Integration of parasite genetic information in malaria transmission modelling

Oliver John Watson

Thesis submitted for PhD examination

Department of Infectious Disease Epidemiology
Imperial College London

2019

Abstract

Mathematical models of malaria transmission are increasingly used to quantify the impact of malaria control efforts and to assist in the development and costing of future initiatives such as the WHO Global Technical Strategy for Malaria 2016-2030. These models have highlighted both the progress made so far, but also how continued investment is needed to reach the milestones required. However, the increase in global malaria cases reported in 2018 suggests that new tools may be required to continue the gains made and to address the growing risk of antimalarial resistance threatening to reverse the recent declines in malaria burden. The proliferation of genetic sequencing and the publication of the *Plasmodium falciparum* reference genome in 2002 has facilitated a greater understanding of the genetic determinants of resistance and molecular tools are subsequently poised to become a routine tool for malaria control. Consequently, integrating parasite genetic information into established models of malaria transmission models can contribute to both our understanding of the drivers and optimum policies for addressing resistance and detailing the potential of molecular tools within malaria control.

Plasmodium falciparum is known to have evolved several times in response to first line antimalarials. However, recent evidence has shown evolution to rapid diagnostic tests. The WHO has consequently issued guidance advising national malaria control programmes to conduct surveillance for *pfhrp2/3* deletions. The timing of this policy recommendation and my previous work modelling *pfhrp2* deletions necessitated a timely extension of our previous model to evaluate the implications of seasonality in malaria transmission on estimates of the prevalence of *pfhrp2/3* deletions.

Recent studies have suggested that malaria genotyping could be a useful tool for epidemiological surveillance. By developing an extended version of an established model of malaria transmission, which now models individual mosquitoes affording the full parasite life cycle to be represented, I characterise the potential utility of malaria genomics for inferring changes in transmission intensity. I conclude that although molecular tools could enable accurate estimation of malaria prevalence, greater attention needs to be placed on the chosen sampling scheme, recording patient metadata and developing the statistical toolkit for analysing polyclonal infected individuals.

In 2015, health ministers in the Greater Mekong Subregion (GMS) adopted the WHO strategy for malaria elimination in the GMS 2016-2030. The strategy was developed to accelerate elimination in South-East Asia, which is currently the best approach to address the growing threat of artemisinin resistance and the emergence of multidrug resistant parasite lineages. In response, I demonstrate how the therapeutic lifespan of the five currently recommended artemisinin combination therapies can be prolonged by reducing antimalarial overprescription by ensuring that all suspected malaria fevers are tested before administering antimalarials. I conclude by comparing different cycling and mixing strategies before reviewing how each strategy can be improved to slow the spread of antimalarial resistance.

Elimination in the GMS is undoubtedly an effective mechanism for preventing the spread of artemisinin resistance to Africa. However, if efforts to eliminate by 2030 have failed it will be imperative to understand the mechanisms with which resistance may continue to spread. To this extent, the capability of resistant strains to invade susceptible populations is evaluated using data from standard membrane feeding assays. Findings are incorporated in the transmission model to quantify the transmission advantage of artemisinin resistance at the population level.

Copyright

Declaration of Originality

I declare that this work is my own, completed under the supervision of Prof Azra Ghani, Dr Lucy Okell and Dr Robert Verity. Much of the work involved in this thesis has been of a collaborative nature. All individuals who have shared data, made suggestions for analysis, and commented on manuscripts have been acknowledged. All sources used in the thesis have been cited appropriately.

- In chapter 6, to estimate the impact of artemisinin on the development of oocysts in experimental feeds, experimental data conducted by Kathrin Witmer, Farah A. Dahalan and Jake Baum was shared with me. The laboratory methods described in this chapter were conducted by the Baum research laboratory, which is clarified again in chapter 6.

- I wrote all the code used within this thesis. However, the code for solving the equilibrium solution for the deterministic implementation of the Imperial College malaria transmission model and implementing vector based interventions in the deterministic model was written alongside Dr Joel Hellewell and Dr Juliette Unwin.

The work in chapter 2 has been published as the following paper (**Watson OJ, Verity R, Ghani AC, Garske T, Cunningham J, Tshefu, A, Mwandagalirwa, MK, Meshnick SR, Parr JB, Slater HC.** Impact of seasonal variations in Plasmodium falciparum malaria transmission on the surveillance of pfhrp2 gene deletions. eLife 2019. *8*, e40339). I conducted the analysis and wrote the first draft:

Oliver Watson September 2019

Contents

List of Figures

Appendix Figures

List of Tables

List of Abbreviations

ACT	Artemisinin combination therapy
AL	Artemether-lumefantrine
Aldo	Aldolase
ANC	Ante-natal care
ASAQ	Artesunate-amodiaquine
COI	Complexity of infection
COU	Coefficient of uniqueness
CRT	Chloroquine resistance transporter
CQ	Chloroquine
DHA-PPQ	Dihydroartemisinin-piperaquine
DHS	Demographic and Health Surveys
DRC	Democratic Republic of Congo
EIR	Entomological innoculation rate
GMS	Greater Mekong Subregion
GMEP	Global Malaria Eradication Programme
HRP2	Histidine-rich protein 2
IBD	Identity by descent
IBS	Identity by state
IRS	Indoor residual spraying
IPTi	Intermittent preventive treatment in infants
IPTp	Intermittent preventive treatment in pregnancy
ITN	Insecticide treated net
K13	Kelch 13
$K13^{mut}$	Kelch 13 mutations
LPF	Late parasitological failure
LDH	Lactate dehydrogenase
LLIN	Long lasting insecticidal net
LOD	Limits of detection
MAE	Mean absolute error
MFT	Multiple first line therapies
MalariaGEN	Malaria Genomic Epidemiology Network
MDA	Mass drug administration
MDR1/2	Multidrug resistance protein 1/2

MQ	Mefloquine
MPAC	Malaria Policy Advisory Committee
MSAT	Mass screen and treat
NMF	Non-malarial fever
OUCRU	Oxford University Clinical Research Unit
PI3k	Phosphatidylinositol-3-kinase
PCR	Polymerase chain reaction
PfEMP1	Pf erythrocyte membrane protein 1
PfPR	*Plasmodium falciparum* prevalence
PPQ	Piperaquine
QT-NASBA	Quantitative nucleic acid sequence based amplification
RBM	Roll Back Malaria
RDT	Rapid diagnostic test
RMSE	Root mean squared error
POC	Reporter operating characteristic
SMC	Seasonal malaria chemoprevention
SMFA	Standard membrane feeding assays
SNP	Single nucleotide polymorphisms
SP	Sulfadoxine/pyrimethamine
SPC	Slow parasite clearance
SSA	Sub-Saharan Africa
TPP	Target product profiles
WHO	World Health Organization
WHO-FIND	WHO Foundation for Innovative New Diagnostics
WWARN	WorldWide Antimalarial Research Network

Acknowledgements

I'd like to start by thanking Bob Verity for his day to day supervision and for always being around to answer all my questions, ranging from the mundane to the downright stupid – and all after I momentarily forgot his first name during my acknowledgements slide at my first ASTMH. There is no uncertainty in my mind that my PhD experience has been significantly improved in every dimension due to his supervision and I cannot thank him enough. I would like to also thank Lucy Okell for her guidance and supervision, especially in the final weeks when everything needed to come together quickly. Her positivity and optimism have been invaluable at times when I've had doubt in my work and her generosity of spirit is something I will aspire to practice at all times in my life. Lastly, I would like to thank my primary supervisor Azra Ghani. If I was able to tell myself four years ago about everything that I have been able to do during my PhD, I would surely have thought that the next four years would involve me severely abusing some powerful hallucinogens. The level of trust that Azra has placed in me has allowed me to do things well above my pay grade while also giving me the freedom to follow my own research interests. I am incredibly grateful for all the opportunities I have had during my PhD and want to thank all three of my supervisors again for this.

Outside of my direct supervisors there are numerous members of the department I need to thank for their help. Most importantly, I want to thank Hannah Slater for being my rock. I was exceptionally fortunate to be have been sat in the right corner of the Monkey Puzzle one evening to hear about a master's project on hrp2 deletions with Hannah, Bob and Azra. Despite not being one of my supervisors for my PhD, it was still Hannah who was so often there to help and deal with my idiotic moments and her generosity in the time she would allocate to supervision is truly remarkable. I would also like to thank everyone in the malaria group in DIDE for their help, with special mention to Ettie, Patrick, Pete and Tom for their help with a number of projects. I would also like to thank Jeff for an incredibly enjoyable collaboration in producing rdhs. Lastly, a big thank you to Rich and Wes for their assistance – I would still be attempting to brain myself with the space bar without their help.

Outside the department, I would like to thank Steve Meshnick. Steve was the first academic I met outside of my department in a work context while at GEM 2016. His kindness and interest in my work at such an early stage is true testament to his mentorship and I am incredibly grateful to have been able to work with him. I also want to thank Jonathan Parr who has been exceptionally generous with his time and direction throughout my collaboration with him on hrp2 deletions. I always look forward to my discussions with Jonathan on this subject (as well as the Great British Bake Off). Lastly, I want to thank everyone in the wider IDEEL group for a number of excellent academic and non-academic

activities, in particular to Andreea, Jon Juliano, Jeff, Oz, Molly and Nick. I also need to thank Nick again for his company and help during the final push.

I have made some terrific friends during my PhD who deserve considerable thanks. Firstly, to the stream – Fin, Izzy, Matt, Robin, Lucia and Jess. The MRes year was a blast (despite the cardboard box) and I could not have asked for a nicer group of people to start with at the department, which undoubtedly meant that I never once felt the usual loneliness associated with doing a PhD. Also a big thank you to Joel, Julia, Amy, Charlie, Claire, James and Lily. The battic lunch times have been a true highlight of the PhD and the sheer ridiculousness of most of the topics of conversation is worthy of its own thesis. I look forward to seeing how everyone of you shines in your future endeavours.

I also want to thank all the individuals who are responsible for me doing a PhD in the first place as it was not something I thought I either wanted to do or could do. Firstly, thank you to my undergraduate pharmacology supervisor Sio Ball for pointing me in the direction of a career in academia. This resulted in my first experience of research with James Murray at Imperial College, which did much to change my thoughts about pursuing research. Next, a large thank you to Nik Cunniffe at Cambridge who again changed my perceptions of what doing a PhD would be like and ultimately gave me my first experience of infectious disease modelling that led to me doing this PhD. Thank you also to the Wellcome Trust for the opportunity to do this degree and to help contribute to malaria research efforts.

I would like to thank Ralph for his intimate knowledge of all aspects of malaria and transmission modelling – his unwavering company during the write up has helped me preserve some aspects of my sanity while also surely robbing me others. I also want to thank my brother and sister. Clare and James, you have always led by example and have unknowingly given me the confidence to pursue my academic interests. Similarly, I want to thank my parents for everything and for the start in life they have given me – it feels correct to recognise how exceptionally privileged I have been and to thank them fully for the investment they have placed in my education and the help they given that has allowed me to invest more time into my studies over the years. Thank you.

Lastly, I want to thank my partner Chloe for her unwavering support, immeasurable compassion and exceptional ability to listen to my mind-numbingly boring and decidedly poor science chat. You have always been there to pick me up during the tougher days of the PhD and I cannot thank you enough for all the proof reading you have done – especially on the drier topics of my study. This brief acknowledgement does not do justice to your contributions and I will endeavour to thank you as often as I can for your support and love.

Chapter 1. Introduction

Malaria is estimated to have caused 435,000 deaths in 2017 and continues to be one of the largest burdens to global health efforts, despite decreasing morbidity and mortality rates over the last 20 years. To continue reductions in the burden of malaria, an understanding of both its epidemiology and the emergence of new threats is required. In this chapter I give an introduction to the epidemiology of malaria and the lifecycle of the parasite and how it, along with external sources of selection, shapes the genetic diversity of the malaria parasite. The introduction also covers genetic approaches to the study of malaria and how they have led to our current understanding of the emergence of resistance, both with respect to antimalarial drugs and rapid diagnostic tests (RDTs). I finish by reviewing how mathematical models of malaria transmission have been previously used to model malaria control, focussing on models of drug resistance. The chapter subsequently serves to give an overview of the current state of malaria transmission modelling and lays the foundation for the work I present within the thesis.

