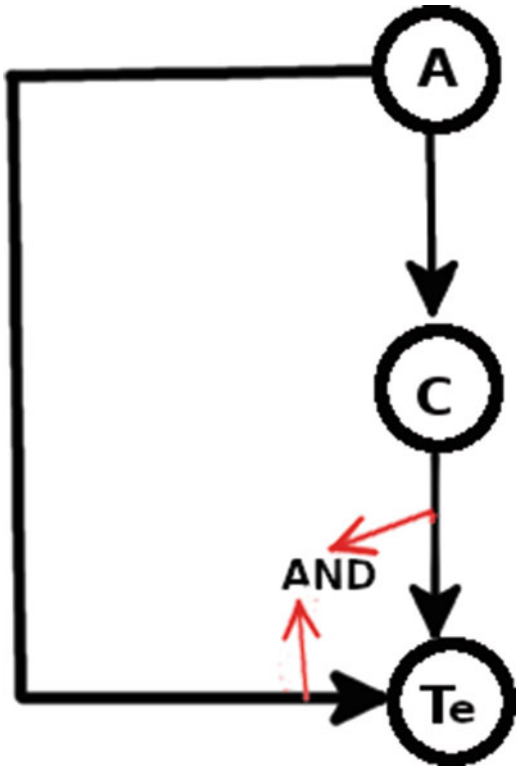


Fig. 1 Structure of the coherent type 1 feedforward loop proposed to govern T effector (T_e) cell activation. The presence of antigen (A) upregulates production of costimulation signals (C). Activation of T_e requires both A and C

threshold for the T cell (denoted K_{CTe}). This delay is the time it takes C to pass the costimulation threshold and is given by

$$T_{ON} = \frac{1}{\alpha_c} \log\left(\frac{C_{st}}{C_{st} - K_{CT_e}}\right), \tag{1}$$

where Cst is the steady-state concentration of the costimulation signal C. From this expression, if $K_{CTe} \rightarrow Cst$, $T_{ON} \rightarrow \infty$. Therefore, the activation threshold K_{CTe} should be smaller than Cst. Also, the larger the value of K_{CTe} , the longer it takes for the T cell to become activated. It is worth noting that, in the light of the assumed necessity of two signals in this work, absence of costimulation will generate no T cell response (i.e. T_e will remain zero if C is zero).

To further illustrate this sign-sensitive delay, we compared the dynamics obtained with a short-lived versus a longer-lived antigen signal (Fig. 3). Notably, long durations of the antigen signal elicit a T cell response after a short delay, whereas short durations of the antigen signal fail to elicit a response. This failure of T cell activation occurs because the short duration of the antigen signal does not provide enough time for the costimulation signal to accumulate past the costimulation threshold. Therefore, noise in the form of short pulses of antigen signal is filtered

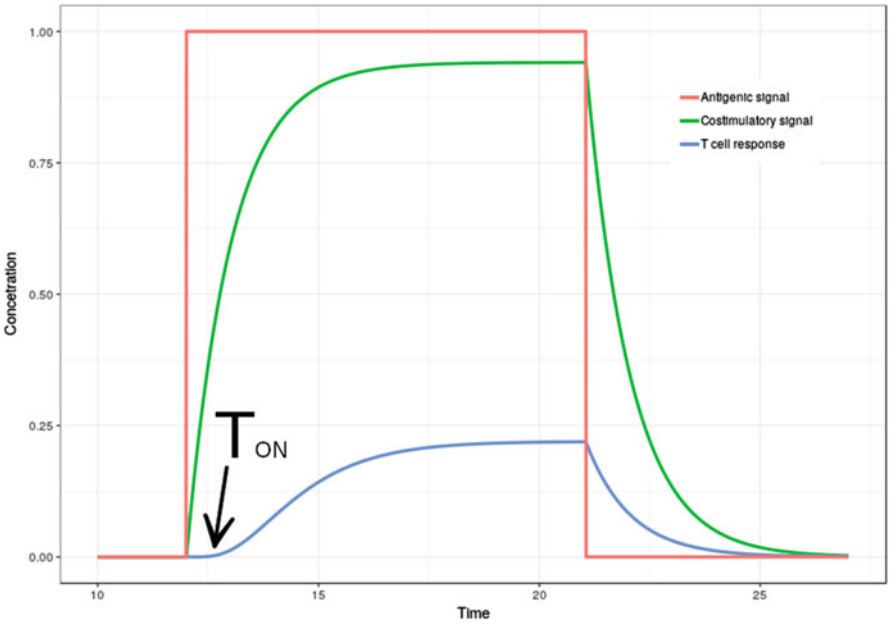


Fig. 2 Dynamics of C1-FFL with an AND gate input function as a sign-sensitive delay element. Activation of T effector cells delays for a duration, T_{ON}. Parameters used were: $\beta_A = \beta_c = \beta_e = \alpha_c = \alpha_e = 1$, $n = 4$, $\lambda = 0.03$, $\mu = 1$, $K_{AC} = 0.5$, $K_{ATe} = 1$, $K_{CTe} = 1$

out, whilst a persistent signal results in T cell activation. Of note, even though a step-like stimulus was used in these simulations, sigmoidal forms yield qualitatively similar dynamics. For example, consideration of an exponentially growing antigen signal $A(t) = e^{\lambda t}$, $\lambda > 0$, reproduces the observed delay in T cell activation (Fig. S1).

2.3 The Two-Signal Requirement and Immune Regulation Together Induce a Coupled Coherent-Incoherent Feedforward Loop (CCI-FFL)

Usually, network motifs do not work in isolation but they are embedded within larger networks in a manner that preserves their dynamical functions [60]. The interaction between T_e cells and T regulatory (T_r) cells regulates the effector and regulatory properties of the immune system. An excessive response by T_e cells is normally prevented by T_r cells via mechanisms such as competition for B7 ligands. Before it can suppress T_e cells, however, a T_r cell must itself be activated following a similar pathway as the T_e cell. Proliferation is further enhanced by interleukin 2 (IL-2) secreted by T_e cells [61–63]. IL-2, however, cannot substitute for CD28 in regulatory T cell activation [36]. Therefore, as shown in Fig. 4,

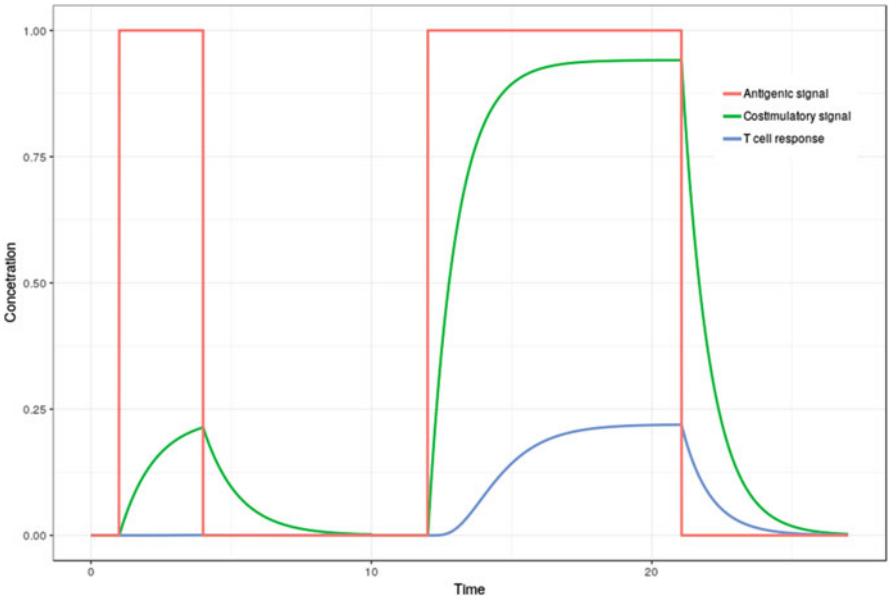


Fig. 3 Dynamics of C1-FFL with an AND gate input function as a noise detector. Short-lived signals are unable to kick-start T_e cell activation but a persistent signal does. Parameter values used in all simulations: $\alpha e = \beta_A = \beta_e = K_{ATe} = K_{CTe} = 1$, $\alpha c = 0.4$, n = 4, $K_{AC} = 0.5$. For the persistent signal, $\beta c = 0.4$, $\lambda = 0.001$, $\mu = 1.5$, whilst for the short-lived signal $\beta c = 0.2$, $\lambda = \mu = 0$. The latter signal was manually turned OFF

the two-signal requirement, combined with immune regulation, induce a coupled coherent-incoherent feedforward loop (CCI-FFL). Notably, this motif has an OR logic gate for the activation and proliferation of T_r cells. This CCI-FFL is a multiple output feedforward loop since the antigen and costimulation signals regulate the activation of more than one type of T cell.

As before, we instantiate the CCI-FFL in a mathematical model (Methods), which we simulate numerically to investigate its dynamics. Although it has been shown that regulatory T cells need not recognise the same antigen as effector T cells [64], we consider the crossregulation model [65] in which regulatory and effector T cells have the same antigen specificity. Simulations show that the CCI-FFL retains the property of sign-sensitive delay (Fig. 5). The delay in activation of each T cell type depends mainly on their respective costimulation threshold parameter, K_{CT}; the larger the value of this parameter, the longer the delay. In particular, if $K_{CTr} > K_{CTe}$, then T_e cells will proliferate first, before later being suppressed by the proliferation of T_r cells, that is, after the magnitude of the costimulation signal has exceeded K_{CTr} [31, 66].

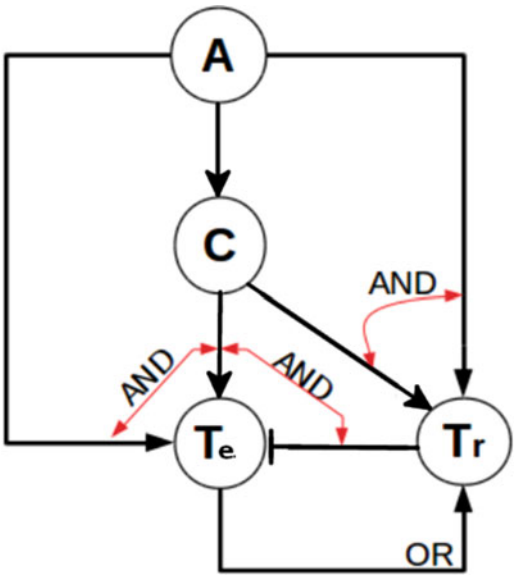


Fig. 4 Network structure of the coupled coherent-incoherent feedforward loop, proposed to govern T_e-T_r interactions. Initially, each T cell type is activated by the two signals with the AND gate. Subsequent activations are influenced by the statuses of the counter cell type

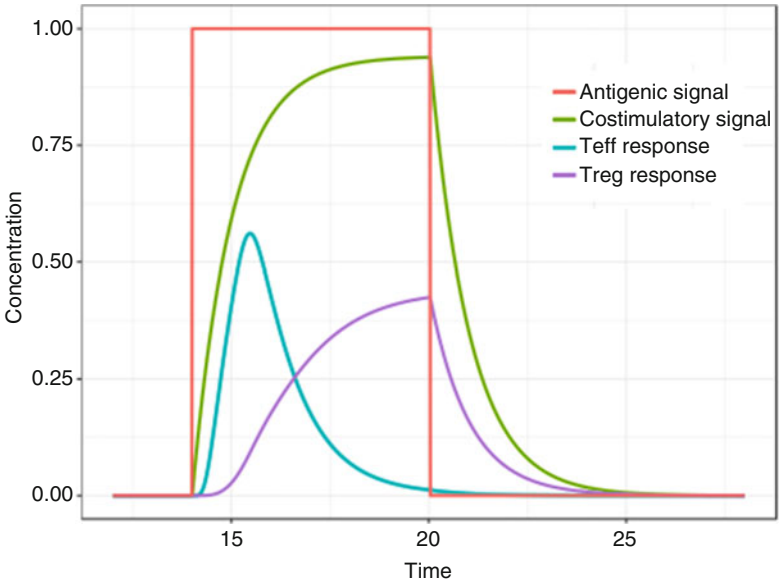


Fig. 5 Dynamics of the coupled coherent-incoherent feedforward loop (CCI-FFL). Although activation of both T cell types delays, in this particular scenario, T_r waits longer. An initial rise in the concentration of T_e is suppressed by the rise of T_r. Parameters used were: $\beta c = 1$, $\beta r = 1$, $\beta e = 1$, $\alpha c = 1$, $\alpha e = 1$, $\alpha r = 1$, $n = 4$, $K_{AC} = 0.01$, $K_{CTr} = 1$, $K_{ATr} = 0.1$, $K_{ATe} = 0.01$, $K_{CTe} = 0.3$, $Ker = 1$, $Kre = 0.1$. $\lambda = 0.001$, $\mu = 1.29$, $\theta = 0.5$

3 Discussion

Understanding the organising principles of T cell activation is key to knowing how the adaptive immune system works, when it will be effective and why it sometimes fails, for example, why it is dampened during cancer. This is also key to improving and devising new T cell-based prophylactics and therapeutics. In this work, we explored the dynamic consequences of the two-signal requirement for T cell activation using a novel approach based on network motifs.

