

Applied Malaria Dynamics

Dynamical Systems for Adaptive Malaria Control

David L. Smith and the RAMP Team

2023-12-03

Contents

Foreword	11
Contributors	13
1 The Study of Malaria in Populations	15
1.1 The Study of Epidemics	15
1.2 The Prevention of Malaria	18
1.3 Models for Research and Policy	19
2 The Ross-Macdonald Model	27
2.1 Transmission Dynamics	27
2.1.1 Alfred J. Lotka	27
2.1.2 George Macdonald	28
2.2 Rebuilding Macdonald's Model	28
2.2.1 Variables	28
2.2.2 Parameters	29
2.2.3 Dynamics	29
2.3 The Basic Reproductive Number	29
2.4 Sensitivity to Parameters	29
2.5 The Classification of Transmission	29
3 Extensible Transmission Dynamics	31
3.1 The Aron-May Model	33
3.1.1 Mosquito Ecology	33
3.1.2 Blood Fed Mosquitoes	33
3.1.3 Infected Mosquitoes	34
3.1.4 Infectious Mosquitoes	34
3.1.5 Infected Humans	35
3.1.6 ... as a System	36
3.2 Solutions	37
3.2.1 Derivatives	39
3.2.2 Initial Values	39
3.2.3 Parameter Values	40
3.2.4 Solving	41

3.2.5	Visualizing	41
3.3	Steady States	43
3.3.1	Mosquito Density	43
3.3.2	EIR	43
3.3.3	Vectorial Capacity	44
3.3.4	Malaria Prevalence & Thresholds	47
3.3.5	Checking our Work	48
3.4	Stable Orbits	50
3.4.1	Thresholds	51
3.4.2	Orbits	52
3.4.3	Average Dynamics	53
3.5	Discussion	53
4	Realism & Complexity	55
4.1	Epidemiology	56
4.1.1	Superinfection	57
4.1.2	Infection and Immunity	58
4.1.3	Gametocytes and Infectiousness	59
4.1.4	Disease	60
4.1.5	Treatment and Chemoprotection	60
4.1.6	The Time Course of an Infection	60
4.1.7	Intrahost Models	60
4.1.8	Synthesis	61
4.2	Heterogeneous Transmission	61
4.3	Stratification	61
4.3.1	Strata in the Ross model	62
4.3.2	Frailties	62
4.3.3	Age	63
4.4	Humans on the Move	64
4.4.1	Travel	64
4.4.2	Mobility	64
4.4.3	Migration	64
4.5	Blood Feeding	64
4.5.1	Search and Risk	65
4.5.2	Search Weights and Availability	65
4.5.3	Functional Response	66
4.6	Environmental Heterogeneity	66
4.7	Mosquito Behavior and Ecology	66
4.7.1	Resource Availability	66
4.7.2	Egg Laying	66
4.7.3	Search and Dispersal	67
4.8	Aquatic Mosquito Ecology	67
4.8.1	Regulation	67
4.8.2	Exogenous Forcing	67
4.8.3	Habitat Dynamics	67
4.9	Space	67

4.9.1	The Mixing Matrix	67
4.9.2	Pathogen Dispersal by Humans	67
4.9.3	Pathogen Dispersal by Mosquitoes	67
4.10	Time	67
4.10.1	The EIP	67
4.10.2	Seasonality	67
4.10.3	Exogenous Forcing	67
4.10.4	Vector Control	67
4.11	Integrated Vector Control	67
4.12	Pharmaceutic Interventions	67
4.13	Context	67
5	Building & Evaluating Models	69
5.1	Primary Model Development	70
5.1.1	Variables	70
5.2	Implementation and Verification	71
5.2.1	Solving	71
5.2.2	Mathematical Analysis	71
5.2.3	Simulation	71
5.3	What is a Model?	71
5.4	Accuracy & Validation	71
5.4.1	Validation Points	71
5.5	The Measure of a Model	71
5.5.1	Limitations	71
5.5.2	Parsimony	71
5.6	Secondary Model Development	72
5.7	Conclusions	72
6	Modularity and Software	73
6.1	Model Building	75
6.2	Modular Computation	76
6.2.1	exDE	76
7	Measuring Malaria	77
7.1	Realistic Bounds	79
7.2	The Local Fraction	79
7.2.1	Travel	80
7.2.2	Mobility	80
7.3	Drug Taking	80
7.4	Seasonality	80
7.5	Frailty	80
7.6	Environmental Heterogeneity	80
7.7	$PfEIR$ vs. $PfPR$ in Data	80
7.8	$PfEIR$ vs. $PfFoI$ in Data	80
7.9	Synthesis	80

8	Measuring Vector Control	81
8.1	Coverage	81
8.2	Effect Sizes	81
I	Transmission	83
9	Heterogeneous Exposure	85
9.1	Overview	86
9.1.1	Age	86
9.1.2	Location	86
9.1.3	House Type	86
9.1.4	Activities	86
9.2	Frailty	86
9.3	Environmental Heterogeneity	86
10	Blood Feeding	87
10.1	Host Availability	87
10.2	Blood Feeding Rates	87
10.3	The Human Fraction	87
10.4	The Mixing Matrix, β	87
11	Spatial Dynamics	89
12	Temporal Dynamics	91
12.1	Exogenous Forcing	91
12.2	Mosquito Survival through the EIP	91
13	Cohort Dynamics	93
13.1	Boxcar Models	93
13.2	Delay	93
14	Demography	95
14.1	Migration	95
15	Stratification	97
II	Human Ecology	99
16	Human Behavior	101
17	Human Mobility	103
18	Human Travel and Malaria Importation	105

III	Epidemiology	107
19	Malaria Infection and Immunity	109
19.1	Infection	109
19.1.1	Duration of a Simple Infection	109
19.1.2	Superinfection	109
19.1.3	Age of Infection (AoI)	109
19.1.4	Parasite Densities	109
19.1.5	Stage of Infection (SoI)	109
19.2	Immunity	109
19.3	Within Host Models	109
20	Age, Exposure and Malaria Immunity	111
20.1	The Garki Model	111
20.2	Stage-Structured Immunity	111
20.3	Strain Specific Immunity	111
20.4	Memory Tracking	111
20.5	Age <i>vs.</i> Prevalence	111
21	Detecting Parasites	113
21.1	Parasite Densities and Detection	113
21.2	Light Microscopy	113
21.3	Biomarkers and RDTs	113
21.4	PCR	113
22	Gametocytes and Infectiousness	115
22.1	Gametocytemia	115
22.2	Anti-Gametocyte Immunity	115
23	Fever and Severe Disease	117
23.1	Fever	117
23.2	Anemia	117
23.3	Severe Disease	117
24	Care Seeking	119
25	Drug Taking	121
25.1	Curing Infections	121
25.2	Chemoprotection	121
25.3	Adherence	121
25.4	Treatment Rates	121
26	Pharmaceutical Interventions	123
26.1	SMC	123
26.2	MDA	123
26.3	Drugs	123
26.4	Vaccines	123

27 Malaria Epidemiology	125
27.1 Age of the Youngest Infection	125
 IV Mosquito Ecology	 127
28 Mosquitoes	129
28.1 Aquatic Dynamics	130
28.1.1 Equations	130
28.1.2 Solutions	130
28.1.3 Regulation	131
28.2 Understanding Mosquito Dynamics	133
29 Behavioral State Models	135
30 Aquatic Ecology	137
31 Mosquito Microecology	139
32 Mosquito Dispersal	141
33 Microsimulation	143
34 Mosquito Ecology	145
34.1 Population Dynamics	145
35 Vector Competence	147
36 Measuring Mosquitoes	149
 V Malaria Control	 151
37 Vector Control	153
38 Insecticide-Treated Bednets	155
39 Indoor Residual Spraying	157
40 Larval Source Management	159
41 Attractive Toxic Sugar Baits	161
42 Integrated Vector Control	163
43 Spatial Control	165
44 Pharmaceutical Interventions	167
44.1 SMC	167

<i>CONTENTS</i>	9
44.2 MDA	167
44.3 Drugs	167
44.4 Vaccines	167
45 Discrete Time	169
46 Stochasticity	171
47 Base Models	173
48 Built-in Analytics	175
49 Spatial Concepts and Connectivity	177
50 Model Libraries	179
VI Supplements	181
51 References	183

Foreword

A large fraction of my time over the past 20 years has been devoted to learning about malaria and mathematical models of malaria epidemiology, transmission dynamics, mosquito ecology, vector control, and the evolution of resistance. All the time I was building and analyzing models, I was looking for a way of organizing, operationalizing, and applying the rich body of theory developed over more than a century of malaria research, starting with Ronald Ross [1,2]. A new framework would integrate the concepts and models that have influenced malaria through the present day [3,4], and it might even serve as a platform for recasting a theory of malaria epidemiology, transmission dynamics, and control [5].

My goal was to build a framework that could serve the task of using evidence to develop malaria policy. Policy discussions might need to use mathematics, but it should focus on the policy questions, the concepts, and evidence but not on the mathematics *per se*. Ideally, mathematical ideas should give rigor to the analysis help structure the discussions without ever drawing attention to itself, except perhaps in the rare circumstance when a mathematical detail could help to clarify something. I knew that building a framework would face some core challenges, but if we could solve those, we could make the math easier to use so that discussions would focus on the issues that mattered. Ideally, the framework could be taken up and used by teams of local experts working in their own countries to reduce the burden of malaria and plan for its elimination. The academic activities supporting all this would meanwhile be taken up on the side.

I wanted the new framework to be *extensible*, with *plug-and-play* modularity (for major dynamical components), with flexibility in the choice of functional forms describing various relationships, and with flexibility on the structural components, including space, time, and population strata. It should be scalable: it should be possible to scale down spatially for fine-grained spatial simulations [6–8], to scale down temporally for fine-grained temporal simulations [9], or to scaling up to understand or analyze regional processes and the emerging patterns [10]. To serve the needs of malaria programs, a framework would need built-in support for exogenous forcing by weather and vector control to model malaria as a changing baseline modified by control. To get integrated vector control right, we went all in on mosquito ecology with an individual-based simulation model

with exquisite biological detail [9], which inspired new ideas about mosquito search and a new base model for mosquito ecology and behavior. To get malaria transmission right, we would need to be able to deal with heterogeneity of all sorts. To serve programmatic needs, we needed algorithms that could address the decay of interventions – coverage could decay over time through loss (*e.g.* bed nets) or waning potency – and the net effect of one unit of vector control in relation to transmission intensity and coverage. To make all the pieces fit together, we needed interfaces that could connect up models in a generic way; in the design phase, we worked with two model families for each major dynamical component – one that was dead simple, and one that had a was highly realistic. In some cases, the interface designs called for development of new algorithms: blood feeding, egg laying, environmental heterogeneity, human mobility, and mosquito dispersal. In making a master list to test the framework’s extensibility, we found that some odd cases that needed to pass information among components – endectocides and auto-disseminated larvicides – but it was easy enough to accommodate these. Beyond the algorithms, these models needed the support of mathematical theory. We wanted the software to help understand thresholds, so we wrote the routines that would compute thresholds for malaria transmission in heterogeneous systems, when appropriate.

Sometime in the fall of 2022, the last few pieces came together. We published the first versions of MicroMoB and exDE at CRAN. We submitted a paper to PLoS Computational Biology [11]. A second paper forced us to confront a set of software design questions in relation to exogenous forcing [**addCite?**]. Then it was time to write this book.

The software lowers the costs of building models that are up to the task of guiding malaria policy. An advantage of using this framework and accompanying software is that it has built-in solutions to a large set of problems that arise when building models that combine many factors. These sorts of pitfalls are inevitable and annoying, and they are usually discovered the hard way. We found design solutions that would help others building models to avoid these sorts of problems.

The framework took longer than expected, in part, because there were more pitfalls than we had anticipated.

This book shows, through examples, how to use the software to build malaria models. Even without the software, it fills a gap for students who have taken an introduction to mathematical epidemiology or infectious disease modeling and want to go on in malaria. What are all the special topics that would need to be covered to build models that could be needed in malaria? This book *could* be the basis for such a course, if there were ever enough students. Since there will probably never be enough graduate students at my university who are interested in applied malaria dynamics, the material is being developed for any student anywhere. I will also be writing and recording some short lectures on a subset of topics.

The premise of this book is thus that the reader has started with a solid back-

ground in infectious disease models and malaria. We assume they've seen the Ross-Macdonald model before, that they've taken a class in mathematical epidemiology, and that they know something about how to construct and analyze models. This book emphasizes concepts and teaches through examples. We have left out a lot of the technical and mathematical details, but we have written some vignettes and lessons to supplement the book. Most of this is found in the documentation for `exDE` or `MicroMoB` or it can be found in the `RAMP-Model-Library`.

While this book should help others build models for malaria policy, but it stops short of applying models to policy. That is covered in another book, ***Robust Analytics for Malaria Policy***.

Contributors

I've done the primary writing for this book. The framework would not exist without the work of Sean L Wu and a few others. The book borrows from the work of others, and we have done our best to give credit through citations. It has been a collaborative process (see Contributors). The errors, however, are mostly mine. If you find mistakes or have questions, please drop me a note by email: `smitdave@gmail.com`.

This is a work in progress, so the list of contributors will change over time.

The software package `MicroMoB` was written by Sean L Wu, Sophie Liebkind, and David L Smith. The software package `exDE` was written by Sean L Wu and David L Smith.

Most of the content so far was written by David L Smith. Contributors from the RAMP Team include:

... please consult Dave if you would like a writing role.

Chapter 1

The Study of Malaria in Populations

Ronald Ross is best known for being the first person to observe malaria parasites infecting a mosquito [12], an event that is now celebrated on *mosquito day*, August 20. In his next paper, Ross presented evidence that mosquitoes transmit the parasites to humans, a step that was required to prove that the parasites cause malaria [13]. A few months later, the life-cycle of human malaria parasites was described by an Italian team, led by Giovanni Battista Grassi [14,15]. Ross and Grassi had closed a loop that Charles Laveran opened in 1880 when he published his discovery of malaria parasites infecting humans [16,17]. After Laveran, several studies described basic features of malaria infections in humans [18]. These discoveries and studies set the stage for the modern study of malaria epidemiology. The life-cycle of human malaria parasites as it is now taught was not described fully until the 1940s, when Garnham and colleagues described the liver stages of malaria parasites [18,19]. The story of how the malaria parasite's life cycle was unravelled over time is a topic that worth taking some time to learn, and several good accounts have been written [18,20–22].

While that story that is of great interest, this book on **Applied Malaria Dynamics** is about the use mathematics, especially dynamical systems, to study malaria and the prevention of malaria in populations. This brings us back to Ronald Ross, who – in the later stages of his career – turned his attention to prevention of malaria and the study of epidemics.

1.1 The Study of Epidemics

In 1899, Ross launched a new career that took a very broad approach to the prevention of malaria [23,24]. Ross was a key player in the early days of malaria research, initiating studies of all sorts in malaria epidemiology and control [2,25].

Among his publications are found the first three mathematical models describing malaria transmission dynamics or control [1,26–28]. For the most part, the end of this phase of his career is marked by publication of the 2nd edition of *The Prevention of Malaria* in 1911,[1].

His next ambitious goal was to describe mathematical foundations for the study of epidemics [29,30]. In his 1916 paper, he wrote

It is somewhat surprising that so little mathematical work should have been done on the subject of epidemics, and, indeed, on the distribution of diseases in general.

Which is not to say that *no* work had been done. It's not clear whether Ross was aware, but he certainly did not cite a mathematical model describing small pox epidemics, published by Bernoulli. He. or flu epidemics by **.

... the principles of epidemiology on which preventive measures largely depend, such as the rate of infection, the frequency of outbreaks, and the loss of immunity, can scarcely ever be resolved by any other methods than those of analysis.

Ross proposes a rudimentary three-tiered classification system for epidemics based on the patterns of fluctuating incidence: 1) leprosy and tuberculosis; 2) measles and malaria; 3) plague and cholera.

To what are these differences due? Why indeed should epidemics occur at all, and why should not all infectious diseases belong to the first group and not always remain at an almost flat rate?

In 1917, he teamed up with Hilda Hudson, a Cambridge mathematician, to write the first major contribution to mathematical epidemiology, or what he called *a priori pathometry* [31,32]. Over 18 years (1899-1917), Ross's ideas laid a solid foundations for the modern study of malaria transmission and theory of malaria control [2].

Despite these heroic efforts, the work by Ross and Hudson on *a priori pathometry* was missing one of the key elements of the mathematical study of epidemics, which is now called *mathematical epidemiology*. That missing piece was published by Martini, in 1920, in the form of a new mathematical model. In Ross's model, recovery from human malaria infections returned an individual to the state of being susceptible to infection once again. In Martini's model, recovery was followed by lifelong immunity. Martini's model would be explored in great detail in a series of papers by Kermack, and Mckendrick. In modern language, Ross's models would be studied as SIS compartmental models, while Martini's equations would be called SIR compartmental models.

The dynamical differences between these two kinds of systems of equations are stark. Ross's equations can be solved exactly. After Ross published the equations in *Nature* in 1911 [28], Alfred Lotka published a closed-form solution in 1912 [33]. In an SIS model, individuals could be infected many times, and prevalence

is a naturally meaningful statistic. There is no closed form solution to the SIR model, and prevalence is an ephemerally changing quantity. The natural summary statistic for SIR epidemic is called S_∞ , the portion remaining uninfected asymptotically. In simple epidemics – an epidemic in a single population with no replenishment of susceptible hosts – the dynamics differ depending on how many cases that first case would tend to generate, called R_0 . If $R_0 < 1$, then there is a closed form solution: $S_\infty = (1 - R_0)^{-1}$. If that first case would generate more than one other case, and the number of cases would initially rise to a peak, when each new case would generate exactly one case. Thereafter, the number of cases would decline. The time course of an SIR epidemic does not have a closed form solution. The difference between SIS and SIR models would provide a mathematical basis for answering Ross’s question, and it points to an important role for the concept of population immunity. In SIS models, there is no immunity, so prevalence tends to approach an endemic equilibrium with very little tendency to cycle. When susceptible populations the SIR models are replenished by birth or migration, there is a natural tendency to cycle.

The SIS model is a useful starting point, but even Ross was aware of evidence that immunity to malaria would develop in humans. The role of immunity is more complex, but a critical feature is that infection with one parasite would not prevent reinfection with another. In places where exposure rates were high, reinfection – also called *superinfection* – was quite common. The first model of malaria that was *not* an SIS model – published by George Macdonald but relying on mathematics by P. Armitage – considered a role for *superinfection*, or reinfection of susceptible individuals. In models with superinfection, an interesting statistic is the multiplicity of infection (MoI). A useful discussion of this superinfection model and its mathematical flaws was written by Paul EM Fine. [34] What is, perhaps, more important is that superinfection in malaria facilitates sexual recombination for the parasites. Sexual recombination facilitates development of parasite diversity, which could partly explain population immunity in malaria.

Sticking with Ross’s core challenge of how to study epidemics mathematically to understand the tendency to cycle, we must acknowledge the complex natural history of population immunity. In malaria, a model for the mathematical talents of Klaus Dietz combined with the epidemiological skills of Louis Molineaux to produce the first model of population immunity to malaria. That model described infection dynamics in a population with partial immunity. In malaria epidemiology, a core challenge is the problem of malaria immunity and its relationship to parasite genetic diversity, disease, and infectiousness. In the mathematical study of malaria epidemiology, the formulation of adequate models is among the most pressing and most difficult to address. Measles and a few other pathogens are well-described by the SIR model. For these acute immunizing infections, infection is followed by a life-long immunity. The dengue viruses has four (or maybe five) functionally distinct viruses with complex patterns of cross immunity functionally distinct viruses with complex patterns of cross immunity. For other pathogens, population immunity is undermined when immune-escape

variantes evolve in the pathogen populations. For the influenza viruses and the coronavirus descendents of the viruses that sparked the global COVID-19 epidemic, escape variants arise at a rate that is high enough to sustain yearly outbreaks.