1.1. Epidemiology of malaria

1.1.1. Global burden of malaria

Malaria is distributed across the tropical and subtropical regions of the world and is caused by the *Plasmodium* parasite (Guerra et al. 2006; Hay et al. 2009; Gething et al. 2011b). Malaria is estimated to have caused 219 million clinical cases in 2017, with an associated 435,000 deaths worldwide (World Health Organization 2018c). Human malaria is caused by six different species of *Plasmodium* parasites: *P. falciparum, P. malariae, P. knowlesi, P. ovale curtisi, P. ovale wallikeri,* and *P.* vivax (Calderaro et al. 2013). Globally, the vast majority of malaria cases are caused by infection with *P. falciparum* and occur in Africa, with approximately 92% of global cases occurring in Africa and 99.7% of the 200 million estimated cases in Africa caused by *P. falciparum* (World Health Organization 2018c).

P. falciparum is not the major infecting species within all endemic malaria regions, however, with *P. vivax* being the predominant parasite species in the WHO defined region of the Americas accounting for 74.1% of malaria cases (World Health Organization 2018c). Similarly, zoonotic cases of malaria caused by *P. knowlesi* are the most common cause of human malaria in Malaysia, increasing from 1600 to over 3600 cases between 2016 and 2017 (Cooper et al. 2019). The spatial heterogeneity in the predominance of different *Plasmodium* species reflects the different life history of the parasites (Price et al. 2007), different mosquito vectors (Medley et al. 1993), and different genetic variation in

the human population (Kwiatkowski 2015) such as the presence of *P. vivax* resistance in Duffy-negative individuals in West Africa (Langhi and Bordin 2006).

The burden of malaria has declined substantially since 2000, despite its still current high burden and inherent heterogeneities and complexities. This decline represents the scale-up of malaria control programmes and the renewed political interest in reducing the global burden of malaria, which has resulted in over 40% reduction in clinical cases between 2000 and 2015 (Bhatt et al. 2015). These reductions have involved a number of countries being certified as malaria free. Algeria and Argentina were both certified as malaria free in May 2019 and Paraguay was certified in June 2018 (World Health Organization). These achievements represent significant milestones with a number of other countries also nearing elimination, with 26 countries in 2017 reporting less than 100 indigenous cases, which provides a strong indication that elimination is within reach (World Health Organization 2018c).

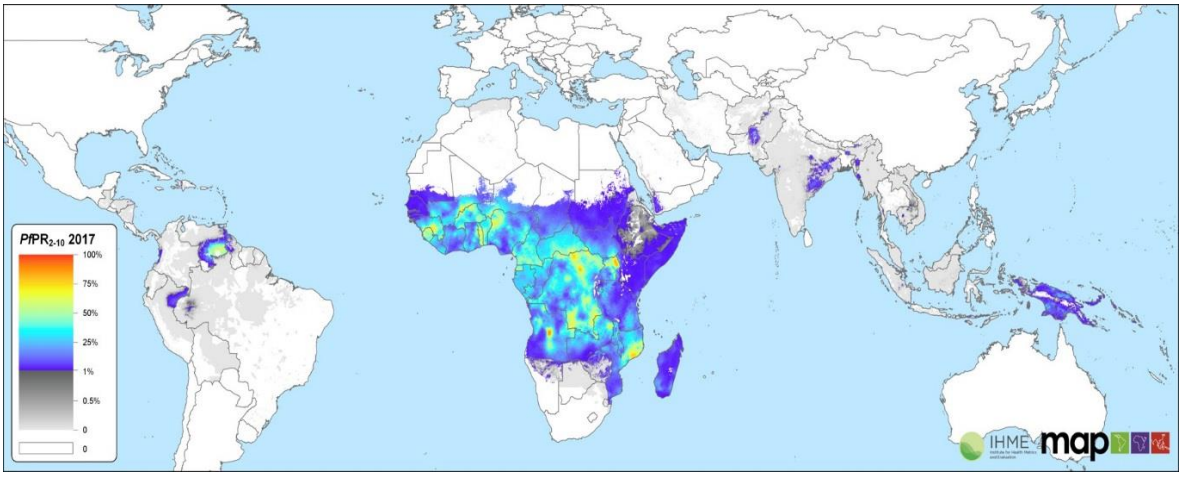


Figure 1.1 The spatial distribution of *P. falciparum* malaria burden in 2017. Regions with no risk of malaria are shown in white, regions with low endemicity (0-1% prevalence in children aged 2-10 ($PfPR_{2-10}$)) in shades of grey, and prevalence greater than 1% shown scaled from blue to red. Sourced from Weiss et al (CC BY 4.0) (Weiss et al. 2019).

1.1.2. Lifecycle of malaria

The lifecycle of *Plasmodium falciparum* is structured in three distinct cycles: the exo-erythrocytic cycle, the erythrocytic cycle and sporogenic cycle (Figure 1.2). The exo-erthrocytic or pre-erythrocytic cycle starts with the delivery of malaria sporozoites from an infected *Anopheles* mosquito. The parasite then travels from the site of the bite until reaching a blood vessel and migrating to the liver. In the liver sporozoites invade hepatocytes, develop into trophozoites before maturing into schizonts, which eventually burst releasing approximately 20,000 merozoites into the blood stream (Meis et al. 1986). However, not all sporozoites make it to the liver. Sporozoites reside in the skin for roughly 1-6 hours, with 20% estimated to migrate to the lymph nodes (Amino et al. 2006) where they are believed to

trigger anti-parasite immune responses (Guilbride et al. 2012). In fact, most sporozoites fail to reach the lymph nodes and are simply cleared at the inoculation site, with only a small proportion reaching a blood vessel and subsequently the liver (Sidjanski and Vanderberg 1997). The precise proportion, however, is not known.

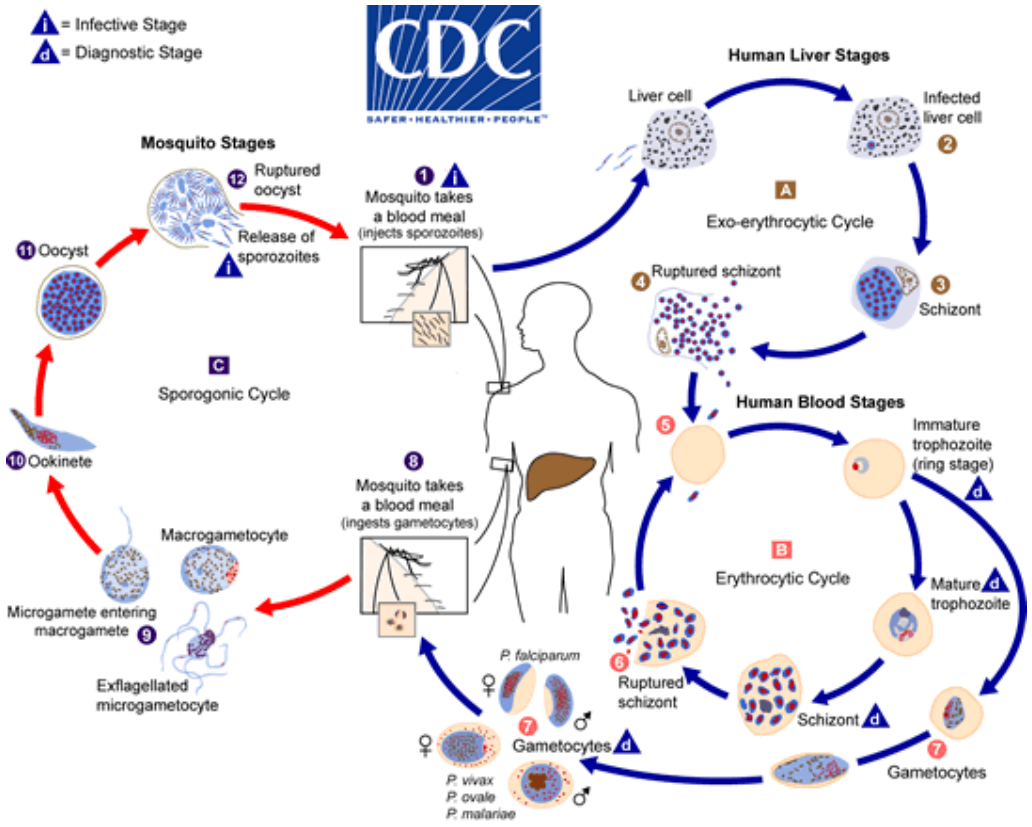


Figure 1.2 Malaria parasite lifecycle in humans and mosquitoes. Public Domain. Centers for Disease Control and Prevention (Centers for Disease Control and Prevention [Public Domain]).

Once merozoites have entered into the blood stream, the erythrocytic cycle starts with merozoites invading erythrocytes and developing into ring stage trophozoites. Ring stage trophozoites mature into schizonts before causing the infected red blood cell to rupture releasing approximately 8-32 merozoites into the blood stream to start the next erythrocytic cycle (Sinden and Gilles 2002). The erythrocytic cycle for *P. falciparum* lasts for two days and is responsible for the periodic fever commonly associated with clinical malaria. A number of merozoites will differentiate into the sexual stages of the parasite, known as gametocytes. Gametocyte sex ratios are usually female-biased, with the male gametocyte producing up to eight male gametes via exflagellation when taken up by a mosquito. However, the sex ratio of malaria parasites has been shown to be adaptive and dependent on a range of factors that affect the hematologic state of the host (Paul et al. 2000). For example, anaemia has been shown to be associated with an increase in male gametocytes in infected children (Robert et al. 2003). Changes in parasite density, either through antimalarial treatment (Stone et al.

2017), immune suppression (Sowunmi et al. 2009) or other environmental factors are suggested to trigger malaria parasites to invest more in their transmission stages by increasing gametocytogenesis or increasing the proportion of male gametocytes (Bradley et al. 2018).

The sporogenic cycle takes part inside the mosquito, starting with male gametes fertilising female gametes to form a zygote. The specific dynamics of mating, however, have been shown to be affected by transmission intensity, with an increased chance of outcrossing occurring at higher transmission intensities, reflecting the increased number of different parasite strains in individuals (Ranford-Cartwright et al. 1991; Paul et al. 1995). The zygote develops into a motile stage known as an ookinete, which invades the mosquito midgut wall before developing into an oocyst. Successful ookinete invasion is dependent on a number of genetic and immune factors (Smith and Barillas-Mury 2016) and consequently only a fraction of the potential oocysts that could form, based on the number of gametes taken up in a feed, go on to form oocysts. Oocysts will then grow before rupturing and, depending on the *Plasmodium* species, release up to thousands of sporozoites that travel to the mosquito salivary glands ready to be inoculated into humans (Vaughan 2007).

1.1.3. Malaria mortality, morbidity and immunity

Malaria infection in individuals with no immunity results in the rapid proliferation of merozoites leading to the development of symptoms from up to seven days after inoculation, however more usually between 9-15 days (Boyd and Kitchen 1937). Symptoms usually include fever, headache and nausea, which result from infected erythrocytes in the spleen triggering an immune response and the production of pro-inflammatory cytokines (Angulo and Fresno 2002). As the infection progresses, acute anaemia can occur due to a reduction in erythrocytes resulting from the intraerythrocytic cycle, which is responsible for the cyclical fever and chills associated with malaria infection (Lamikanra et al. 2007). A proportion of clinical cases will also develop severe malaria. Symptoms of severe malaria include cerebral malaria, respiratory distress, acute renal failure, pulmonary edema and severe anaemia (World Health Organization 2015a). Rapid development of any number of these symptoms can occur leading to death within hours or days (Trampuz et al. 2003).

Infection with malaria results in an immune response to the emergence of blood stage parasites. The immune response includes both an immediate non-specific response and a slower malaria-specific response. The innate response acts to quickly hinder parasite development and limit the initial phase of parasite replication, which allows the specific adaptive immune response to be developed and enable the infection to be eventually cleared (Stevenson and Riley 2004). The adaptive immune response targets multiple stages of the parasite life cycle. Antibodies are generated that target

sporozoites and merozoites providing protection against parasites invading hepatocytes and merozoites respectively (Bull and Marsh 2002; Schwenk et al. 2003). Interferon-mediated responses are also recruited to inhibit parasite development in hepatocytes and in the activation of macrophages to target infected erythrocytes and merozoites in the blood stream for phagocytosis (Miller et al. 2014).

Acquired immunity to malaria builds up sequentially in response to repeated infection. For example, the risk of clinical disease increases from birth with infants becoming susceptible to malaria after 3-4 months, likely due to the waning of maternally-derived immunity (Doolan et al. 2009). After this period immunity to severe malaria develops rapidly, with the frequency of clinical cases resulting in mortality falling between the ages of two to five years depending on the transmission setting (Doolan et al. 2009). Immunity against non-severe clinical malaria is acquired more slowly and is developed over the course of multiple infections. Consequently, the rate at which acquired clinical immunity develops is dependent on the transmission intensity, with adults in high transmission regions rarely developing clinical symptoms of malaria due to high levels of acquired immunity (Langhorne et al. 2008). However, the diversity of *P. falciparum* parasites and their ability to switch which proteins are expressed on the surface of erythrocytes act to delay the rate at which malaria immunity can be acquired. The different types of malaria immunity and the need for multiple infections to develop effective acquired immunity cause the clinical burden of malaria to be highest in children, with the majority of malaria morbidity and mortality occurring in children under the age of five in sub-Saharan Africa (Murray et al. 2012).