The two-signal theory posits that T cells require at least two signals to become activated. One signal is in the form of an antigen (A) and the other in the form of costimulation (C). Examining the nature of T cell activation via these two signals, we showed that the two-signal theory induces a coherent type 1 feedforward loop. This network motif behaves as a sign-sensitive delay element that responds differently to signals depending on their sign (Fig. 2). A delay, T_{ON}, in the ON step enables the T cell to decide whether or not to launch a response to potential health threats. This delay depends inversely on the costimulation threshold, K_{CT} as shown in Eq. (1). The higher the costimulation threshold, the longer the delay. On the other hand, absence of a delay when the antigen signal is switched off ensures there is no wastage of resources and limits the tissue damage that would otherwise result from a prolonged response.

As a signal-processing element, the sign-sensitive delay rejects any transient activation signals in the form of antigen (A) and responds only to persistent signals, whilst allowing a rapid termination of the immune response (Figs. 2 and 3). This is important because persistent signals are more likely to represent real health threats compared to short-lived signals. T cell responses consume the body's resources (e.g. ATP) and can be extremely inefficient if not required. Our results indicate that the two-signal model for T cell activation provides a molecular mechanism for ensuring that immune responses occur only when they are required and for the correct length of time.

When the two-signal model was combined with the interaction between the T_e and T_r cells, we found that they induce a coupled coherent-incoherent feedforward loop (CCI-FFL). The CCI-FFL forms a multiple output feedforward loop, since antigen (A) and costimulation (C) signals regulate the activation of more than one type of cell. As the activation/inactivation signals vary with time, they pass through different activation/inactivation thresholds. The CCI-FFL can thus generate temporal orders of T cell activation/inactivation depending on the respective thresholds. We predict a similar temporal regulation of T effector and T regulatory cells by antigen and costimulation, which can prevent immune exhaustion during chronic infections like HIV/AIDS.

In Fig. 5, we observe a delay in the initial activation of both T_e and T_r cells since both induce a C1-FFL individually. This is because T_e and T_r have different costimulation thresholds, and hence they wait for different lengths of time after the antigen signal has been turned on to become activated. In remarkable concordance with this theoretical result are experimental data showing that the costimulation

threshold for T_r is much higher than that for T_e [36], implying that T_e would normally become activated first. The temporal separation between the activation of T_e versus T_r should increase as the difference $K_{CTr} - K_{CTe}$ increases. The existence of this temporal separation makes sense biologically because T_e should not be suppressed by T_r before it has had the chance to clear an infection. Therefore, our results indicate that the difference between T_{ON} (T_e) and T_{ON} (T_r) defines a critical time window within which T_e can function unopposed by T_r. Our analysis thus illuminates an important functional consequence of the previously observed [36] difference in costimulation threshold between T_e and T_r cells.

In conclusion, we applied network science and mathematical modelling to investigate the dynamical behaviour of T cell regulation. We focused on the classical version of the two-signal theory of T cell activation, leaving the more complex consideration of coinhibition for the future. For now, coinhibitory receptors can be encapsulated in the existing model as costimulation threshold setters. We started by analysing a novel structural representation of the two-signal requirement, in terms of C1-FFL. We then analysed the effect of coupling the two-signal requirement to immune regulation, which induced a coupled coherent-incoherent feedforward loop. Mathematical analysis of these regulatory structures provided new insight about how the two-signal requirement ensures that T cell activation is both targeted and efficient.

4 Methods

4.1 Mathematical Model of the C1-FFL with AND Logic Gate

We model mathematically the T cell activation dynamics governed by the C1-FFL motif depicted in Fig. 1. The scope of our simple model spans both molecular (A and C) and cellular (T_e, T_r) levels. As recent studies [67] have shown, at a single cell level, T cell response is mostly all-or-none, with minimal dependency on antigen concentration. In these studies, the strength of internally perceived signal correlates mainly with antigen affinity. As a result, we use a digital (ON-OFF) antigenic signal for C and T_e, implemented as a Heaviside function of the prevailing antigen concentration. The changes with time for concentrations of A, C and T_e are described by the following system of ordinary differential equations:

$$\frac{dA}{dt} = \lambda A - \mu A T_e$$

$$\frac{dC}{dt} = \beta_c \frac{[H(A)]^n}{K_{AC}^n + [H(A)]^n} - \alpha_c C$$

$$\frac{dT_e}{dt} = \beta_{T_e} \left(\frac{[H(A)]^n}{K_{AC}^n + [H(A)]^n} \frac{C^n}{K_{CT_e}^n + C^n} \right) - \alpha_{T_e} T_e, \tag{2}$$

where *H(A)* could be set to any reasonable function of A such as the modified Heaviside function,

$$H(A) = \beta_A \begin{cases} 0, A < m_A \\ 1, A \geqslant m_A \end{cases}. \tag{3}$$

Parameters K_{AC} and K_{ATe} are the activation coefficients of C by A and T_e, respectively, K_{CTe} is the activation coefficient of T_e by C, and βc and βT_e are the maximal production rates of the costimulatory signal and the T cell, respectively. m_A represents the T cell detectable antigen concentration, whilst β_A is the maximum magnitude of antigenic signal generated following TCR engagement with cognate antigen. Before T cell activation and response, the antigen grows uncontrollably at a rate λ. After activation, the antigen is cleared by interactions with the generated immune response at a rate μ, hence the mass action term $A\ T_e$. However, since A affects both C and T_e through a form of the Hill function, its effect on the T_e response is bounded, and so is the effect of varying λ on where the T_e response eventually converges (Fig. S1). We also assume that the production of C and T_e is balanced by the decay rate αc and the cell death rate αT_e, respectively. Consider a case when the antigen signal is ON for a time period $t_1 \leq t < t_2$. If this is the initial instance of an infection, then until $t = t_1$ the costimulation signal C is off, i.e. $C = 0$. At $t = t_1$, A switches ON ($A = 1$), whilst at $t = t_2$, A switches off ($A = 0$ again). Therefore, the dynamics of C are given by

$$C = \frac{\beta_c}{\alpha_c}\left(\frac{1}{K_{AC}^n + 1}\right)\left(e^{\alpha_c t_2} - e^{\alpha_c t_1}\right) e^{\alpha_c t}. \tag{4}$$

The functional form $f(A) = [H(A)]^n / \left(K_{AC}^n + [H(A)]^n\right)$ found in Eq. (2) is called the Hill function and represents the probability of activation of C for a given amount of A. The parameters K_{AC} and n represent the activation coefficient and the Hill coefficient, respectively. The activation coefficient corresponds to the amount of A that generates a 50% chance of activation, whilst the Hill coefficient represents the effective number of molecules of A required to activate C. When $H(A) >> K_{AC}$, f (A) $\rightarrow$ 1 and when $H(A) << K_{AC}$, f (A) $\rightarrow$ 0. The higher the value of n, the more the Hill function behaves like a step function. In this scenario, K_{AC} behaves like a threshold value for the ON and OFF regulatory influence of A. When n is close to zero, however, the Hill function can still be approximated by the logic function $f(A) = \theta(A > K_{AC})$, which is equal to one when the amount of A is greater than K_{AC}, and zero otherwise.

4.2 Mathematical Model of the CCI-FFL

We use the same approach as above to model the T cell activation dynamics governed by the CCI-FFL. The rates of change of T_e and T_r are described by the

following system of ordinary differential equations:

$$\frac{dT_e}{dt} = \beta_{T_e}\left(\frac{[H(A)]^n}{K_{AT_e}^n + [H(A)]^n}\frac{C^n}{K_{CT_e}^n + C^n}\frac{K_{re}^n}{K_{re}^n + T_r^n}\right) - \alpha_{T_e}T_e$$

$$\frac{dT_r}{dt} = \beta_{T_r}\left(\frac{[H(A)]^n}{K_{AT_r}^n + [H(A)]^n}\frac{C^n}{K_{CT_r}^n + C^n} + \theta\frac{T_e^n}{K_{er}^n + T_e^n}\right) - \alpha_{T_r}T_r. \quad (5)$$

Equation (5) takes into account the OR input function for the activation of T_r cells by T_e-produced IL-2; IL-2 is inadequate for maximal activation, hence the parameter $0 < \theta < 1$. Parameters β represent the maximal production rate of each entity, K_{ij} is the activation/suppression coefficient of entity j by activating/suppressing factor i and α are the respective decay rates.

A.1 Electronic Supplementary Material

Fig. S1 Dynamics of C1-FFL with an AND gate input function and an exponentially increasing antigen signal. Activation of T effector cells waits until the costimulation signal passes the activation threshold. As expected, based on Eqs. (1) and (5), the waiting time decreases as the growth rate of the antigen signal increases due to a corresponding increase in the costimulation signal. In addition, the point at which the T_e response converges increases with the growth rate up to a maximum. Parameter values used were: $\beta c = \beta T_e = \alpha c = \alpha e = K_{AC} = K_{ATe} = K_{CTe} = \mu = 1$, $n = 4$. λ was set to 0.1 (**a**), 0.5 (**b**) and 0.9 (**c**) (PNG 55 kb)

References

1. Davidson, A., Diamond, B.: Autoimmune diseases. N. Engl. J. Med. **345**, 340–350 (2001)
2. Lee, K.A.: Linking immune defenses and life history at the levels of the individual and the species. Integr. Comp. Biol. **46**, 1000–1015 (2006)
3. Rauw, W.M.: Immune response from a resource allocation perspective. Front. Genet. **3**, 267 (2012)
4. Yanagi, Y., Yoshikai, Y., Leggett, K., Clark, S.P., Aleksander, I., Mak, T.W.: A human T cell-specific CDNA clone encodes a protein having extensive homology to immunoglobulin chains. Nature. **308**, 145–149 (1984)
5. Schwartz, R.H.: Costimulation of T lymphocytes: the role of Cd28, CTLA-4, and B7/BB1 in interleukin-2 production and immunotherapy. Cell. **71**, 1065–1068 (1992)
6. Bretscher, P., Cohn, M.: A theory of self-nonself discrimination. Science. **169**, 1042–1049 (1970)
7. Lafferty, K.J., Cunningham, A.: A new analysis of allogeneic interactions. Aust. J. Exp. Biol. Med. Sci. **53**, 27–42 (1975)