Adding to the complexity, immunity is not the only factor affecting a tendency to cycle. For arthropod-transmitted pathogens, fluctuations in arthropod population densities and behaviors driven by weather affect transmission rates. For directly transmitted pathogens, the tendency for populations to congregate amplifies transmission. In malaria, fluctuating mosquito one factor that must *a priori* force transmission, and since population immunity probably has a weak effect, seasonal malaria is largely driven by fluctuating mosquito populations. In SIR models, seasonal fluctuations are sustained by an interaction between environmental factors and the depletion of susceptible individuals.

The narrative in this book is often historical, which serves two purposes. First, it provides an excuse to cite and discuss old papers that should not get overlooked. Second, a historical narrative allows us to introduce malaria complexity in an ordered way, starting from principles that are simple and abstract, and progressing towards ideas that are non-linear, messy, or subtle. We will often use history as a way of structuring discussions, even as our attention turns increasingly to the mathematics we need to deal with the biological complexity of malaria in populations.

1.2 The Prevention of Malaria

Ross's interest in malaria control traces to 1899,[23,24] just a few months after mosquito day:

... in order to eliminate malaria wholly or partly from a given locality, it is necessary only to exterminate the various species of insect which carry the infection. This will certainly remove the malaria to a large extent and will almost certainly remove it altogether. It remains only to consider whether such a measure is practical [23].

Through the first half of the 1900s, Ross promoted malaria control and research on malaria control in various ways. He engaged in public debates about malaria control hosted by the *British Medical Journal*, and he published a book, *Mosquito Brigades and How to Organize Them* [35].

Many of the early attempts to control malaria were implemented by the British Military. In the first few years, there were some successes, but there were also some failures [36]. Heterogeneity in the responses to malaria control efforts seem to have turned Ross to mathematics. In 1904, he presented a paper at the International Congress of Arts and Science in St. Louis, Missouri that applied diffusion models to larval source management (published also in *Science*[26]).

Ross's transmission models appear to trace back to his work consulting with

national malaria programs in Mauritius and Greece [27,37]. While he was writing about the nuts and bolts of control, he was also grappling with mathematical formulas that could help him understand malaria control quantitatively. How would reducing mosquito densities change the prevalence of malaria? Was there a critical population density of mosquitoes required to sustain transmission? The result was the first mathematical model to describe malaria transmission dynamics [27]. A short time later, Ross would reformulate the model in the 2nd edition of his book, *The Prevention of Malaria* [1]. For Ross, the mathematical models were a logical next step of trying to understand malaria control in rigorous quantitative terms.

By 1915, when Ross turned his attention to the broader study of epidemics, his intent was to establish a discipline, that he called *a priori pathometry*. The problem of understanding epidemic patterns, as we've described above, is quite complex. After his years of studying malaria control, he realized the difficulty of understanding malaria epidemics, so he argued for an integrated approach. In proscribing a method for the study of epidemics, he wrote:

The whole subject is capable of study by two distinct methods which are used in other branches of science, which are complementary of each other, and which should converge towards the same results – the a posteriori and the a priori methods. In the former we commence with observed statistics, endeavor to fit analytical laws to them, and so work backwards to the underlying cause (as done in much statistical work of the day); and in the latter, we assume a knowledge of the causes, construct our differential equations on that supposition, follow up the logical consequences, and finally test the calculated results by comparing them with the observed statistics.

1.3 Models for Research and Policy

Malaria is complex and heterogeneous, which makes it difficult to study and manage. A core challenge in both science and policy is the availability of information. Mathematical models can help us understand and analyze all that complexity and make informed decisions despite the data gaps.

Policy and science are different in many ways, so there are good reasons why we might want to use different models and methods for basic research and policy analytics. In policy, decisions *must* be made in a timely way, and they *should* use all available evidence, even if it's weak. Basic research is epistemologically conservative, by design. Studies that are published in peer review must be repeatable, and in anticipation of criticism, the scientists aim to do things so well that they are unassailable. This usually has the effect of narrowing and controlling the conditions under which the study was conducted. When we operate in policy settings, we *should* design our studies in a different way, and we will thus need to deal with the uncertainty differently. The studies that inform

policies will need to take an approach that is broader – the models should be realistic enough to address the question of interest. Such studies will often need to make compromises and decisions that are poorly informed by the evidence, which raises the question of how much a policy maker could trust it. One strategy for making policies trustworthy is to repeat the supporting analysis using every reasonable approach, so that we can be reasonably sure our policy recommendations would not change. The idea of fully propagating uncertainty is the essential feature of *robust analytics*. If we make the effort, we can identify key sources of uncertainty, identify priority data needs, and collect new data that could help resolve some of the most important sources of uncertainty. Building models to do this is challenging for practical reasons, and it requires drawing heavily on basic research. In giving advice, we must give different weights to the uncertainty than we would in research.

In basic research, we develop mechanistic models to understand malaria as a biological process. In malaria epidemiology, the states and parameters describe infection, immunity, infectiousness, disease, and drug taking in response to exposure. Scientists focus on basic biological mechanisms in order to understand differences in malaria across spectrum of transmission. Immunity and drug-taking are important factors to consider, but it may be that differences in epidemiology and disease across settings arise from differences in the local parasite populations. The models are a way of summarizing knowledge in a quantitative form – something like a complex hypothesis. A test of a model's adequacy is whether it can describe malaria accurately after accounting for differences in drug taking patterns and pattern of exposure.

We study mosquito ecology and blood feeding to understand malaria transmission and develop theory for malaria control. Transmission models couple parasite infection dynamics in humans and mosquitoes through blood feeding. Mosquito populations are shaped by the aquatic habitats for immature mosquito populations. These habitats are standing water bodies, and they are shaped by topography, hydrology, land use, and the water chemistry, which is affected by surrounding rocks, soils, vegetation and pollution. These habitats are filled (exogenously forced) by rainfall and after some eggs are laid, the mosquito dynamics are affected by crowding, predation, and other endogenous dynamics. Larval development and parasite development rates are modified by temperature. Adult mosquito activity rates are affected by temperature, relative humidity, and vector control. Indoor residual spraying (IRS) kills mosquitoes when they rest on a sprayed surface, usually after blood feeding or during the process of searching for a host. Insecticide treated nets (ITNs) protect humans from biting and kill some mosquitoes. By reducing the availability of potential blood hosts, nets can slow blood feeding in some contexts. Larval source management (LSM) reduces immature population densities.

By studying mosquito ecology and malaria transmission dynamics, we can start to understand malaria as a changing baseline that has been modified by malaria control. This is the problem confronted daily in malaria programs, but dealing

with the evidence requires having the tools available to synthesize data describing different parts of malaria. The models help translate evidence into information that can be used to make decisions, to make strategic plans, and to mark progress against national plans. The models encapsulate information about transmission in context, so it is possible to study how malaria persists in a place over time, and how various factors have modified (or could modify) mosquito population dynamics and blood feeding and thereby suppress transmission. Transmission models help us to set intervention coverage targets based on an understanding of malaria connectivity to surrounding regions and local thresholds.

In policy, we use these models with the expectation that – if we fit the models by adjusting parameters that affect how malaria works in some particular place – they *should* help us understand transmission in some particular context and make good decisions about what to do.

Frustratingly, the heterogeneity and the complexity conspire against us. We would like to be sure about how malaria works across settings before we start using the models to stratify populations, tailor interventions to context, or targeting the interventions. Instead, we must admit that we don't know everything we'd like to, and we probably never will. We must proceed with policy without having satisfactory answers to some basic questions. In policy, we will use the models to evaluate the consequences of having missing information, but we will also use the models to help us prioritize missing data so we can fill in the gaps. What missing data would reduce our uncertainty about what to do about malaria? How do we fill the critical knowledge gaps.

To understand malaria or to give policy advice, we must start simple and then add complexity, layer on layer. To deal with missing information, we start with generic models, and then add details to address concerns about some of the details that we hope to identify by studying the systems as we intervene. This approach – starting simple and then layering on complexity – makes it possible to learn as we go. A question is when it stops making sense to add realism to a model. A model that is too simple and abstract might help us understand the basic dynamics and give generic advice, but we would question the model's adequacy if it could not reproduce the patterns we cared about in some particular place at some particular time. As a rule of thumb, a model should be just complex enough to *describe* the patterns we care about and *weigh* the relevant options to give advice. Practically speaking, it's hard to know you've gone far enough unless, at some point, it's clear that you've gone a bit too far.

Over the past few years, we developed a new framework for building models that would make it possible to start simple and then build models of malaria transmission at any level of complexity. We wanted to be able to build in realism by adding complexity one feature at a time. Through this process we can create nested, hierarchical models in branching chains. At the ends of the chains, we might find highly realistic models that are, perhaps, overfit. (The cautions against overfitting play out differently in policy given the urgency of acting in a timely way, but it is also possible to go out and collect new data.) We call

the framework’s ability to do this **scalability** and the resulting swarms have **scalable complexity**. The iterative attempt to make plans, weigh evidence, quantify uncertainty, gather new data to reduce uncertainty, and then restart the annual cycle, is called **adaptive malaria control**.

To make this possible, we needed a way of building models that would keep the focus on the policy questions and on a dialogue between malaria managers and the analytical support team. We thus sought to design modular software with *plug-and-play* functionality and a high degree of structural flexibility. We needed the framework to be extensible. After making a lot of mistakes, the primary design phase is over, and the algorithms have been published in two software packages. We are currently extending the library of *base models*, which includes some simple or classical models that are instructive or of historical interest. We are also fine-tuning the design requirements for models as we develop protocols that streamline fitting models to data. The software avoids the mistakes we made over the past few years, reuses models, and streamlines the model building process. We hope this software has dramatically lowered the costs of building and analyzing these complex, realistic models.

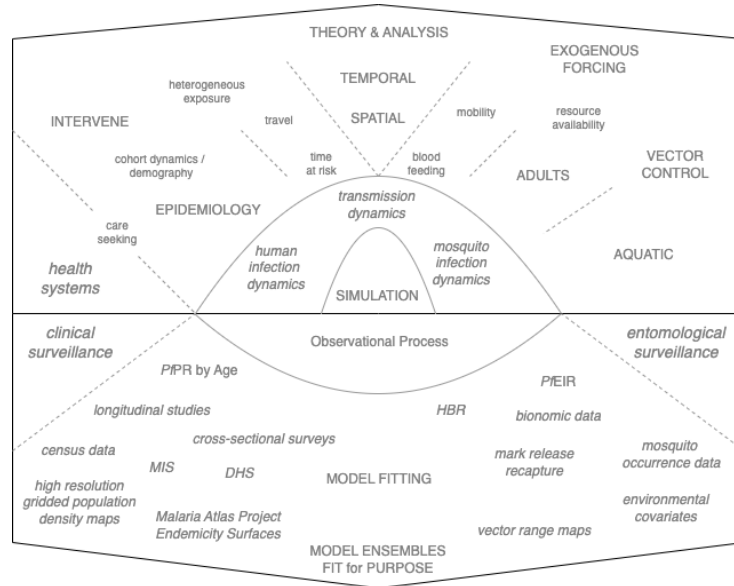


Figure 1.1: A schematic diagram of the elements in the framework (top half) and the process of model building and model fitting (bottom half)

This book has been written to introduce the features of the framework (see Figure 1.1). The book itself is embedded in the RAMP-Model-Library, which was set up during the primary design phase. The RAMP-Model-Library is where we made all our design mistakes: was the software truly plug-and-play, and was the framework truly extensible? As the primary design phase came to a close,

the library that was once the laboratory became a classroom and a museum. The library is being transformed into a resource for any developer who wants to add new base models to the library or add functionality. Most of all, it is being set up for the end user, someone in a malaria program or working with a malaria program who wants to use simulation based analytics to analyze policies. This book is structured into a set of lessons that teach concepts. Some of the concepts build on one another, and others take on new challenges. We combine these lessons into some examples where we show some algorithms to build models fit for purpose. When a topic deserves a deeper dive, we have supplemented this book with vignettes or lessons.

In malaria epidemiology (narrowly defined as a study of infection and disease in humans), the relationship between exposure, infection, immunity, disease, and infectiousness changes in populations as they age, and it is affected by drug taking. This picture grows more complex as we consider intervening with vaccines or monoclonal antibodies, or as we look at interactions with anemia, nutritional status, and human genetics. Our models need to interface with data from clinical settings and research, so they will need to consider diagnostics, parasite counts, detection, and transmission. Combining these factors can give rise to an overwhelming amount of complexity. Later, we will introduce new models and show how it possible to simplify all this complexity and make sense of malaria.

We are interested in using these models to guide policy, which requires both solid computation and good communication. In this book, we lay a foundation for understanding the complexity by studying some simple compartmental models. We will review classical queuing models for superinfection and the multiplicity of infection (MoI); new models for the age of infection (AoI) or stage of infection (SoI); immunity; parasite densities, fever, disease, and detection; gametocytes and transmission, and drug taking. To end up with models that can handle all the complexity, we build probabilistic models that combine these factors. In doing so, we find that we can do some powerful analysis, and we can map the states in these models onto outcomes that matter for research and policy: test positivity, parasite counts, infectiousness, and disease. With patience, we can combine these factors and develop a framework for understanding malaria in populations that match the features of individual-based simulation models. We end up with a sensible understanding malaria epidemiology as ontogeny – development of immunity as a part of an organisms history. We back this view with some very usable models that capture the changing character of malaria in cohorts of humans as they age.

We are interested in understanding malaria control in context, which requires delving into mosquito ecology and behavior. In this book, we start with a simple model for mosquito ecology and parasite infection dynamics in mosquitoes. We add aquatic population dynamics, mosquito population regulation, and exogenous forcing by weather. Later, we worry about adult mosquito behavioral states such as mating, sugar feeding, and egg laying. We introduce the concept of resource

availability, and we develop an understanding of mosquito search and movement in response to resource availability. We take some deep dives to understand how mosquito spatial dynamics work at a fine spatial grain, and then we scale up to understand mosquito populations on landscapes.

At first, we describe mosquito blood feeding and transmission with a few simple parameters. Later, we develop a new model for mosquito blood feeding in a dynamically changing host population with parameters that allow host strata to be more or less available. We also modify our understanding of heterogeneous exposure to biting. We develop a methods for modeling environmental heterogeneity, heterogeneous exposure by age, and a generalized way of handling faulty—other sources of heterogeneous biting – through stratification.

We must take a detour to understand how to handle the effects of temperature on the parasite’s extrinsic incubation period (EIP). We need a way of dealing with mosquito survival and dispersal through the EIP. This problem has been effectively solved.

To round out this picture, we need a way of dealing with other aspects of human ecology that affect malaria transmission dynamics, including human mobility, human demography, bed net usage, adherence to drugs, and care seeking. Differences among humans call for a synthesis of studies that have identified traits that affect malaria, stratification, and simulation to identify useful ways of propagating the heterogeneity through analyses.

To go along with a theory of transmission, we need a theory of control. We compute effect sizes and evaluate area effects. We develop a generalized concept of effect modification that considers the total effect of a single unit of control. We modify basic processes by including the effects of vector control and mass medical interventions (*e.g.* seasonal malaria chemoprotection, mass drug administration, vaccines, and monoclonal antibodies). Relying on behavioral state models and the concept of resource availability, we develop a models for integrated vector control.

In doing all this, we are building on an enormous body of work that started with Ronald Ross. While Ross is better known for identifying malaria parasites in a mosquito gut, which proved that malaria is mosquito transmitted, we are more interested in the academic work that followed.

After winning the Nobel Prize in 1902, Ross was instrumental in building solid quantitative foundations for malaria transmission and its measurement. Ronald Ross wrote the first models describing malaria transmission. In his writings from 1899 to 1908, it’s clear that he was searching for quantitative way of saying something simple – if there are not enough mosquitoes, the malaria transmission can’t be sustained. There must be a critical mosquito density, above the cutoff malaria transmission would be sustained, and below it malaria would be eliminated. Ross was looking for a formula that encapsulated his intuition:

how were thresholds related to the fact that it took two bites for a mosquito to complete its life cycle? Eventually, Ross wrote down some systems of equations that would describe malaria. The ideas, mathematics, and identification of parameters and processes were extended by other scientists later, most notably Alfred Lotka and George Macdonald.

It seems that the challenge of malaria control was what pushed Ross toward modeling. Ross's first model was a discussion of adult mosquito movement to guide larval source management [26]. The first model describing malaria transmission appeared in a book, *The Prevention of Malaria in Mauritius* [27]. When it came to thinking through control, Ross found it useful to do the math.

This is a book about how to do the math that is required for malaria programs. The goal is to use all the data available, but especially the data generated by malaria programs, to paint a clear picture of malaria transmission as a changing baseline that has been modified by control. The software is structured into three major domains: the humans and malaria epidemiology, including the effects of treating malaria with drugs; the mosquitoes and the way they have been changed by weather and vector control; and parasite transmission through mosquito blood feeding. Within each domain, there are multiple sub-domains, and there are built in design features to deal with heterogeneity and other features for malaria control. After a 140 years of studying malaria, there's a lot of detail that could be important in some way. Part of what we need to do is sort through all that detail to find what is most relevant.

We have organized the concepts in this book around a narrative that allows us to introduce the core concepts – those that make modular computation possible – in an order that minimizes the need to draw on unfamiliar concepts. We start with the Ross-Macdonald model, but our next task is to update the model for mosquito blood feeding.

Our philosophy has been to design a framework for model building that can be used by programs. The material in this book is designed to be used by non-experts too, so in this context, *model building* means applying a set of tools to computational tasks we wish our brains could do.

The software we have developed is meant to lower the costs of building and using models. We want programs to be focused on the decisions, the data, the concepts, and the analysis. As a metaphor, some students learn a numerical method for approximating $\sqrt{2}$ in school, but after learning it once, they stop worrying about *how* it is computed and they punch buttons into a calculator. Knowing how to compute something is sometimes useful, but worrying about how to compute it each time would interrupt the process that called for computing it. Instead, we punch the formula into a scientific calculator or any software that does computation confident that the machine knows how to do it. In applying models, the same kind of logic applies. People need to understand the concepts, but like a calculator, the tools should hide the technical details that

don't add to a discussion. The software we have developed is a reliable interface for calculations designed to support policy.

To learn how to use that software, we need to get through a lot of material. The background material in the following presentation is fairly sparse. We are trying to introduce just *enough* mathematics to teach users the critical concepts so they know what the software can do. We assume that the work will be done by teams that include a few people who understand the mathematics, who can guide others through the process. To fill in some of the gaps and technical, we have written (or can write) vignettes. On occasion, the text includes links to these vignettes for those who might find them useful. Please send suggestions about new vignettes to smitdave@gmail.com.

The first model we present is a Ross-Macdonald model.

Chapter 2

The Ross-Macdonald Model

2.1 Transmission Dynamics

The development of the Ross-Macdonald model has been discussed elsewhere [2]. There is not a single paper where a system of equations was published by Macdonald. Instead, the model traces back to several different publications that present equations and formulas describing malaria transmission. We have already mentioned that Ronald Ross published two models of malaria. The first described malaria transmission dynamics using difference equations [27]; it was reviewed, analyzed, and critiqued by H. Waite [38]. The second was formulated as a system of two ordinary differential equations, in the 2nd edition of *The Prevention of Malaria* [1]. Ross also published the same model in *Nature* [28] in this form:

$$\begin{aligned} dz/dt &= k'z'(p - z) + qz \\ dz'/dt &= kz(p' - z') + q'z' \end{aligned}$$

At the time, there was not much evidence from malaria research to frame the problem. A fully parameterized model, using biologically meaningful parameters, would have to wait for a few decades of research and a synthesis by Macdonald.

2.1.1 Alfred J. Lotka

- While Alfred J. Lotka is more famous for his work in demography and ecology, he took an interest in Ross's work on malaria and he made some important contributions to mathematical malaria epidemiology:
- In 1912, he published a set of solutions to Ross's equations [33];
- In 1923, Lotka published an analysis of malaria dynamics in five parts. The first two parts reformulate Ross's models [39,40]. The third part tackles

numerical issues, which includes a photograph of a clay model of the phase plane as a surface [41], and the fifth part is a concise summary [42].