1.1.4. Malaria Diagnosis

Diagnosis of malaria is complicated due to the commonality of symptoms and similar clinical presentation to a number of common diseases. For example, in 2014 35.7% of self-reported fevers in sub-Saharan Africa were estimated to be accompanied by a malaria infection (Dalrymple et al. 2017). As a result, previous guidance to carry out presumptive treatment and syndromic management has led to both overtreatment of patients, with between 47%-95% of patients presenting with a non-malarial febrile illness being presumptively treated with antimalarials (Rao et al. 2013). Consequently, the recommendation made by the WHO in 2010 to test, treat and track all malaria infections has necessitated the deployment of quick and accurate diagnostic tests (World Health Organization 2010a). The two main diagnostic tests used regularly are microscopy-based diagnosis and rapid diagnostic tests (RDTs).

Microscopy has been the most commonly used diagnostic test for the past century, affording detection thresholds of up to 4-20 parasites per microliter in laboratory controlled settings (Wongsrichanalai et al. 2007). In the absence of highly skilled microscopists, however, detection thresholds of 100-200 parasites per microliter are more commonly observed in field settings, which is substantially higher than limits of detection of molecular diagnostic methods that can achieve <5 parasites per microliter (Payne 1988; Tangpukdee et al. 2009). Diagnosis by microscopy requires the preparation of blood smears which enable trained microscopists to visually assess for the presence of malaria parasite life stages. It is also possible for different parasite species to be determined using microscopy based diagnosis, however, this is more prone to incorrect diagnoses and is more dependent on parasite densities (McKenzie et al. 2003). Microscopy based diagnosis is inexpensive but is not feasible in all settings. The need for training and basic equipment, as well as the time taken for staining, can impair diagnosis feasibility in remote areas and high endemic settings where the demand for diagnostics exceeds the capacity afforded by microscopy (Payne 1988; Ochola et al. 2006). Poor specificity in microscopy can also occur due to a number of factors, such as poor quality reagents, inadequate film preparation and human error relating to both microscopist training and their workload (Durrhelm et al. 1997; Houwen 2002).

Malaria RDTs were designed to be useful in settings where microscopy based diagnosis is not feasible, and now possess comparable detection sensitivities to field microscopy (Wu et al. 2015). The ease of use and absence of additional equipment allows malaria to be diagnosed with minimal training in rural settings (Carrara et al. 2006). The simplicity of using malaria RDTs and their widespread adoption has resulted in an increase in diagnostic testing for suspected malaria. In 2007, less than 10% of individuals treated for malaria had a confirmed diagnosis (World Health Organization 2008), whereas the median percentage of cases that were tested before receiving antimalarial treatment in 2015-2017 was 74% (IQR: 51–81%) (World Health Organization 2018c). This increase helps to both reduce expenditure on antimalarials and reduces the level of over-prescription, which has been shown to increase the probability of drug resistance emerging in other infectious diseases (Llor and Bjerrum 2014). Additionally, it can help clinicians rule out malaria as the cause of an undiagnosed fever, which in turn increase the chance that the correct diagnosis and treatment is provided for the non-malarial fever (Odaga et al. 2014).

RDTs are immunochromatographic lateral flow devices that use antibodies to detect malarial antigens. Different RDTs are able to detect different malaria species depending on the targeted antigen. RDTs have been developed to detect either single species (*P. falciparum* or *P. vivax*) or multiple species (*P. falciparum, P. vivax, P. malariae* and *P. ovale*), with some able to distinguish between *P. falciparum* and non-falciparum infection. The most widely used RDTs target histidine-rich protein 2 (HRP2), which

is unique to *Plasmodium falciparum* (World Health Organization 2015b). Other targeted antigens include plasmodium lactate dehydrogenase (LDH) and aldolase (ALDO), although most RDTs being tested target HRP2 with 44 out of 46 RDTs tested in round 7 of the WHO Foundation for Innovative New Diagnostics (WHO-FIND) Malaria RDT Evaluation Programme targeting PfHRP2 (Organization 2017). The HRP2 antigen, however, has a long half-life in the human body and subsequently has a slow clearance rate, with the antigen remaining in the bloodstream for weeks after successful clearance of the parasite, leading to a false-positive RDT result for a median of 35-42 days (Grandesso et al. 2016). The duration of being HRP2 positive leads to complications when determining if treatment has been successful or whether lingering HRP2 concentrations are indicative of a recrudescent event. However, an analysis of HRP2 concentrations from patients in Angola, Tanzania and Senegal showed that comparisons of day three and day seven antigen concentrations against pre-treatment levels can be useful for predicting recrudescence (Plucinski et al. 2018a). Day three levels taken alone provide sufficient information to determine a drug failure, with an area under the receiver operating characteristic (ROC) curve of 0.86 (95% confidence interval, 0.73–0.99). In contrast, RDTs that detect LDH or ALDO have shorter half-lives and the majority of LDH and ALDO are cleared concurrently with parasite clearance reducing the possibility of a false-positive RDT result (Plucinski et al. 2018b, 2019). However, RDTs that target ALDO and LDH generally possess higher limits of detection (LOD) and lower sensitivities (Iqbal et al. 2002; Murray et al. 2008) as well as being less thermostable (Chiodini et al. 2007; Singh et al. 2013).

The development of highly sensitive nucleic acid based detection methods, such as polymerase chain reaction (PCR) (Johnston et al. 2006) and quantitative nucleic acid sequence based amplification (QT-NASBA), has enabled sub-patent infections to be detected (Tangpukdee et al. 2009; Hofmann et al. 2015). However, the costs of conducting PCR based diagnosis means that PCR analysis is not feasible for widespread detection and is largely used for research purposes to test the diagnostic performance of less sensitive diagnostics (Mappin et al. 2015), or for studies of parasite genetics (Ranford-Cartwright et al. 1991). For example, a comparison of malaria diagnosis by PCR, microscopy and RDT revealed that only 41% of infections positive by PCR were RDT positive (Wu et al. 2015). Similar diagnostic sensitivities were found for microscopy based diagnosis, with 87% of infections detected by RDT also detected by microscopy (Wu et al. 2015). However, a recent analysis of the discordance between RDT and microscopy diagnosis in sub-Saharan Africa revealed an excess of false-negative RDT results within community-based surveys (Watson et al. 2019a). This finding suggested that RDT performance in community based surveillance is less sensitive than microscopy due to increased false-negative RDT results in areas with lower malaria prevalence, which have been shown to exhibit lower parasite density infections on average (Slater et al. 2019). Lastly, there are several reports of

individuals with high density *P. falciparum* infections which are detected using microscopy but for whom PfHRP2-based RDT results have been negative (Cheng et al. 2014). These occurrences have been attributed to the part or whole deletion of the *pfhrp2* gene, resulting in no HRP2 expression. These *pfhrp2*-deleted mutants may still possess a functioning *pfhrp3* gene, however the cross reactivity of the PfHRP3 epitope is such that PfHRP2-based RDTs may be only partially effective at very high parasitaemia (Baker et al. 2005). The presence of *pfhrp2*-deleted mutants has been confirmed and mapped to several locations in South America, with the scale of *pfhrp2*-deletions resulting in recommendation against the use of PfHRP2-based RDTs in certain regions with high false-negative rates (Akinyi et al. 2013; Abdallah et al. 2015). Confirmed *pfhrp2*-deleted mutants were historically rarer in Africa, however in the last four years there has been an increase of confirmed cases, prompting the WHO to issue guidance for the prevalence of *pfhrp2/3* deletions to be estimated in key areas in sub-Saharan Africa (World Health Organization 2018b).

1.1.5. Malaria Interventions

Malaria control interventions can be broadly summarised as either targeting the malaria parasite or the mosquito vector. Multiple interventions have been developed that target different life stages of the parasite and vector, attempting to increase the probability of preventing onwards transmission.

1.1.5.1. Parasite Based Interventions

Antimalarial drugs are used to both prevent and treat clinical cases of malaria by targeting blood-stage parasites. Chloroquine (CQ) was the first antimalarial used widely to control malaria, however, its extensive use as part of the Global Malaria Eradication Programme (GMEP) led to the evolution of CQ resistance (Payne 1987) and was attributed in part to the failure of the GMEP (Cohen et al. 2012a). CQ was subsequently replaced as a first line drug by sulfadoxine/pyrimethamine (SP), however, resistance to SP shortly emerged (Nair et al. 2003; Roper et al. 2003), with Southeast Asia appearing to be the centre of resistant malaria emergence (Roper et al. 2004; Anderson and Roper 2005). The discovery of artemisinin towards the end of the 20th century has been instrumental in renewed efforts to reduce the global burden of malaria. Artemisinin was briefly used as a monotherapy before artemisinin combination therapies (ACTs), an artemisinin derivative combined with a longer lasting partner drug, were developed. The decision to administer artemisinin within an ACT is both due to the shorter half-life of artemisinin, which reduces the total parasite killing of the drug compared to other antimalarials despite its comparatively higher rate of parasite clearance (Hastings et al. 2016), and to decrease the speed at which artemisinin resistance emerges (Group 2004). Additionally, treatment with an ACT has

been shown to reduce the probability of onward transmission through killing sexual stage parasites (Okell et al. 2008).

Recent recommendations have also seen the expanded use of antimalarials for prophylaxis. In 2010, the WHO recommended the use of intermittent preventive treatment in infants (IPTi) in areas of moderate to high transmission (World Health Organization 2010b). This involves treatment being given three times during the first year of life, regardless of whether the child is infected with malaria (Aponte et al. 2009). This style of treatment is known as chemoprevention and is designed to provide protection by providing a period of prophylaxis. In 2012 this was extended to include intermittent preventive treatment in pregnancy (IPTp) (World Health Organization 2012b), which involves administering SP (ACTs are not recommended due to safety considerations) to protect pregnant women from their second trimester onwards (Holtz et al. 2004). Additionally, in 2012 seasonal malaria chemoprevention (SMC) was officially recommended (World Health Organization 2012c), in which three month drug courses are given to children under the age of 5 years in areas with highly seasonal transmission (Greenwood 2006). In 2015, the WHO also recommended the use of mass drug administration (MDA) within near elimination settings (World Health Organization 2015c), in which antimalarial drugs are given to the whole population regardless of infection status (Gosling et al. 2011). This was also extended to include time-limited MDA as an initial response for malaria epidemic control and in emergency settings such as during the recent Ebola outbreak (Walker et al. 2015). Although not currently recommended, mass screen and treat (MSAT) policies have been explored that use antimalarials in combination with RDT to screen a population for malaria parasite prevalence before administering ACTs to parasite positive individuals (Tiono et al. 2013).

A number of additional interventions that target parasite stages of infection are in development. This includes new antimalarials in clinical trials such as artefenomel–ferroquine and lumefantrine-KAF156 (Ashley and Phyo 2018), as well the potential first malaria vaccine. Malaria vaccine development was initially believed to be feasible after the observation that immunisation of mice with irradiated sporozoites yielded a level of protective immunity (Nussenzweig et al. 1967). Since then a number of potential vaccine candidates have been developed although none have been shown to offer sufficient protection. In 2019, however, pilot introductions of the RTS,S/AS01 vaccine, which is the only vaccine to date that has shown a protective effect in a Phase III trial, have started in three African countries as part of the WHO Malaria Vaccine Implementation Programme (MVIP) (2018). RTS,S/AS01 is a pre-erythrocytic vaccine that aims to protect against the parasite forming liver stages. Potential vaccines candidates that offer protection against blood stage parasites are yet to reach Phase III trials (Miura 2016), and transmission blocking vaccines, which have seen recent interest as a tool for interrupting malaria transmission (malERA Consultative Group on Vaccines 2011), are just entering Phase II trials.