8. June, C.H., Ledbetter, J.A., Gillespie, M.M., Lindsten, T., Thompson, C.B.: T cell proliferation involving the CD28 pathway is associated with cyclosporine-resistant interleukin 2 gene expression. Mol. Cell. Biol. **7**, 4472–4481 (1987)
9. Linsley, P.S., Clark, E.A., Ledbetter, J.A.: T cell antigen CD28 mediates adhesion with B cells by interacting with activation antigen B7/BB-1. Proc. Natl. Acad. Sci. **87**, 5031–5035 (1990)
10. Acuto, O., Michel, F.: CD28-mediated co-stimulation: a quantitative support for TCR signalling. Nat. Rev. Immunol. **3**, 939–951 (2003)
11. Shahinian, A., Pfeffer, K., Lee, K.P., Kundig, T.M., Kishihara, K., Wakenah, A., et al.: Differential t cell costimulatory requirements in CD28-deficient mice. Science. **261**, 609–613 (1993)
12. Hutloff, A., Dittrich, A.M., Beier, K.C., Eljaschewitsch, B., et al.: ICOS is an inducible T cell co-stimulator structurally and functionally related to CD28. Nature. **397**, 263–266 (1999)
13. So, T., Lee, S.-W., Croft, M.: Immune regulation and control of regulatory T cells by OX40 and 4-1BB. Cytokine Growth Factor Rev. **19**, 253–262 (2008)
14. Chambers, C.A., et al.: CTLA-4-mediated inhibition in regulation of T cell responses: mechanisms and manipulation in tumor immunotherapy. Annu. Rev. Immunol. **19**(1), 565–594 (2001)
15. Keir, M.E., et al.: PD-1 and its ligands in tolerance and immunity. Annu. Rev. Immunol. **26**, 677–704 (2008)
16. Watanabe, N., et al.: BTLA is a lymphocyte inhibitory receptor with similarities to CTLA-4 and PD-1. Nat. Immunol. **4**(7), 670 (2003)
17. Rudd, C.E., Taylor, A., Schneider, H.: CD28 and CTLA-4 coreceptor expression and signal transduction. Immunol. Rev. **229**, 12–26 (2009)
18. Nishimura, H., Okazaki, T., Tanaka, Y., Nakatani, K., Hara, M., Matsumori, A., et al.: Autoimmune dilated cardiomyopathy in PD-1 receptor-deficient mice. Science. **291**, 319–322 (2001)
19. Waterhouse, P., Penninger, J.M., Timms, E., Wakeham, A., Shahinian, A., Lee, K.P., et al.: Lymphoproliferative disorders with early lethality in mice deficient in CTLA-4. Science. **270**, 985–988 (1995)
20. Sakaguchi, S., Yamaguchi, T., Nomura, T., Ono, M.: Regulatory T cells and immune tolerance. Cell. **133**, 775–787 (2008)
21. Tang, Q., Bluestone, J.A.: The Foxp3+ regulatory T cell: a jack of all trades, master of regulation. Nat. Immunol. **9**, 239–244 (2008)
22. Möller, G.: Do suppressor T cells exist? Scand. J. Immunol. **27**, 247–250 (1988)
23. Sakaguchi, S., Wing, K., Miyara, M.: Regulatory T cells – a brief history and perspective. Eur. J. Immunol. **37**, S116–S123 (2007)
24. Lenschow, D.J., Walunas, T.L., Bluestone, J.A.: Cd28/b7 system of T cell costimulation. Annu. Rev. Immunol. **14**, 233–258 (1996)
25. Mueller, D.L., Jenkins, M.K., Schwartz, R.H.: Clonal expansion versus functional clonal inactivation: a costimulatory signalling pathway determines the outcome of T cell antigen receptor occupancy. Annu. Rev. Immunol. **7**, 445–480 (1989)
26. Takahashi, T., Kuniyasu, Y., Toda, M., Sakaguchi, N., Itoh, M., Iwata, M., et al.: Immunologic self-tolerance maintained by CD25+CD4+ naturally anergic and suppressive t cells: induction of autoimmune disease by breaking their anergic/suppressive state. Int. Immunol. **10**, 1969–1980 (1998)
27. Thornton, A.M., Shevach, E.M.: CD4+CD25+ immunoregulatory T cells suppress polyclonal T cell activation in vitro by inhibiting interleukin 2 production. J. Exp. Med. **188**, 287–296 (1998)
28. Salomon, B., Lenschow, D.J., Rhee, L., Ashourian, N., Singh, B., Sharpe, A., et al.: B7/CD28 costimulation is essential for the homeostasis of the Cd4+CD25+ immunoregulatory T cells that control autoimmune diabetes. Immunity. **12**, 431–440 (2000)
29. Sansom, D.M., Walker, L.S.: The role of CD28 and cytotoxic T-lymphocyte antigen-4 (CTLA-4) in regulatory T cell biology. Immunol. Rev. **212**, 131–148 (2006)

30. Tang, Q., Henriksen, K.J., Boden, E.K., Tooley, A.J., Ye, J., Subudhi, S.K., et al.: Cutting edge: CD28 controls peripheral homeostasis of CD4+CD25+ regulatory T cells. J. Immunol. **171**, 3348–3352 (2003)
31. Bour-Jordan, H., Salomon, B.L., Thompson, H.L., Szot, G.L., Bernhard, M.R., Bluestone, J.A.: Costimulation controls diabetes by altering the balance of pathogenic and regulatory T cells. J. Clin. Invest. **114**, 979–987 (2004)
32. Rossini, A.A., Greiner, D.L., Mordes, J.P.: Induction of immunologic tolerance for transplantation. Physiol. Rev. **79**, 99–141 (1999)
33. Adams, A.B., Ford, M.L., Larsen, C.P.: Costimulation blockade in autoimmunity and transplantation: the CD28 pathway. J. Immunol. **197**, 2045–2050 (2016)
34. Bouguermouh, S., Fortin, G., Baba, N., Rubio, M., Sarfati, M.: CD28 co-stimulation down regulates Th17 development. PLoS One. **4**, e5087 (2009)
35. Vogel, I., Verbinnen, B., Maes, W., Boon, L., Van Gool, S.W., Ceuppens, J.L.: Foxp3+ regulatory T cells are activated in spite of B7-Cd28 and Cd40-CD401 blockade. Eur. J. Immunol. **43**, 1013–1023 (2013)
36. Hombach, A.A., Kofler, D., Hombach, A., Rappl, G., Abken, H.: Effective proliferation of human regulatory T cells requires a strong costimulatory CD28 signal that cannot be substituted by IL-2. J. Immunol. **179**, 7924–7931 (2007)
37. Chen, L., Flies, D.B.: Molecular mechanisms of T cell co-stimulation and co-inhibition. Nat. Rev. Immunol. **13**, 227–242 (2013)
38. Smith-Garvin, J.E., Koretzky, G.A., Jordan, M.S.: T cell activation. Annu. Rev. Immunol. **27**, 591–619 (2009)
39. Manickasingham, S.P., et al.: Qualitative and quantitative effects of CD28/B7-mediated costimulation on naive T cells in vitro. J. Immunol. **161**(8), 3827–3835 (1998)
40. Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D.T., Bar-Joseph, Z., Gerber, G.K., et al.: Transcriptional regulatory networks in saccharomyces cerevisiae. Science. **298**, 799–804 (2002)
41. Mangan, S., Zaslaver, A., Alon, U.: The coherent feedforward loop serves as a sign- sensitive delay element in transcription networks. J. Mol. Biol. **334**, 197–204 (2003)
42. Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., Alon, U.: Network motifs: simple building blocks of complex networks. Science. **298**, 824–827 (2002)
43. Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., et al.: Essential cell biology. Garland Science, New York (2013)
44. White, J.: Neuronal connectivity in caenorhabditis elegans. Trends Neurosci. **8**, 277–283 (1985)
45. Althaus, C.L., De Boer, R.J.: Dynamics of immune escape during HIV/SIV infection. PLoS Comp. Biol. **7**, e1000103 (2008)
46. Asquith, B., McLean, A.R.: In vivo CD8+ T cell control of immunodeficiency virus infection in humans and macaques. Proc. Natl. Acad. Sci. U. S. A. **104**(15), 6365–6370 (2007)
47. Ganusov, V.V., De Boer, R.J.: Estimating in vivo death rates of targets due to CD8 T cell-mediated killing. J. Virol. **82**(23), 11749–11757 (2008)
48. Perelson, A.S.: Modelling viral and immune system dynamics. Nat. Rev. Immunol. **2**(1), 28 (2002)
49. Saez-Rodriguez, J., et al.: A logical model provides insights into T cell receptor signaling. PLoS Comp. Biol. **3**(8), e163 (2007)
50. Beyer, T., et al.: Integrating signals from the T cell receptor and the interleukin-2 receptor. PLoS Comp. Biol. **8**, e1002121 (2011)
51. Figueroa-Morales, N., Leon, K., Mulet, R.: Stochastic approximation to the T cell mediated specific response of the immune system. J. Theor. Biol. **295**, 37–46 (2012)
52. García-Martínez, K., Leon, K.: Modeling the role of IL-2 in the interplay between CD4+ helper and regulatory T cells: assessing general dynamical properties. J. Theor. Biol. **262**(4), 720–732 (2010)
53. Groß, F., Metzner, G., Behn, U.: Mathematical modeling of allergy and specific immunotherapy: Th1–Th2–Treg interactions. J. Theor. Biol. **1**, 70–78 (2011)

54. Khailaie, S., et al.: A mathematical model of immune activation with a unified self-nonself concept. Front. Immunol. **4**, 474 (2013)
55. Saeki, K., Iwasa, Y.: Optimal number of regulatory T cells. J. Theor. Biol. **263**(2), 210–218 (2010)
56. Sontag, E.D.: A dynamic model of immune responses to antigen presentation predicts different regions of tumor or pathogen elimination. Cell Syst. **4**, 231–241 (2017)
57. Nakagawa, T.Y., Brissette, W.H., Lira, P.D., Griffiths, R.J., Petrushova, N., Stock, J., et al.: Impaired invariant chain degradation and antigen presentation and diminished collagen-induced arthritis in cathepsin S null mice. Immunity. **10**, 207–217 (1999)
58. Abbas, A.K., Janeway, C.A.: Immunology: improving on nature in the twenty-first century. Cell. **100**, 129–138 (2000)
59. Nijkamp, F.P., Parnham, M.J.: Principles of immunopharmacology. Springer Science & Business Media, New York (2006)
60. Alon, U.: Biological networks: the tinkerer as an engineer. Science. **301**, 1866–1867 (2003)
61. Guo, F., Iclozan, C., Suh, W.-K., Anasetti, C., Yu, X.-Z.: CD28 controls differentiation of regulatory T cells from naive CD4 T cells. J. Immunol. **181**, 2285–2291 (2008)
62. Refaeli, Y., Van Parijs, L., London, C.A., Tschopp, J., Abbas, A.K.: Biochemical mechanisms of IL-2–regulated FAS-mediated T cell apoptosis. Immunity. **8**, 615–623 (1998)
63. Savir, Y., Waysbort, N., Antebi, Y.E., Tlusty, T., Friedman, N.: Balancing speed and accuracy of polyclonal T cell activation: a role for extracellular feedback. BMC Syst. Biol. **6**, 111 (2012)
64. Alpan, O., Bachelder, E., Isil, E., Arnheiter, H., Matzinger, P.: 'Educated' dendritic cells act as messengers from memory to naive T helper cells. Nat. Immunol. **5**, 615–622 (2004)
65. Leon, K., Perez, R., Lage, A., Carneiro, J.: Modelling T cell-mediated suppression dependent on interactions in multicellular conjugates. J. Theor. Biol. **207**, 231–254 (2000)
66. Bour-Jordan, H., Bluestone, J.A.: Regulating the regulators: costimulatory signals control the homeostasis and function of regulatory T cells. Immunol. Rev. **229**, 41–66 (2009)
67. Au-Yeung, B.B., et al.: A sharp T cell antigen receptor signaling threshold for T cell proliferation. Proc. Natl. Acad. Sci. U. S. A. **111**(35), 3679–3688 (2014)

Modeling Ebola Transmission Dynamics with Media Effects on Disease and Isolation Rates

Erick Oduniyi, Brad Gibbons, Myunghyun Oh, and Folashade B. Agusto

1 Introduction

The Ebola Virus Disease (EVD) is a reoccurring epidemic that affects both human and non-human species. EVD has appeared infrequently since the late 1970s where the first Ebola outbreak emerged in 1976 near remote villages in Central and East-Central Africa; now considered the Democratic Republic of Congo and South Sudan, respectively [39]. There are currently four identified species of Ebola that cause EVD in humans: Zaire Ebola virus; Sudan virus; Taï Forest Ebola virus; and Bundibugyo Ebola virus. Of these four, the Zaire Ebola virus (also known as EBOV) has the highest case fatality rate and is responsible for the most recent 2014–2016 West Africa Ebola epidemic [21]. The EVD is transmitted by direct contact with bodily fluids (e.g., blood, saliva, sweat, vomit, etc.), or semen of an infected or deceased individual [10]. Infected individuals are confirmed through laboratory diagnostics testing, in which EVD is identified by reverse transcriptase-polymerase chain reaction (RT- PCR), immunostaining techniques, or by antigen-capture enzyme-linked immunosorbent assay (ELISA). Today, diagnostic test makes use of RT-PCR together with ELISA to provide successful identification of EVD [27, 31, 38].

After the 2014–2016 West African Ebola epidemic, the World Health Organization (WHO) estimated more than 28,000 confirmed cases and 11,300 confirmed deaths. According to the Centers for Disease Control and Prevention (CDC), this was the largest Ebola outbreak in history and prompted health organizations and communities to reassess disease surveillance systems and public health infrastructure [6, 17, 36, 39]. Not until recently, there was no specific treatment or federally approved drug or vaccine used in the treatment or prevention of the disease, as

E. Oduniyi · B. Gibbons · M. Oh · F. B. Agusto (✉)
Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS, USA

M. I. Teboh-Ewungkem, G. A. Ngwa (eds.), *Infectious Diseases and Our Planet*,
Mathematics of Planet Earth 7, https://doi.org/10.1007/978-3-030-50826-5_10

such, stopping EVD is nontrivial. The standard practice from CDC and WHO is to prescribe control measures (or intervention packages) as a contingency plan for future Ebola epidemics. Such measures include detailed case management and contact tracing, quarantine procedures, safe burials practices, and more recently intense media campaigns and other forms of community engagement.