- In the 4th, which was led by Sharpe, a new model was introduced that included delays for the latent periods [43].
- Most importantly, Lotka developed the concept of the basic reproductive number in his work in human demography, which was defined as the expected number of females that would be born to a newborn female.

2.1.2 George Macdonald

- In the 1950s, George Macdonald published a set of papers describing malaria transmission dynamics and control, culminating in publication of a book *The Epidemiology and Control of Malaria* [44]. An important caveat is that Macdonald relied on the mathematical talents of Armitage for the equations describing superinfection and the sporozoite rate [45].
- In 1950, Macdonald published a synthetic review of malaria epidemiology [46].
- In 1950, Macdonald also published a new model of malaria epidemiology with superinfection [47].
- In 1952, Macdonald published a synthetic review of medical entomology [48].
- In 1952, Macdonald described malaria transmission dynamics in a paper that combined his other work, in which he introduced R_0 [49].

It's highly likely that Macdonald was aware of Lotka's demographic concept of R_0 , as well as his work with Sharpe on delay differential equations [43], though I have been unable to find a single line where Macdonald gives Lotka the credit. For these reasons, and others, the equations ought to be described as the Ross-Lotka-Macdonald model.

2.2 Rebuilding Macdonald's Model

The following is a candidate for what the Ross-Macdonald model might have looked like, if Macdonald had ever written it down.

2.2.1 Variables

The model has two variables:

- x is the fraction of humans who are infected;
- y is the fraction of mosquitoes who are infected.

2.2.2 Parameters

The model has several parameters:

- g : mosquitoes die at a constant rate, g . This is equivalent to assuming that the mosquito lifespan is exponentially distributed with a mean $1/g$. Macdonald's equation use p , the fraction surviving one day, $p = e^{-g}$.
- τ : it takes τ days for parasites to mature and reach the salivary glands, called the EIP. The fraction surviving the EIP is $p^\tau = e^{-g\tau}$.
- m : there are m mosquitoes per human;
- a : mosquitoes blood feed on humans at the rate a ;
- b : a fraction of bites by *infectious* mosquitoes causes an infection;
- r : human malaria infections clear at the rate r ;
- c : a fraction of bites on infected humans infect a mosquito, denoted c

2.2.3 Dynamics

In this formulation of the model, we ignore the delay for the EIP but we count the mortality:

$$\begin{aligned} dx/dt &= bma z(1-x) - rx \\ dy/dt &= cax(1-y) - gy \\ dz/dt &= cax_\tau(1-y_\tau) - gz \end{aligned}$$

this.

2.3 The Basic Reproductive Number

2.4 Sensitivity to Parameters

2.5 The Classification of Transmission

Chapter 3

Extensible Transmission Dynamics

If we want to use mathematical models to evaluate policies, we must acknowledge the limitations of the Ross-Macdonald model. Its simplicity is useful as a concise way of understanding malaria transmission through the parasite's life-cycle. There is, however, a great deal of complexity that the Ross-Macdonald model does not address, and some of it matters a great deal in policy. We must extend the model using new mathematical constructs to deal with several features of malaria:

- malaria immunity and disease in cohorts of humans as they age;
- anti-malaria drugs and other pharmaceutical interventions;
- heterogeneous exposure
- spatial dynamics and malaria importation;
- temporal dynamics including seasonality, perhaps exogenously forced by weather;
- mosquito ecology and behavior;
- and vector control.

In the following chapters, we will introduce and critically review the models that have been developed to address these topics in some detail. As we take on these new features and add complexity, we will (in effect) be going around the parasite and mosquito life-cycles over and over again, each time pursuing a new theme.

We want to pursue the idea of scaling complexity in dynamic models ranging from simple and abstract to highly realistic. We want to arrive at a happy compromise – a model that is (as Einstein probably never exactly said [50]) “as

simple as possible, but no simpler.” It might be hard to convince anyone that a model of intermediate complexity is good enough unless at least *some* of the models have clearly gone too far. We don’t want to add new features to our models unless they would affect some important aspect of malaria transmission dynamics and control. Fortunately, we can pursue many of these questions about what details are biologically relevant through model-model comparison. We think of all these models as crude approximations to a messy reality – the question we want to address is whether the approximations are good enough to use as a basis for giving policy advice. Building very complex models that are also mathematically and biologically sound and meaningful is a daunting task, but we have developed a modular system for building models that streamlines the model building process.

To get started, we will reintroduce the Ross-Macdonald model in a form that is easier to extend than the one Macdonald made famous [2]. We will start with a model that first appeared in a 1982 book chapter written by Joan Aron and Robert May [51]. It is a good starting point because it is *extensible*.

We chose it because the variables represent population densities, which are used to compute proportions, like *prevalence*. In the version of the Ross-Macdonald we presented before (in Rebuilding Macdonald’s Model), the variables are proportions. In some models, we would like the *total* number of hosts to change over space or time, but if our state variables are proportions, then these densities appear in the denominators, which would make them much more difficult to modify. By writing down equations using densities, it is far easier to add complexity.

The model is extensible because, as we will see, mosquito population density is a variable, not a term. A parameter in the equation describing changes in mosquito density provides a link to aquatic mosquito ecology, which makes it possible to develop models for mosquito population dynamics, when that is an important consideration. In this version, we find that the formulas describing R_0 and vectorial capacity tell the story of parasite transmission through two blood meals. We will thus update the classical formulas. While Macdonald’s analysis and formulas are familiar to many, they were incomplete [52,53]. We develop a formula for vectorial capacity that is consistent with the intent of the original, but our analysis of sensitivity to parameters includes effects on mosquito ecology [54]. (A lengthy and philosophical discussion of the history and its failings is planned.)

We chose it because it is *realistic*. Time does not appear in most versions of the Ross-Macdonald model: the equations are *autonomous*. These equations use time to drive a seasonal pattern: they are *non-autonomous*. Since we know we are interested in dealing with exogenous forcing, we start out with a model that is forced.

While the following model is basic, we recommend reading it, if only because we introduce concepts and conventions that are important for the software design.

3.1 The Aron-May Model

We define the following variables:

- M is the density of mosquitoes.
- Y is the density of infected mosquitoes.
- Z is the density of infectious mosquitoes.
- X is the density of infected humans.

In dynamical systems, we ask how the variables (*i.e.* M , Y , Z , and X) change over time. In the following, we describe the changes on variable about a time.

3.1.1 Mosquito Ecology

For our first equation, we start with adult, female mosquito populations. (It is tiresome to repeat *adult, female* each time, and we're ignoring male mosquitoes at this point anyway, so *mosquito* hereafter means *adult, female mosquito*, unless we say otherwise.) The number of mosquitoes is changing as new adults emerge from aquatic habitats or die.

To model changes in M , we assume the following:

- mosquitoes emerge from aquatic habitats at the rate of $\Lambda(t)$ adults, per day;
- mosquitoes die at a constant rate, g . This is equivalent to assuming that the mosquito lifespan is exponentially distributed with a mean $1/g$. The fraction surviving one day is e^{-g} .

Our first equation describes changes in the number of mosquitoes:

$$\frac{dM}{dt} = \Lambda(t) - gM \quad (3.1)$$

3.1.2 Blood Fed Mosquitoes

At this point, we will take a detour and define a variable describing the density of mosquitoes that have blood fed at least once, V . After blood feeding, a mosquito is either gravid or *parous*, meaning its ovaries are distended from laying an egg batch. We do this, in part, because the fraction of mosquitoes that are parous is routinely collected, and because it gives us a chance to focus on blood feeding.

To describe *blood feeding*, we assume the following:

- mosquitoes blood feed at the rate f , per mosquito, per day; in this model, this implies that the waiting time to a blood meal is $1/f$ days.
- a fraction of all mosquito blood meals, q , is taken on humans; we call this the *human fraction*

- the human blood feeding rate is the product of these two parameters, $f q$, which is defined as the number of human blood meals, per mosquito, per day.

The number of human blood meals by a population of vector mosquitoes, per person, per day is called the human biting rate (HBR). In this model, HBR is given by a formula:

$$\text{HBR} = \frac{f q M}{H}$$

Later, we discuss the correspondence between the HBR in models and data.

$$\frac{dV}{dt} = f q (M - V) - g V \quad (3.2)$$

We won't use V to describe the dynamics of infection, but we might find it useful to understand how parity changes in mosquito populations.

3.1.3 Infected Mosquitoes

Mosquitoes become infected after blood feeding on an infectious human. To model changes in Y , we extend the model of blood feeding to include infection. We need to know what fraction of blood meals end up infecting a mosquito that has not already been infected.

To model changes in Y , we need to describe *infection rates*. We assume the following:

- a fraction of human blood meals, infects mosquitoes. We call this quantity *net infectiousness* (NI) and (for reasons that we will discuss in a moment), we give it a name, κ :

$$\kappa(t) = c \frac{X(t)}{H} \quad (3.3)$$

- infected mosquitoes die at the same rate as uninfected mosquitoes.

We can now write down our second equation describing changes in the number of infected mosquitoes:

$$\frac{dY}{dt} = f q \kappa (M - Y) - g Y \quad (3.4)$$

3.1.4 Infectious Mosquitoes

To become infectious, a mosquito has to become infected and then survive through the extrinsic incubation period (EIP). We assume:

- mosquitoes become infectious after a fixed delay, τ days, called the EIP. The fraction of mosquitoes that survive through the EIP is $e^{-g\tau}$.
- infectious mosquitoes die at the same rate as other mosquitoes.

For a mosquito to become infectious, it must have become infected τ days ago and survived through τ days with probability $e^{-g\tau}$. To write this in equations, we use a subscripted τ to denote the value of a variable (M , Y or X) or term (κ) at time $t - \tau$. For example X_τ is the number of people who were infected and infectious at time $t - \tau$, and M_τ is the number of mosquitoes at time $t - \tau$.

The number of infectious mosquitoes that are added to the population at a point in time includes all the mosquitoes that became infected at time $t - \tau$ and survived the EIP. This is our third equation describing changes in the number of infectious mosquitoes:

$$\frac{dZ}{dt} = fq\kappa_\tau(M_\tau - Y_\tau)e^{-g\tau} - gZ \quad (3.5)$$

Here, Z represents the number of mosquitoes with *sporozoites* in their salivary glands. The *fraction* of mosquitoes with sporozoites in their salivary glands has been called the *sporozoite rate* (SR), which in our notation is

$$z = \frac{Z}{M}$$

The number of bites by vector mosquitoes, per person, per day is called the entomological inoculation rate (EIR). It is defined as the product of the HBR and the SR:

$$\text{EIR} = \text{SR} \times \text{HBR}$$

In our notation, the EIR is:

$$\text{EIR} = z \frac{fqM}{H} = \frac{fqZ}{H}$$

As with the HBR, we would like to know how to connect estimated values of the EIR to our formulas. Since that's *really* complicated, we've spent a lot of time in the following sections discussing it.

3.1.5 Infected Humans

Humans become infected after being bitten by an infectious mosquito. We assume the following:

- A fraction b of all bites by infectious mosquitoes cause an infection.

- The hazard rate for infection, also called the *force of infection* (FoI) and denoted h is $b \times$ EIR:

$$h = fq b \frac{Z}{H}$$

- Infections clear at the rate r , per infection, per day (the average time to clear is $1/r$ days), and after clearing an infection a person becomes susceptible to infection again.

We can now write down our fourth equation describing changes in the number of infected humans:

$$\frac{dX}{dt} = h(H - X) - rX \quad (3.6)$$

3.1.6 ... as a System

While we presented these equations one at a time, they work as a system. To see it all at once, we write it here as a system with four equations and two terms:

$$\begin{aligned} \frac{dM}{dt} &= \Lambda(t) - gM \\ \frac{dY}{dt} &= fq\kappa(M - Y) - gY \\ \frac{dZ}{dt} &= fq\kappa_\tau(M_\tau - Y_\tau)e^{-g\tau} - gZ \\ \frac{dX}{dt} &= h(H - X) - rX \end{aligned} \quad (3.7)$$

$$\begin{aligned} \kappa &= c \frac{X(t)}{H} \\ h &= bfq \frac{Z(t)}{H} \end{aligned}$$

These equations describe processes in three domains (Figure 2.1):

- adult mosquito ecology (M , and perhaps V);
- parasite infection dynamics in mosquito populations (Y and Z);
- parasite infection dynamics in human populations (X).

The equations describing parasite infections in mosquito populations also include the variable M , so the mosquito infection dynamics are coupled to the mosquito population dynamics. The way we've written the equations, each compartment has an input term (*i.e.*, Λ , κ , or h) that depends on something else. We've passed Λ as a parameter. For the infection dynamics, the terms κ and h couple two separate systems. For adult mosquito dynamics, emergence is passed to the model as a parameters.

There are, of course, more compact ways of writing these equations. We have written the equations this way to emphasize a few things. First, the terms make it clear exactly how the equations in one domain are connected to another. Second, if we wanted to start *changing* some of the assumptions, these terms

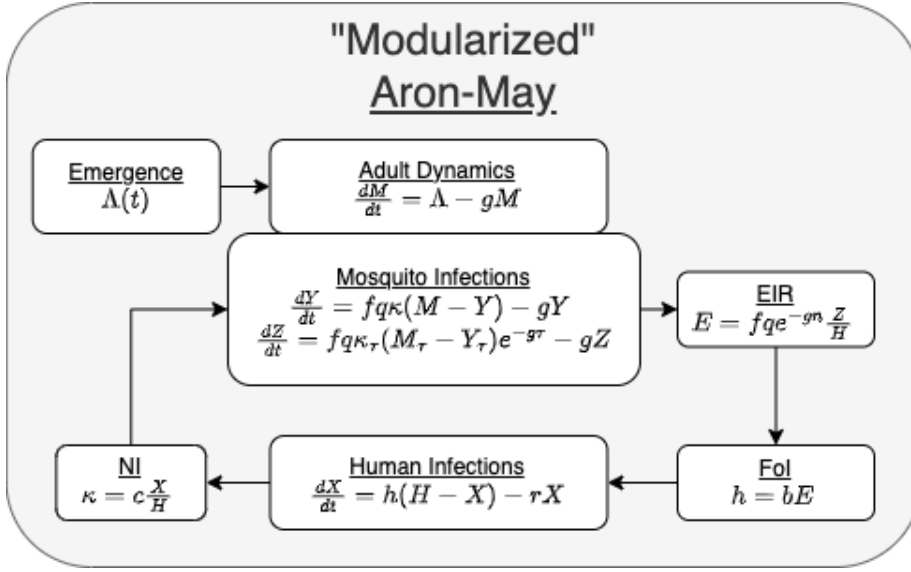


Figure 3.1: A diagram of the a version of the Ross-Macdonald model, using equations from Aron and May [51]

help to isolate the parts we might like to change. By writing the equations in this modularized form, we can start to understand how we might be able to write software that would allow us to represent mosquito infection dynamics with different systems of equations.

The next step is to find solutions.

NOTE: We don't introduce `exDE` or `MicroMoB` until Modularity and Software.

3.2 Solutions

What does a **solution** to these equations look like?

Solutions to these equations are values of the variables over time ($M(t), Y(t), Z(t), X(t)$) that satisfy the system of four equations described above. We call these solutions *orbits*. To put it another way, if we took the derivatives of the orbits for any variable at any point in time using the basic definition

$$\lim_{h \rightarrow 0} \frac{x(t+h) - x(t)}{h},$$

and then we used the values of the variables at time t to compute dM/dt , dY/dt , dZ/dt , and dX/dt (*i.e.*, using the formulas), we would get the same values.

It is important that these orbits are unique: after specifying the *initial values* of the variables, there is one and only one set of orbits that solves the equations.

When we solve the equations, we usually produce solutions from a starting point into a future, but the orbits are defined for all time – *i.e.* the process implies the existence of solutions far back into the past. These are deterministic equations, after all.

As written, the equations do not define a *model*. Instead, the equations define a process or a **model family**. A model is something that *can* produce orbits. A model is defined only after assigning specific values to the parameters. Informally, we will often slip and use the “model” to describe a model family. It’s easy to slip up, and sometimes we can get by with being sloppy, but we need to remember the distinction. When we say that the software is *modular*, we mean that it is easy to swap out one *model family* for another.

To find solutions of equations we use an R software package called `deSolve`. Because of the delay for the EIP, these are called *delay differential equations*, which are handled using a function called `dede`. An important step in solving delay differential equations is a function `lagvalue()` that computes and returns the values of variables at a time lag, ℓ . In these equations, the lag is set by the EIP, τ , so we must evaluate `lagvalue(t-tau)`.

In solving *ordinary differential equations*, we must pass initial conditions. To solve a delay differential equations with a maximum lag ℓ , we must specify the initial conditions for the interval $[-\ell, t_0)$, where t_0 is the point in time when we start computing solutions. In these equations, since the equation for dZ/dt looks back τ units, we must specify values of $M(t)$, $Y(t)$, and $X(t)$ for all values of $t \in [-\tau, t_0)$. This forces an awkward choice, since we don’t know the solutions backwards in time, but would need to know those solutions to use them. What is typically done – and we’ve done it here – is to specify a constant set of initial values and moving on.

Doing this introduces a little *numerical slop*. By slop, we mean that these values are *not* what we would get if we ran the equations backwards in time. In these equations, it won’t affect our analysis most of the time, so we’re happy to acknowledge this little problem and find ways around it. It’s a little thing, but we should never forget it, because we might find that it *is* affecting our analysis at some point.

With `deSolve`, solving differential equations is not difficult – it just involves following a few steps. In the following, we walk through these steps:

- Write a function that computes the derivatives;
- Define initial conditions;
- Define the values of the parameters;
- Define a mesh on time;
- Call a function that solves the equations, such as `dede` for delay differential equations.

Many users will find that reading this code is like learning how to compute $\sqrt{2}$. If so, feel free to learn it once and then skip it.

3.2.1 Derivatives

The first step is to write down the equations to compute the derivatives. The solver expects a function with three required arguments (in this order):

- `t` is time
- `y` is the list of variables
- `params` is a set of parameters

The derivatives are computed and returned in the same order as ‘`y`’ in a `list`. To make code that is easy to read, we make `params` as a `list` with parameter names (see below), so that inside the function `with(params,{...})`, the parameter names are visible.

```
dAronMay = function(t, y, params){with(params,{
  # Variables
  if(t<=tau) ylag<-y0 else ylag <- lagvalue(t-tau)
  M=y[1]; M_tau = ylag[1]
  Y=y[2]; Y_tau = ylag[2];
  Z=y[3];
  X=y[4]; X_tau = ylag[4]

  # Terms
  kappa = c*X/H; kappa_tau = c*X_tau/H
  h = b*f*q*Z/H

  # Dynamics
  dM = Lambda(t) - g*M
  dY = f*q*kappa*(M-Y) -g*Y
  dZ = f*q*kappa_tau*(M_tau-Y_tau)*exp(-g*tau) -g*Z
  dX = h*(H-X)-r*X

  return(list(c(dM, dY, dZ, dX)))
})}
```

3.2.2 Initial Values

To run the model, we must supply initial values. If you were writing code yourself, it would be important to remember that the initial values and the return value for the derivatives must occur in the same order.

A useful convention in `{R}` is to pass the initial values as a named list. Later, we can turn the outputs into a data frame, and then we can retrieve the variables

by name.

```
y0= c(M=60, Y=0, Z=0, X=1)
```

The object `y0` is a named list – the names are attached but invisible.

```
y0
```

```
## M Y Z X
## 60 0 0 1
```

When we turn it into a list, with `as.list`, the names are attached to the values:

```
as.list(y0)$M
```

```
## [1] 60
```

If we use `with`, we create an environment where we can simply use the names:

```
with(as.list(y0), {
  M
})
```

```
## [1] 60
```

3.2.3 Parameter Values

We pass the parameters as a list. It might seem like overkill, but we have written a function `makeParams()` that takes default values and generates a list. This makes it easy to generate a new set of parameter values with alternative values, and it also helps us to write and pass function $\Lambda(t)$ with parameters we like. By passing the parameter as a list, the parameter values are available to the function `dAronMay` when we use `with(params, {})`.