1.1.5.2. Vector Based Interventions

Vector based interventions work by either preventing contact between the mosquito and the human or through directly killing the mosquito. Insecticide treated bed nets (ITNs) provide both a chemical and physical barrier between the mosquito and the human, and are used to provide a sleeping area that is protected from night time feeding mosquito activity (Lengeler 2004). ITNs confer both a direct protection to the individual using the net, preventing potential malaria transmission from infectious mosquito bites, as well as indirect protection by preventing uninfected mosquitoes from infection by feeding on an infected human (Howard et al. 2000; Maxwell et al. 2002). Long lasting insecticidal nets (LLINs) are improved ITNs, which directly incorporate insecticide into the net material increasing the duration of chemical protection to up to three to five years (Kilian et al. 2011), although estimates suggest that effective chemical protection is often less than three years (Gnanguenon et al. 2014). Although LLINs are primarily effective against night feeding vectors, they have been shown to confer some protection against other mosquito species that feed primarily during the day*, likely through* a killing action due to resting on treated nets (Lenhart et al. 2008). LLINs are attributed to have prevented 68% of the 665 million malaria cases that were prevented as part of the scale up of malaria intervention control between 2000 – 2015 (Bhatt et al. 2015). However, as with antimalarials, the continued use of pyrethroids as the only insecticide class in LLINs has resulted in the emergence of insecticide resistance to pyrethroids (Strode et al. 2014). This has led to the WHO encouraging the development of new types of LLINs. One example is the use of piperonyl butoxide (PBO), which is used as a chemical synergist in pyrethroids-treated LLINs, which has been shown in recent randomised control trials in Tanzania to result in an increased in the effectiveness of LLINs compared to nets treated with only pyrethroids (Protopopoff et al. 2018).

Indoor residual spraying (IRS) is another widely used form of vector based intervention, in which insecticides are directly sprayed onto the walls of houses to either kill mosquitoes that land after feeding or to prevent mosquitoes from entering the house and feeding in the first place (Pluess et al. 2010). The use of IRS has been attributed to preventing only 13% of malaria cases between 2000 – 2015 (Bhatt et al. 2015), which reflects the reduced levels of IRS used compared to ITNs, reduced levels of personal protection compared to ITNs as well as difficulties within communities of achieving effective coverage and ensuring repeat rounds of IRS are carried out to yield continued protection (Sherrard-Smith et al. 2018). Larviciding can also be used to target known larval breeding sites with insecticides in an attempt to kill aquatic larval stages of the mosquito (Walker and Lynch 2007). However, larviciding is only cost-effective for malaria control in settings where vector breeding sites

are few, fixed and findable and are thus is not recommended in many areas in sub-Saharan Africa (World Health Organization 2012a).

There are a number of new vector control tools in development that promise to increase vector based intervention options in light of insecticide resistance. Mosquito gene drive is a technique in which genetically modified mosquitoes are introduced into native mosquito populations in an effort to reduce mosquito abundance or alter their effectiveness as a malaria vector (Burt 2003). In a similar approach, sterile insect techniques have been proposed as a method for reducing mosquito population sizes by releasing an overwhelming number of sterilised mosquitoes (Alphey et al. 2010). Other approaches include using adult mosquitoes as a method of disseminating insecticides to larval habitats (Devine et al. 2009), the deployment of attractive targeted sugar baits (Mueller et al. 2010; Zhu et al. 2015) and the use the ivermectin, which possesses mosquitocidal properties (Smit et al. 2018), impairs parasite development (de Carvalho et al. 2019) and has been shown to be effective when used as part of repeated MDA for reducing malaria incidence (Foy et al. 2019).

1.1.6. Malaria Transmission

1.1.6.1. Transmission Metrics

The discovery by Ross and Grassi that mosquitoes transmitted malaria resulted in the start of attempts to model its transmission and understand the factors driving the burden of malaria within populations (Cox 2010). The burden of malaria in a population can be measured by the parasite prevalence (PfPR for *P. falciparum)*, which is the proportion of the sampled population that is positive for malaria. PfPR will vary depending on both the age of the cohort sampled and the method of diagnosis used. PCR PfPR will typically be higher than PfPR measured using microscopy due to their lower limits of detection (Okell et al. 2009, 2012). Any measure of PfPR will depend on how representative the individuals samples are of the population as a whole. For example, active case detection through community health workers is unlikely to provide unbiased estimates of the prevalence in the entire population, with focus usually placed in areas of known higher transmission. On the other hand, passive case detection methods, usually in the form of individuals arriving at treatment clinics or hospitals, allows for estimates of the clinical incidence of malaria, defined as the average number of clinical malaria episodes per person per year, to be made. Clinical incidence is an easier measure of malaria burden to make, relying on largely passive case detection schemes. Efforts have been made to use measures of clinical incidence to infer the PfPR in the population. This process, however, is complicated by reporting errors, symptomatic individuals not reporting at health facilities and differences in the proportion of malaria infections that develop symptoms based on historic malaria

prevalence and population level immunity (Cameron et al. 2015). Cluster designed household surveys provide a method of capturing individuals in the population who may be unaccounted for via passive case detection, either through having an asymptomatic infection or choosing to seek care for a fever through private health services that are not reported to national malaria control programmes (Demographic and Health Surveys Program 2018).

The entomological innoculation rate (EIR) is one metric for the intensity of malaria transmission, which measures exposure to infectious mosquitoes by reporting the average number of infectious bites per person in a year. Attempts to measure EIR require some form of counting the number of mosquitoes that attempt to take a feed from a human adult in one day and multiplying it by the proportion of the mosquitoes that are infectious. The gold standard way to measure this is using human landing catches, in which volunteers are required to reside outside during mosquito feeding time and count the attempted feeds (Gimnig et al. 2013). However, given the human cost inherent in human landing catches, a number of other approaches have been used to estimate EIR. These approaches use traps to capture mosquitoes, including light traps (Wong et al. 2013), tent traps (Govella et al. 2009) and window exit traps (Müller et al. 2017). An infectious mosquito bite does not always lead to an infection, however, as infections can either be suppressed by an immune-response (Schwenk and Richie 2011) or may simply fail to progress to an infection by chance (Rickman et al. 1990). The proportion of infectious bites that do yield a blood-stage infection can be multiplied by the EIR to yield the force of infection.

1.1.6.2. Transmission Heterogeneity

The variation in malaria transmission is also important to consider. Variation can occur both spatially and temporally. Spatial heterogeneity in malaria transmission can be characterised in terms of the spatial of autocorrelation, with metrics such as Moran's I being used to assess for the scale spatial heterogeneity (Mogeni et al. 2017). Fine-scale spatial heterogeneity, occurring at the village or household level, can arise due to local proximity to mosquito larval breeding sites (Smith et al. 2004), differential ownership and effective usage of bed nets (Atiele et al. 2011) and variation in the quality of building materials used in housing (Lindsay and Snow 1988; Tusting et al. 2019). Spatial heterogeneity at larger scales occurs due to differences in urbanicity (Donnelly et al. 2005; Kamya et al. 2015) and economic development (Charchuk et al. 2016), climactic differences often associated with differences in altitude (Gething et al. 2011a; Cairns et al. 2015) and mosquito and parasite species distributions (Guerra et al. 2008; Hay et al. 2010; Battle et al. 2019; Weiss et al. 2019). Heterogeneity can even exist within the same household due to human factors related to an individual's age, body

temperature and surface area, sleeping behaviour and olfactory cues related to differing respiratory levels (Burkot 1988; Knols et al. 1995; Rodriguez-Barraquer et al. 2016).

Temporal variation in malaria transmission is largely driven by seasonality in rainfall and temperature, which impacts the total mosquito population size in an area (Craig et al. 1999). Interactions between climactic effects can also occur, for example heavy rainfall accompanied with high temperatures is observed to cause sharper increases in malaria transmission than in colder regions (Teklehaimanot et al. 2004). Malaria regions can be broadly defined by their level of seasonality and endemicity. Regions possessing stable annual temperatures are more capable of sustaining mosquito populations throughout the year, leading to perennial transmission (Hay et al. 2008). Most regions, however, experience some degree of seasonality and can be characterised by the periodicity of transmission, with transmission increasing in the rainy season and falling during the dry season (Cairns et al. 2012).

1.1.6.3. Transmission targets

Declines in malaria transmission resulting from human intervention can be classed with respect to their long term reduction targets. Malaria control is associated with the reduction of burden such that the immediate public health concern is removed. Malaria elimination involves the interruption of local malaria transmission, defined as zero locally contracted cases, knowing that people may acquire malaria infections while travelling to malaria regions (Cohen et al. 2010; Tatem et al. 2017). This definition allows for malaria cases acquired outside the region but diagnosed in the region concerned, known as an imported case, as well as introduced cases, which are first-generation local transmissions that are epidemiologically linked to proven imported cases (World Health Organization and Global Malaria Programme 2017). This distinction was recently introduced to adapt the requirements for a country to be certified as malaria eliminated by the WHO, however, methodology for classifying an introduced case through epidemiological methods is not 100% accurate, with genetic measures often required to determine a true introduced case (Hemingway et al. 2016). Lastly, malaria eradication is the substantially more complicated challenge of achieving worldwide permanent reduction of malaria infection to zero (Alonso et al. 2011). Eradication was previously viewed as an unrealistic goal after previous failed eradication programmes, however, there has been recent optimism and renewed efforts to achieve eradication, largely resulting from a call for eradication made by the Bill and Melinda Gates Foundation in 2007 (Roberts and Enserink 2007). The specific timelines to eradication are contentious (Enserink 2019), with a lack of consensus between the WHO and a recent Lancet commission on malaria eradication (Feachem et al. 2019). However, it has been shown that local elimination is possible in many parts of the world, although highly endemic regions are currently unlikely to achieve elimination with currently available tools (Griffin et al. 2010).

1.2. Genetic approaches within malaria

In 1996, an International Consortium of Scientists set a goal to determine the sequence of the *P. falciparum* genome because the human genome project was already progressing at a satisfying speed (Hoffman et al. 1997). Seven years later, the first genome sequence for *P. falciparum* was published. In the seventeen years after this our understanding of the biology of the malaria parasite and how it interacts with both the human and the mosquito vector has increased substantially. Large-scale population genomic analyses have enabled malaria genomics to be used to inform malaria control, largely through the identification of the genetic factors associated with the parasite's response to host and vector immune responses and the discovery of novel antimalarial targets and the origins of emerging antimalarial resistance.

1.2.1. Malaria Genome Project

In 2002, the first reference genome for *P. falciparum* was published (Gardner et al. 2002). The *P. falciparum* 3D7 reference genome started a series of efforts to sequence the genomes of other known malaria parasites, including the genome of the rodent malaria species, *P. yoelii*, which was also published in 2002 (Carlton et al. 2002). The increasing use of Sanger sequencing in the 21st Century resulted in the sequencing of multiple *Plasmodium* species genomes, which recently included the publication of the *P. malariae* and *P. ovale* genomes in 2017, representing the last major human malaria species' genome to be sequenced (Rutledge et al. 2017). The publication of all *Plasmodium* species genomes to date, as well as key technical developments for conducting genetic engineering are shown in Figure 1.3.

The publication of the *P. falciparum* 3D7 genome and efforts to ensure its maintenance and improve its annotation have created a hugely beneficial tool for the global malaria research community by enabling comparative genome studies. Comparative genome studies are a major research area and use the growing collection of genome sequences collected across malaria endemic regions to characterise genetic diversity, enabling genomic epidemiology to be applied for the study of malaria (Amato et al. 2016). Underpinning these efforts is the creation of international research networks, such as the Malaria Genomic Epidemiology Network (MalariaGEN)). Development of consortia of this nature helps to developed standards of openness by which samples collected from multiple research institutions and studies can be sequenced to the same level of accuracy and disseminated publically for the wider research community to use in follow up comparative studies. For example, studies have identified how the parasite genome is shaped both by human genetic variation (Kwiatkowski 2015)

and drug pressure leading to the identification of the origins of artemisinin resistance (Miotto et al. 2015).

The publication of malaria genomes has also triggered the development of other omics related approaches that harness the increasing computational power to study the network of ways in which a parasite's genetics are translated through to the observed phenotype of the parasite. Similarly, improved forwards and reverse genetic approaches, such CRISPR-Cas9 gene editing (Wagner et al. 2014), have been used to increase the functional annotation of the *P. falciparum* genome, leading to the number of genes with an unknown function approximately halving between the original 3D7 genome publication in 2002 and version 3.2 of the genome published in 2019 (Böhme et al. 2019). Central to both of these again is the open development of community tools and data sources, such as PlasmoDB, which has furthered progress in functional genomics research across multiple *Plasmodium* species (Aurrecoechea et al. 2009).