Public health researchers and mathematical epidemiologist commonly utilize compartmental models to assess various control measures on their effectiveness in reducing the final epidemic size. Compartmental models of EVD take into account the biology of the disease, and, environmental characteristics like population size and temperature. In the end, researchers work to reveal insights on future intervention strategies for Ebola [18, 36]. For example, incorporating quarantine or isolation compartments can be used to understand the impact of strict quarantine [15], and sexually-infectious compartments can be used to account for the viruses persistence within sexual organs [1, 3]. Additionally, epidemiological research also considers how media campaigns work to engage and inform the population about Ebola, potentially influencing the epidemic's outcome [11, 20, 23, 28, 36, 37]. In these models, media campaigns are commonly defined as some decreasing function [20, 24].

In this paper, we develop a system of ordinary differential equations to describe the transmission dynamics of EVD; the model incorporates sexual transmission, and media campaigns as a means of community engagement. Our goal is to investigate the consequences of treatment facilities with the incorporation of an isolation compartment with limited beds, as well as, the influence of mass media effects. Accordingly, we assess the effectiveness of these constraints and control measures as a means to highlight the potentially positive role of media in modern-day disease transmission and we only carry out basic theoretical stability analysis.

2 Model Formulation

The compartmental model we develop in this paper is composed of the following groups, which are subdivided by the diseases status of individuals within the population: susceptible individuals, $S(t)$, exposed individuals, $E(t)$, infectious individuals, $I_1(t)$, sexually-infectious individuals, $I_2(t)$, isolated individuals, $L_1(t)$, individuals that are not yet isolated $L_2(t)$, recovered individuals, $R(t)$, and deceased individuals due to the infection, $D(t)$. The total population size is given by $N(t)$, where

$$N(t) = S(t) + E(t) + I_1(t) + L_1(t) + L_2(t) + I_2(t) + R(t) + D(t).$$

The total population $N(t)$ consists of the total living population, $N_L(t)$, and individuals in the population who are deceased due to the infection, $D(t)$. These deceased individuals are highly infectious, the virus can live in and on the deceased

for up to 7 days [16, 25]. Thus, $N(t) = N_L(t) + D(t)$, where $N_L(t) = S(t) + E(t) + I_1(t) + L_1(t) + L_2(t) + I_2(t) + R(t)$.

Furthermore, our model takes into account individuals in the population who are infectious and have recovered but still have the virus in their sexual organs. The model also considers individuals that are infectious but isolated in treatment centers, and deceased individuals that are infectious. Formally, the model describing the transmission dynamics of EVD is given by the following system of equations

$$\begin{aligned}
\frac{dS}{dt} &= \pi - \lambda(t)S(t) - \mu S(t) \qquad (1)\\
\frac{dE}{dt} &= \lambda(t)S(t) - (\sigma + \mu)E(t)\\
\frac{dI_1}{dt} &= \sigma E(t) - (\varepsilon + \mu + \delta_1)I_1(t)\\
\frac{dL_1}{dt} &= p\varepsilon I_1(t) + \omega_2 L_1(t) - (\omega_1 + \kappa_1 + \mu + \delta_2)L_1(t)\\
\frac{dL_2}{dt} &= (1-p)\varepsilon I_1(t) + \omega_1 L_2(t) - (\omega_2 + \kappa_2 + \mu + \delta_1)L_2(t)\\
\frac{dI_2}{dt} &= \kappa_1 L_1(t) + \kappa_2 L_2(t) - (\gamma + \mu)I_2(t)\\
\frac{dR}{dt} &= \gamma I_2(t) - \mu R(t),\\
\frac{dD}{dt} &= \delta_1[I_1(t) + L_2(t)] + \delta_2 L_1(t) - \nu D(t),
\end{aligned}$$

where $\lambda(t)$ is the force of infection which accounts for the risk of infection from the living and the deceased population and is given by

$$\lambda(t) = \frac{\beta_1[I_1(t) + L_2(t)] + \beta_2 I_2(t) + \beta_3 L_1(t) + \beta_4 D(t)}{N(t)}.$$

The force of infection, $\lambda(t)$, is weighted by the total population $N(t)$ instead of the total living population $N_L(t)$ since the deceased individuals contribute to the infection in the community. The parameters β_1, β_2, β_3, β_4 in the force of infection $\lambda(t)$ are the effective transmission probability per contact due to infectious, sexually-infectious, isolated, and the deceased individuals. The parameter σ is disease progression rate at which exposed individuals become infectious, ε is the isolation rate due to examination of an infected individual. The parameter p denotes the fraction of infectious individuals that are isolated or quarantined, and $(1 - p)$ is the proportion of individuals who are not isolated. The parameter ω_1 is the rate at which individuals leave isolation and enter the non-isolated population. The parameter ω_2 is the rate at which individuals enter isolation from the non-isolated population; during the outbreak, some individuals were known to have left the care

units [3]. This could be due to efforts at contact tracing or simply by voluntary efforts of the individual due to ill health. The parameters κ_1 and κ_2 are the rates at which individuals become sexually-infectious. The parameter κ_1 is the rate at which isolated individuals move into the sexually-infectious class, we assume that $\kappa_2 < \kappa_1$. The disease-induced death rates are denoted by δ_1, and δ_1. The rate δ_2 is the death rate due to EVD while in isolation, we assume that $\delta_2 < \delta_1$. Individuals in the sexual-infectious class only have the virus in their sexual organs and do not die from the disease. The parameter γ is the recovery rate of the EVD infectious who recover from the infection; finally, the rate at which the infectious deceased are cremated or buried in the population is ν. The model variables and parameters descriptions are given in Table 1 and the model flow diagram is shown in Fig. 1.

Table 1 Description of the variables and parameters for the Ebola model (1)

Variable	Description
$S(t)$	Susceptible individuals
$E(t)$	Exposed individuals
$I_1(t)$	Infectious individuals
$L_1(t)$	Isolated individuals
$L_2(t)$	Individuals that are not isolated
$I_2(t)$	Sexually-infectious individuals
$R(t)$	Recovered/removed individuals
$D(t)$	Deceased individuals
π	Natural birth rate
β_1	Transmission probability due to infectious individuals
β_2	Transmission probability due to sexually-infectious individuals
β_3	Transmission probability due to isolated-infectious individuals
β_4	Transmission probability due to deceased-infectious individuals
σ	Disease progression rate
ε	Isolation rate
κ_1	Rate at which infectious individuals in isolation become sexually-infectious individuals
κ_2	Rate at which unisolated infectious individuals become sexually-infectious individuals
p	Fraction of population isolated
ω_1	Rate of isolated individuals moving into non-isolated population
ω_2	Rate of non-isolated moving into isolated population
h	Fraction of symptomatic individuals who recovered
γ	Recovery rate
μ	Natural death rate
δ_1	Disease-induced death rate
δ_2	Disease-induced death rate in isolation
ν	Burial or cremation rate of infectious dead

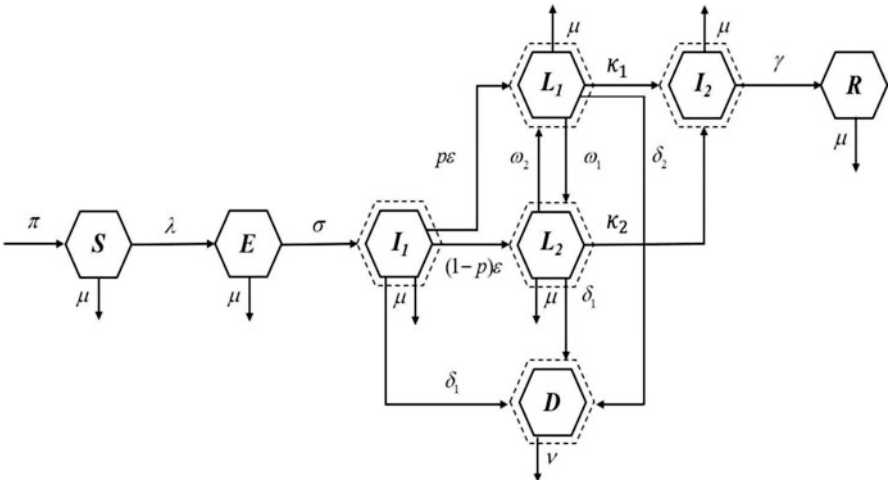


Fig. 1 Flow diagram of the Ebola model (1)

2.1 Model Analysis

The basic qualitative properties such as the positivity and boundedness of solutions, including the analysis of the biologically-feasible invariant region, are stated in Appendix. In the next subsection, the conditions for the existence and stability of the equilibria of the model (1) are stated.

2.1.1 Stability of the Disease-Free Equilibrium (DFE)

The Ebola model (1) has a disease-free equilibrium (DFE). The DFE is obtained by setting the right-hand sides of the equations in the model (1) to zero, which is given by

$$\mathcal{E}_0 = (S, E, I_1, L_1, L_2, I_2, R, D) = \left(\frac{\pi}{\mu}, 0, 0, 0, 0, 0, 0, 0\right).$$

The stability of $\mathcal{E}_0$ can be established using the next generation operator method on system (1). Taking E, I_1, L_1, L_2, I_2, and D as the infected compartments and then using the notation in [32], the Jacobian F and V matrices for new infectious terms and the remaining transfer terms, respectively, are defined as:

$$F = \begin{pmatrix} 0 & \beta_1 & \beta_3 & \beta_1 & \beta_2 & \beta_4 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad V = \begin{pmatrix} g_1 & 0 & 0 & 0 & 0 & 0 \\ -\sigma & g_2 & 0 & 0 & 0 & 0 \\ 0 & -p\,\varepsilon & g_3 & -\omega_2 & 0 & 0 \\ 0 & -(1-p)\,\varepsilon & -\omega_1 & g_4 & 0 & 0 \\ 0 & 0 & -\kappa_1 & -\kappa_2 & g_5 & 0 \\ 0 & -\delta_1 & -\delta_2 & -\delta_1 & 0 & \nu \end{pmatrix}.$$

Therefore, the reproduction number is

$$\begin{aligned} R_0 &= \rho(FV^{-1}) \\ &= \frac{\sigma[\varepsilon\omega_1 p+\varepsilon(1-p)g_3+(g_3g_4-\omega_1\omega_2)]\beta_1}{g_1g_2(g_3g_4-\omega_1\omega_2)} \\ &\quad + \frac{\sigma\varepsilon[p(\kappa_1g_4+\kappa_2\omega_1)+(1-p)(\kappa_2g_3+\kappa_1\omega_2)]\beta_2}{g_1g_2(g_3g_4-\omega_1\omega_2)g_5} + \frac{\sigma\varepsilon[pg_4+\omega_2(1-p)]\beta_3}{g_1g_2(g_3g_4-\omega_1\omega_2)} \\ &\quad + \frac{\sigma[p\varepsilon(\delta_1\omega_1+g_4\delta_2)+(1-p)\varepsilon(g_3\delta_1+\delta_2\omega_2)+\delta_1(g_3g_4-\omega_1\omega_2)]\beta 4}{g_1g_2(g_3g_4-\omega_1\omega_2)\nu}. \end{aligned}$$

where ρ is the spectral radius and $g_1 = \mu+\sigma$, $g_2 = \varepsilon+\mu+\delta_1$, $g_3 = \omega_1+\kappa_1+\mu+\delta_2$, $g_4 = \omega_2+\kappa_2+\mu+\delta_1$, $g_5 = \gamma+\mu$.

The reproduction number, $\mathcal{R}_0$, is the number of secondary infections in completely susceptible population due to infections from one introduced infectious individual with Ebola. The reproduction number measures the average number of secondary infections from four different groups, the infectious (in the infected and non-isolated class), the sexually-infectious, isolated, and deceased individuals. Further, using Theorem 2 in [32], the following result is established.

Lemma 2.1 *The disease-free equilibrium (DFE) of the Ebola model* (1) *is locally asymptotically stable (LAS) if* $\mathcal{R}_0 < 1$ *and unstable if* $\mathcal{R}_0 > 1$.

The implication of the reproduction number is that when $\mathcal{R}_0 < 1$, the disease will be eliminated from the community, while if $\mathcal{R}_0 > 1$, the disease will persist and continue to spread within the population.