Note that we have also attached the initial values of the variables as a parameter set, which are the return values for `lagvalue(t)` when $t < 0$.

```
makeParams = function(y0,
  g=1/12, f=1/2.5, q=0.95,
  c=0.15,
  b=0.55, r=1/200, H=1000,
  m=.05, ss=1,
  tau=10
){
  ss = min(1,max(0, ss))
  Lambda = function(t){m*H*(1 + ss*sin(2*pi*t/365))}
  return(list(y0=y0,g=g,f=f,q=q,c=c,
    H=H,m=m,tau=tau,b=b,r=r,Lambda=Lambda))
}
params = makeParams(y0)
```


To make it absolutely clear, we are assuming:

- $g = 1/12$: mosquitoes live about 12 days, on average
- $f = 1/2.5$: mosquitoes feed every 2.5 days, on average
- $q = 0.95$: the human fraction is 95%; mosquitoes feed on humans 95% of the time
- $c = 0.15$: about 15% of bites on infectious humans infect a mosquito
- $b = 0.55$: about 55% of bites by infective mosquitoes cause an infection
- $r = 1/200$: human infections last about 200 days, on average
- $H = 1000$: we're simulating transmission in a population of a thousand humans
- $\tau = 10$: the extrinsic incubation period is about 10 days
- For emergence, we tune the average value using m and it is scaled to H :
 - The parameter m in the function above has been set to 0.05 by default.
 - The parameter ss affects the amplitude of the fluctuations. We force it to take on values between 0 and 1.
 - Emergence is modeled as a sinusoidal function with a yearly cycle.

$$\Lambda(t) = mH \left(1 + \sin \left(\frac{2\pi t}{365} \right) \right)$$

3.2.4 Solving

We define a mesh over time – the points in time when we would like to know the values of the variables:

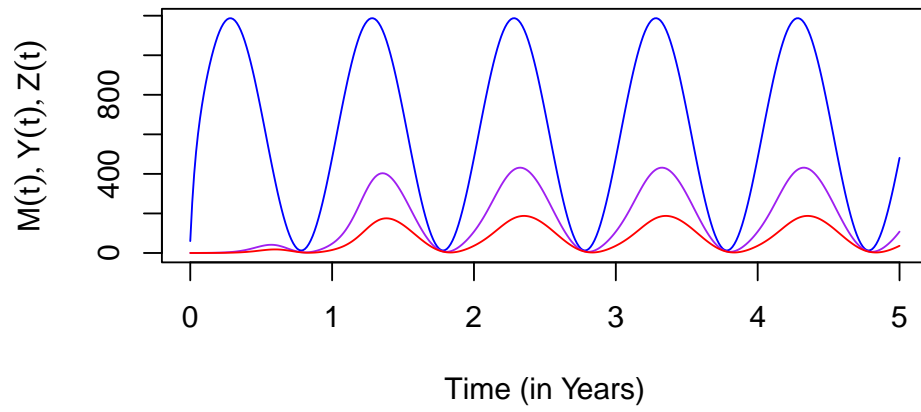
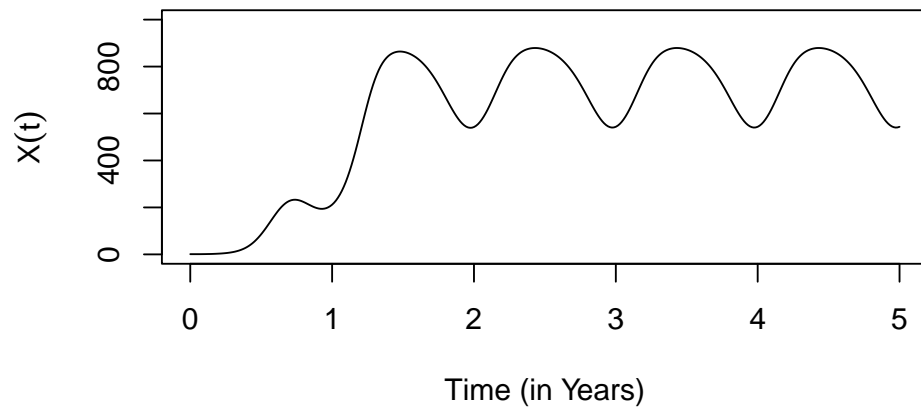
```
tt = seq(0, 5*365, by=5)
```

This code solves the equations:

```
require(deSolve)
yout <- dede(y=y0, times=tt, func=dAronMay, parms=params)
```

3.2.5 Visualizing

When we plot out the solutions, they look like this.

Mosquitoes**Humans**

3.3 Steady States

Here, we analyze the system of equations in a narrow case when there is *no seasonality*, and the system reaches a *steady state*. To do so, we set the parameter $\mathbf{ss}=1$, so that $\Lambda(t)$ is a constant; the resulting system is *autonomous*. We do this, in part, because the resulting system is easier to understand. We can develop intuition that can be applied (albeit with caution) to more complex systems. To be clear, we are dealing with this system:

$$\begin{aligned} \frac{dM}{dt} &= \Lambda - gM \\ \frac{dY}{dt} &= fq\kappa(M - Y) - gY \\ \frac{dZ}{dt} &= fq\kappa_\tau(M_\tau - Y_\tau)e^{-g\tau} - gZ \\ \frac{dX}{dt} &= h(H - X) - rX \end{aligned} \tag{3.8}$$

$$\begin{aligned} \kappa &= c \frac{X(t)}{H} \\ h &= bfq \frac{Z(t)}{H} \end{aligned}$$

As before, we have put the equations in their modularized form above, and the connecting terms below.

The first thing to note is that M affects Y and Z , which affect X ; but M is not affected by Y or Z . Mosquito population density is *exogenous* to malaria dynamics.

3.3.1 Mosquito Density

We can thus treat it separately in the analysis:

$$\frac{dM}{dt} = \Lambda - gM \tag{3.9}$$

Since emergence rates are steady, mosquito population density reaches a steady state when $dM/dt = 0$, which occurs at:

$$\bar{M} = \frac{\Lambda}{g} \tag{3.10}$$

3.3.2 EIR

Next, we note that at a steady state, the delayed values of variables and terms don't change, so from dY/dt , we get:

$$g\bar{Y} = fq\kappa(\bar{M} - \bar{Y}) \tag{3.11}$$

If we substitute the formula for \bar{M} and solve for \bar{Y} , we get:

$$\bar{Y} = \frac{fq\kappa}{g + fq\kappa} \frac{\Lambda}{g} \quad (3.12)$$

and since at the steady state, any variable at time $t + \tau$ is equal to its value at time t , we substitute the formula for $g\bar{Y}$ into dZ/dt to get:

$$g\bar{Y}e^{-g\tau} = g\bar{Z} \quad (3.13)$$

Solving for \bar{Z} we get:

$$\bar{Z} = \frac{fq\kappa}{g + fq\kappa} \frac{\Lambda}{g} e^{-g\tau} \quad (3.14)$$

At the steady state,

$$\text{EIR} = fq \frac{\bar{Z}}{H}.$$

In field studies, the EIR is the product of the HBR and the sporozoite rate (SR). The sporozoite rate (SR, z) is given by:

$$\bar{z} = \frac{Z}{M} = \frac{fq\kappa}{g + fq\kappa} e^{-g\tau} \quad (3.15)$$

So we can understand the EIR as having two parts:

$$\text{EIR} = \text{HBR} \times \text{SR} \quad (3.16)$$

or equivalently

$$\text{EIR} = \frac{fq\Lambda}{H} \times \frac{fq\kappa}{g + fq\kappa} e^{-g\tau} \quad (3.17)$$

This formula for the SR (albeit with slightly different notation) was originally derived as part of the Ross-Macdonald model [45,48]. Also, Smith and McKenzie (2004) have written a useful discussion of mosquito demography [52].

3.3.3 Vectorial Capacity

Here, we pause to define a term that describes the number of human blood meals each mosquito would take over its whole life:

$$S = \frac{fq}{g}.$$

Since $1/g$ is the mosquito lifespan in days, and fq is the human blood feeding rate, S is the number of human bloodmeals a mosquito would take over its lifespan. Intuitively, it makes sense that this *should* be what drives transmission, since it takes two human blood meals to transmit malaria parasites.

If we rearrange the terms a bit, we can rewrite out the expression for the EIR:

$$\text{EIR}(\kappa) = fq \frac{\bar{Z}}{H} = \frac{\Lambda}{H} S^2 e^{-g\tau} \frac{\kappa}{1 + S\kappa} \quad (3.18)$$

This formula for the EIR has two parts. We call the first part *vectorial capacity* (V):

$$V = \frac{\Lambda}{H} S^2 e^{-g\tau} \quad (3.19)$$

The second part is an expression that involves mainly κ .

$$\frac{\kappa}{1 + S\kappa} \quad (3.20)$$

The relationship between VC and EIR at a steady state is a product:

$$\text{EIR}(\kappa) = V \frac{\kappa}{1 + S\kappa} \quad (3.21)$$

Vectorial capacity describes the slope of the EIR when κ is small:

$$d \frac{\text{EIR}(\kappa)}{d\kappa} \bigg|_{\kappa=0} = V \quad (3.22)$$

We say that VC describes *potential* transmission, even if the parasites are absent. Another way to say the same thing is that when κ is small, then:

$$\text{EIR}(\kappa) \approx V\kappa \quad (3.23)$$

We can interpret vectorial capacity (V) in simple terms. It describes *the number of infective bites that would arise from all the mosquitoes biting a single human on a single day but only if all those mosquitoes became infected*. Vectorial capacity tells the story of potential parasite transmission by mosquitoes in four steps, which highlights the fact that two human blood meals are required for the parasite to be transmitted and complete its life-cycle.

$$\begin{array}{c|c|c|c|c}
\Lambda/H & \rightarrow & S\kappa & \rightarrow & e^{-g\tau} \\
\text{Mosquito} & & \text{Parasite} & & \text{Mosquito} \\
\text{Emerges} & & \text{Infects} & & \text{Survives} \\
& & \text{Mosquito} & & \text{EIP} \\
& & & & \rightarrow \\
& & & & S \\
& & & & \text{Parasite} \\
& & & & \text{Infects} \\
& & & & \text{Human}
\end{array} \quad (3.24)$$

As a reminder, while Eq. (3.24) includes κ , the formula for VC, in Eq. (3.19), assumes that $\kappa = 1$: the VC describes transmission as if humans were perfectly infectious. It was defined this way on purpose: it was meant to include mosquito parameters and exclude human factors. We can think of VC as defining something like a conditional expectation, a maximum, or (as we have already said) a measure of *potential transmission by mosquitoes* that is independent of human factors.

While κ (the numerator in Eq.(3.20) accounts for *most* of the difference between the EIR and the VC, the rest of the difference is due to the denominator in Eq. (3.20), $1 + S\kappa$, which traces back to the formula from dY/dt , which assumes that mosquitoes are either infected or not. The denominator is a measure of saturation – the fraction of mosquitoes that get *superinfected* with parasites. The main point here is that as κ increases, saturation increases. If we set S to the values in the previous plots, we can isolate the relationship:

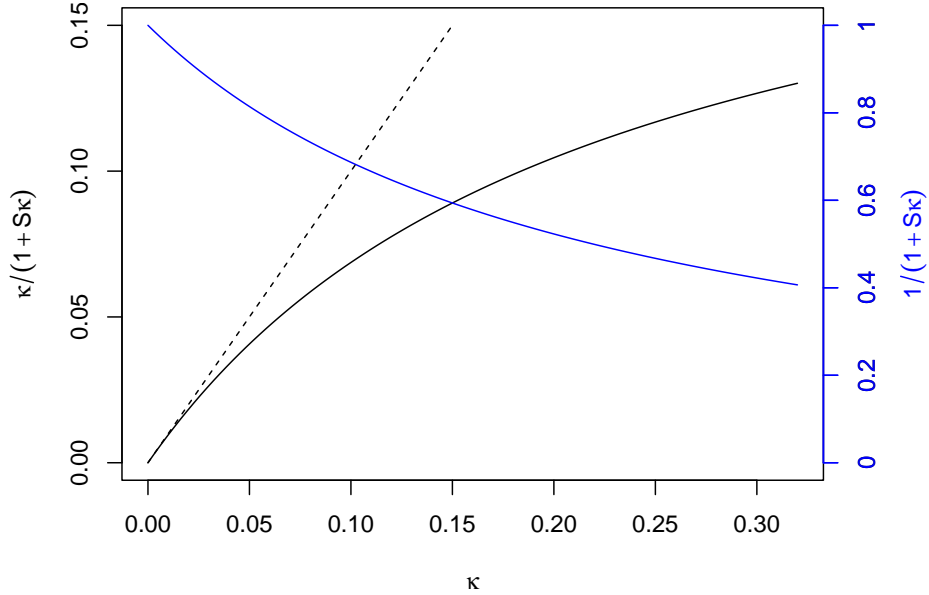


Figure 3.2: The effect (compare the solid and dashed black lines) and effect size of saturation (blue), graphically.

In these formulas, the measure of saturation is exactly $1 + S\kappa$. We could rewrite the relationship between the EIR and VC in a way that tells us something about how we might be underestimating a parasite's reproductive success:

$$\text{EIR} \times (1 + S\kappa) = V\kappa$$

which suggests that each infectious bite is passing along an excess $S\kappa$ bites.

3.3.4 Malaria Prevalence & Thresholds

We let x denote infection prevalence:

$$x = \frac{X}{H} \quad (3.25)$$

so $\kappa = cx$, and

$$\frac{dx}{dt} = \frac{1}{H} \frac{dX}{dt} = h(1 - x) - rx \quad (3.26)$$

We can also define the basic reproductive number:

$$R_0 = \frac{bcV}{r}. \quad (3.27)$$

It is the product of four terms:

- Vectorial capacity, V , the number of infective bites, per person, per day;
- The number of days a person would remain infectious, $1/r$;
- The fraction of infectious bites that would infect a human, b ;
- The fraction of blood meals on infectious humans that would infect a mosquito, c

After taking their product, we can interpret R_0 as a measure of the parasite's reproductive success after a single generation. It only depends on where we start counting. It could be one of the following:

- the number of infected mosquitoes that would arise from a single infected mosquito;
- the number of infectious mosquitoes that would arise from a single infectious mosquito;
- the number of infected and infectious humans that would arise from a single infected and infectious human.

Here, R_0 plays an important role in these equations if we start with dX/dt ; then transform it to dx/dt ; then replace h with bE ; then replace κ with cx ; then divide by r ; and rearrange:

$$\frac{1}{r} \frac{dx}{dt} = x \left[R_0 \left(\frac{1-x}{1+cSx} \right) - 1 \right] \quad (3.28)$$

Since x is the prevalence, it is always in the interval $[0, 1]$. When x is very close to 0, then

$$\frac{1-x}{1+cSx} \lesssim 1. \quad (3.29)$$

and as x grows very small:

$$\lim_{x \rightarrow 0} \frac{1-x}{1+cSx} = 1. \quad (3.30)$$

It follows that when x is small, $dx/dt > 0$ if and only if $R_0 > 1$. Depending on R_0 , only one of two possibilities can hold:

- either $R_0 < 1$, so that $x = 0$ is the steady state;
- or $R_0 > 1$, and the steady state is:

$$\bar{x} = \frac{R_0 - 1}{R_0 + cS} \quad (3.31)$$

Since at the steady state, $\kappa = c\bar{x}$, we can plug this back into the formulas above to get \bar{Y} and \bar{Z} .

What we've learned about these equations is that if mosquito population densities are constant, then malaria reaches a steady state: if $R_0 > 1$, then there is a positive endemic equilibrium, and if $R_0 < 1$, then malaria is absent from the system. The system is said to be stable – in fact, is globally asymptotically stable, which means that all the orbits end up converging to the steady state. This statement has been proved many times in many papers, and since this book is focused on policy, we'll let others worry about proofs.

3.3.5 Checking our Work

An advantage of working in this environment is that we can check our work. One way we could solve these equations would be to run them for a very long time:

We make a parameter set that defines the model:

```
y0=c(M=100, Y=10, Z=5, X=200)
paramsSteady = makeParams(y0, ss=0)
```

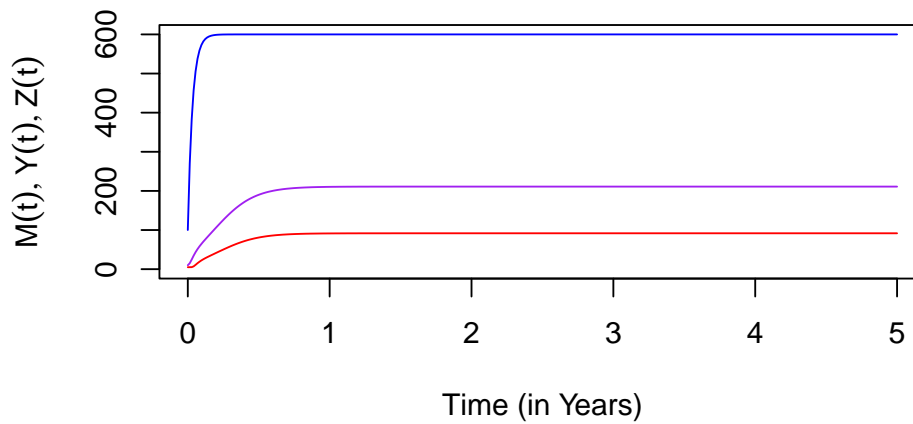
```
dede(y0, times=tt, func=dAronMay, parms=paramsSteady) -> yout
tail(yout, 1)[-1] -> eq1
eq1
```



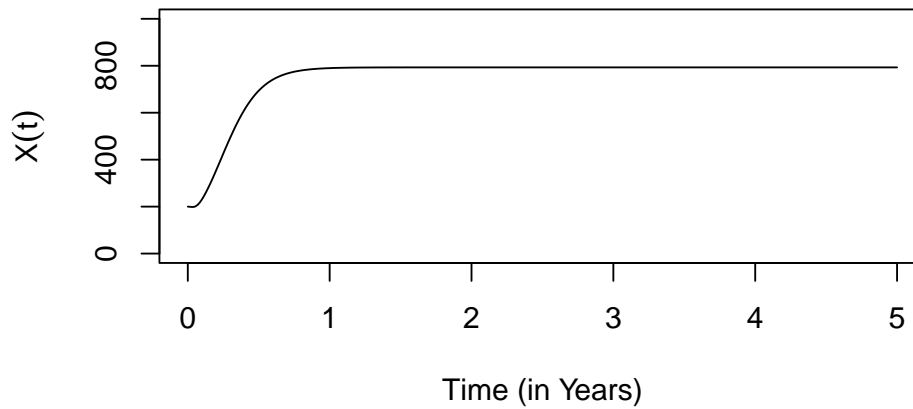
```
## [1] 600.00000 211.01706 91.70764 793.10541
```

By plotting it out, we can check to see if we've run it for long enough:

Mosquitoes



Humans



We can check our algebra by computing the same quantities, and R_0 and other quantities we care about:

```
steadyStates_AronMay = function(params){with(params,{
  Lambda = m*H
  Meq = Lambda/g
  S = with(paramsSteady, f*q/g)
  V = m*S^2*exp(-g*tau)
  R0 = b*c*V/r
```

```

x = ifelse(R0>1,(R0-1)/(R0+c*S), 0)
Xeq = x*H
kappa = c*x
Yeq = S*kappa/(1+S*kappa)*Meq
Zeq = Yeq*exp(-g*tau)
EIR = f*q*Zeq/H
FoI = b*EIR
aEIR = 365*EIR
aFoI = 365*FoI
extra=c(S=S, V=V, R0=R0, x=x, kappa=kappa,
        EIR=EIR, FoI=FoI)
annual =c(aEIR = aEIR, aFoI=aFoI)
list(std=c(M=Meq, Y=Yeq, Z=Zeq, X=Xeq),
     extra=signif(extra, 3),
     annual = signif(annual,3))
})}
steadyStates_AronMay(paramsSteady) -> eq2
eq2

```

```

## $std
##           M           Y           Z           X
## 600.00000 211.01706  91.70764 793.10541
##
## $extra
##           S           V           R0           x kappa           EIR           FoI
## 4.5600 0.4520 7.4600 0.7930 0.1190 0.0348 0.0192
##
## $annual
## aEIR aFoI
## 12.7  7.0

```

Now, we can compare directly:

```

rbind(eq1=eq1, eq2=eq2$std)

```

```

##           M           Y           Z           X
## eq1 600 211.0171 91.70764 793.1054
## eq2 600 211.0171 91.70764 793.1054

```

3.4 Stable Orbits

If emergence rates vary seasonally, how much of the analysis that we did to understand *steady states* still holds? Obviously, if conditions are changing seasonally, the model does not reach a steady state. In fact, after modification to suit the context, many of the same principles translate. The steady state analysis provides a good qualitative guide, but that the answers will look different. Here,

we illustrate by solving systems to illustrate some basic points, which is easy enough. Analysis of the resulting dynamics can be quite difficult; it is covered in Temporal Dynamics.