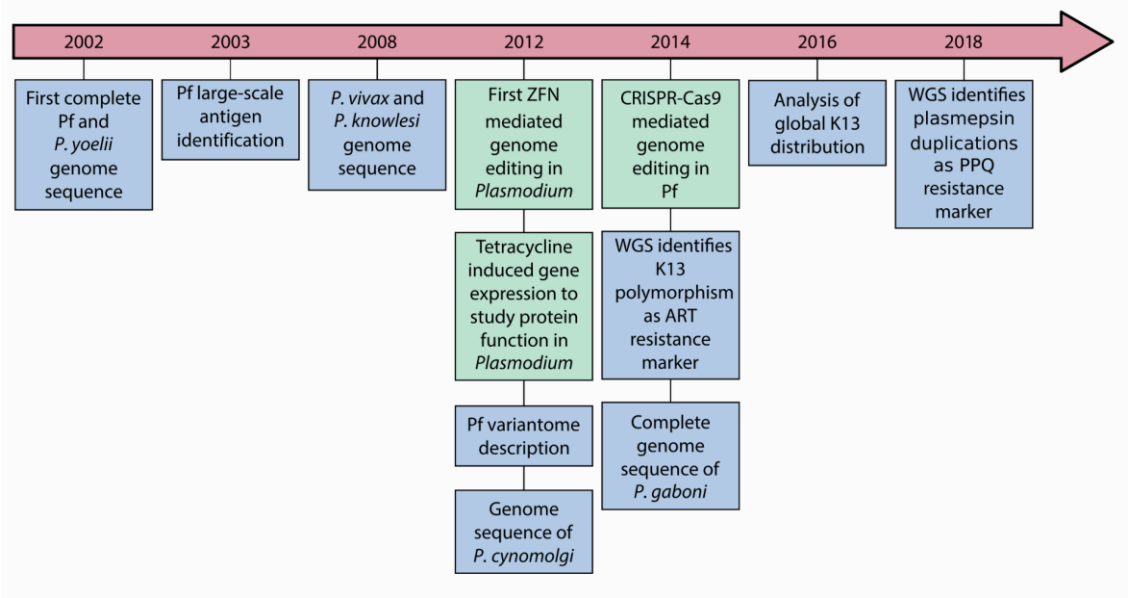


Figure 1.3 Advances in malaria genomics and related fields. Genome sequences and key resistance and malaria pathogenesis studies are shown in blue, with advancements in genetic engineering shown in green. Adapted from Kirchner et al (Open Access CC BY 4.0) (Kirchner et al. 2016).

1.2.2. Implications of parasite genomics for malaria control

1.2.2.1. Parasite virulence and immune responses

P. falciparum has evolved numerous mechanisms in response to host immune responses (Ferreira et al. 2004). The human immune system primarily targets the cell surface antigen Pf erythrocyte membrane protein 1 (PfEMP1). PfEMP1 is encoded by the exclusive expression of one of roughly 60

different *var* genes, with the parasite able to switch expression between different *var* genes and thereby to adapt between different virulence phenotypes associated with different disease severities and symptoms (Recker et al. 2011; Almelli et al. 2014). Immune evasion via *var* gene expression is exclusive to *P. falciparum*, with the characterisation of other *Plasmodium* species' adaptations to immune responses less defined (Neafsey et al. 2012). There are a number of other cell surface antigens expressed by *P. falciparum*, which include multi-copy gene families of RIFIN and STEVOR proteins, the latter of which has been associated with the severity of malaria symptoms within individuals with blood group A (Frech and Chen 2013; Bachmann et al. 2015).

Within the mosquito, the malaria parasite is also subject to immune pressures, which has also resulted in the evolution of adapted immune systems that are expressed in the sexual stages of parasite development. Pfs47 has been shown to be expressed on the surface of the motile ookinetes stage and is expressed by the female gamete which forms the body of the ookinetes after fertilisation (Molina-Cruz et al. 2013). Experimental studies using parasite strains expressing globally distinct genetic variants of Pfs47 were used to demonstrate that *P. falciparum* is adapted to its local vector, with non-adapted parasites less able to successfully evade the mosquito immune system and develop oocysts (Molina-Cruz et al. 2015).

1.2.2.2. Resistance identification

Genetic approaches have been used in the identification of the molecular markers of antimalarial resistance. Genetic crossing between CQ-resistant and CQ-sensitive parasites was used to identify a 36kb region in chromosome 7 of *P. falciparum* as the source of CQ resistance (Wellems et al. 1991). Further studies built on this to identify mutations in the gene encoding for the Pf chloroquine resistance transporter (CRT) that were significantly associated with the resistance phenotype (Fidock et al. 2000; Thomas et al. 2002). Similar approaches were used to identify point mutations in dihydropteroate synthetase and dihydrofolate reductase that are responsible for SP resistance (Peterson et al. 1988; Wang et al. 1997). In light of emerging artemisinin resistance, the first scientific goal of MalariaGEN focussed on analysing genetic variation using whole genome sequencing to understand how antimalarial resistance emerges and spreads. By combining genome sequences with the clinical phenotype of the parasite with respect to drug treatment failure, single nucleotide polymorphisms (SNPs) associated with drug treatment failure were identified (Takala-Harrison et al. 2013). These efforts resulted in the identification of Pf Kelch13 (K13) as the gene responsible for rising levels of artemisinin resistance in the Greater Mekong Subregion (Cheeseman et al. 2012). This discovery has greatly helped efforts to control emerging artemisinin resistance, with control efforts now able to sequence for the molecular markers associated with drug resistance. The WorldWide

Antimalarial Research Network (WWARN), for example, is an international network that uses molecular markers, as well as clinical markers, of resistance to provide maps and databases to help document the current prevalence of antimalarial resistance and to track its continued emergence and spread (WorldWide Antimalarial Research Network (WWARN)).

Recent developments in genetic engineering of *P. falciparum* have enabled the role of K13 mutations in conferring artemisinin resistance to be confirmed (Straimer et al. 2015). More recently, CRISPR-Cas9 engineering has been used to assess for differential fitness phenotypes associated with different K13 SNPs (Nair et al. 2018), in an effort to explain why resistant strains with C580Y mutations in K13 have recently emerged amongst multiple different K13 SNPs (Anderson et al. 2016). However, it is increasingly likely that the apparent benefits associated with C580Y mutations are reflective of the genetic background on which C580Y strains emerged, suggesting that a number of compensatory mutations in other genes offset fitness costs associated with K13 mutations. Possible sources for these beneficial mutations include genes encoding for ferredoxin, multidrug resistance protein 2 (MDR2) and CRT (Ménard et al. 2016). Alternatively, it has been suggested that K13 SNPs interact with inositol phosphate paths, causing reduced binding efficacies of Pf phosphatidylinositol-3-kinase (PI3K) (Neafsey et al. 2012). Understanding of the involvement of PI3K and other associated genes, however, has been greatly helped by the availability of longitudinally sampled genome sequences that enabled emergence times of novel artemisinin resistance targets to be compared against K13 emergence (Cerqueira et al. 2017).

1.2.3. Factors shaping parasite population genetics

The lifecycle of *P. falciparum* has been the topic of substantial research (Aly et al. 2009; Gerald et al. 2011), with the debate concerning the existence of a sexual stage and recombination within the mosquito (Maccallum 1897) resulting in numerous investigations to characterise the specifics of the obligate sexual stage (Ranford-Cartwright et al. 1991; Paul et al. 1995). In addition, malaria transmission between both the human and the mosquito involves a series of population bottlenecks (Vaughan 2007; Churcher et al. 2010), which combined with the rapid sexual stage involving a single two-step meiotic division (Bennink et al. 2016) have important implications upon the population genetics of *P. falciparum* (McKenzie et al. 2001). This is also extenuated by the increasing evidence of the transmission of multiple clonally-related parasites within one bite (Wong et al. 2017), within host mediated immune responses (Barry et al. 2007; Portugal et al. 2011) and density-dependent regulation of superinfection (Bruce et al. 2000; Pinkevych et al. 2013). These factors, along with the external selective pressures placed upon the parasite by antimalarial drugs, have resulted in a complicated network of processes driving the genetic diversity of the parasite population. In addition,

frequent polyclonal infection makes it harder to distinguish the actual parasite genotypes within an infected individual. This complexity led to the description of polyclonally infected individuals in terms of their complexity of infection (COI), defined as the total number of genetically distinct parasites within an individual. However, recently developed computational methods (Galinsky et al. 2015; Ken-Dror and Hastings 2016; Chang et al. 2017) have improved the ability to predict higher COIs, allowing for the prevalence of minor clones associated with lower parasitaemia to be estimated. This greater understanding of the parasite genetic lifecycle and population dynamics, combined with recombination maps of the *P. falciparum* genome (Mu et al. 2010; Jiang et al. 2011) and the availability of longitudinal genomic data increase the potential to create and importantly parameterise new mathematical models of malaria transmission that integrate parasite genetic information.

1.3. Mathematical models of malaria transmission

Models of malaria transmission are mathematical tools that describe how malaria is transmitted within a population and allow for the formulation of mathematical definitions for the prevalence of malaria with respect to key quantities and parameters. These tools are able to be used to quantify the benefit of different malaria control programmes and can help guide the deployment of future initiatives and how they can be optimised for malaria control.

1.3.1. The Ross-MacDonald model

The suite of modern models of malaria transmission are the end result of numerous iterative modelling developments taking place in the 20th century. Ronald Ross published the first mathematical model of malaria transmission in 1908 when he was tasked with characterising the different methods of preventing malaria in Mauritius (Ross 1908). In this model he described how the ratio of mosquitoes to humans, denoted as m, is related to malaria infection incidence. In continuation, Ross defined the proportion of mosquitoes that would need to be killed to result in a mosquito density, denoted m', below which transmission would be interrupted. The key parameters involved in Ross' are summarised in Table 1.1 (Smith et al. 2012a). From this, Ross's insight can be summarised by the formula:

$$m' > \frac{gr}{a^2 bce^{-gv}}.$$

During the GMEP in 1955-1969 the original Ross model was developed further by George Macdonald (MacDonald 1957). These developments incorporated needed biological processes such as the potential for individuals who are already infected to be reinfected, known as superinfection (MacDonald 1950), allowing the accumulation of multiple parasites within an infected individual to be captured. The developed Ross-MacDonald model was applied to provide insight into the efficacy of

using the insecticide DDT alongside treatment with CQ for malaria control. Although the eradication efforts of the GMEP ultimately failed, the modelling approaches developed helped increase our understanding of the importance of mosquito life expectancy on malaria transmission, which was used to explain the non-linear relationship between increasing mosquito death rates and decreasing sporozoite positivity rates (MacDonald 1952b). MacDonald's involvement in the GMEP helped to develop a new quantitative understanding of how entomological control impacts the transmission of malaria, which both led to the first attempts to define the basic reproductive number, R_0 (MacDonald 1952a), for malaria and subsequent modelling developments to better understand how to model the implementation of different malaria interventions to control malaria (MacDonald 1956).

Table 1.1 Human Infection State Transitions. Reproduced from Smith et al. (Open Access CC BY) (Smith et al. 2012a).

Parameter	Description
m	The ratio of mosquitoes to humans
g	The instantaneous death rate of a mosquito
r	The daily rate each human recovers from infection
a	The rate at which a mosquito takes human blood meals
b	The probability that a bite by an infectious mosquito infects a human
c	The probability a mosquito becomes infected after biting an infected human
v	The number of days from infection to infectiousness in the mosquito

1.3.2. Using Models for Intervention Evaluation

After the failure of the GMEP, a WHO sponsored investigation in collaboration with a Nigerian government research team was undertaken in the Garki district of Nigeria between 1966-1976 (Molineaux 1980). The investigation differed from previous efforts that had attempted to test the feasibility of eradication campaigns, but rather focussed primarily on the collection of measures of malaria transmission. Malaria measures included serological, parasitological and entomological measures with the view of building relationship between malaria metrics and transmission intensity (Molineaux and Gramiccia 1980). In addition, the Garki project aimed to measure the effect of IRS in combination with an MDA campaign using sulfalene-pyrimethamine. From the outset though, the Garki project also wanted to use the controlled set up to incorporate the collected malaria data and intervention data within a constructed malaria model that could be used going forward to evaluate malaria control efforts. Although malaria transmission returned after the end of the Garki project, it

highlighted the temporal and spatial heterogeneity in malaria endemicity and affirmed the importance of incorporating the collection of epidemiological data as part of any planned malaria control programme (Nájera et al. 2011).

The shift away from eradication to malaria control after the GMEP and the Garki project resulted in models of malaria transmission needing to incorporate measurable entities that could be used to assess the efficacy of interventions. This subsequently allowed for the cost-effectiveness of interventions to be calculated when combining malaria modelling predictions with the costs of disease burden. The shift in modelling focus also occurred alongside, or arguably as a result of, computational advancements that enabled originally deterministic models to be extended to more completely capture heterogeneity in the human population, including age-dependent malaria risk (Dietz et al. 1988; Anderson and May 1991), acquired immunity (Aron 1988) and variability in transmission intensity (Hasibeder and Dye 1988; Gupta et al. 1994) and the mosquito biting behaviour (Dye and Hasibeder 1986). As a result of continued extensions, mathematical models of malaria transmission are now capable of assessing the impact of intervention strategies and of simulating novel intervention strategies and tools.