2.2 Sensitivity Analysis

In order to determine the robustness of the model in relation to each parameter, a sensitivity analysis was performed. Analysis of parameters can give insight into the uncertainty an input may have and how this will affect the outcome of the model. To determine system sensitivity to its parameters, the normalized forward sensitivity index [7, 12, 14] given in Eq. (2) is used

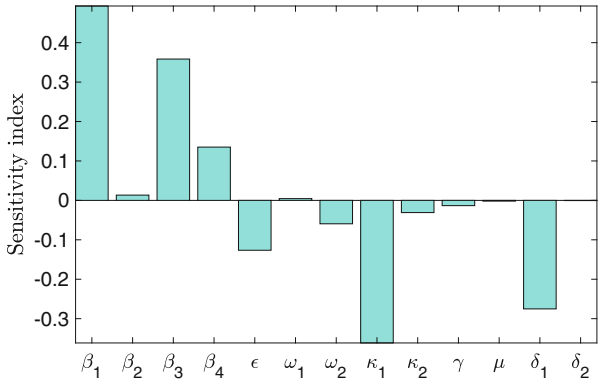


Fig. 2 Sensitivity index of the reproduction number, $\mathcal{R}_0$, of the Ebola model (1)

$$\Upsilon_q^{\mathcal{R}_N} = \frac{\partial \mathcal{R}_0}{\partial q} \times \frac{q}{\mathcal{R}_N}, \tag{2}$$

where q represent the parameter of interest and $\mathcal{R}_0$ is the reproduction number of the Ebola model (1). The sensitivity index is plotted in Fig. 2 using the parameter values in Table 2. The figure shows β_1 having the highest positive sensitivity index, indicating that increase in β_1 will increase $\mathcal{R}_0$ and subsequently, the disease burden. This is followed by β_3 which has a positive impact on $\mathcal{R}_0$. The parameter κ_1 has a negative sensitivity index, which will reduce the disease burden as the value of κ_1 increases. The other parameters impacting $\mathcal{R}_0$ are δ_1, β_4, ε, ω_2, and κ_2. The parameters ω_2, and κ_2, although have a small negative sensitivity index, increase in their values will still lead to a decrease in disease burden. On the other hand, the parameters β_2 and ω_1, also small in value, will increase the disease burden. These parameters can be impacted by the actions of the individuals in the community.

Of these parameters, we can directly influence the transmission probabilities β_1, β_3, β_4 and the transition rate ω_2 *via* educational media campaigns to reduce infection and to encourage infected individuals to seek treatment. Similarly, the parameter ω_1 can be influenced *via* educational media campaigns to reduce infection by encouraging infected individuals to stay in treatment facilities. The disease-induced death rate on the other hand, can be influenced by ensuring adequate treatment measures to prevent death. The transition rates κ_1 and κ_2 cannot, however, be directly influence by the actions of individuals in the community since this is a natural course of the infection.

We explore in the next section the impact of media on the parameters ω_1, ω_2, and ε. First we show in Fig. 3 the impact of varying these parameters on $\mathcal{R}_0$, as expected, as values of ε and ω_2 increases, the reproduction decreases. On the other hand as the values of ω_1 increase, $\mathcal{R}_0$ increases as indicated from the sensitivity index. Furthermore, Fig. 3 unlike Fig. 2 shows the trend in $\mathcal{R}_0$ values as these parameter values varies from 0–1. We expect that the inclusion of media will promote the

Table 2 Parameters values of Ebola model (1) and Ebola media model (5)

Parameter	Definition	Values	References
π	Natural birth rate	0.0001149	[2]
β_1	Transmission probability due to infectious	0.30	[5]
β_2	Transmission probability due to sexually-infectious	0.001	[1, 5]
β_3	Transmission probability due to isolated-infectious	0.080	[26]
β_4	Transmission probability due to deceased-infectious	0.0078	[35]
σ	Disease progression rate	0.0250	[22]
μ	Natural death rate	0.0001149	[4, 22]
ε	Rate of isolation	0.5	[23, 35]
κ_1	Rate of infectious in isolation becoming sexually-infectious	0.05	Variable
κ_2	Rate of infectious in isolation becoming sexually-infectious	0.005	[35]
p	Fraction of population isolated	0.75	[18]
ω_1	Transition rate from isolated into the non-isolated class	0.00024	[22]
ω_2	Transition rate from non-isolated into isolated class	0.0133	[22]
h	Fraction of infectious individuals who recover	0.48	[4]
γ	Recovery rate	0.017	[4, 11]
δ_1	Disease-induced death rate	0.050	[18]
δ_2	Death rate while in isolation	0.001	[18]
ν	Burial or cremation rate of the infectious dead individuals	0.005	Variable
b	Number of beds	Variable	[2]
m	Media effect	$[1.2 \times 10^{-8}, 5]$	[29]
β_{10}, β_{11}	Pre-media and post-media transmission probabilities	0.2, 0.8	[11, 29]
β_{20}, β_{21}	Pre-media and post-media transmission probabilities	0.2, 0.8	[11, 29]
β_{30}, β_{31}	Pre-media and post-media transmission probabilities	0.2, 0.8	[11, 29]
β_{40}, β_{41}	Pre-media and post-media transmission probabilities	0.2, 0.8	[11, 29]
$\mu_{0\omega_1}, \mu_{1\omega_1}$	Pre-media and post-media number of transition rate into non-isolated class	0.15, 0.25	[11, 22]
$\mu_{0\omega_2}, \mu_{1\omega_2}$	Pre-media and post-media number of transition rate into isolated class	0.15, 0.75	[22, 29]
$\mu_{0\varepsilon}, \mu_{1\varepsilon}$	Pre-media and post-media isolation rate	0.27, 0.4	[22, 29]

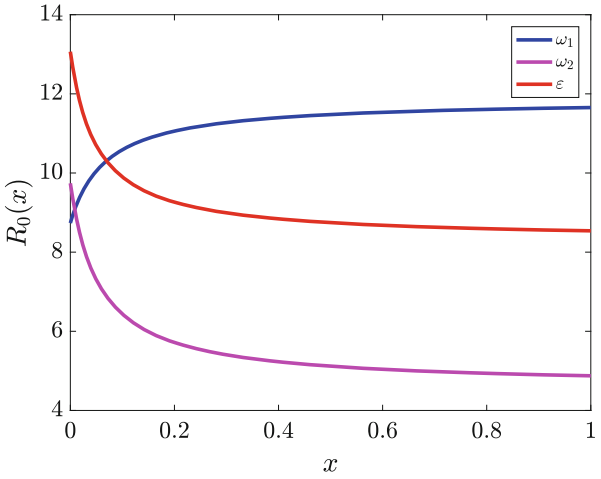


Fig. 3 Simulation of the reproduction number, $\mathcal{R}_0$, of the Ebola model (1) while varying ω_1, ω_2, and ε

actions (either positively or negatively) of these parameters on the reproduction number $\mathcal{R}_0$.

3 Ebola Model with Media Coverage Effects

One of the critical factors that exacerbated the recent West African Ebola crisis was lack of media coverage during the early months of the outbreak. This initial failure of media campaigns and health communication in countries like Guinea, Liberia, and Sierra Leone is believed to be the result of underestimating the difficulty in containing Ebola. Consequently, perceptions of Ebola as a local and global health threat were dismissed until it was too late. Fortunately, by late 2015 various international and national government agencies realized the importance of providing awareness about the disease during the epidemic through timely, accessible, high-quality media coverage [36, 37].

This national and global realization generated fears about infected individuals, where even people outside of the West African region, such as the USA, may have been unjustifiably concerned and eager to learn about the disease [28, 30, 40]. On the other hand, this realization highlights a media campaigns ability to reduce the overall epidemic size through sensationalizing stories. Clearly, media coverage in modern-day disease transmission is an important control measure to incorporate into epidemiological models. Further, due to the worlds increasingly interconnected nature [13], not only are media outlets dispensing news to the public through television and radio broadcast, but individuals have the capability of quickly reporting events as they arise via social media. To account for these media

dynamics, researchers have proposed many different media formulations. Still, these formulations usually take the form of either modifying the contact structures of the disease model as a function of media effects [8, 9, 19, 20, 29, 33, 34], or introducing into the model explicit media compartments that track individuals under the presence of media [23], or by the use of statistical models involving media and disease prevalence, which take advantage of large data sets from Twitter or Google Analytics, or some combination of each of these component [20, 24].

Mass media has tremendous effect on the society; in regard to disease, its effect maybe be positive resulting in a reduction in disease spread since individuals in the community are avoiding contact with a possible infectious individual [8]. While on the other hand, it effect maybe be negative leading to panic in the community [8]. Thus, in our model, we modify the contact structure using a function with media effects. First, we make the simplifying assumption that media is platform independent and can be grouped together as mass media. Following Shen *et al.* [29], we represent mass media and its influences by first augmenting the force of infection using an exponentially decaying function given by

$$\lambda_M = \frac{\beta_1(I_1 + L_2) + \beta_2 I_2 + \beta_3 L_1 + \beta_4 D}{N}, \tag{3}$$

where

$$\beta_1 = \beta_{10} + (\beta_{11} - \beta_{10})e^{-m(C_I+C_D)}, \quad \beta_2 = \beta_{20} + (\beta_{21} - \beta_{20})e^{-m(C_I+C_D)}$$
$$\beta_3 = \beta_{30} + (\beta_{31} - \beta_{30})e^{-m(C_I+C_D)}, \quad \beta_4 = \beta_{40} + (\beta_{41} - \beta_{40})e^{-m(C_I+C_D)},$$

where β_{i0} and β_{i1}, $i = 1, 2, 3, 4$ are the pre-media and post-media infection rates. That is the infection rates before and after the effect of media are apparent in the community. C_I and C_D are cumulative number of infectious and deceased individuals in the community, these are determined from the following equations.

$$\frac{dC_I}{dt} = \sigma E + \varepsilon I_1$$
$$\frac{dC_D}{dt} = \delta_1(I_1 + L_2) + \delta_2 L_1.$$

Note that C_I and C_D are not epidemiological variables. Furthermore, the non-negative parameter m induces the effect of media reported cumulative numbers of infected cases and deaths in the community. If $m = 0$ or relatively small, the infection rates and β_i, $i = 1, 2, 3, 4$ are equal or close to the constant β_{i1}. On the other hand, if $m > 0$, there is increased awareness about the disease in the community and the infection rate could be decreased to β_{i0} ($< \beta_{i1}$) as the number of accumulated infected cases C_I or deaths C_D increases as shown in Fig. 4a, similar behaviors are observed for β_2, β_3, β_4 and are not shown here. Then, the force of infection λ_M incorporates the effects of media in reducing the contact

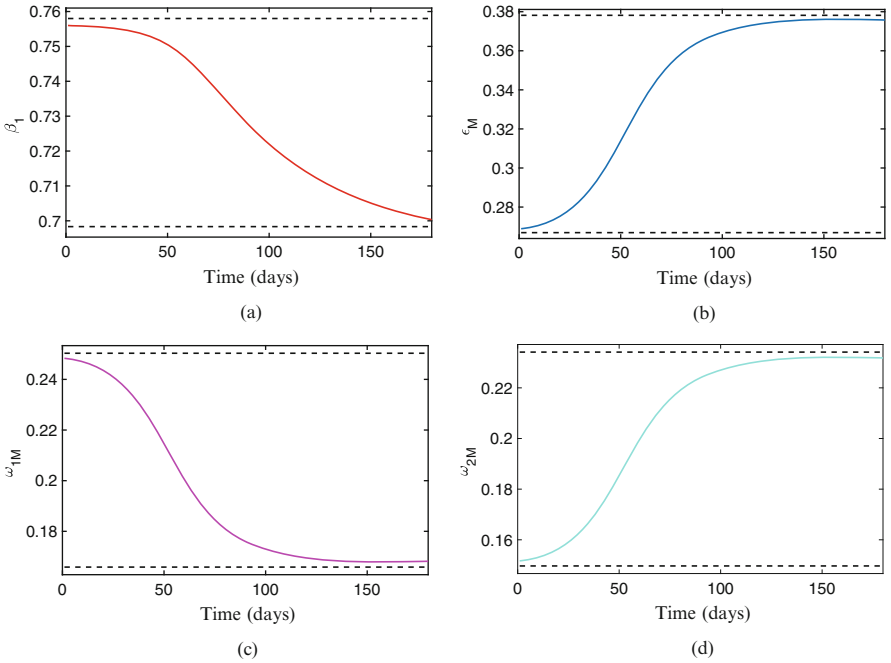


Fig. 4 Simulation of the functions showing media effect. (**a**) Transmission probability, β_M, with media effect; and (**b**) Isolation rate, ε_M, with media effect (**c**) Transition rate from isolated, ω_{1M} with media effect (**d**) Transition rate from non-isolated, ω_{1M}, with media effect

between the healthy susceptible individuals and the infectious in the community as the cumulative number of infected and deceased increases in the community. Figure 4a shows the behavior of β_1 as it decreases with media awareness in the community.