3.4.1 Thresholds

There is a threshold condition $R_0 > 1$ that determines whether malaria is endemic, but the formula for R_0 depends on the form of $\Lambda(t)$. If we set $R_0 = 1$, we can show that the threshold for persistence in a seasonal environment is $R_0 > \sigma > 1$ (see Figure 1.1). The math to compute threshold conditions in seasonal environments is in Temporal Dynamics.

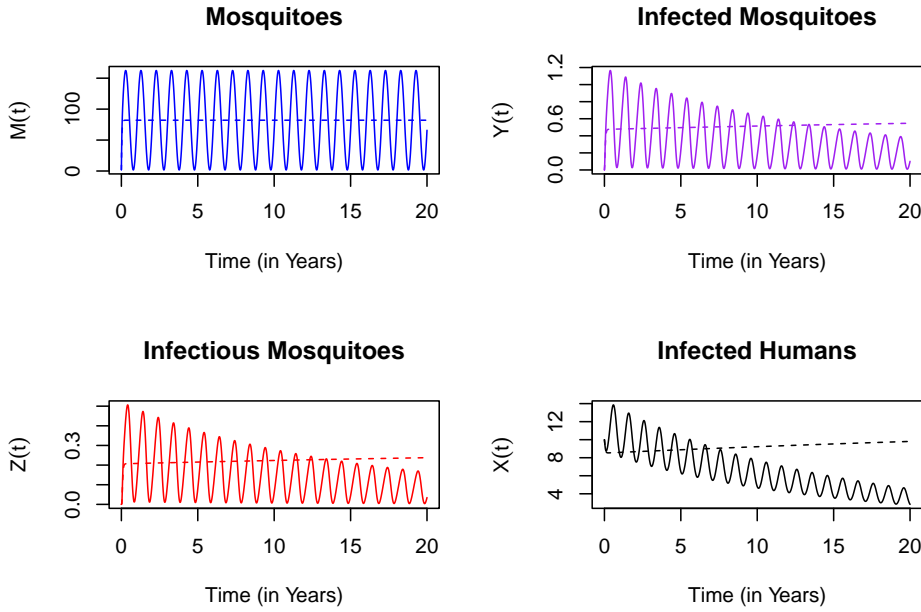


Figure 3.3: Here, we set $R_0 = 1.02$ for the model with constant emergence, and we show that malaria persists. For the same parameters and for the same *average* emergence rate, malaria declines with seasonality.

3.4.2 Orbits

If $R_0 > 1$, then all orbits converge to a set of *stable orbits* (See Figure 1.1). If $\Lambda(t)$ has an annual cycle, then after the orbits converge:

- $M(t+365) = M(t)$;
- $Y(t+365) = Y(t)$ and $Z(t+365) = Z(t)$;
- $X(t+365) = X(t)$.

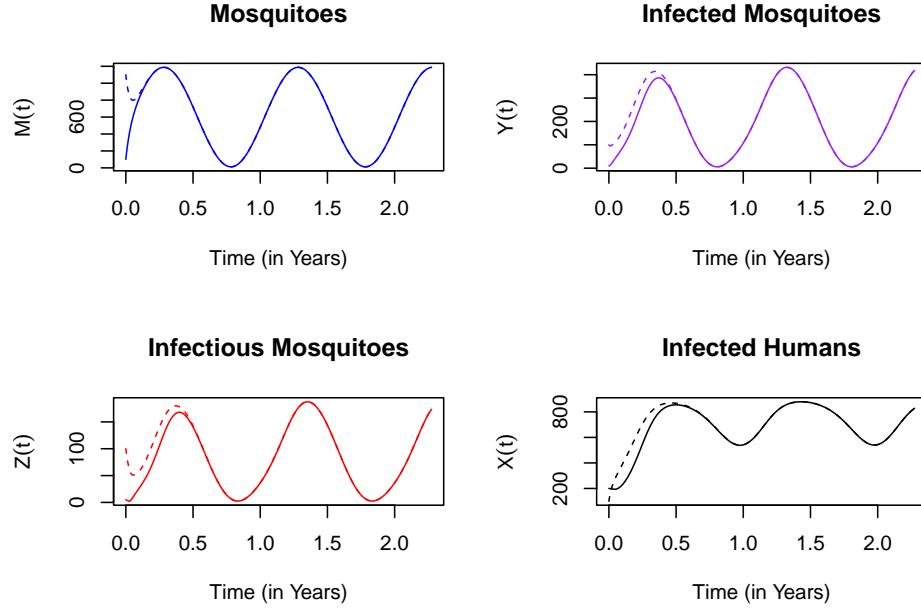


Figure 3.4: With different initial values, the orbits converge and eventually lie on top of one another.

3.4.3 Average Dynamics

If $R_0 > 1$ and malaria is endemic, the *average* prevalence of malaria infection is variable in a seasonal environment. While the prevalence is higher at the peak, the average for the whole year tends to be lower.

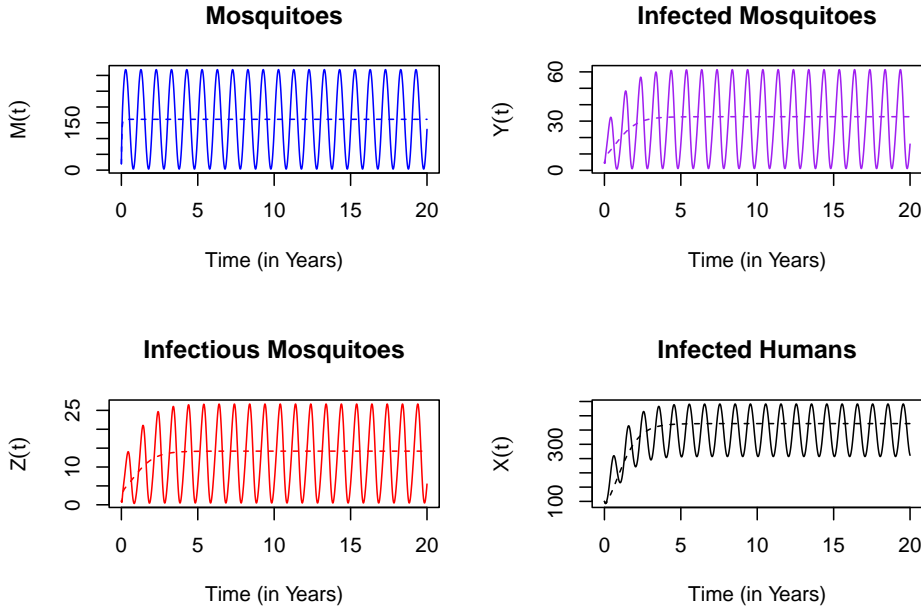


Figure 3.5: Here, we set $R_0 = 2$ for the model with constant emergence, and we show that the prevalence of malaria is similar in the seasonal environment, but it's higher as transmission peaks, lower in the off-season, and lower overall.

3.5 Discussion

If we were simply learning the math, . . .

When we analyze these equations to determine their *stability* and to identify *threshold conditions*, we focus on threshold conditions and the behavior of these systems when malaria is rare. In most places, malaria is endemic so we need to be concerned about malaria immunity and its effects on transmission; malaria is under some level of control; and because of weather and other factors, the baseline conditions change from year to year.

Chapter 4

Realism & Complexity

There are several reasons why models for policy will tend to be more complex than models developed for science. In science, the questions are narrow and the study designs are focused enough to form tests of ideas that have clear outcomes. When models are used to guide policy, the same level of scientific rigor should be applied, but weighing all the evidence requires a broder synthesis.

To build models that can guide malaria policies, the models must be *realistic* enough to be compelling, and they *ought* to reflect the knowledge and experience accumulated over years of studying and controlling malaria. Policy advice should be checked for consistency across studies. The models must be complex enough to serve many purposes all at once. To weigh tradeoffs, policy requires broad, synthetic models that allow for comparisons across subject matter domains.

To carry a conversation forward, the models used to guide discussions will need to retain a memory of what has been learned already, so they will tend to add features and grow more complex. Given the uncertainty, policy should be based on model swarms that propagate the uncertainty. The predictions of those models must be specific enough to be proven wrong, so that over time *some* of the models can be trusted over others. The same models can be used to identify which missing data would have the greatest impact on a policy, and ideally, studies can be conducted to gather this data. Over time, the advice should shift from *generic* advice to *specific* advice as more evidence is gathered. This is, in a nutshell, how daptive management works.

It might take a lot of work to build a model that has been fit to all the evidence describing malaria in a management unit over the recent past, and it might cut against the instincts we have as scientists to add all that realism, but it's worth it to make the effort if it helps communicate with malaria managers.

In designing a software solution to the problem of building realistic models, we

designed a framework for building models and a toolbox to build model swarms that would address the concerns of malaria programs. In the chapters that follow, we'll show the features of this framework by constructing examples. Even if we're principled about adding complexity, a cost of doing so is *computational complexity*. That is something the software was designed to manage. For the moment, we thus want to set aside concerns about *realism* vs. *abstraction*, about *parsimony*, and about *error propagation*, and we want to simply ask the question of how to build models with the features we want.

This chapter is an overview of the historical development of malaria models and an introduction to the toolbox. We'll cover the same material in much greater detail in the chapters that follow, and we'll construct examples using `exDE` and `MicroMoB`.

4.1 Epidemiology

A major challenge for malaria dynamics is how to define an state space describing malaria infection and immunity in human populations that captures the essential elements of malaria dynamics well enough to trust for making policies. There are features of malaria infections that have been identified and studied in the past: superinfection; the complex time course of an infection – including fluctuating parasite densities – and the problem of detection; gametogenesis, gametocyte maturation and gametocyte dynamics; fever and disease; development of immunity with exposure including its effects on infection, disease, and infectiousness; treatment, adherence to drug regimens, chemoprotection and infection curing. Over time, these issues have been addressed in various models. We need model that is good enough for policy, but this also means developing a common understanding of malaria that can serve as a basis for discussion.

To get to that point, we must start simple and add complexity.

The model for malaria infection that we presented in [Malaria Dynamics] was developed by Ross. In today's vernacular, it would be called an SIS compartment model. The model is very simple, and it is probably inadequate for every task, but it is useful and it has been used. The model assumes that malaria infections clear at a constant rate regardless of the age of infection or other factors. The persistence of malaria infections over decades tells us that this assumption is clearly false, but it is good enough for some programmatic needs. During the GMEP, Ross's model was used to characterized the response timelines for the *PfPR* after the interruption of transmission. Drawing on multiple sources, Macdonald estimated that the duration of infection was around 200 days, which was good enough to use as a basis for monitoring and evaluating the interruption of transmission [55]. The simpler model was used even though Macdonald had already proposed an alternative model that considered superinfection [47]. Despite the simplicity, the model was adequate to the task [56]. An important lesson is that the simplicity has some advantages, and the models that get used

in policy tend to be very simple.

The question is how to develop models that are simple and yet are up to policy tasks, which means that the models must (at some point) get validated against research data. Doing so means having sufficient complexity to deal with exposure, infection, detection, immunity, disease, infectiousness, care seeking, and drug taking. Whatever model is selected as a basis for policy, it should be simple enough to understand and yet complex enough to capture the *gist* of malaria epidemiology. The models, however chosen, must get it right. Sorting through all the complexity to get a model that is good enough is a daunting task. This introduction is mainly historical, but we use it to preview some of the themes. In the following history, we discuss some of the important innovations.

4.1.1 Superinfection

From early on in malaria epidemiology it was clear that exposure to malaria differed among populations, and that in some places, the rate of exposure was far higher than the rate of clearance. Ross emphasized a need to measure exposure both entomologically, through metrics that are known today as the EIR and the FoI, and parasitologically, through the prevalence of infection by light microscopy (or more commonly today, through RDTs), which was called the malaria *parasite rate*. There was no good reason to believe that people in highly malarious areas would be exposed faster than they would clear infections, so they would carry infections that could be traced back to many infectious mosquitoes [57]. This phenomenon was called superinfection.

Macdonald was the first to develop a mechanistic model of superinfection [47], but the mathematical formulation was at odds with his description [34]. It is an interesting bit of history for a different time.

A mathematical basis for understanding superinfection was worked out as a problem in the study of stochastic processes as part of *queueing theory*. This may seem strange, but understanding how many people are queueing involves understanding how people come in and how fast they are processed. One of these queueing models has become a mainstay of malaria epidemiology; in queueing theory, it is called $M/M/\infty$.

The model tracks the **multiplicity of infection** (MoI). It assumes that infections arrive through exposure at a rate h (the FoI), and that they clear independently. Without clearance, the MoI, denoted ζ , would just go up. The model assumes that each infection clears at the rate r ; if the MoI were 3 then infections would clear at the rate $3r$. Regardless of how fast infections arrive, the fact that the pressure for the MoI to go down increases with MoI means that the MoI will reach a stable state. The mean MoI is h/r . The following diagram illustrates and provides the equations:

$$\zeta_0 \xrightleftharpoons[r]{h} \zeta_1 \xrightleftharpoons[2r]{h} \zeta_2 \xrightleftharpoons[3r]{h} \zeta_3 \xrightleftharpoons[4r]{h} \dots$$

$$\begin{aligned} d\zeta_0/dt &= -h\zeta_0 + r\zeta_1 \\ d\zeta_i/dt &= -(h+r)\zeta_i + h\zeta_{i-1} + r(i+1)\zeta_{i+1} \end{aligned}$$

If one is willing to abandon compartment models, then it is possible to formulate more elegant solution using hybrid models. The mean MoI, m changes according to the equation:

$$\frac{dm}{dt} = h - rm.$$

Using queuing models, it is easy to show that the distribution of the MoI is Poisson, and in these hybrid models, if the initial distribution is not Poisson, then it will converge to the Poisson distribution asymptotically. The complex dynamics of superinfection can thus be reduced to this simple equation.

Unfortunately, things become more complex if we add simple features such as treatment with drugs, or heterogeneous exposure. The distribution of the MoI is no longer Poisson [58]. Superinfection is an important part of malaria epidemiology, and we will use these models for superinfection in developing some adequate models for infection and immunity.

In the Garki Model (see below), the waiting time to clear an infection used these queuing models to formulate an approximate clearance rate:

$$\frac{h}{e^{h/r} - 1}$$

4.1.2 Infection and Immunity

The biggest failing of Ross's model, perhaps, was that it did not make any attempt to grapple with acquired immunity to malaria. It had always been clear that immunity to malaria was important because the prevalence of infection declined throughout adolescence and was consistently lower in adults, and because disease and severe disease were common in young children. The data accumulated through years of studying malaria, done as part of malaria therapy, provided supporting evidence for immunity. There was a difference in outcomes from being exposed to the same parasite (homologous challenge) compared with a different parasite (a heterologous challenge). Immunity had something to do with the number of different parasites that a person had seen.

The first model to grapple with immunity was the Garki Model [59]. The main idea in the Garki Model was that it would be possible to understand malaria dynamics by expanding the number of compartments: the population was sub-divided into two non-immune or semi-immune. Infection dynamics were tracked separately within each immune category: the infections would clear

faster from semi-immune individuals, they were not infectious, and they are less likely to test positive if they were infectious. Some features of the Garki model seem odd in retrospect: there were two infected states for non-immunes (y_1 and y_2), but only one for semi-immunes; there was no way to lose immunity; and the assumption that semi-immunes are not infectious.

The Garki Model has had a powerful influence on malaria modeling. Several models since then have expanded on various themes. Several compartment models have been developed that replicate infection dynamics across immune stages: we call this *stage-structured immunity*.

In the Garki Model, we can simulate the immuno-epidemiology of cohorts as they age. Eventually, the cohort *would* settle to an equilibrium. At that point, everyone is semi-immune, a sizable fraction remains non-immune after a century. By the time the cohort reaches the steady state, everyone in the cohort has died. If we focus on the dynamics in the first two decades of life, prevalence rises as people become infected, and then it falls as people become semi-immune. The changing epidemiology as cohorts age is an important feature of malaria. In models like this, the concept of a *steady state* teaches us something, but the models draw attention to the sharp changes in malaria that occur through the first 20 years of life. We can adapt the idea of steady state to suit our needs – under constant exposure, cohorts trace out *stable orbits* as they age. These stable orbits are a basis for understanding malaria dynamics *vs.* age.

One application of these stable orbits is to understand the relationship between age and infection prevalence as a function of exposure. Curiously, the Garki Model captures the basic shape of age-*PfPR* curves, but it does not get the details right. When we start to look at the factors affecting the *PfPR* by age in populations, we must acknowledge the need to add other features: drug taking and chemoprotection; differences in exposure that arise for a number of reasons; anemia, perhaps; seasonality. Not everything is about immuno-epidemiology.

4.1.3 Gametocytes and Infectiousness

In Ross's models, everyone who is infected is also infectious. This is clearly wrong, but it may not be a terrible assumption under most circumstances. The Garki Model made the extreme assumption that semi-immune individuals are not infectious at all. There is now copious evidence that adults *do* transmit parasites to mosquitoes, but they are not as infectious. This decline in infectiousness occurs for two reasons: first, the densities of asexual-stage parasites in adults are controlled by immunity, so they are lower. Since a fraction of asexual parasites becomes gametocytes, the densities of gametocytes are also lower in adults. Second, gametocyte densities are modulated by an immune response that affects malaria parasites in mosquitoes, which is called *gametocyte-stage transmission blocking* immunity. The dynamics of gametocyte-stage immunity change with age and exposure, and we will need to understand how this form of immunity waxes and wanes.

There are some other important details about malaria infections that might be relevant in some contexts. First, *P. falciparum* gametocytes take 8-12 days to mature. When combined with the 6 days in the liver, we must acknowledge that the latent period is at least 2 weeks. Because gametocytes must reach densities high enough to be transmitted, the effective latent period for humans is probably closer to 20 days. Since the parasites also need 10 days or more to mature in mosquitoes, the shortest parasite generations are probably at least a month long.

Another feature of gametocytes that matters is that gametocyte populations are not always affected by anti-malarial drugs, so after taking drugs that clear all the asexual-stages, some people will remain infectious to mosquitoes for quite a while after being treated with some drugs.

Ross's assumption may serve most needs, but the models must be good enough to guide policies, such as MDA or malaria elimination, when details about gametocytes and infectiousness can affect the outcomes of policies.

4.1.4 Disease

Disease was not incorporated into most mechanistic models of malaria until recently.

4.1.5 Treatment and Chemoprotection

It is impossible to understand malaria infection dynamics without accounting for treatment with anti-malarial drugs and a brief period of chemo-protection that follows. The first model for drug treatment was developed to understand MDA [60]. In developing models for policy, we must be careful about drug taking and its effects because it modifies the relationship between exposure (the EIR) and infection.

4.1.6 The Time Course of an Infection

The time course of infections is complex, and we will need to develop some models that relate parasite densities. In the chapters that follow, we introduce two main kinds of models:

- AoI
- SoI

4.1.7 Intrahost Models

There are two kinds of models we will discuss, but we would like to avoid them in making policy if possible.

- In host models;
- Individual-based models.