Table 1.2 WHO recommended malaria interventions supported by malaria modelling.

Intervention	References
Increasing Treatment Coverage	(Okell et al. 2011; Johnston et al. 2014)
Seasonal Malaria Chemoprevention (SMC)	(Cairns et al. 2011)
Intermittent Preventive Treatment in pregnancy (IPTp)	(Walker et al. 2017)
Mass Drug Administration (MDA) for elimination	(Walker et al. 2015; Brady et al. 2017)
Insecticide-treated nets (ITNs)	(Le Menach et al. 2007; Chitnis et al. 2010; Okumu and Moore 2011; White et al. 2011)
Indoor residual spraying (IRS)	(Chitnis et al. 2010; Okumu and Moore 2011; White et al. 2011)
Larval Control	(Killeen et al. 2002; Eckhoff 2011a)

In 1988, the WHO, World Bank, UNICEF and UNDP founded the Roll Back Malaria (RBM) Partnership in an effort to control malaria and halve the burden by 2010 (Nabarro 1999). As highlighted earlier, much of the progress achieved as part of RBM is attributed to the scale-up of ITNs (Bhatt et al. 2015), which demonstrates that the focus on reducing adult female mosquito populations in the Ross-MacDonald model has been shown to be effective. Malaria models have been used since to evaluate and confirm the efficacy of a number of other interventions, with a number of intervention strategies

and WHO guidance and policy recommendations based on and incorporating findings from malaria modelling efforts (Table 1.2). Given the direct and indirect benefit afforded by malaria interventions, modelling has been used to explore both the individual and population benefits of different intervention combinations and policies.

From a population perspective, modelling has been used to review how cost-effective combinations of malaria interventions are and to review different proposed strategies. For example, modelling was used as part of the creation of the WHO Global Technical Strategy for Malaria 2016–2030, helping to cost different strategies and predict their long-term impacts on malaria prevalence (Griffin et al. 2016). This style of exercise also highlights how previous scale-up of malaria interventions has contributed to the declines in malaria observed during the RBM partnership and how continued investment in malaria prevention is required to prevent rebounds in malaria prevalence, in particular in response to potential cuts to international funding agencies such as the President's Malaria Initiative (Jakubowski et al. 2017). These types of study are important for highlighting that the current malaria burden in any country reflects its intervention coverage, with models highlighting the non-linear relationship between malaria exposure and malaria prevalence in the context of current treatment coverage and how it relates to the potential impact of increased treatment coverage (Penny et al. 2015). Different mathematical models have also been used to model malaria elimination strategies at the country-level (Maude et al. 2012), how malaria interventions should be deployed in response to global health emergencies such as Ebola outbreaks (Walker et al. 2015) and in the optimisation of malaria control and the importance of tailoring intervention packages to the ecological landscape of the region concerned (Walker et al. 2016).

Models can also be used to evaluate novel intervention tools and provide target product profiles (TPPs) to help guide policy-decision making and industrial research effort when designing new vaccines, antimalarials, diagnostics and vector control tools. For example, by considering the potential efficacy of different vaccine candidates in different transmission settings, mathematical models have been used to characterise the minimally effective profile of pre-erythrocytic vaccines (Maire et al. 2006c) and potential blood-stage and transmission blocking vaccines (Penny et al. 2008). These studies have highlighted the need for vaccine development to be considered in the context of the transmission settings they will be deployed in and how feasible mass vaccination campaigns will be that require follow up booster doses. TPPs for antimalarials have been used to help screen new drugs with lower risks of antimalarial resistance emerging (Ding et al. 2012). Mathematical models have been used in creating TPPs for next generation RDTs with increased sensitivity in an effort to increase the utility of MSAT programmes by increasing the proportion of asymptomatic infections that would be treated

(Slater et al. 2015a). These approaches can then be used to prevent wasted research and development costs in testing new RDTs if preliminary tests are not sufficiently sensitive. Similar approaches have been used when reviewing novel vector control tools that could be used to slow the spread of insecticide resistance. Spatial models of individual mosquitoes have been used to evaluate the optimum use of attractive toxic sugar baits (Marshall et al. 2013), and to guide mosquito gene drive release frequencies and schedule in relation to seasonality (Eckhoff et al. 2016). Lastly, the inclusion of ivermectin within mass treatment strategies has been modelled to consider the impact on mosquito mortality and reductions in the probability of infected mosquitoes developing sporozoites thus preventing onwards transmission (Slater et al. 2014).

1.3.3. Consensus Modelling Approaches

Inherent in the strength of any model is its ability to reproduce historic data and make accurate forward predictions. The former relies on model fitting, which has taken advantage of increasingly powerful methods for statistical inference and the increased quality and availability of data collected from field studies. Importantly, the ability to conduct multiple simulations of a model is increasingly easier due to decreased computation resource costs and distributed computing services. This allows model predictions to better capture and report model uncertainty, which is crucial when using models to support policy decisions. In addition to this, it is beneficial to also combine the predictions of different models to assess whether the same recommendations are made.

The recent evaluation of the RTS,S/AS01 vaccine used a consensus approach of multiple mathematical models to compare the cost-effectiveness and public-health benefit of the RTS,S/AS01 vaccine (Penny et al. 2016). The use of an ensemble of models that were each individually calibrated to the Phase III trial data for RTS,S/AS01 allowed the different models, and importantly their individual model's assumptions, to be brought within one framework to provide recommendations regarding the prevalence ranges that would most benefit from the RTS,S/AS01 vaccine as well as reviewing different dosing schedules. The different dosing schedules is worth further discussion as consensus was not achieved across all four models used, with one model only predicting a marginal benefit of moving from a three-dose to a four-dose schedule. However, by reviewing the components of each model and using sensitivity analysis to review the impact of a given model's assumptions, assumptions related to the estimated waning of protection against infection afforded by the vaccine were identified. This point highlights the strength of correctly setting up a consensus approach, in which only the parameters of the settings being modelled and not the modelling assumptions have been harmonised. At this stage recommendation can also be made about the need for further data to clarify how immunity wanes given its identified importance. However, if consensus had been achieved regardless

of different modelling assumptions then attention could be focussed elsewhere. In essence, the strength of the consensus approach is the differences between model predictions because it also strengthens the conclusion of findings when consensus is reached in other settings. Due to the strength of consensus approaches they have been used in reviewing mass drug administration (Brady et al. 2017) and in evaluating the benefit of increased sensitivity in RDTs (Slater et al. 2015b).

1.3.4. Current Malaria Models

The substantial range of malaria interventions that have been modelled and discussed above has been largely due to the work of a few research groups, each of which has developed a different mathematical model of malaria transmission. I now cover in brief the structure of four models of malaria transmission that have been used either in the evaluation of intervention strategies or considering the emerging threat of antimalarial resistance, before comparing their differences.

1.3.4.1. The Intellectual Ventures model

A team at the Institute for Disease Modelling published a model of malaria transmission in 2011 that focussed initially on capturing the dynamics of mosquito populations in response to seasonal effects and the implementation of intervention strategies (Eckhoff 2011b). The model explores the different daily outcomes for a given mosquito, depending on whether the mosquito feeds on livestock or humans and if the latter how interventions may cause mosquitoes to be repelled or killed as a result of interventions. The model was extended in 2012 to include an individual-based human population, which included an immune component that tracked a human's exposure to different parasite variants in terms of their expressed variant surface antigens, with exposure to up to 50 PfEMP-1 variants included (Eckhoff 2012). Modelling immunity in this way allows for the gradual acquisition of immunity and was used to show how increasingly accurate predictions of the relationships between malaria prevalence and age across transmission intensities can be achieved by increasing the number of unique antigenic components included in the model (Eckhoff 2012). The model has been used to also estimate the size of the infectious reservoir (Gerardin et al. 2015b) of malaria and evaluate potential malaria vaccines (Wenger and Eckhoff 2013) and MDA campaigns (Gerardin et al. 2015a), both independently and as part of the consensus exercises highlighted earlier evaluating the benefit of the RTS,S/AS01 vaccine (Penny et al. 2016).

1.3.4.2. OpenMalaria

In 2006, researchers at the Swiss Tropical and Public Health Institute published a series of papers exploring different components and functionality of a micro-simulation model of the epidemiology of malaria transmission (Smith et al. 2006a). The additional components of the models published at the

same time detail multiple aspects of the human component of the transmission model, which include the dynamics of asexual parasitaemia in non-immune individuals (Dietz et al. 2006), the acquisition of blood-stage immunity (Maire et al. 2006b) and the relationship between parasite density and the infectiousness of humans to mosquitoes (Killeen et al. 2006; Ross et al. 2006a). In addition, the model considers a variety of clinical symptoms and disease outcomes and was also used to explore the dynamics of clinical and severe malaria (Ross et al. 2006b; Smith et al. 2006b) as well as malaria-associated anaemia (Carneiro et al. 2006) and neonatal mortality (Ross and Smith 2006). The model was initially developed to evaluate pre-erythrocytic vaccines (Maire et al. 2006c, 2006a) and has consequently been used for cost-effectiveness studies of both case management and incorporating vaccines in expanded immunisation programmes (Tediosi et al. 2006). The focus on vaccine evaluation continued by predicting in 2012 that any vaccine similar to RTS,S/AS01 included in immunisation programmes would substantially decrease malaria morbidity and mortality (Smith et al. 2012b). However, it was shown to be dependent on transmission intensity with little effect predicted at transmission intensities prevalent in sub-Saharan Africa. Further vaccine studies have also been conducted to assess the utility of blood-stage and transmission-blocking vaccines (Penny et al. 2008). Outside of vaccine studies, OpenMalaria has been used in evaluating MSAT (Crowell et al. 2013) and reactive case detection strategies (Reiker et al. 2019) as well as modelling the threats of insecticide resistance (Briët et al. 2013) and changing mosquito feeding behaviour on vector control (Briët and Chitnis 2013). It has also been used in both consensus modelling exercises highlighted earlier.

1.3.4.3. The Imperial College model

An individual-based simulation model for *P. falciparum* transmission has been developed by the malaria modelling group at Imperial College London, which was initially developed for the evaluation of the impact of different intervention packages for malaria control (Griffin et al. 2010). Parameter estimates were fit using a mixture of clinical and parasitological data from multiple transmission settings (Griffin et al. 2014) and has been used to guide the costing of the Global Technical Strategy for malaria 2016-2030 (Griffin et al. 2016). Three different mosquito species (*An. gambiae*, *An. arabiensis* and *An. funestu*) can be modelled for the transmission of malaria. Potential interventions that can be modelled include changing treatment coverage with multiple antimalarials, IRS and ITNs. Similarly to the other models, it has been used to independently model and cost evaluate the benefit of mass drug administration in intervention packages (Walker et al. 2016) and the implementation of the RTS/S vaccine (Winskill et al. 2017), as well as taking part in the highlighted consensus exercises modelling these areas.

1.3.4.4. The OUCRU Malaria Sim

An individual-based microsimulation of malaria transmission was developed by individuals at Penn State and the Oxford University Clinical Research Unit (OUCRU). The resultant model, OUCRU Malaria Sim., was developed initially to assess the use of multiple first line therapies as a method to slow the emergence of ACT resistance (Nguyen et al. 2015). This version of the simulation includes transmission between hosts without explicit implementation of mosquitoes but rather using a population of mosquitoes and tracking the prevalence of exposed and infectious adult mosquitoes. The model also tracks within host parasite development and explicitly models parasite densities of multi-clonal infections and the impact on the relative infectiousness of different within host parasite strains. Additionally, the model includes a pharmacokinetic and pharmacodynamics component as well as mutation events to multiple types of drug-resistant genotype. The model also reproduces heterogeneity in malaria prevalence and symptoms, which involves the modelling of symptomatic and asymptomatic infections, age-dependent immune acquisition, and individual and age-dependent variation in host biting rates. The model was not included in the mentioned consensus exercises, but has been involved in simulating antimalarial resistance threat in response to MDA as part of the WHO Evidence Review Group on MDA for malaria (Malaria Policy Advisory Committee 2019).