More often than not, Ebola-affected communities have limited resources and failing health infrastructures [36, 37]. As a result, the available health system is unable to cope with outbreak-related emergencies; in particular, there are inadequate number of hospital beds. We endeavor to reflect such a constraint by specifying a limit on how many people the health-care units can isolate. Therefore, we assume there is some minimum and maximum number of individuals the health-care units can sustain due to limited infrastructure regarding hospital beds, denoted by b. Further, we assume that these limitations in the number of available beds affect the rate at which individuals are quarantined (ε) and affect the transitions rates ω_1 and ω_2. From the results of the sensitivity analysis, we see from Fig. 2 that these parameters impact the reproduction number. We further assume that media can impact these rates. Since ε and ω_2 decrease $\mathcal{R}_0$, we propose an increasing media related function that will reduce $\mathcal{R}_0$. Similarly, for ω_1, we propose a decreasing media related function that will reduce $\mathcal{R}_0$. Thus, based on these assumptions, we use the following functions to capture media related isolation function ε_M,

and media related transition functions ω_{1M} and ω_{2M} which are represented by the following functions:

$$\omega_{1M} = \mu_{0\omega_1} + (\mu_{1\omega_1} - \mu_{0\omega_1})\left(\frac{b}{b+N_I}\right)e^{-m(I_1+L_2+I_2+L_1+D)} \tag{4}$$

$$\omega_{2M} = \mu_{1\omega_2} + (\mu_{0\omega_2} - \mu_{1\omega_2})\left(\frac{b}{b+N_I}\right)e^{-m(I_1+L_2+I_2+L_1+D)}$$

$$\varepsilon_M = \mu_{1\varepsilon} + (\mu_{0\varepsilon} - \mu_{1\varepsilon})\left(\frac{b}{b+N_I}\right)e^{-m(I_1+L_2+I_2+L_1+D)},$$

where $N_I = I_1 + I_2 + L_1 + L_2 + D$. The dynamic behavior of functions ε_M and ω_{2M} are depicted in Fig. 4b, d, and they are similar since we want to increase the number of individuals going into the isolation units. The dynamics of ω_{1M} is given in Fig. 4c and it is similar to the dynamics of β_1, since our goal is to reduce the number of individuals leaving the isolation units.

Note that $\omega_{1M} > 0$ for $N_I > 0, b > 0$. Furthermore, for arbitrarily small number of infectious individuals N_I, the media related transition function ω_{1M} converges to $\mu_{1\omega_1}$, that is,

$$\lim_{N_I \to 0} \omega_{1M} = \mu_{1\omega_1} > 0,$$

the maximum transition rate out of the isolated class before media manifests in the community. Also, as the number of infectious individuals N_I grows, the transition function ω_{1M} converges to ω_{1M}, that is,

$$\lim_{N_I \to \infty} \omega_{1M} = \mu_{0\omega_1} > 0,$$

the minimum transition rate out of the isolation class as media manifests in the community.

Consequences for an arbitrarily small number of infectious individuals N_I, the media related isolation function ε_M converges to $\mu_{0\varepsilon}$, that is,

$$\lim_{N_I \to 0} \varepsilon_M = \mu_{0\varepsilon} > 0,$$

the minimum isolation rate before the onset of media coverage. Also, as the infectious individuals N_I get larger, ε_M converges to $\mu_{1\varepsilon}$

$$\lim_{N_I \to \infty} \varepsilon_M = \mu_{1\varepsilon} > 0,$$

the maximum number of individuals that are isolated as a result of media coverage.

Similarly, the media related transition function, ω_{2M}, back into the isolation class, converges to $\mu_{0\omega_2}$, that is,

$$\lim_{N_I \to 0} \omega_{2M} = \mu_{0\omega_2} > 0,$$

the minimum transition rate for small number of infectious individuals N_I before the onset of media coverage. And $\lim_{N_I \to \infty} \omega_{2M} = \mu_{1\omega_2} > 0$, the maximum transition rate for large number of infections individuals N_I as media effect manifest.

Now, incorporating the media and limited infrastructure related function (3) and (4) into the Ebola model (1), we have the following system of differential equations

$$\begin{aligned}
\frac{dS}{dt} &= \pi - \lambda_M S(t) - \mu S(t) \\
\frac{dE}{dt} &= \lambda_M S(t) - (\sigma + \mu) E(t) \\
\frac{dI_1}{dt} &= \sigma E(t) - (\varepsilon_M + \mu + \delta_1) I_1(t) \\
\frac{dL_1}{dt} &= p\varepsilon_M I_1(t) + \omega_{2M} L_1 - (\omega_{1M} + \kappa_1 + \mu + \delta_2) L_1(t) \\
\frac{dL_2}{dt} &= (1-p)\varepsilon_M I_1(t) + \omega_{1M} L_2(t) - (\omega_{2M} + \kappa_2 + \mu + \delta_1) L_2(t) \\
\frac{dI_2}{dt} &= \kappa[L_1(t) + L_2(t)] - (\gamma + \mu) I_2(t) \\
\frac{dR}{dt} &= \gamma I_2(t) - \mu R(t) \\
\frac{dD}{dt} &= \delta_2[I_1(t) + L_2(t)] + \delta_1 L_1(t) - \nu D(t).
\end{aligned} \tag{5}$$

Ebola model (5) with media effect is analyzed in a biologically-feasible region $\Gamma \subset \mathbb{R}^8_+$,

$$\Gamma = \left\{ (S(t), E(t), I_1(t), I_2(t), L_1(t), L_2(t), R(t), D(t)) \in \mathbb{R}^8_+ : \ N(t) \leq \frac{\Pi}{\mu} \right\}.$$

Using the approach in Lemma A.2 in Appendix, we can state the following result:

Lemma 3.1 *The region $\Gamma \subset \mathbb{R}^8_+$ is positively-invariant for the model (5) with non-negative initial conditions in $\mathbb{R}^8_+$.*

Furthermore, following Lemma A.3 in Appendix, the biologically-feasible region, Γ_L of all the living individuals in the population is positively-invariant. This region is defined as

$$\Gamma_L = \left\{ (S(t), E(t), I_1(t), L_1(t), L_2(t), I_2(t), R(t)) \in \mathbb{R}^7_+ : \ N_L(t) \leq \frac{\Pi}{\mu} \right\}.$$

The reproduction number, $\mathcal{R}_{0_M}$, of the Ebola model (5) is given as

$$\begin{aligned}\mathcal{R}_{0_M} = & \frac{\sigma[u_{0_\varepsilon} pV_{33} + u_{0_\varepsilon} u_{1_{\omega_1}}(1-p) + (V_{44}V_{33} - u_{1_{\omega_1}} u_{0_{\omega_2}})]\beta_{11}}{V_{22}g_1(V_{44}V_{33} - u_{1_{\omega_1}} u_{0_{\omega_2}})} \\ & + \frac{\sigma u_{0_\varepsilon}[p(\kappa_2 V_{33} + \kappa_1 u_{0_{\omega_2}}) + (1-p)(V_{44}\kappa_1 + \kappa_2 u_{1_{\omega_1}})]\beta_{21}}{V_{22}g_5 g_1(V_{44}V_{33} - u_{1_{\omega_1}} u_{0_{\omega_2}})} \\ & + \frac{\sigma u_{0_\varepsilon}[u_{0_{\omega_2}} p + V_{44}(1-p)]\beta_{31}}{V_{22}g_1(V_{44}V_{33} - u_{1_{\omega_1}} u_{0_{\omega_2}})} \\ & + \{\sigma[pu_{0_\varepsilon}(V_{33}\delta_1 + u_{0_{\omega_2}}\delta_2) + (1-p)u_{0_\varepsilon}(V_{44}\delta_2 + \delta_1 u_{1_{\omega_1}}) \\ & + \delta_1(V_{33}V_{44} - u_{0_{\omega_2}} u_{1_{\omega_1}})]\beta_{41}\}/V_{22}g_1\nu(V_{44}V_{33} - u_{1_{\omega_1}} u_{0_{\omega_2}}),\end{aligned}$$

where $g_1 = \sigma + \mu,\ g_2 = \mu + \delta_1,\ g_3 = \kappa_1 + \mu + \delta_2,\ g_4,\ g_4 = \kappa_2 + \mu + \delta_1,\ g_5 = \gamma + \mu,\ V_2 = u_{0_\varepsilon} + g_2,\ V_3 = u_{1_{\omega_1}} + g_3,\ V_4 = u_{0\omega_2} + g_4$

The production number, $\mathcal{R}_{0_M}$, is the number of secondary infections in completely susceptible population due to infections from one introduced infectious individual with Ebola. The reproduction number measures in a community with media awareness the average number of secondary infections from four different groups, the infectious (in the infected and non-isolated class), the sexually-infectious, isolated, and deceased individuals. Further, using Theorem 2 in [32], the following result is established.

Lemma 3.2 *The disease-free equilibrium (DFE) of the Ebola model* (5) *is locally asymptotically stable (LAS) if* $\mathcal{R}_{0_M} < 1$ *and unstable if* $\mathcal{R}_{0_M} > 1$.

The implication of the reproduction number is that when $\mathcal{R}_{0_M} < 1$, the disease will be eliminated from the community, while if $\mathcal{R}_{0_M} > 1$, the disease will persist and continue to spread within the population.

4 Numerical Simulations

To illustrate the impact of media on the disease transmission the following values were taken for the initial conditions $S(0) = 1000,\ E(0) = 0,\ I_1(0) = 3,\ L_1(0) = 0,\ L_2(0) = 0,\ I_2(0) = 0,\ R(0) = 0$, and $D(0) = 0$. The parameter values are shown in Table 2. It should be pointed out that the initial conditions here are only of theoretical sense to illustrate the impact of media.

Setting the same initial conditions and parameter values for both the non-media Ebola model (1) and Ebola media model (5) we study the effect of media coverage on disease burden in the community using the infectious, $I_1(t)$, isolated, L_1, non-isolated, L_2, and the sexually-infectious, $I_2(t)$, classes.

As shown in Fig. 5, we see clearly the effect of media on disease burden in the community. In Fig. 5a we see a higher number of infected at the peak infection period in the model without media and infection peak earlier compared to the

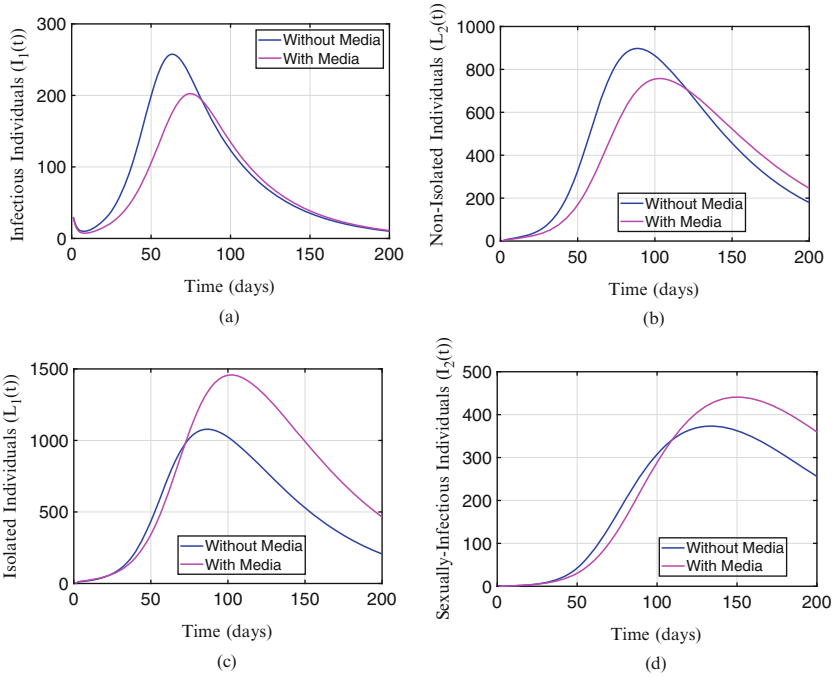


Fig. 5 Simulation of the Ebola model (5) without and with media effect (using media effect $m = 1$). (**a**) Infectious individuals; (**b**) Non-isolated individuals; (**c**) Isolated individuals; (**d**) Sexually-infectious individuals. Parameter values used are as given in Table 2

model with media effect. This dynamics is also reflected in the non-isolated class in Fig. 5b. Figure 5c, d on the other hand shows higher number of individuals in the isolated and sexually-infectious classes in the model with media compared to the non-media class. Similarly, we observe a shift in the peaks of these compartments unlike the peaks in the infectious class. We also observed the same dynamics for the susceptible, exposed and recovered classes are they are not shown here. This figure shows the effect of public awareness in the community; the more people are aware of the dangers inherent about a disease, the more careful they will become. Thus, public awareness and media coverage will subsequently lead to a lower number of infected in the community. This figure (Fig. 5) clearly shows the effect of media in the community, it is, however, difficult to tease out the effect on every single part of the system. In the next figure (Fig. 6), we will explore the effect of media on the isolated and non-isolated classes since we have modeled some of the transition rates into and out of these classes are functions of media.