4.1.8 Synthesis

In the end, we do not need perfect models of malaria infection and immunity, but we do need a sound understanding of several things to make policy:

- The prevalence of infection by age as a function of exposure and drug-taking;
 - ... in a cross-section of the population;
 - ... in the care-seeking population.
- The incidence of malaria by severity and by age;
- The fraction of malaria that is promptly treated by severity and by age;
- The net infectiousness of a population of humans to mosquitoes.

In the chapters that follow, we will develop some models that based on a new concept – the age of the youngest infection – that combine many of the ideas in the chapters above.

4.2 Heterogeneous Transmission

4.3 Stratification

Human populations are heterogeneous. Some kinds of heterogeneity affect how we understand malaria and what we should do, including who to target. To deal with heterogeneity in models, we will often need to *segment* a human population into sub-populations, or *strata*. When we talk about *stratification*, we mean it the narrow sense of segmenting a human population (*i.e.* not subdividing landscapes spatially¹), because the model predictions made by creating strata that are more homogeneous should be more accurate. The guiding principle is that our analytics will should strive to be more accurate, and that we should thus identify and remove those sources of heterogeneity that would affect policy advice, whether it affects estimating the impact of interventions in the past or projecting those impacts into the future. We acknowledge that models are approximations, and that our approximations don't have to be perfect. The goal is to find ways of propagating uncertainty that are *good enough* for the task at hand.

In malaria epidemiology, *some* kinds of endogenous heterogeneity *could* be built into the *epidemiological state space*. Other kinds of heterogeneity, including consistent differences in exposure, differences in care seeking and drug taking, and differences created by malaria control (*e.g.* net ownership or vaccination),

¹In a broader sense, stratification is also about subdividing landscapes into a set of spatial domains that share relevant features in order to *tailor interventions to context*. That is a topic we take up in a separate book, (**Robust Analytics for Malaria Policy**).

usually require stratification. The decision about how to strike the right balance depends on the model and the purpose of a study.

The framework and supporting software offer a toolbox for stratification. It is designed to stratify populations in a principled way, so that we can *understand* how the heterogeneity affects transmission or outcomes that we care about, but we can also *combine* effects. We want to stratify populations by applying rules that *split* populations when the differences are large enough. (If we started with complex models, we might choose to *join* populations if the differences were small.) By so doing, we can *compare* the behaviors of models that differ from each other in only one way. If the differences are not too large, or if the differences in dynamical behaviors we care about are not too large, we might decide not to split the strata, and use the average. Because of the way models are encoded, it's easy to build models that split the strata in multiple, independent ways.

4.3.1 Strata in the Ross model

As a simple example, consider a simple Ross-style model for infection with exposure and recovery (described in Section 3.1.5):

$$\frac{dX}{dt} = h(H - X) - rX$$

If exposure is heterogeneous, we could split this population into two strata and add subscripts (*i.e.*, indexed by $i \in \{1, 2, \dots\}$):

$$\frac{dX_i}{dt} = k_i h(H_i - X_i) - rX_i$$

We hold the *average* FoI constant by constraining the values of k_i :

$$\frac{\sum_i k_i H_i}{H} = 1$$

Stratification is important if the differences are large. With two strata, it would not make sense to stratify if $k_1 \approx k_2$, but if $k_2 \gg k_1$ then it might change our expectations, or it might change what we recommend.

4.3.2 Frailties

We will introduce segmentation first through models of Heterogeneous Exposure to malaria, where we consider various sources of *frailty* – proportional differences in the average hazard rate for infection (k_i , in the example above). These differences in exposure can arise because of age, house type, risky behaviors, other factors. Frailty that is attributable to location (*e.g.* proximity of home to aquatic habitats) can be dealt with by sub-dividing space into *patches*, a topic

that is taken up in Space below and Spatial Dynamics. Depending on the size of the patches, some differences in average rates of exposure due to location can persist, and these could be dealt with by generic stratification into high *vs.* low exposure strata.

Some of the heterogeneous traits that we care about change dynamically, so we will also need to consider population *flows* among strata, which change the sizes of the strata. We would like to deal with these flows in a principled way. Bed net ownership and use are among the human behaviors that matters most for programs. In some cases, we will want to understand dynamic changes in bed net ownership, the patterns of use among those who own a net, personal protection, and community effects. Later, we show how to construct an example that *describes* all of these aspects of bednets.

Segmentation is what we need to build models of pharmaceutical interventions with waning effectiveness, such as mass vaccination. Among the most important factors in malaria is age. We have defined algorithms to model [Aging] and other demographic change, the loss of bednets, waning protection or changing housing quality.

4.3.3 Age

Immunity to malaria develops with age and exposure. The development of immunity is probably changing throughout life, so it makes sense to think of malaria epidemiology as ontogeny.

For systems described generically by the state space, \mathcal{X} , the dynamics we care about have the form:

$$\frac{\partial \mathcal{X}(a, t)}{\partial a} + \frac{\partial \mathcal{X}(a, t)}{\partial t}$$

We might want to deal with malaria differently if we are studying malaria in cohorts. In a population where the FoI over time is $h(t)$, we might want to follow a birth cohort, so we define $h_d(a) = h(t - a)$ for all $t > d$. We can then solve:

$$\frac{d\mathcal{X}}{da}$$

which produces states in cohorts as they age, $\mathcal{X}(a|h)$.

When we simulate malaria transmission dynamics in populations for policy, we will want to put a mesh on age and segment the population. The dynamics are define for age strata, where the FoI is defined differently for each age stratum:

$$\frac{d\mathcal{X}_a}{dt}$$

which produces age-dependent states over time, $\mathcal{X}_a(t|h)$.

Our algorithms should guarantee that the epidemiological states over time provide an accurate match for the epidemiological states over age.

4.4 Humans on the Move

The notion of a spatially distributed risk for humans and the modalities of human travel.

- Humans move around, so we develop a model of *time spent*. Time spent is sub-divided into three parts:
 - time spent at home;
 - time spent traveling, when a night is spent away from home;
 - human mobility, which describes time spent around home when not traveling.
- For travel, we estimate a travel FoI.
- For time at home and mobility, after weighing time spent and mosquito diurnal activity patterns by time of day, we modify time spent to get a notion of *time at risk*
- After modifying time at risk by search weights, mosquito blood meals are distributed among all hosts according to their availability.

4.4.1 Travel

4.4.2 Mobility

4.4.3 Migration

4.5 Blood Feeding

The second topic we must tackle is blood feeding, which is an interaction between mosquitoes and humans. It is an asymmetric relationship – mosquitoes search for blood hosts, select a host, and blood feed. Humans, for their part, attract mosquitoes from a distance, move around, and spend time in places when mosquitoes are biting. Humans can wear protective clothing (or not), use bed nets (or not), or do other things that make them more or less available to humans. Despite all this, humans are often unaware that they have been bitten.

Transmission occurs during blood feeding, and models of blood feeding *should* be able to take all this heterogeneity into account. If the models do a proper accounting, then the total number of human blood meals taken by mosquitoes would equal the number of blood meals received by humans. In doing so, we find no inspiration from Macdonald, whose description of human blood feeding was simple and phenomenological: a single parameter described the human blood

feeding rates. After Garrett-Jones described the human blood index, drawing on decades of work, the one parameter was split into an overall blood feeding rate (f) and a human fraction (q). The question left unaddressed by Macdonald was how these rates vary by context, and the consequences for exposure. To do this, we reformulated the algorithm describing blood feeding [11].

Over the past two decades, several papers have drawn attention to the way blood feeding behaviors are or ought to be constrained by the availability of vertebrate hosts. It may be fine to assume that the density of vertebrate hosts doesn't change, but *something* should change when a large fraction of people are using bednets. Even with static parameters, we should think through the limiting cases: if there are no vertebrate hosts, then there blood feeding should not occur (*i.e.*, $f = 0$); if there are no human hosts, then there should be no human blood meals ($q = 0$); and if there are no alternatives to humans, all blood meals should be on humans ($q = 1$).

The concepts we devised for blood feeding must, therefore, integrate the notion of frailty with the process of mosquito search. On the one hand, the mosquitoes should blood feed at a slower rate if hosts are unavailable. On the other hand, human biting should become heterogeneous. To arrive at an adequate description, we need to formalize this notion of host availability. The logic is that mosquitoes *search* for humans. Differences among humans in their attractiveness are represented by a *search weight*. Mosquito search in a place depends on the amount of time spent by humans, but also by daily mosquito activity patterns; from these, we develop a notion of *time at risk* that characterize the way human activities expose them to mosquitoes. The mosquitoes add up all the time at risk spent by all the humans, which gives a measure of their *availability*. Availability describes humans as well as other vertebrate hosts, which are modified by mosquito preferences. The overall feeding rates and the human fraction are computed from availability using *functional responses*.

To complete the picture, we consider how the expected rate of exposure could have a distribution in the population, which we call environmental heterogeneity.

4.5.1 Search and Risk

4.5.2 Search Weights and Availability

To deal with heterogeneous exposure and many other phenomena, we need a sensible way of segmenting humans into population **strata**. Stratification makes it possible to deal with population heterogeneity.

A new model of **blood feeding** is based on a model of blood feeding as the endpoint of a search for a blood host [11].

- Each sub-population has a *search weight* (w), and the total *availability* of humans for blood feeding (W) is the sum of the sizes of the strata weighted by their search weights.

- We also consider the availability of alternative vertebrate species for blood feeding (O).

4.5.3 Functional Response

- Mosquito blood feeding rates are computed using a *functional response* to total availability of vertebrate hosts ($f = F_f(B)$).
- To compute total availability, we add a scaling parameter on alternative hosts, because mosquito preferences can translate into different patterns of search; total availability is $B = W + O^\zeta$.
- The human fraction is proportional to the relative availability of hosts $q = W/B$.

4.6 Environmental Heterogeneity

- The *search weights* thus translate into a kind of **Frailty**, which is one component of *heterogeneous exposure*. Important sources of frailty include bednet use, housing type, and age.
- We also want to consider *variability* in exposure within a stratum – what is the distribution of the *expected* number of bites over time? We have already discussed frailties, so this is a different kind of heterogeneous exposure that we call **Environmental Heterogeneity**. This helps us to align models with data: mosquito counts data tend to be described well by *negative binomial* distributions, so it is likely that the distribution of infectious bites also follows a negative binomial distribution. We introduce a function that translate the EIR into the FoI:

$$h = F_h(E)$$

In the Ross-Macdonald model, the underlying assumption is consistent with a Poisson distribution, but we have also derived *negative binomial hazard rates*. Environmental heterogeneity can arise from two sources:

- the aggregated distributions of mosquitoes in micro-habitats, and the redistribution of mosquito populations by wind and weather;
- random movements of humans around mosquito micro-habitats that affect their risk in a way that doesn't tend to change the mean;

4.7 Mosquito Behavior and Ecology

4.7.1 Resource Availability

4.7.2 Egg Laying

4.7.3 Search and Dispersal

4.8 Aquatic Mosquito Ecology

4.8.1 Regulation

4.8.2 Exogenous Forcing

4.8.3 Habitat Dynamics

4.9 Space

Space is big, so we start by drawing boundaries around a part of the world we want to study, that we call the *spatial domain*.

4.9.1 The Mixing Matrix

4.9.2 Pathogen Dispersal by Humans

4.9.3 Pathogen Dispersal by Mosquitoes

To describe mosquito spatial dynamics, we

4.10 Time

4.10.1 The EIP

4.10.2 Seasonality

4.10.3 Exogenous Forcing

4.10.4 Vector Control

4.11 Integrated Vector Control

4.12 Pharmaceutic Interventions

4.13 Context

Chapter 5

Building & Evaluating Models

There must have been a time, in the beginning, when we could have said – no. But somehow we missed it. (Guildenstern)

– from *Rosencranz and Guildenstern are Dead*, Tom Stoppard

If we want to scale up simulation-based analytics to support malaria policy through eradication, it will inevitably involve capacity building, but what sort of training is needed? Certainly, we hope that an analytics team would include someone who has the skills required to critically evaluate models, but this is not as easy as you think. What makes a model good? What makes an analysis useful? It is easy to imagine training up a generation who could do some kind of analysis, but we don't want to train automata – people who would keep doing some analysis, over and over, just because it is what has been done before. We ought to have a reason for doing the analysis. It ought to serve some purpose. Success in policy analytics requires people who have skills in thinking creatively and critically to solve problems. We need people who can build and evaluate models.

If we want to discuss the models we might use to analyze malaria policies, and if we want to have a discussion about that process, it will probably help us to be more careful about how we use terms like *model* and *model building*. In the service of defining the skill sets required to be a good analyst, this chapter discusses the life-cycle of a model, including the process of building and evaluating models from the beginning to the end (or in math from α to ω).

5.1 Primary Model Development

Most models started with a question: something like *Why is the sky blue?* or *How many mosquitoes does it take to sustain malaria transmission?* Blue sky questions don't go anywhere unless there's a basis for understanding the underlying process. If there is some kind of basis, the next step is to try represent the process in some logically compelling form, such as a diagram with boxes and arrows.

In the case of malaria, the first step looked something like this:

$$\left[\begin{array}{lcl} \# \text{ People Infected Today} & = & \# \text{ People Infected Yesterday} - \\ & & \# \text{ Who Cleared Infections} + \\ & & \# \text{ Uninfected Who Got Infected} \end{array} \right]$$

This description ignores infected people who died, infected people who emigrated, and infected people who immigrated. It's not perfect, but it's a starting point. The biggest problem with this description is that it lacks mosquitoes. So we write down:

$$\left[\begin{array}{lcl} \# \text{ Mosquitoes Infected Today} & = & \# \text{ Mosquitoes Infected Yesterday} - \\ & & \# \text{ that Died} + \\ & & \# \text{ Uninfected Who Got Infected} \end{array} \right]$$

What's important here is that infectious mosquitoes are infecting people when they bite, and infectious people are infecting mosquitoes when the mosquito blood feeds. This process is clarified by a diagram:

5.1.1 Variables

What is critical here is that the uninfected people got infected when they were bitten by an infectious mosquito, and the uninfected mosquitoes who got infected yesterday must have blood fed on an infected person.

To build a model, we need to be quantitative, so we need to start being very precise:

- Let H be the number of people, and we assume that number doesn't change
- Let X_t be the number of people who are infected at time t
- Let M be the number of mosquitoes, and we assume that number doesn't change
- Let Y_t be the number of mosquitoes who are infected at time t

So now we 0

- **Variables:** Defining the state space

- **Parameters:** transitions / rates
- equations / algorithms

5.2 Implementation and Verification

5.2.1 Solving

- **Fitted Parameters:**
- **Initial Conditions:**

5.2.2 Mathematical Analysis

- steady states
- stability analysis
- algorithms
- primary analysis

5.2.3 Simulation

5.3 What is a Model?

5.4 Accuracy & Validation

- Model features
- Functional range
- Heterogeneity

5.4.1 Validation Points

5.5 The Measure of a Model

If we want to evaluate whether a model is good

5.5.1 Limitations

- **Functional Range** the maximum granularity

5.5.2 Parsimony

5.6 Secondary Model Development

- **elaboration**

Model building is a core part of applied dynamics, but if we are to othere are many different activities

What does it mean to build a model?

5.7 Conclusions

Chapter 6

Modularity and Software

Hundreds of publications have described new models of malaria [3,4]. The challenge we have taken on is to find a new way of building models for malaria that draws from all those good ideas to build models at any level of complexity. We want to do this with reusable, professional quality software. Ideally, the models that we develop would be sufficiently complex to address policy questions, yet remain amenable to analysis. To get there, we took a step back to try and understand *malaria models*, and to put this into a birds-eye view of the process of model building.

From Ross's first published model in 1905 to the first draft of this book, 117 years have passed. The story of malaria models can be summarized in three epochs.

Ross's models, and contributions to mathematical study of malaria made by Alfred J Lotka (1912-1923), George Macdonald (1950-1968), and Garrett-Jones (1964-1970) take us to the end first epoch, which is marked by the end of the Global Malaria Eradication Programme (GMEP, 1955-1969). As part of the GMEP, Macdonald's formulas were extended by Garrett-Jones into the concept of *vectorial capacity* and a rudimentary theory of vector control. By 1970, the *Ross-Macdonald* model was more than just a set of equations. It was a theory for malaria dynamics and control supported by a well-developed set of concepts, parameters and metrics [2].

Over that same period of time, mathematical theory for directly transmitted diseases took a parallel path, with important mathematical contributions from Kermack and McKendrick, NTJ Bailey, and Bartlett. Sometime around 1980, mathematical epidemiology began a period of innovation and synthesis, particularly after the publications of Robert May and Roy Anderson made it a mainstream activity in departments of ecology.

In malaria and mosquito-borne diseases, Klaus Dietz publications span the second epoch (1971-2006), including development of a mathematical model with immunity for the Garki Project [59], work on the dynamics of malaria under treatment by drugs [60], seasonality [61], and heterogeneous biting [62,63]. During this time, theory developed for malaria borrowed concepts and methods. In spatial dynamics, the patch models of Yorke and ** were modified to by Dye and Hasibeder to describe mosquito-borne pathogens [64,65].

The last epoch of malaria, which starts around 2006, is marked by two major developments: a maturing theory of malaria control; and the rise of branded, individual-based models.

The publication of *OpenMalaria* in 2006 marks the beginning of the last epoch of malaria. Some important antecedents were Dana Fochs models for *Aedes* dynamics and dengue virus transmission, as *CIMSiM* and *DENSiM*. In malaria, several within-host models had been developed [66,67]. *OpenMalaria* traces its history back to an intrahost model developed by Dietz and Louis Molineaux [67]. After *OpenMalaria*, two other branded individual-based models were developed. One was developed by a team at Imperial College called *Malaria Tools*. Another was developed by a team at the Institute for Disease Modeling called *eMod*. The fact that the models were named and branded was significant – the authors had developed software that they would maintain and that they were willing to stand behind. The models had finally dealt with *disease* in a serious way, and through publications, the fitted models demonstrated a fidelity to evidence. The branding signaled continuity and consistency.

Around 2007, new models of vector control began to appear that related intervention coverage levels to effect sizes. Macdonald's work had focused on sensitivity to parameters, and the GMEP emphasized technical efficiency to achieve very high coverage (with IRS). Garrett-Jones developed vectorial capacity as a way of understanding vector control and effect modification by insecticide resistance. The new models extended Garrett-Jones ideas. The need for new models was motivated, at least in part, by the goal of achieving universal coverage with ITNs. What were reasonable coverage targets? The new generation of vector control models introduced the concept of an effect size on transmission as a function of intervention coverage levels, where coverage had one definition for operations (*e.g.* something like ownership) and another for effect sizes (*e.g.* related to vector contact rates with interventions). The goal of achieving very high coverage with ITNs bumped into the reality that nets are not durable, so new models have been devised to look at intervention coverage in relation to distribution schemes and product durability. While these concepts had been considered during the GMEP design phase, they did not appear in Macdonald's models.

If we want to take advantage of all the research that has been done, we need a way of understanding malaria models and the whole business of model building.

6.1 Model Building

Model building is a fairly involved process that includes several unavoidable steps:

- There must be some motivation for building a model, which usually starts with a conversation, boxes and arrows drawn on paper or a chalkboard or whiteboard. The process involves refining the questions, until there's a well-formed idea – a reason for building a model.
- The idea gets translated into mathematics. The boxes get translated into variables, the arrows are rate parameters, a mathematical formalism is selected.
- The model gets analyzed. In some cases, when the model complexity exceeds a very low threshold on complexity, this is done with pencil and paper. It is only possible to analyze individual parts of the model this way.
- The model gets translated into pseudo-code, and then it gets implemented as software that can produce output. This is followed by a long and painful process of verifying that the software does what the pseudo-code says it *should* do. After awhile, the software is trusted, and it's time to use it.
- Some thought is given to the correspondence between the variables in a model, observable quantities, and the observational process itself. This process can be a part of what happens above, but at some point, the models need to be fitted to data.
- The software produces output and then: the outputs are visualized; models are fitted to data; graphs are made; papers or reports are published; and careers advance.