1.3.4.5. Model Differences

The four models described above differ in terms of the level of detail each model places on different aspects of the transmission of malaria, what sources of data were used to parameterise the models and the level of stochasticity in each model. These differences are not an indication of any model being better or worse than another model and are largely reflective of the specific questions the models were initially designed to answer. For example, OpenMalaria was initially developed to evaluate pre-erythrocytic vaccines and consequently includes more biological realism with respect to within host parasite densities and how they relate to the build-up of acquired immunity, with immunity reflecting both previous numbers of infections as well as total parasite density exposure (Maire et al. 2006b), and extended to include waning immunity with modelling ensembles (Smith et al. 2012b). Both the Intellectual Ventures and OUCRU models also model absolute parasite density, with the former model largely focussing on biological realism and the latter model requiring it to capture fitness costs and differential parasite clearance rates resulting from antimalarial resistance. The Imperial College Model, on the other hand, does not track parasite density because the initial model development did not necessitate its inclusion to produce a model that was able to both reproduce prevalence and incidence dynamics during model fitting and was developed initially to contrast different intervention strategies. Aspects of malaria transmission that are related to parasite density, such as the probability

of developing a fever, being detected with a given diagnostic and the probability of onwardly infecting a mosquito, are modelled using relationships related to acquired immunity as a proxy for parasite density. These assumptions do, however, give computational benefits related to how quickly model simulations can run. Additionally, despite these differences the three models used within consensus exercises gave comparable predictions, although there are key differences when comparing the incidence against prevalence profiles for different age groups (Cameron et al. 2015), as well as differences in the amplitude of malaria prevalence resulting from seasonal transmission (Brady et al. 2017).

Each model handles uncertainty in model parameters and incorporates stochasticity within model predictions in different ways. For the 2015 analysis of ACT deployment strategies, the OUCRU model was parameterised using either known values resulting from direct measurements from experimental or field data. A similar approach was used across all models where possible, with the malaria therapy data being commonly used to estimate the duration of infection in nonimmune individuals (Eyles and Young 1951). Where parameters were unavailable, parameters were fit by reproducing known malaria relationships from field data, such as relationships between prevalence and EIR in different age groups. Lastly, sensitivity analysis was employed where parameters were unknown and simulation repetitions with different random seeds were used to provide summary statistics for metrics of antimalarial resistance such as drug treatment failures and mutation frequencies. Consequently, stochasticity in these prediction stems from the inherent stochasticity of the model with most parameters remaining fixed. In contrast, the Intellectual Ventures model often uses incremental mixture importance sampling to parameterise their model. This approach was used to reproduce prevalence and incidence data when calibrating the within host parasite component of the model to data collected from sites in Tanzania and Senegal (McCarthy et al. 2015), as well as in estimating the annual reproductive rate when investigating how malaria parasite genetic barcode data could be used for inferring transmission intensity (Daniels et al. 2015). The Imperial College model relies on the posterior distribution of model parameters to draw parameter sets for simulations, thus allowing model prediction uncertainty to arise from uncertainty in the model parameterisation when the model was fit to incidence data from 23 sites in Africa (Griffin et al. 2010). Lastly, the OpenMalaria model incorporates uncertainty using different modelling assumptions, which are combined within an ensemble of sub-models that acknowledge uncertainty in how processes such as treatment access and immune decay occur (Smith et al. 2012b).

The comparison of model differences above highlights the need to consider the question that is being answered when developing a model. In the work presented in this thesis, I will be producing a new model based on the Imperial College model that is capable of simulating parasite genetics.

Consequently, the consideration of how other models have chosen to model within host parasite dynamics is of particular merit. For example, within host parasite densities will have an impact on which parasite strains are taken up in a mosquito feed, which directly impacts the generation of new genetic diversity as a result of potential recombination in the sporogenic cycle. Additionally, tracking parasite strain types will need to be incorporated so that the advantages and fitness costs of drug resistance can be included. There are, however, other models of parasite dynamics that are not full transmission models but give insight into how parasite genetic diversity occurs and also how resistant strains may be expected to emerge under different conditions.

1.4. Models of malaria genetics

Mathematical models of malaria parasite genetics have largely been developed to answer questions related to the emergence of drug resistance. However, a handful of studies have built models aimed at understanding the occurrence of neutral genetic variation and often the spatial distribution of genetic variance. These models largely draw from modelling approaches within the field of population genetics, rather than using epidemiological models. Despite the limited number of studies using population genetics approaches, genetic methods are increasingly thought of as a necessary tool for malaria surveillance going forwards. For example, a technical consultation was proposed in October 2018 to the WHO Malaria Policy Advisory Committee to assess the role of genetics in optimising surveillance programs (Noor and Ringwald 2018).

1.4.1. Population Genetic Models

Population genetics seeks to use models to explain the genetic diversity within and between different populations. Its application in malaria has largely focussed on understanding and detecting selection within populations, which has immediate implications for antimalarial resistance, which are discussed in the next section. The other main applications of population genetics for malaria include evolutionary analyses and demographic inference, including inferring how population sizes have changed over time and how related different populations are to one another in space.

One application of population genetics in malaria has been to explain the evolutionary history of the *Plasmodium* genus. These approaches have used samples from different *Plasmodium* species to estimate the genetic distances between species before constructing phylogenetic trees to infer the relationship between species. Phylogenetic approaches were used in the reclassification of *P. ovale* as two distinct species*, P. ovale curtisi* and *P. ovale wallikeri* (Sutherland et al. 2010), which has been verified more recently with further phylogenetic studies (Ansari et al. 2016). These analyses also offer insight into possible zoonotic origins of different malaria species and help identify invasion related

genes, such as the RBP family (Rutledge et al. 2017). Another use of phylogenetics is to link it with epidemiological models, often based on coalescent theory (Kingman 2000) or birth-death models (Volz 2012), to infer the effective population size. This has particular benefits for the study of infectious diseases as it promises to be able to allow us to infer the historic prevalence of the disease. For malaria this would be very useful, due to the difficulties in surveillance efforts resulting from missed low density parasite infections (Okell et al. 2009) and an inability to conduct surveillance effectively in some regions. However, the application to malaria is complicated by frequent recombination, which is poorly handled in most models incorporated in phylodynamic approaches. One of only a few examples in malaria, which used microsatellite models of genetic distances, failed to detect changes in the effective population size in a population in Turbo, Columbia over an 8-year period of increased treatment and decreased cases of malaria (Chenet et al. 2015). These issues of the unsuitability of commonly used genetic models for malaria was highlighted in a study investigating unusual patterns of SNPs (Chang et al. 2013). Under the Wright-Fisher model, one would expect that changes in gene frequencies due to random drift in finite populations to reduce natural selection, a process which is expected to occur due to competition between strains and selection towards a virulent phenotype. However, the frequent population bottlenecks occurring at both transmission stages appears to impact the commonly accepted dogma of a trade-off existing between genetic drift and selection, with both processes being enhanced in the life cycle.

The other use of population genetics approaches in malaria is to assess for population structure in space and to assess for links between different populations. Approaches using genetic distances based on microsatellites have been used to quantify movement of malaria in the DRC, revealing a complicated landscape of both close genetic ties between geographically distant parasites and isolation by distance (Carrel et al. 2015). In lower prevalence countries, however, alternative measures are required to quantify spatial connectivity due to parasites appearing increasingly similar. Measures of genetic distances based on identity by descent (IBD), which unlike measures of identity by state (IBS), explicitly account for the genetic markers used by conditioning on allele frequencies, were shown to be able to relate genetic distance to spatial distances in Thailand (Taylor et al. 2017). A recent opinion article highlighted the differences needed in genetic approaches with respect to malaria prevalence. The article suggested that these approaches will increasingly be used as malaria control efforts change the landscape of malaria prevalence away from one of high prevalence homogeneity to one where hotspots and source-sink dynamics of connected populations are more prevalent (Wesolowski et al. 2018). With these changes, different approaches to malaria control can be taken that focus interventions and ensure that imported cases of malaria vs local cases are detected accordingly as has been used in Namibia (Tessema et al. 2019) and Bangladesh (Chang et al. 2019) by

combining genetic data with mobile phone and travel survey data. Central to these approaches, however, are the developed statistical approaches for detecting spatial patterns within proximal populations. However, analyses of malaria at larger scales has used more standard methodology to understand malaria biology (Volkman et al. 2012), with analysis of microsatellites used to identify the source of outbreaks of malaria (Patel et al. 2014). At these spatial scales model-based Bayesian methods, such as STRUCTURE (Pritchard et al. 2000; Lawson et al. 2012), have been used to assess population structure data and helped to identify the sources of resistance to both RDTs (Akinyi et al. 2013; Murillo Solano et al. 2015) and antimalarial drugs (Roper et al. 2004; Miotto et al. 2013).

1.4.2. Antimalarial Resistance Models

The development of antimalarial drugs is frequently guided by mathematical models of within host pharmacokinetics and pharmacodynamics of antimalarial compounds (Simpson et al. 2014). These models have been used to assess the therapeutic efficacy (Tarning et al. 2013) and decay of antimalarial compounds to inform both the optimisation of dosing schedules (WWARN DP Study Group 2013) for available antimalarials and to guide antimalarial drug development (Slater et al. 2017). However, the occurrence of antimalarial resistance has necessitated considering the implications of drug-based control strategies upon the emergence of antimalarial resistance, especially with the increase in the use of drugs within SMC and IPT (O'Meara et al. 2006). Initially, mathematical models of malarial genetics were utilised to understand both the emergence and spread of antimalarial resistance (Mackinnon 2005). The former models the emergence of resistant mutations using within host parasite dynamics under different assumptions concerning the fitness and selective advantage of emerging strains (Antao and Hastings 2011). The latter, however, often assumes that the resistant strain has emerged and seeks to determine the impact of population demographic elements, such as disease prevalence, host immunity and drug treatments, upon the spatiotemporal dynamics of the spread of resistance (Smith et al. 2010b). Recent efforts have been made to incorporate both modelling scales within one framework with the concomitant modelling of resistance evolution both within and between hosts yielding important insights into the suitability of different strategies to preserve the therapeutic longevity of available antimalarials (Nguyen et al. 2015). However, the realism of either the transmission models or the genetic evolutionary process has been limited in these models, with the representation of the parasite lifecycle within the mosquito often simplified.

1.4.2.1. Within-host models

Within-host models choose to focus on the dynamics of the parasite population within an individual and not focus on the transmission of resistant parasites to mosquitoes. As a result, these models focus

on identifying the factors that accelerate the emergence of a new resistant parasite within a single infection resulting from mutation. For example, by considering the number of distinct parasite generations and the numbers of parasites generated at each stage of the parasite life cycle, the de novo probability of resistance arising within each of three stages of a parasite life cycle can be calculated (Pongtavornpinyo et al. 2009). Alternatively, considering the timing of treatment with respect to the stage of infection has been used to reveal interactions between how the timing of host immunity and drug pressure may increase the likelihood of developing viable drug-resistant populations (Gatton et al. 2001).

As mentioned earlier, these models often use pharmacokinetic and pharmacodynamic approaches to address resistance emergence (Saralamba et al. 2011) and have been used to recommend increasing assessment of resistance emergence to greater than 28 days after treatment (Simpson et al. 2000). Lastly, approaches similar to those used in population genetics have been used by considering the frequency of parasites that would survive from one infection in order to populate the population frequency of resistant alleles over time. These can be used to address questions of emergence (Hastings 1997), as well as the transmission of emergent resistance, for example demonstrating the necessity of using antimalarials within combination therapies and that adding second drugs to already failing antifolate drugs is unlikely to work and will lead to partner drug resistance (Watkins et al. 2005).

1.4.2.2. Population models

Population models focus on modelling how resistant strains spread within a population of individuals and consequently place more focus on how control interventions impact these processes. One major area of research is in the review of how different long-term drug treatment strategies can be used to slow the spread of resistance. For example, compartmental models of malaria epidemiology have been used to contrast using multiple first line therapies (MFT) against cycling available drug treatments (Boni et al. 2008). In this study a number of evaluation metrics were also introduced that enabled strategies to be assessed based on treatment failure rates as well as population allele frequencies related to resistance. This study predicted that MFT has reduced treatment failures over a 20-year period compared to cycling strategies, explaining that MFT creates a varied drug pressure environment that slows the path to resistance emergence and that the rate at which resistance is lost is not sufficient to cause resistance frequencies to return to pre-emergence levels. Similar compartment models have also been used to look at whether increased or decreased drug pressure slows the spread of resistance, while also looking at how MFT can be used to slow the increase in the fitness of resistant parasites in response to drug pressure (Smith et al. 2010b). A version of the Imperial College model was also used to explore resistance, but primarily from the point of view of its impact

on malaria prevalence, suggesting that partner drug resistance is likely to cause higher increases in morbidity than artemisinin resistance (Slater et al. 2016).