In Fig. 6, we compare the difference between the isolated and non-isolated class with the Ebola model (1) and the Ebola media model (5). We see (in Fig. 6a) that the difference in the non-media model is smaller (about 180.8 individuals) than the difference (about 700.5 individuals) in the media model (see Fig. 6b). This difference shows the impact of media in encouraging non-isolated individuals in

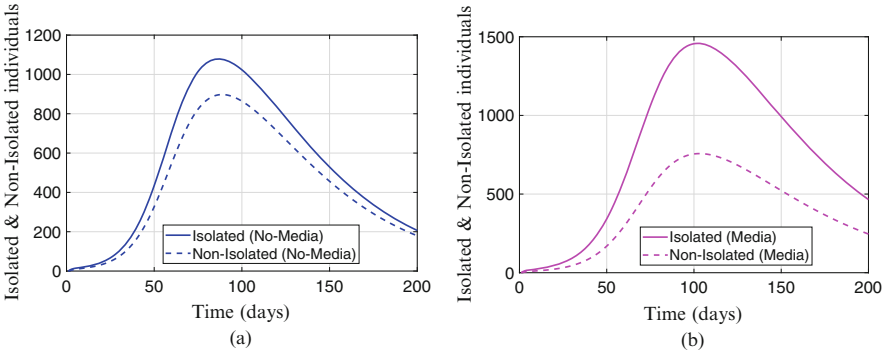


Fig. 6 Simulation of the Ebola model (5) using media effect $m = 1$. (**a**) Isolated and Non-Isolate individuals without media effect; (**b**) Isolated and Non-Isolate individuals with media effect. Parameter values used are as given in Table 2

the community to seek treatment, while encouraging those in isolation units to stay and complete their treatments in the treatment facilities. In order words, fewer individuals will seek treatment when there is weak media effect compared to when there are strong media effects where more people seek treatment and remain in the treatment units.

Public health is both a concern of governments and individuals. For this reason, mass media will necessarily attempt to disseminate disease outbreak-related information during an epidemic. For instance, although media coverage during the 2014–2016 EVD crisis was nonuniform across many national and regional communities, eventually the presence of media manifested [28, 36, 37]. Hence, we turn from investigating the impact of introducing media coverage into a community, to evaluating the impact of varying the level of media coverage on disease burden. Figure 7 shows the comparison between two values media effect parameter m, that is, $m = 1$ and $m = 1.2 \times 10^{-5}$. These values represent strong and weak media effect. We observe in Fig. 7a that the number of infectious individuals peak early when media effect is weak compared to when media effect is strong where the disease peak later; furthermore, the number infectious is much smaller with strong media effect. We observe the same dynamics in Fig. 7b among non-isolated individuals, isolated, and sexually-infectious individuals. The number of individuals in isolated and sexually-isolated groups are more when media effects is stronger $m = 1$ compared to when media effect is weak $m = 1.2 \times 10^{-5}$.

Next, we also compare the difference between the isolated and non-isolated classes using the Ebola media model (5) with weak and strong media effect, that is, $m = 1$ and $m = 1.2 \times 10^{-5}$ (see Fig. 8). We see, in Fig. 8a, that the difference in the infectious class under weak media effect is smaller (448.8 individuals) than the difference (700.5 individuals) in the infectious under strong media (see Fig. 7b). This difference as earlier observed shows the impact of media in encouraging non-isolated individuals in the community to seek treatment, while encouraging those

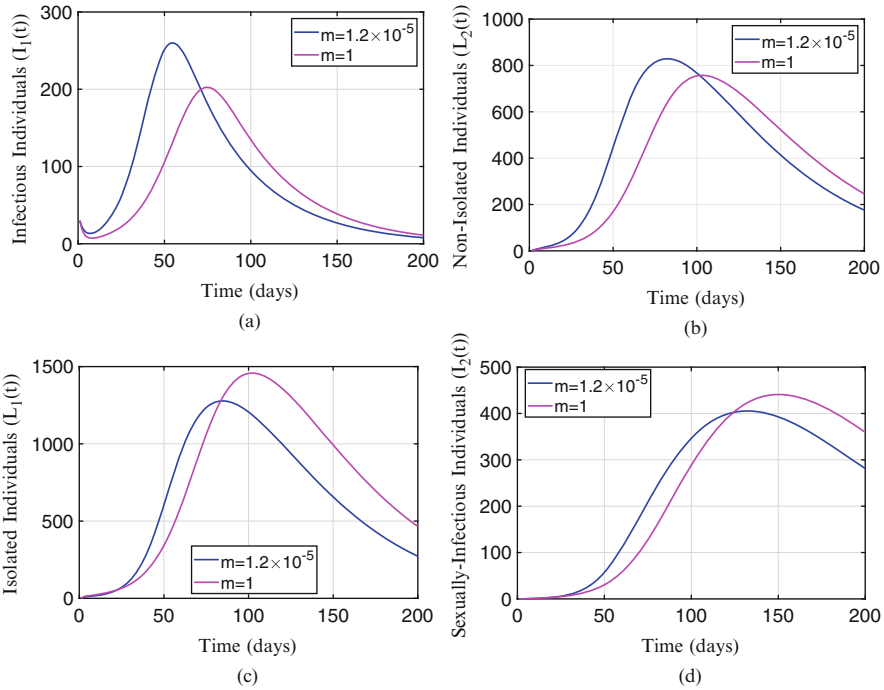


Fig. 7 Simulation of the Ebola model (5) using weak and strong media effect $m = 1.2 \times 10^{-5}$ and $m = 1$. (**a**) Infectious individuals; (**b**) Non-isolated individuals; (**c**) Isolated individuals; (**d**) Sexually-infectious individuals. Parameter values used are as given in Table 2

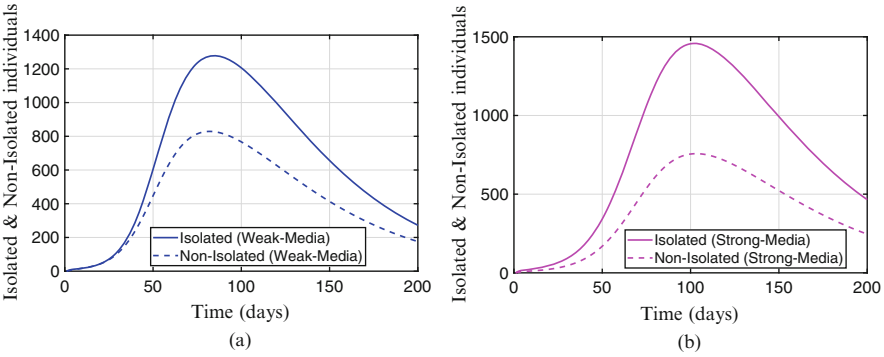


Fig. 8 Simulation of the Ebola model (5) using weak ($m = 1.2 \times 10^{-5}$) and strong media effect ($m = 1$). (**a**) Isolated and Non-Isolated individuals with weak media effects; (**b**) Isolated and Non-Isolated individuals with strong media effects. Parameter values used are as given in Table 2

in isolation units to stay and complete their treatments in the treatment facilities. In order words, fewer individuals will seek treatment when there is weak media effect compared to strong media effects where we see more people seeking treatment.

5 Conclusion and Discussion

In this paper, we formulated a deterministic model that uniquely incorporate isolation of the infectious in the community, and non-isolated individuals. We assume these individuals can self-discharge themselves from the isolation units, as was observed during the 2014 West African Ebola outbreak. Following the life-history of the disease, we include the sexually-infectious individuals who have recovered from the infection and cleared the virus from their blood; however, the virus remains in their sexual and other vital organs of their body and is infectious to susceptible individuals following any sexual contact. Results from our basic theoretical analysis show that the Ebola model is locally asymptotically stable when the reproduction number is less that one and unstable otherwise.

We carried out a sensitivity analysis to determine the impact of each parameter on the reproduction. We found that the parameters β_1, β_3, β_4, ε, ω_2, κ, and δ_1 have the highest impact on the reproduction number. These parameters have positive and negative sensitivity index that will increase or reduce $\mathcal{R}_0$ depending on the sign of the sensitivity index. Of these parameters, ω_1, ω_2, and ε can be directly influenced by educational media campaigns. Thus, we uniquely incorporate into the Ebola model the effect of media on these parameters.

Simulations of this model broadly showcase how the presence of media campaigns in a community reduces the disease burden. In particular, our model shows that as the quantity of media campaigning increases, the peak of the epidemic is delayed, and produces fewer number of infectious individuals, and an increase in the number of isolated individuals; subsequently leading to fewer EVD-related deaths (not shown in our simulation). On these accounts, our model lends itself to interpreting media as a competent control measure on EVD. More specifically, media campaigning works to reduce the rate at which people become infected and increase the rate at which people seek help and treatment, and therefore the rate they recover from the infection. Furthermore, media campaign also encourages individuals in the treatment units or facilities to stay and complete their treatment.

Media coverage, whether it is through social media or radio and television broadcast, informs the populace about disease symptoms and treatments, which is vital to understanding the transmission of the disease. Indeed, the 2014–2016 Ebola epidemic showcased the importance of prompt and quality health care communication at the global, national, and local levels. As was observed during the 2014 outbreaks some communities initially underestimated the extent and seriousness of the disease, but as the CDC and WHO began to highlight the case fatalities, the seriousness became apparent and the importance of spreading information about the epidemic increased.

Thus, we have assumed that mass media encourages healthier behavior and thereby reducing the number of confirmed cases. However, most Ebola epidemic models do not account for the effect of media in their formulation. Hence, to adequately represent the impact of media coverage it is vital to incorporate into these models data on Ebola-related tweets, Ebola segments on the nightly news, and the amount of airtime given Ebola during the 2014–2016 epidemic. Additional model insights may even be driven by inquiries related to how different West African countries utilize media platforms and their respective preferences for disseminating epidemic related information.

In our model we have assumed that media coverage has a positive effect on reducing disease transmission and encouraging positive behavior towards seeking treatment and remaining in treatment facilities. However, that is not always the case, media could have a counter negative effect; for instance, media could promote misinform leading to mistrust around government health engagement. Hence, it will be worthwhile to incorporate the negative consequences of mass media into these epidemic models.

Finally, this current study has relied on existing literature to parameterize the model. In a future work we will explore ways to fit this model to data so as to determine vital model parameters, particularly the media related parameters.

Appendix: Basic Qualitative Properties

Positivity and Boundedness of Solutions

For the Ebola model (1) to be epidemiologically meaningful, it is important to prove that all its state variables are non-negative for all time. In other words, solutions of the system (1) with non-negative initial data will remain non-negative for all time $t > 0$.

Lemma A.1 *Let the initial data $F(0) \geq 0$, where $F(t) = (S(t), E(t), I_1(t), L_1(t), L_2(t), I_2(t), R(t), D(t))$. Then the solutions $F(t)$ of the Ebola model (1) are non-negative for all $t > 0$. Furthermore*

$$\limsup_{t\to\infty} N(t) \leq \frac{\pi}{\mu},$$

and

$$N(t) = S(t) + E(t) + I_1(t) + L_1(t) + L_2(t) + I_2 + R(t) + D(t).$$

Proof Let $t_1 = \sup\{t > 0 : F(t) > 0 \in [0, t]\}$. Thus, $t_1 > 0$. It follows from the first equation of the system (1) that

$$\frac{dS}{dt} = \pi - \lambda(t)S - \mu S$$

which can be re-written as

$$\frac{d}{dt}\left\{S(t)\ exp\left(\int_0^{t_1}\lambda(\zeta)d\zeta+\mu t\right)\right\}\ =\ \pi\ exp\left(\int_0^{t_1}\lambda(\zeta)d\zeta+\mu t\right).$$

Hence,

$$S(t_1)\ exp\left(\int_0^{t_1}\lambda(\zeta)d\zeta+\mu t_1\right)-S(0)=\int_0^{t_1}\pi\ exp\left(\int_0^{p}\lambda(\zeta)d\zeta+\mu p\right)dp$$

so that,

$$\begin{aligned}S(t_1) &= S(0)\ exp\left[-\left(\int_0^{t_1}\lambda(\zeta)d\zeta+\mu t_1\right)\right]\\ &\quad+exp\left[-\left(\int_0^{t_1}\lambda(\zeta)d\zeta+\mu t_1\right)\right]\int_0^{t_1}\pi\ exp\left[\left(\int_0^{p}\lambda(\zeta)d\zeta+\mu p\right)\right]dp\\ &\quad>0.\end{aligned}$$

Similarly, it can be shown that $F > 0$ for all $t > 0$.