That's the simple story of model development. What happens next is could be one of the following:

- Someone re-examines an existing model and notices it is inadequate in some way: it is missing some features, or it might make an assumption that ought to be modified. Simple models become spatial models, single populations are structured.
- Someone decides to implement the model in a different way, perhaps with a different mathematical formalism. Continuous time models are translated into discrete time models. Deterministic models become stochastic. Autonomous processes become non-autonomous.

Through this process, hundreds of malaria models were published.

A problem with this process has been that the software is often developed for bespoke tasks (*i.e.* to publish a paper). The software is often poorly documented and difficult to reuse. The costs of building a model for one task limited the complexity of the model. It was difficult to combine elements of one model

developed for one purpose, with someone else's model developed for another purpose.

In malaria, this *ad hoc* process of writing new models was found to be inadequate to serve the broad range of policy questions. One way of dealing with the complexity was to build individual-based models, but individual-based models have some of the same limitations as reality.

6.2 Modular Computation

Before *OpenMalaria*, most models of malaria modified the Ross-Macdonald model in one way [3]. The innovation was focused on specific themes or questions: how long would an infection last in models with superinfection?

6.2.1 exDE

We have written the software that solves these equations in a package called `exDE`.

Chapter 7

Measuring Malaria

If we want to use evidence to make policy, we need to understand transmission and disease in its local context, which we learn about from assembling and analyzing the evidence. In malaria, the evidence is usually one or more of a core set of metrics. One useful metric is the *PfPR* from cross sectional surveys; these are available from research studies, but they are also available from demographic health surveys (DHS) or malaria indicator surveys (MIS). If we are lucky, we will also have mosquito catch counts data (by some method), which we scale by effort and report as something like the *PfHBR*. We estimate the *PfSR* from some of these caught mosquitoes, and we take their product, the *PfEIR*. From other designs, we might have estimates of the *PfFoI*, or if serology was done, we might know the *PfSCR*.

From routine clinical surveillance, we could get case counts or the *PfSPR*. We can learn more if we stratify by age, or if we follow populations over time.

If we are to make sense of malaria as a changing baseline that has been modified by control, then we must know something about how transmission has been modified by control. First and foremost, we must understand anti-malarial drug taking, including the frequency of drug taking (for any purpose), adherence to prescribed regimens, and drug taking by severity. To understand transmission, we must have data describing intervention coverage and its timing. This would include patterns of bed net ownership and use over time, IRS and its coverage. We can't make sense of vector control coverage data without knowing something about the patterns of insecticide resistance. We discuss Measuring Vector Control in the next chapter.

For planning, we would like to know the burden of malaria, and we would like to have an estimate of transmission – something like a malaria reproductive number.

In most places, we will have routine surveillance data. We would like to piece

together a coherent and accurate picture of transmission from this and from any other data we may have. To do so, we would like to know how these statistics compare in settings where two or more of them have been measured at the same place and at the same time. From studies like this, we can start to understand the **scaling relationships** among metrics, and the factors that influence them. What we have learned from studying these relationships is that there the relationship are highly non-linear [68], and that some of the factors affecting this relationship are context dependent. We have already mentioned drug taking, and we must acknowledge the difficulty of estimating drug taking by age, in part, because it is often highly heterogeneous in populations. Other factors that have been implicated are heterogeneous biting, environmental heterogeneity, seasonality, travel and mobility.

Sitting between us and the truth is an observational process, and we must attempt to quantify bias and error in this sampling process if we are to have any hope of getting an accurate picture of transmission in context.

To get there, we need to organize our algorithms:

- The $PfPR$ varies in a population by age, sex, and the diagnostic method; so we simulate malaria in cohorts as they age, and we output the predicted $PfPR$ by age, sex, and diagnostic method.
- The $PfPR$ is modified by drug-taking; the effect modification varies in a population by age, access to care, and adherence to drug regimens. In our algorithms, we must consider various patterns of drug taking and chemoprotection.
- Exposure is *seasonal*
 - seasonality *would* affect the true $PfPR$ and the true $PfEIR$ over time;
 - seasonality *could* be a source of error in the estimated $PfPR$ (especially with respect to age) and estimated $PfEIR$, particularly if the $PfPR$ is reported in young children, and mismatched even slightly with $PfEIR$.
- Heterogeneous exposure:
 - exposure varies by age;
 - there are other identifiable sources of frailty that we can deal with through stratification (*e.g.* house type, location, bed net use), and there are many frailties we will not be able to identify;
 - the relationship between exposure and infection is affected by *environmental heterogeneity*; because of variability in mosquito populations at a fine spatial grain over time and space, and the trajectories of individual humans who are put at risk,
- The $PfEIR$ near home is obscured by human mobility through time at risk and exposure away from home;

- Travel and imported malaria can increase the *PfPR*, even if there is no local exposure.
- Measurement errors in the *PfEIR*
 - There are *house effects* that bias the estimation of the *PfEIR*
 - Different methods of catching mosquitoes give different estimates for the HBR.

These effects are incorporated into the framework. Frailties are dealt with by stratifying by age ($\omega(a)$) and other traits and assigning *biting weights*; when combined with mobility, these become part of the mixing matrix, β . Environmental heterogeneity modifies the relationship between the *PfEIR* by stratum and the *PfFoI* (through F_h). Seasonality in mosquito density ($Z(t)$) translates into seasonality in exposure ($S(t)$). Drug taking and immunity modify the *PfPR* after exposure.

We would like to understand mosquito populations, but we will rely on whatever data we have to *estimate* the *PfEIR* (see Figure 4.1). This is a core algorithm for simulation-based analytics, and the framework was designed to serve these ends.

In the following sections, we use the framework to illustrate the effects and effect sizes of these factors. The task of estimating is for another book, **Robust Analytics for Malaria Policy**. In this study, we make full use of modularity. We start by *stratifying* our human population by age, and possibly by other factors. We write algorithms that pass the density of blood feeding mosquitoes in a place, and our algorithms predict mosquito catch counts, depending on the method. From this, we compute the total FoI, which includes the *local* FoI (at and around home), and the *travel* FoI. These exposure rates, and algorithms describing drug taking, are translated into estimates of the true *PfPR*. The models output the predicted *PfPR* by age, sex, and method; and we also output the predicted patterns of care seeking and associated metrics at health facilities in an HMIS, by age and sex.

7.1 Realistic Bounds

An important reality check is that we can set sensible expectations about upper bounds on the *PfPR* in relation to exposure. The maximum value of the *PfPR* will tend to be observed in untreated populations with homogeneous exposure. To identify these bounds, we assume that we are passing the *true PfEIR*, but not necessarily the *local PfEIR*.

7.2 The Local Fraction

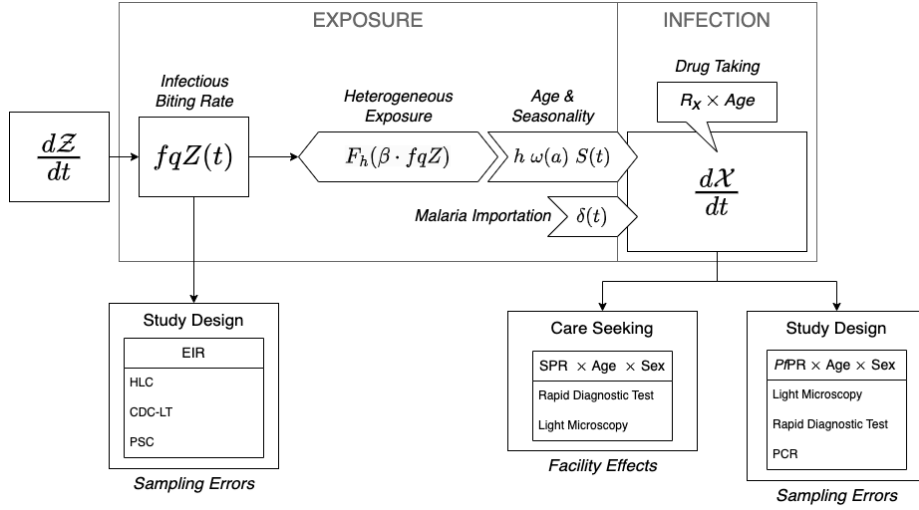


Figure 7.1: Figure 4.1: A generalized framework for understanding the relationship between the *PfEIR* and the *PfPR*:

7.2.1 Travel

7.2.2 Mobility

7.3 Drug Taking

7.4 Seasonality

7.5 Frailty

At low intensity, frailty can affect the prevalence

7.6 Environmental Heterogeneity

7.7 *PfEIR* vs. *PfPR* in Data

7.8 *PfEIR* vs. *PfFoI* in Data

7.9 Synthesis

Chapter 8

Measuring Vector Control

In the previous chapter, we consider adult mosquito population density as a process closely related to, but also exogenous to the process of parasite transmission.

8.1 Coverage

8.2 Effect Sizes

Part I

Transmission

Chapter 9

Heterogeneous Exposure

For humans, exposure to malaria means exposure to the bites of infectious mosquitoes. A problem that we'll have to deal with sooner or later is that exposure risk differs among humans over space and time. While this might seem like an odd thing to introduce so early, we will have to tackle the topic sometime. The discussion of [Heterogeneous Biting], in the previous chapter, showed that heterogeneity plays an important in understanding transmission and thresholds. This discussion of heterogeneous exposure (*i.e.*, looking at heterogeneous biting from the human side) is a good way of introducing some of the core concepts that are built into the framework:

- [Heterogeneous Biting] is one way of getting around a conundrum. In models with homogeneous biting, the relationship between *average* mosquito density and the prevalence of infection would lead us to make quantitative predictions about the likely effects of vector control.
- We discuss two different kinds of heterogeneous exposure: frailty, and environmental heterogeneity. In a nutshell, frailty multiplies the mean hazard rate for a sub-population by some amount k . Environmental heterogeneity does not affect the mean, but it changes the distribution of the mean.
- We introduce the idea that we can deal with frailties in human populations by segmenting the population into strata.
- We set the stage for a new model of mosquito **blood feeding** that we introduce in the next chapter.
- In a chapter on [Approximation], we use these models to discuss the problem of model-based inference.

9.1 Overview

Some reasons heterogeneous exposure to malaria have been documented in hundreds of studies. This is an overview.

9.1.1 Age

- Port, Boreham
- Carnevale

9.1.2 Location

9.1.3 House Type

9.1.4 Activities

9.2 Frailty

In general, we define frailty as a multiplicative factor on the FoI. If the average FoI in the population is h , then the FoI in a stratum is hk . The size of the stratum, p_k , is constrained such that:

$$\int_0^\infty k p_k dk = 1$$

With this constraint, the mean FoI in the population is h .

Continuous distributions are difficult to extend, but we can stratify a population to accomplish some of the same effects.

9.3 Environmental Heterogeneity

Chapter 10

Blood Feeding

The endpoint

10.1 Host Availability

10.2 Blood Feeding Rates

10.3 The Human Fraction

10.4 The Mixing Matrix, β

Chapter 11

Spatial Dynamics

Chapter 12

Temporal Dynamics

12.1 Exogenous Forcing

12.2 Mosquito Survival through the EIP

Chapter 13

Cohort Dynamics

We need a way of incorporating age into our models.

13.1 Boxcar Models

13.2 Delay

Chapter 14

Demography

14.1 Migration

Chapter 15

Stratification

Part II

Human Ecology

Chapter 16

Human Behavior

Chapter 17

Human Mobility

Chapter 18

Human Travel and Malaria Importation

Part III

Epidemiology

Chapter 19

Malaria Infection and Immunity

In the following sections, we walk through several models for the dynamics of malaria infection and immunity in humans. We cover infection and detection, immunity, infectiousness, disease, drug taking, and cohort dynamics.

19.1 Infection

19.1.1 Duration of a Simple Infection

19.1.2 Superinfection

19.1.3 Age of Infection (AoI)

19.1.4 Parasite Densities

19.1.5 Stage of Infection (SoI)

19.2 Immunity

19.3 Within Host Models

Chapter 20

Age, Exposure and Malaria Immunity

Immunity to malaria has been described as being *weak* and immunity has a *poor memory*. In quantitative terms, this means that almost every aspect of infection is affected by the history of exposure, and exposure accumulates with age.

In developing our models so far, we have ignored age as a factor in malaria epidemiology. In analyzing the models, we looked for steady states or stable temporal orbits. These analyses give a misleading impression of malaria.

The first model to examine malaria immunity was developed for the Garki Project [59]. We will present the model, for historical reasons, and then we will it is much easier to formulate models for malaria infection dynamics in cohorts as they age.

20.1 The Garki Model

20.2 Stage-Structured Immunity

20.3 Strain Specific Immunity

20.4 Memory Tracking

20.5 Age *vs.* Prevalence

Chapter 21

Detecting Parasites

21.1 Parasite Densities and Detection

21.2 Light Microscopy

21.3 Biomarkers and RDTs

21.4 PCR

Chapter 22

Gametocytes and Infectiousness

22.1 Gametocytemia

22.2 Anti-Gametocyte Immunity

Chapter 23

Fever and Severe Disease

23.1 Fever

23.2 Anemia

23.3 Severe Disease

Chapter 24

Care Seeking

Chapter 25

Drug Taking

25.1 Curing Infections

25.2 Chemoprotection

25.3 Adherence

25.4 Treatment Rates

Chapter 26

Pharmaceutical Interventions

26.1 SMC

26.2 MDA

26.3 Drugs

26.4 Vaccines

Chapter 27

Malaria Epidemiology

27.1 Age of the Youngest Infection

Part IV

Mosquito Ecology

Chapter 28

Mosquitoes

In a previous chapter, we consider adult mosquito population density as a process closely related to, but also exogenous to the process of parasite transmission. We formulated the models for mosquito density in terms of an emergence rate, $\Lambda(t)$. In many studies, this might be good enough. There are some challenges in vector control, however, that call for a deeper understanding of adult mosquito population dynamics in relation to the population dynamics of immature mosquito populations. It's hard to imagine giving any advice about LSM without a model of mosquito ecology. Adult vector control will reduce mosquito population densities, but it would also reduce egg laying affecting immature mosquito populations [54]. Mosquito populations have their own thresholds for persistence, and their own spatial dynamics.

While it could be argued that these are not the primary concerns for malaria programs – we would not disagree – our mathematical framework must be strong enough to support analytics for integrated vector control, including LSM. Immature mosquito dynamics are one of the core dynamical components in our framework. We introduce the topic here, using a very basic model for mosquito ecology [69].

To extend the previous model, we need to make several changes and additional assumptions:

- We define terms that describe egg-laying by adult mosquitoes;
- We write a basic equation to describe the dynamics of immature mosquitoes in aquatic habitats as a function of egg-laying. These models describe how eggs hatch, how mosquito larvae develop and mature in aquatic habitats, pupate, and then emerge as adults. The models do not need to be complex, but they could be.
- We replace the parameter $\Lambda(t)$ from the Aron-May model with a term that describes emergence of adults from aquatic habitats.

One of the big themes we want to introduce here is that the outcomes of vector control can vary substantially by context because of differences in mosquito ecology.

28.1 Aquatic Dynamics

28.1.1 Equations

Here, we start with the equation for adult ecology from before, leaving out parasite transmission dynamics. We modify the basic slightly by adding one extra term that describes delayed maturation, a different response to crowding [69]. The mosquito maturation rate, $\psi e^{-\sigma L}$ defines the average time from egg to emergence, and we add one parameter that slows down larval development in response to crowding. The per-capita mortality rate, $\psi + \theta L$, has two parts: a density independent term ψ and a response to mean crowding θL . We assume that adults lay χ eggs per feeding cycle.

$$\begin{aligned} \frac{dM}{dt} &= \Lambda - gM \\ \frac{dL}{dt} &= \eta - (\psi e^{-\sigma L} + \phi + \theta L)L \\ \Lambda &= \frac{1}{2} \psi L e^{-\sigma L} \\ \eta &= \chi f M \end{aligned} \tag{28.1}$$

28.1.2 Solutions

```
BasicAquatic_dML = function(t, y, params){with(c(params, as.list(y)),{  
  
  # Terms  
  Lambda = psi*exp(-sigma*L)*L/2  
  eta = chi*f*M  
  
  # Dynamics  
  dM = Lambda - g*M  
  dL = eta - (psi*exp(-sigma*L) + phi + theta*L)*L  
  
  return(list(c(dM, dL)))  
})}  
  
makePar_BasicAquatic = function(f=1/2.5, g=1/10, chi = 50, psi = 1/10, phi = 1/10, theta=1/10, sigma=1/10){  
  list(f=f, g=g, chi=chi, psi=psi, phi=phi, theta=theta, sigma=sigma)  
}
```

```

tt = seq(0,100, by=1)
y0 = c(M=1, L=1)
params = makePar_BasicAquatic()
params1 = makePar_BasicAquatic(phi=1)
params2 = makePar_BasicAquatic(theta=1/10)
params3 = makePar_BasicAquatic(sigma=1/100, theta=0)

```

This code solves the equations:

```

require(deSolve)
yout <- lsode(y=y0, times=tt, func=BasicAquatic_dML, parms=params)
yout1 <- lsode(y=y0, times=tt, func=BasicAquatic_dML, parms=params1)
yout2 <- lsode(y=y0, times=tt, func=BasicAquatic_dML, parms=params2)
yout3 <- lsode(y=y0, times=tt, func=BasicAquatic_dML, parms=params3)

plotML = function(out, llwd=1, clrL = "blue", clrM="red"){with(data.frame(out),{
  plot(time, L, type = "l", lwd=llwd, col = clrL, xlab = "Time (Days)", ylab = "Density")
  lines(time, M, col = clrM, lwd=llwd)
}})

addML = function(out, llty = 2, llwd=1, clrL = "blue", clrM="red"){with(data.frame(out),{
  lines(time, L, col = clrL, lty=llty, lwd=llwd)
  lines(time, M, col = clrM, lty=llty, lwd=llwd)
}})

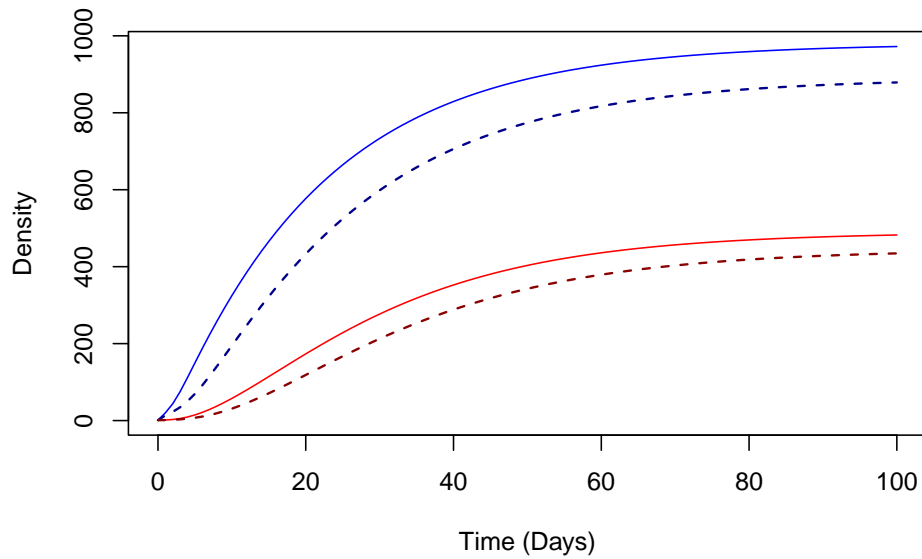
```

The first thing to point out is that changing the dynamics take some time to approach equilibrium – a couple of months, in this model. The second is that increasing mortality rates for mosquitoes in aquatic habitats by a factor of 10 affects the mosquito densities in minor ways.

```

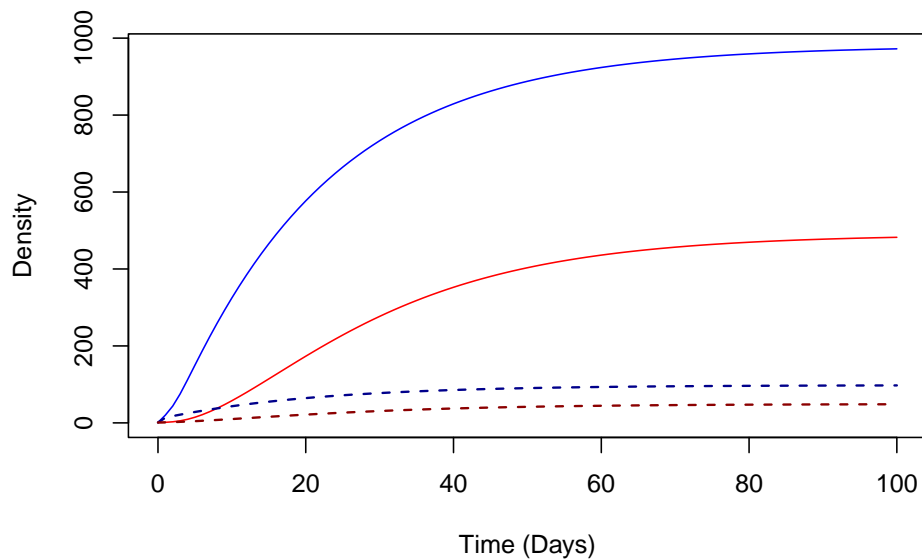
plotML(yout)
addML(yout1, clrL = "darkblue", clrM="darkred", llwd=1.5)

```



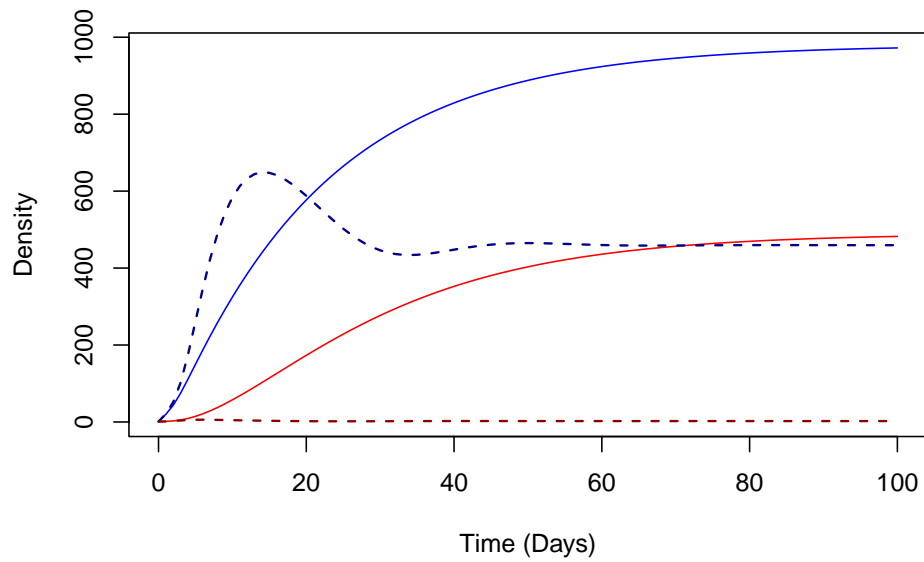
The third is that increasing the response to mean crowding has a very strong effect on the population dynamics.

```
plotML(yout)
addML(yout2, clrL = "darkblue", clrM="darkred", llwd=1.5)
```

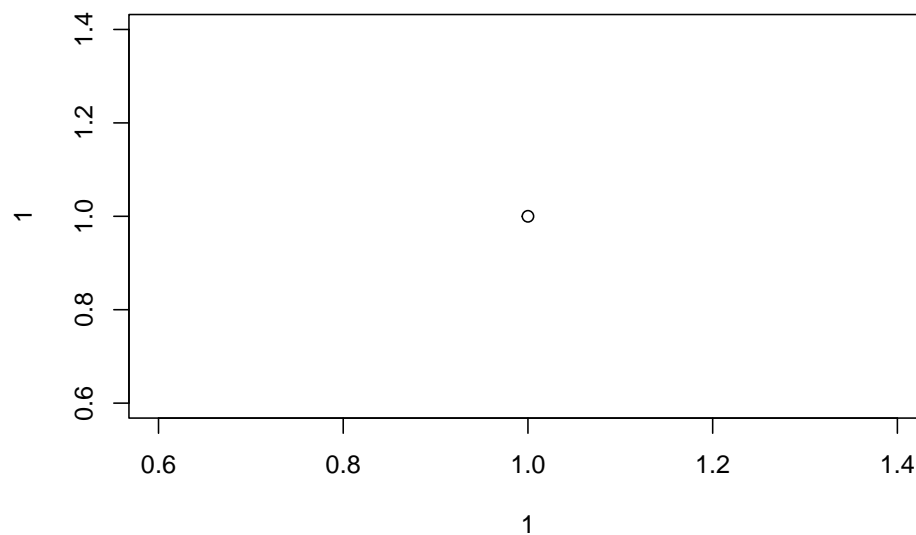


The third point is that density dependent responses to mean crowding cause much larger changes to this system than density-dependent mortality. Aquatic dynamics have a strong tendency to oscillate, but more critically, delayed maturation strongly affects the ratio of adult mosquitoes to adults.

```
plotML(yout)
addML(yout3, clrL = "darkblue", clrM="darkred", llwd=1.5)
```



```
plot(1,1)
```



28.2 Understanding Mosquito Dynamics

This

Chapter 29

Behavioral State Models

Chapter 30

Aquatic Ecology

Chapter 31

Mosquito Microecology

Modeling mosquito population dynamics on point sets.

Chapter 32

Mosquito Dispersal

Chapter 33

Microsimulation

Chapter 34

Mosquito Ecology

34.1 Population Dynamics

Chapter 35

Vector Competence

Chapter 36

Measuring Mosquitoes

Part V

Malaria Control

Chapter 37

Vector Control

Towards a theory of vector control.

Chapter 38

Insecticide-Treated Bednets

Chapter 39

Indoor Residual Spraying

Chapter 40

Larval Source Management

Chapter 41

Attractive Toxic Sugar Baits

Chapter 42

Integrated Vector Control

Chapter 43

Spatial Control

Chapter 44

Pharmaceutical Interventions

44.1 SMC

44.2 MDA

44.3 Drugs

44.4 Vaccines

Chapter 45

Discrete Time

Chapter 46

Stochasticity

Chapter 47

Base Models

Chapter 48

Built-in Analytics

Chapter 49

Spatial Concepts and Connectivity

Chapter 50

Model Libraries

Part VI

Supplements

Chapter 51

References

If you want a PDF and can't find it at the link provided, let us know and we can help you find a copy.

1. Ross R. The Prevention of Malaria. 2nd ed. London: John Murray; 1911.
2. Smith DL, Battle KE, Hay SI, Barker CM, Scott TW, McKenzie FE. Ross, Macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. PLoS Pathog. 2012;8: e1002588. doi:10.1371/journal.ppat.1002588
3. Reiner RC Jr, Perkins TA, Barker CM, Niu T, Chaves LF, Ellis AM, et al. A systematic review of mathematical models of mosquito-borne pathogen transmission: 1970-2010. J R Soc Interface. 2013;10: 20120921.
4. Smith NR, Trauer JM, Gambhir M, Richards JS, Maude RJ, Keith JM, et al. Agent-based models of malaria transmission: A systematic review. Malar J. 2018;17: 299. doi:10.1186/s12936-018-2442-y
5. Smith DL, Perkins TA, Reiner RC Jr, Barker CM, Niu T, Chaves LF, et al. Recasting the theory of mosquito-borne pathogen transmission dynamics and control. Trans R Soc Trop Med Hyg. 2014;108: 185–197. doi:10.1093/trstmh/tru026
6. Carter R. Spatial simulation of malaria transmission and its control by malaria transmission blocking vaccination. International Journal for Parasitology. 2002;32: 1617–1624. doi:10.1016/S0020-7519(02)00190-X
7. Gu W, Killeen GF, Mbogo CM, Regens JL, Githure JI, Beier JC. An individual-based model of *Plasmodium falciparum* malaria transmission on the coast of Kenya. Trans R Soc Trop Med Hyg. 2003;97: 43–50. doi:10.1016/s0035-9203(03)90018-6
8. Perkins TA, Scott TW, Le Menach A, Smith DL. Heterogeneity, mixing, and the spatial scales of mosquito-borne pathogen transmission. PLoS Comput Biol. 2013;9: e1003327.

9. Wu SL, Sánchez C HM, Henry JM, Citron DT, Zhang Q, Compton K, et al. Vector bionomics and vectorial capacity as emergent properties of mosquito behaviors and ecology. *PLoS Comput Biol*. 2020;16: e1007446. doi:10.1371/journal.pcbi.1007446
10. Tatem AJ, Smith DL. International population movements and regional *Plasmodium falciparum* malaria elimination strategies. *Proc Natl Acad Sci U S A*. 2010;107: 12222–12227. doi:10.1073/pnas.1002971107
11. Wu SL, Henry JM, Citron DT, Ssebuliba DM, Nsumba JN, C HMS, et al. Spatial dynamics of malaria transmission. *PLOS Computational Biology*. 2023;19: e1010684. doi:10.1371/journal.pcbi.1010684
12. Ross R. On some peculiar pigmented cells found in two mosquitos fed on malarial blood. *British medical journal*. 1897;2: 1786–1788. Available: <http://www.bmj.com/content/2/1929/1786.full.pdf>
13. Ross R. The role of the mosquito in the evolution of the malarial parasite. *Lancet*. 1898;152: 488–490. Available: https://scholar.archive.org/work/vejsrzltgva5nmudubd5oebye4/access/ia_file/crossref-pre-1909-scholarly-works/10.1016%252Fs0140-6736%252801%252981394-5.zip/10.1016%252Fs0140-6736%252801%252981400-8.pdf
14. Grassi B, Bastianelli G, Bignami A. Ulteriori ricerche sul ciclo dei parassiti malarici umani sul corpo del zanzarone. *Rend Acad Lincei*. 1898.
15. Grassi B. Studi di uno zoologo sulla malaria. Roma : R. Accademia dei lincei,; 1901. doi:10.5962/bhl.title.37999
16. Laveran A. Note sur un nouveau parasite dans le sang de plusieurs malades atteints de fièvre palustre. Note communiquée a l'Académie de Medicine, seance du. 1880;23.
17. Laveran A. Traité des fièvres palustres. Octave Doin; 1884.
18. Garnham PCC. History of Discoveries of Malaria Parasites and of Their Life Cycles. *History and Philosophy of the Life Sciences*. 1988;10: 93–108. Available: <https://www.jstor.org/stable/23329001>
19. Shortt HE, Garnham PCC. The pre-erythrocytic stage of human malaria, *Plasmodium vivax*. *Br Med J*. 1948;1: 547. doi:10.1136/bmj.1.4550.547
20. Bruce-Chwatt LJ. History of malaria from prehistory to eradication. In: Wernsdorfer WH, McGregor IA, editors. *Malaria: Principles and Practice of Malariology*. New York: Churchill Livingstone; 1988. pp. 1–59. Available: <http://www.cabdirect.org/abstracts/19900861468.html>
21. Service MW. A short history of early medical entomology. *Journal of Medical Entomology*. 1978;14: 603–626. Available: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=357723&retmode=ref&cmd=prlinks>

22. Cox FEG. History of the discovery of the malaria parasites and their vectors. *Parasit Vectors*. 2010;3: 5. doi:10.1186/1756-3305-3-5
23. Ross R. Extermination of Malaria. *Ind Med Gaz*. 1899;34: 231–232. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5145409/>
24. Ross R. Inaugural Lecture on the Possibility of Extirpating Malaria from Certain Localities by a New Method. *British medical journal*. 1899;2: 1–4. Available: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=20758555&retmode=ref&cmd=prlinks>
25. Bockarie MJ, Gbakima AA, Barnish G. It all began with Ronald Ross: 100 years of malaria research and control in Sierra Leone (1899–1999). *Ann Trop Med Parasitol*. 1999;93: 213–224. doi:10.1080/00034983.1999.11813416
26. Ross R. The logical basis of the sanitary policy of mosquito reduction. *Science*. 1905;22: 689–699. doi:10.1126/science.22.570.689
27. Ross R. Report on the Prevention of Malaria in Mauritius. London: Waterlow; 1908.
28. Ross R. Some quantitative studies in epidemiology. *Nature*. 1911;87: 466–467.
29. Ross R. Some *a priori* pathometric equations. *Br Med J*. 1915;i: 546–547. doi:10.1136/bmj.1.2830.546
30. Ross R. An application of the theory of probabilities to the study of *a priori* pathometry. Part I. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science. 1916;92: 204–230.
31. Ross R, Hudson HP. An application of the theory of probabilities to the study of *a priori* pathometry. Part II. Proceedings of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences. 1917;93: 212–225.
32. Ross R, Hudson H. An application of the theory of probabilities to the study of *a priori* pathometry. Part III. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science. 1917;93: 225–240. Available: <http://adsabs.harvard.edu/abs/1917RSPSA..93..225R>
33. Lotka AJ. Quantitative Studies in Epidemiology. *Nature*. 1912;88: 497–498. doi:10.1038/088497b0
34. Fine PEM. Superinfection - a problem in formulating a problem. *Tropical Diseases Bulletin*. 1975;75: 475–488.
35. Ross R. Mosquito brigades and how to organise them. London: George Philip & Son; 1902.
36. Ross R. The prevention of malaria in British possessions, Egypt, and parts of America. *The Lancet*. 1907; 879–887.
37. Ross R. Malaria in Greece. Washington, DC: Government Printing Office; 1909 pp. 697–710.

38. Waite H. Mosquitoes and malaria. A study of the relation between the number of mosquitoes in a locality and the malaria rate. *Biometrika*. 1910;7: 421–436. doi:10.2307/2345376
39. Lotka AJ. Contribution to the Analysis of Malaria Epidemiology. I. General Part. *American Journal of Epidemiology*. 1923;3: 1–36. doi:10.1093/oxfordjournals.aje.a118963
40. Lotka AJ. Contribution to the Analysis of Malaria Epidemiology. II. General Part (continued). Comparison of Two Formulae given by Sir Ronald Ross. *American Journal of Epidemiology*. 1923;3: 38–54. doi:10.1093/oxfordjournals.aje.a118965
41. Lotka AJ. Contribution to the Analysis of Malaria Epidemiology. III. Numerical Part. *American Journal of Epidemiology*. 1923;3: 55–95. doi:10.1093/oxfordjournals.aje.a118966
42. Lotka AJ. Contribution to the Analysis of Malaria Epidemiology. V. Summary. *American Journal of Epidemiology*. 1923;3: 113–121. doi:10.1093/oxfordjournals.aje.a118964
43. Sharpe FR, Lotka AJ. Contribution to the Analysis of Malaria Epidemiology. IV. Incubation lag. *American Journal of Epidemiology*. 1923;3: 96–112. doi:10.1093/oxfordjournals.aje.a118967
44. Macdonald G. The epidemiology and control of malaria. Oxford university press; 1957. Available: <https://www.cabdirect.org/cabdirect/abstract/19582900392>
45. Armitage P. A note on the epidemiology of malaria. *Trop Dis Bull*. 1953;50: 890–892.
46. Macdonald G. The analysis of malaria parasite rates in infants. *Tropical diseases bulletin*. 1950;47: 915–938.
47. Macdonald G. The analysis of infection rates in diseases in which superinfection occurs. *Trop Dis Bull*. 1950;47: 907–915. Available: <https://www.ncbi.nlm.nih.gov/pubmed/14798656>
48. Macdonald G. The analysis of the sporozoite rate. *Trop Dis Bull*. 1952;49: 569–586. Available: <https://www.ncbi.nlm.nih.gov/pubmed/14958825>
49. Macdonald G. The analysis of equilibrium in malaria. *Trop Dis Bull*. 1952;49: 813–829. Available: <https://www.ncbi.nlm.nih.gov/pubmed/12995455>
50. Robinson A. Did Einstein really say that? *Nature*. 2018;557: 30–30. doi:10.1038/d41586-018-05004-4
51. Aron JL, May RM. The population dynamics of malaria. In: Anderson RM, editor. *The Population Dynamics of Infectious Diseases: Theory and Applications*. Boston, MA: Springer US; 1982. pp. 139–179. doi:10.1007/978-1-4899-2901-3_5

52. Smith DL, McKenzie FE. Statics and dynamics of malaria infection in *Anopheles* mosquitoes. *Malaria Journal*. 2004;3: 13. doi:10.1186/1475-2875-3-13
53. Smith DL, Musiime AK, Maxwell K, Lindsay SW, Kiware S. A New Test of a Theory about Old Mosquitoes. *Trends Parasitol*. 2021;37: 185–194. doi:10.1016/j.pt.2020.10.011
54. Brady OJ, Godfray HCJ, Tatem AJ, Gething PW, Cohen JM, McKenzie FE, et al. Adult vector control, mosquito ecology and malaria transmission. *International Health*. 2015;7: 121–129. doi:10.1093/inthealth/ihv010
55. Macdonald G, Goeckel GW. The malaria parasite rate and interruption of transmission. *Bull World Health Organ*. 1964;31: 365–377. Available: <https://www.ncbi.nlm.nih.gov/pubmed/14267746>
56. Smith DL, Hay SI. Endemicity response timelines for *Plasmodium falciparum* elimination. *Malaria Journal*. 2009;8: 87. doi:10.1186/1475-2875-8-87
57. Walton GA. On the control of malaria in Freetown, Sierra Leone; I. *Plasmodium falciparum* and *Anopheles gambiae* in relation to malaria occurring in infants. *Ann Trop Med Parasitol*. 1947;41: 380–407. doi:10.1080/00034983.1947.11685341
58. Henry JM. A hybrid model for the effects of treatment and demography on malaria superinfection. *Journal of Theoretical Biology*. 2020;491: 110194. doi:10.1016/j.jtbi.2020.110194
59. Dietz K, Molineaux L, Thomas A. A malaria model tested in the African savannah. *Bull World Health Organ*. 1974;50: 347–357. Available: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=4613512
60. Dietz K. Models for parasitic disease control. *Bull Int Stat Inst*. 1975;46: 531–544.
61. Dietz K. The incidence of infectious diseases under the influence of seasonal fluctuations. In: Berger J, Bühler WJ, Repges R, Tautu P, editors. *Mathematical models in medicine*. Berlin, Heidelberg: Springer Berlin Heidelberg; 1976. pp. 1–15.
62. Dietz K. Models for vector-borne parasitic diseases. In: Barigozzi C, Levin SA, editors. *Vito Volterra Symposium on Mathematical Models in biology*. Berlin: Springer-Verlag; 1980. pp. 264–277.
63. Dietz K, Haderer KP. Epidemiological models for sexually transmitted diseases. *J Math Biol*. 1988;26: 1–25. Available: <https://www.ncbi.nlm.nih.gov/pubmed/3351391>
64. Dye C, Hasibeder G. Population dynamics of mosquito-borne disease: Effects of flies which bite some people more frequently than others. *Trans R Soc Trop Med Hyg*. 1986;80: 69–77. doi:10.1016/0035-9203(86)90199-9

- 65. Hasibeder G, Dye C. Population dynamics of mosquito-borne disease: Persistence in a completely heterogeneous environment. *Theor Popul Biol.* 1988;33: 31–53. doi:10.1016/0040-5809(88)90003-2
- 66. Molineaux L, Dietz K. Review of intra-host models of malaria. *Parassitologia.* 1999;41: 221–231. Available: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=10697860&retmode=ref&cmd=prlinks>
- 67. Molineaux L, Diebner HH, Eichner M, Collins WE, Jeffery GM, Dietz K. *Plasmodium falciparum* parasitaemia described by a new mathematical model. *Parasitology.* 2001;122: 379–391. doi:10.1017/s0031182001007533
- 68. Hay SI, Smith DL, Snow RW. Measuring malaria endemicity from intense to interrupted transmission. *Lancet Infect Dis.* 2008;8: 369–378. doi:10.1016/S1473-3099(08)70069-0
- 69. Smith DL, Perkins TA, Tusting LS, Scott TW, Lindsay SW. Mosquito Population Regulation and Larval Source Management in Heterogeneous Environments. *PLOS ONE.* 2013;8: e71247. doi:10.1371/journal.pone.0071247