For the second part of the proof, note that $0 < S(0) \leq N(t), 0 \leq E(0) \leq N(t), 0 \leq I_1(0) \leq N(t), 0 \leq I_2(0) \leq N(t), 0 \leq L_1(0) \leq N(t), 0 \leq L_2(t) \leq N(t), 0 \leq R(t) \leq N(t), 0 \leq D(t) \leq N(t)$.

Adding components of the model (1) gives

$$\begin{aligned}\frac{dN(t)}{dt} &= \pi-\mu N-(\nu-\mu)D \qquad (A.1)\\ &\leq \pi-\mu N.\end{aligned}$$

Hence,

$$\limsup_{t\to\infty}N(t)\leq\frac{\pi}{\mu}$$

as required.

Invariant Regions

The Ebola model (1) will be analyzed in a biologically-feasible region as follows. Consider the feasible region

$$\Omega\subset\mathbb{R}_+^8,$$

with,

$$\Omega = \left\{(S(t), E(t), I_1(t), L_1(t), L_2(t), I_2(t), R(t), D(t)) \in \mathbb{R}_+^8 : \ N(t) \leq \frac{\Pi}{\mu}\right\}.$$

Lemma A.2 *The region* $\Omega \subset \mathbb{R}_+^8$ *is positively-invariant for the model* (1) *with non-negative initial conditions in* $\mathbb{R}_+^8$.

Proof It follows from summing equations of model (1) that

$$\begin{aligned}\frac{dN(t)}{dt} &= \pi - \mu N - (\nu - \mu)D \\ &\leq \pi - \mu N.\end{aligned}$$

Hence, $\frac{dN(t)}{dt} \leq 0$, if $N(0) \geq \frac{\pi}{\mu}$. Thus, $N(t) \leq N(0)e^{-\mu t} + \frac{\pi}{\mu}(1 - e^{-\mu t})$. In particular, $N(t) \leq \frac{\pi}{\mu}$.

Thus, the region Ω is positively-invariant. Furthermore, if $N(0) > \frac{\pi}{\mu}$, then either the solutions enter Ω in finite time, or $N(t)$ approaches $\frac{\pi}{\mu}$ asymptotically. Hence, the region Ω attracts all solutions in $\mathbb{R}_+^8$.

Following [22], we can also show that the biologically-feasible region, Ω_L of all the living individuals in the population, is positively-invariant. This region is defined as

$$\Omega_L = \left\{(S(t), E(t), I_1(t), L_1(t), L_2(t), I_2(t), R(t)) \in \mathbb{R}_+^7 : \ N_L(t) \leq \frac{\Pi}{\mu}\right\}.$$

This leads to the following lemma:

Lemma A.3 *The region* $\Omega_L \subset \mathbb{R}_+^7$ *is positively-invariant for the model* (1) *with non-negative initial conditions in* $\mathbb{R}_+^7$.

Proof Summing the equations for the living individuals in model (1) we have

$$\begin{aligned}\frac{dN_L(t)}{dt} &= \pi - \mu N_L - \delta(I_1 + L_1 + L_2) \\ &\leq \pi - \mu N_L,\end{aligned}$$

where $\delta = \min\{\delta_1, \delta_2\}$.

Hence, $\frac{dN_L(t)}{dt} \leq 0$, if $N_L(0) \geq \frac{\pi}{\mu}$. Thus, $N_L(t) \leq N_L(0)e^{-\mu t} + \frac{\pi}{\mu}(1 - e^{-\mu t})$. In particular, $N_L(t) \leq \frac{\pi}{\mu}$.

If $N_L(0) > \frac{\pi}{\mu}$, then either the solutions enter Ω_L in finite time, or $N_L(t)$ approaches $\frac{\pi}{\mu}$ asymptotically. Hence, the region Ω_L is positively-invariant and attracts all solutions in $\mathbb{R}^7_+$.

References

1. Abbate, J., Murall, C., Richner, H., Althaus, C.: Potential impact of sexual transmission on Ebola virus epidemiology: Sierra Leone as a case study. PLoS Neglect. Trop. Dis. **10**(5), e0004676 (2016)
2. Abdelrazec, A., Bélair, J., Shan, C., Zhu, H.: Modeling the spread and control of dengue with limited public health resources. Math. Biosci. **271**, 136–145 (2016)
3. Agusto, F., Bewick, S., Fagan, W.: Mathematical model for zika virus dynamics with sexual transmission route. Ecol. Complex. **29**, 61–81 (2017)
4. Agusto, F., Teboh-Ewungkem, M., Gumel, A.: Mathematical assessment of the effect of traditional beliefs and customs on the transmission dynamics of the 2014 Ebola outbreaks. BMC Med. **13**(1), 96 (2015)
5. Althaus, C.: Estimating the reproduction number of Ebola virus (EBOV) during the 2014 outbreak in West Africa. PLoS Curr. **6** (2014)
6. Centers for Disease Control and Prevention (CDC). 2014–2016 Ebola outbreak in West Africa (2017). https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/index.html. Accessed 20 Jan 2019
7. Chitnis, N., Hyman, J., Manore, C.: Modelling vertical transmission in vector-borne diseases with applications to rift valley fever. J. Biol. Dyn. **7**(1), 11–40 (2013)
8. Cui, J., Sun, Y., Zhu, H.: The impact of media on the control of infectious diseases. J. Dyn. Diff. Equ. **20**(1), 31–53 (2008)
9. Cui, J., Tao, X., Zhu, H.: An SIS infection model incorporating media coverage. Rocky Mountain J. Math. **38**(5), 1323–1334 (2008)
10. Frieden, T., Damon, I.: Ebola in west Africa – CDC's role in epidemic detection, control, and prevention. Emerg. Infect. Dis. **21**(11), 1897 (2015)
11. Gani, S., Halawar, S.: Optimal control for the spread of infectious disease: the role of awareness programs by media and antiviral treatment. Optimal Control Appl. Methods **39**4, 1407–1430 (2018)
12. Hamby, D.: A review of techniques for parameter sensitivity analysis of environmental models. Environ. Monit. Assess. **32**(2), 135–154 (1994)
13. Haythornthwaite, C.: Social networks and internet connectivity effects. Inform. Commun. Soc. **8**(2), 125–147 (2005)
14. Helton, J., Iman, R., Brown, J.: Sensitivity analysis of the asymptotic behavior of a model for the environmental movement of radionuclides. Ecol. Model. **28**(4), 243–278 (1985)
15. Hethcote, H., Zhien, M., Shengbing, L.: Effects of quarantine in six endemic models for infectious diseases. Math. Biosci. **180**(1–2), 141–160 (2002)
16. Janvier, F., Delaune, D., Poyot, T., Valade, E., Mérens, A., Rollin, P., Foissaud, V.: Ebola virus RNA stability in human blood and urine in West Africa's environmental conditions. Emerg. Infect. Dis. **22**(2), 292 (2016)
17. Kaner, J., Schaack, S.: Understanding Ebola: the 2014 epidemic. Global. Health **12**(1), 53 (2016)
18. Legrand, J., Grais, R., Boelle, P.-Y., Valleron, A.-J., Flahault, A.: Understanding the dynamics of Ebola epidemics. Epidemiol. Infect. **135**(4), 610–621 (2007)
19. Li, G., Zhang, Y.: Dynamic behaviors of a modified SIR model in epidemic diseases using nonlinear incidence and recovery rates. PloS One **12**(4), e0175789 (2017)

20. Mitchell, L., Ross, J.: A data-driven model for influenza transmission incorporating media effects. R. Soc. Open Sci. **3**(10), 160481 (2016)
21. Na, W., Park, N., Yeom, M., Song, D.: Ebola outbreak in Western Africa 2014: what is going on with Ebola virus? Clin. Exp. Vaccine Res. **4**(1), 17–22 (2015)
22. Ngwa, G., Teboh-Ewungkem, M.: A mathematical model with quarantine states for the dynamics of Ebola virus disease in human populations. Computat. Math. Method. Med. **2016** (2016)
23. Njankou, D., Diane, S., Nyabadza, F.: Modelling the potential role of media campaigns in Ebola transmission dynamics. Int. J. Diff. Equ. **2017** (2017)
24. Pawelek, K., Oeldorf-Hirsch, A., Rong, L.: Modeling the impact of twitter on influenza epidemics. Math. Biosci. Eng. **11**6, 1337–1356 (2014)
25. Prescott, J., Bushmaker, T., Fischer, R., Miazgowicz, K., Judson, S., Munster, V.: Postmortem stability of Ebola virus. Emerg. Infect. Dis. **21**(5), 856 (2015)
26. Rivers, C., Lofgren, E., Marathe, M., Eubank, S., Lewis, B.: Modeling the impact of interventions on an epidemic of Ebola in Sierra Leone and Liberia. PLoS Curr. **6** (2014)
27. Saijo, M., Niikura, M., Ikegami, T., Kurane, I., Kurata, T., Morikawa, S.: Laboratory diagnostic systems for Ebola and Marburg hemorrhagic fevers developed with recombinant proteins. Clin. Vaccine Immunol. **13**(4), 444–451 (2006)
28. Sell, T., Boddie, C., McGinty, E., Pollack, K., Smith, K., Burke, T., Rutkow, L.: Media messages and perception of risk for Ebola virus infection, United States. Emerg. Infect. Dis. **23**(1), 108 (2017)
29. Shen, M., Xiao, Y., Rong, L.: Modeling the effect of comprehensive interventions on Ebola virus transmission. Sci. Rep. **5**, 15818 (2015)
30. The Editorial. The medium and the message of Ebola. The Lancet (2014) https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(14)62016-X/fulltext. Accessed 20 Jan 2019
31. Towner, J., Rollin, P., Bausch, D., Sanchez, A., Crary, S., Vincent, M., Lee, W., Spiropoulou, C., Ksiazek, T., Lukwiya, M., et al.: Rapid diagnosis of Ebola hemorrhagic fever by reverse transcription-PCR in an outbreak setting and assessment of patient viral load as a predictor of outcome. J. Virol. **78**(8), 4330–4341 (2004)
32. Van den Driessche, P., Watmough, J.: Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math. Biosci. **180**(1-2), 29–48 (2002)
33. Wang, L., Liu, Z., Zhang, X.: Global dynamics for an age-structured epidemic model with media impact and incomplete vaccination. Nonlinear Anal. Real World Appl. **32**, 136–158 (2016)
34. Wang, L., Zhou, D., Liu, Z., Xu, D., Zhang, X.: Media alert in an SIS epidemic model with logistic growth. J. Biol. Dyn. **11** Sup 1, 120–137 (2017)
35. Webb, G., Browne, C., Huo, X., Seydi, O., Seydi, M., Magal, P.: A model of the 2014 Ebola epidemic in West Africa with contact tracing. PLoS Curr. **7** (2015)
36. Wilkinson, S. Using media and communication to respond to public health emergencies: lessons learned from Ebola. London: BBC Media Action. (2016). https://www.bbc.co.uk/mediaaction/publications-and-resources/policy/practice-briefings/ebola. Accessed 9 Jan 2020
37. Wittels, A.: Exploring the Role of Communication in Community Health in Sierra Leone. BBC Media Action, London (2016)
38. World Health Organization (WHO). Case definition recommendations for Ebola or Marburg virus diseases: interim guideline. World Health Organization (2014). http://www.who.int/iris/handle/10665/146397
39. World Health Organization (WHO). Ebola data and statistics. World Health Organization (2015). http://apps.who.int/gho/data/node.ebola-sitrep. Accessed 20 Jan 2019
40. Yusuf, I., Yahaya, S., Qabli, S.: Role of media in portraying Ebola in and outside Africa. J. Trop. Dis. **3**(152), 2 (2015)

Index

M. I. Teboh-Ewungkem, G. A. Ngwa (eds.), *Infectious Diseases and Our Planet*, Mathematics of Planet Earth 7, https://doi.org/10.1007/978-3-030-50826-5