Population models have also been developed to explore questions related to the role of sexual recombination and multiplicities of infection in resistance. One investigation sought to focus on the relationship between transmission intensity, superinfection and fitness costs to highlight that fitness costs must both consider the impacts on the length of an infection as well as the probability of onwards transmission within mixed infections (Klein et al. 2012). With regards to the role of recombination, a population genetics approach that included recombination occurring from multiply infected individuals in the population highlighted that whether MFT was better than cycling was dependent on transmission intensity (Antao and Hastings 2012). Additionally, the same model was used to demonstrate that conclusions drawn concerning resistance emergence need to be made by models including any known epistatic relationships between the mechanisms of resistance (Antao and Hastings 2011). For example, there is evidence to suggest that some genes are involved in resistance to multiple drugs, such as *P. falciparum* multidrug resistance gene-1 (*pfmdr1*) (Okell et al. 2018a) and members of the ATP-binding cassette family (Koenderink et al. 2010), with notable difference in drug failure rates dependent on different combinations of particular SNPs and copy number variations (Nwakanma et al. 2014). Consequently, including both synergistic and antagonistic behaviour between genetic loci associated with resistance is key in any resistance modelling.

1.4.2.3. Within-host and Population Models

The inclusion of both within host dynamics and population level transmission enables the realism of both modelling scales to be included within one framework, which enables the interaction between the two scales to occur. For example, one model was developed specifically for this purpose by comparing the influence of transmission dynamics and within host evolution of resistance. The process demonstrated the need for accurate parameterisation of both processes, highlighting that both treatment efficacy and coverage are related to time until resistance emergence (Legros and Bonhoeffer 2016). Similarly, both modelling scales were included in a different model that focussed on the predicting the waiting time for resistance to emerge, be onwardly transmitted and lead to a recombination event that produces a doubly-resistant parasite (Kim et al. 2014). Other examples have been used to highlight the differential impact of chronic asymptomatic infections vs short-lived symptomatic treated infections on the speed of resistance emerging (Chang et al. 2016). Lastly, the OUCRU model that was discussed in the review of transmission models of malaria includes both realistic within host modelling, tracking parasite densities and multiple strains, as well as a full transmission model that reproduces malaria epidemiological dynamics (Nguyen et al. 2015). This

model was used to explore MFT vs cycling, with three different ACTs being used. Due to parasite genotypes being modelled, the shared resistance mechanism associated with artemisinin resistance could be accurately captured. Resistance was, however, assumed to always lead to treatment failure rather than conferring partial resistance and probabilistic treatment failure, which may impact findings related to the superior performance of MFT vs cycling.

1.4.3. Brief lessons from antibiotic resistance modelling

In Smith et al. the question was raised as to whether increased or decreased drug pressure is advantageous in preventing emergence of antimalarial resistance (Smith et al. 2010b). Unsurprisingly, the same question is under active consideration within the wider antimicrobial resistance community. Historic belief is that high drug pressure is key in preventing resistant strains emerging, with imperfect adherence to drug regimens and inappropriate antimicrobial prescription increasing the probability of resistance (Llor and Bjerrum 2014). However, mathematical models have highlighted that this is largely dependent on the relationship between the strength of the resistance mechanism, i.e. complete vs partial resistance, and the strength of inter-strain competition, and that subsequently the best strategy is to either use the highest tolerable drug dose or the lowest clinically effective dose (Day and Read 2016). Modelling of antibiotic resistance has also explored questions of cycling drugs vs MFT. Initial modelling efforts suggested that MFT is superior to drug cycling, with cycling failing to reduce antimicrobial resistance in hospitals (Bergstrom et al. 2004), with similar explanations as to why as suggested in Boni et al (Boni et al. 2008). However, more recent modelling that explored a greater suite of plausible cycling vs MFT strategies, which also included greater variation in the performance and properties of the available antibiotics, suggested it is not as simple as that, with a number of cycling strategies shown to outperform MFT (Beardmore et al. 2017). As before, key questions related to the properties of the drugs in question, the inter-strain dynamics and fitness advantages and costs associated with resistance can cause conclusions to lean each way. This appreciation prompts me to consider two paths going forwards when extending malaria transmission models for simulating resistance dynamics: 1) use parameter sets and modelling assumptions that closely resemble our understanding of resistance, or 2) in the absence of known quantities exhaustive sensitivity analysis and parameter sweeping is required to review all possibilities and assign confidence to the settings that are most probably in an ensemble fashion. Path 2) is undoubtedly most likely, however, with each additional parameter to explore, the computational tractability of this approach diminishes and the merits of a hybrid approach should be considered.

1.4.4. *pfhrp2* resistance models

Mathematical models of malaria transmission have also been developed that consider within host parasite profiles in response to reports of false-negative RDT results due to *P. falciparum hrp2/3* gene deletions (Parr et al. 2016). This approach has been used to characterise the drivers and potential spread of *P. falciparum* hrp2 gene deletions at a population level (Watson et al. 2017), whereas alternative modelling methodologies have focussed on explicitly modelling parasite antigen levels within individuals to consider the potential impact on malaria morbidity (Gatton et al. 2017). Both scales of modelling offer differing insights into the impact of changes in parasite genetics on the efficacy of malaria control interventions and are similarly used when modelling the impact of antimalarial resistance upon control interventions.

1.4.5. Models for inferring malaria transmission intensity

The inclusion of neutral genetic information within population level transmission modelling of malaria has been largely unexplored, although such methods are increasingly being developed for modelling other pathogens (Rasmussen et al. 2014; Kuhnert et al. 2016). One notable example, however, was conducted by Daniels et al. that incorporated parasite genetic barcodes in an agent-based model to infer the effect of intervention strategies upon malaria transmission intensity within Thiès, Senegal (Daniels et al. 2015). This study illustrated the potential for genetic data to add value to other surveillance methods in areas with less developed surveillance systems (Greenhouse and Smith 2015). However, before this can happen the generalisability and accuracy of these methods needs to be validated.

1.5. Thesis Aims

The aim of this thesis is to use mathematical models of malaria transmission that track parasite populations and enable the simulation of parasite genetics to address a range of issues – including how parasite genetics could be used within surveillance contexts as well as exploring the dynamics of antimalarial and diagnostic resistance. Alongside gaining insight in these areas, I aimed to support health policy by considering how best to implement surveillance programs to estimate the prevalence of gene deletions responsible for diagnostic resistant malaria, as well as exploring how current treatment strategies could be adapted to slow the emergence of antimalarial resistance. Specific objectives of each chapter are:

- **Chapter 2:** Extend my previous work looking at *pfhrp2* gene deletions to identify the optimum sampling windows in support of new WHO guidance on estimating the prevalence of *pfhrp2/pfhrp3* gene deletions.
- **Chapter 3:** Develop a new model of malaria transmission for flexible modelling of parasite genetic traits and apply the model to explore the role of superinfection and co-transmission events towards within host parasite genetic diversity.
- **Chapter 4:** Use the model developed in chapter 3 to define the useful limits of malaria genetics for inferring changes in malaria prevalence in response to the scale up of intervention strategies and conclude by developing a statistical model for translating parasite sequence data into estimates of malaria prevalence.
- **Chapter 5:** Adapt the model developed in chapter 3 to explore the dynamics of antimalarial resistance emergence, conducting an analysis of how the reduction in the over-prescription of antimalarials could be used to improve the therapeutic lifespan of current first line antimalarials.
- **Chapter 6:** Analyse laboratory data using generalised linear mixed-effects models to characterise the contribution of artemisinin resistance towards reducing the impact of artemisinin on the parasite's ability to infect mosquitoes.. Incorporate these findings into the transmission model to characterise if a subsequent transmission advantage is conferred by artemisinin resistance when resistant parasites invade susceptible populations.

Throughout this thesis I aim to demonstrate that mathematical models are a useful tool for understanding how parasite genetic information can be used to inform emerging challenges within malaria control, and the policy implications of these results in light of resistance emergence.

Chapter 2. Implications of seasonal malaria transmission for the detection of *pfhrp2/3* gene deletions

The emergence of "diagnostic resistant" malaria resulting from the deletion of the genes that encode for histidine-rich proteins poses a new threat to the control of *P. falciparum* malaria. Eleven countries have now reported *pfhrp2/pfhrp3* gene deletions since the first observation of *pfhrp2*-deleted parasites in 2012. In this chapter I present an extended version of my previous study that characterised the drivers selecting for *pfhrp2/3* deletions and mapped the regions in Africa with the greatest selection pressure (Watson et al. 2017). I explore how the timing of surveillance programmes recommended by the World Health Organization in February 2018 could cause premature decisions to switch to alternative RDTs.

This chapter has been published as: **Watson OJ, Verity R, Ghani AC, Garske T, Cunningham J, Tshefu, A, Mwandagalirwa, MK, Meshnick SR, Parr JB, Slater HC.** Impact of seasonal variations in Plasmodium falciparum malaria transmission on the surveillance of pfhrp2 gene deletions. eLife 2019. *8*, e40339 (Watson et al. 2019b).

2.1. Introduction

Diagnostic testing of suspected malaria cases has more than doubled in the last 15 years, with 75% of suspected cases seeking treatment from the public health sector receiving a diagnostic test in 2017 (World Health Organization 2018d). Much of this progress reflects the increased distribution of rapid diagnostic tests (RDTs), with the most commonly used RDTs targeting the *P. falciparum* protein HRP2 (PfHRP2). In 2014, a review of published reports of *pfhrp2/3* deletions was conducted and included a critical assessment of the comprehensiveness of the diagnostic investigation (Cheng et al. 2014). The findings of this review highlighted a need for a harmonized approach to investigating and confirming or excluding *pfhrp2/3* deletions and called for further studies to determine the prevalence and impact of *pfhrp2/3* gene deletions. Since that review, false-negative RDT results due to *pfhrp2/3* gene deletions have been reported in 10 countries in sub-Saharan Africa (SSA) (World Health Organization 2018a). The frequency of *pfhrp2/3* deletions varies across SSA, with the highest burden observed in Eritrea where 80.8% of samples from Ghindae Hospital were both *pfhrp2*-negative and *pfhrp3*-negative in 2016 (Berhane et al. 2018).

Mathematical modelling has predicted that the continued use of only PfHRP2 RDTs will quickly select for parasites without the *pfhrp2* gene (Gatton et al. 2017). This selection pressure occurs due to the

misdiagnosis of infections caused by parasites lacking the *pfhrp2* gene, which will subsequently contribute more towards onwards transmission than wild type parasites that are correctly diagnosed due to the expression of *pfhrp2.* In 2017, I conducted an analysis of the drivers of *pfhrp2* gene deletion selection, identifying the administrative regions in SSA with the greatest potential for selecting for *pfhrp2-deleted* parasites (Watson et al. 2017). The regions identified were areas with both a low prevalence of malaria and a high frequency of people seeking treatment and being treated on the basis of PfHRP2-based RDT diagnosis. The precise strength of selection, however, is not known with other factors such as the rate of non-malarial fevers and non-adherence to RDT outcomes likely to impact the number of misdiagnosed cases receiving treatment.

In February 2018, the World Health Organization (WHO) issued guidance for national malaria control programmes on how to investigate suspected false-negative RDTs with an emphasis on *pfhrp2/3* gene deletions (World Health Organization 2018b). The primary study outcome to be calculated in the guidance is as follows:

$$\text{Proportion of } \textit{P. falciparum} \text{ cases with false-negative HRP2 RDT results due to } \textit{pfhrp2/3} \text{ deletions} = \frac{\text{\# of confirmed } \textit{falciparum} \text{ patients with } \textit{pfhrp2/3} \text{ gene deletions and HRP2 RDT negative results}}{\text{\# of confirmed } \textit{P. falciparum} \text{ cases (by either RDT or microscopy)}}$$

The guidance recommends that a national change to non PfHRP2-based RDTs should be made if the estimated proportion of *P. falciparum* cases with false-negative HRP2 RDT results due to *pfhrp2/3* deletions is above 5%. If the estimated proportion is less than 5% the country is recommended to establish a monitoring scheme whereby the study is repeated in two years if the 95% confidence interval does not include 5%, or one year if it does include 5%. The 5% threshold approximates the point at which the number of cases missed due to false-negative PfHRP2-based RDTs caused by *pfhrp2/3* deletions may become greater than the number of cases that would be missed due to the decreased sensitivity of non PfHRP2-based RDTs. The guidance also specifies a sampling scheme to be used when estimating the prevalence of *pfhrp2/3* gene deletions. Samples are to be collected from at least 10 health facilities per province to be tested, with sampling focussed on symptomatic *P. falciparum* patients presenting at the health facilities. All sampling is to be ideally completed within an 8-week period.

The 8-week interval permits for a rapid turnaround and allows for efficient investigations and policy responses. However, the timing of the 8-week interval chosen within a transmission season is important. The chosen interval could lead to estimates of the proportion of *P. falciparu*m cases with false-negative HRP2 RDT results due to *pfhrp2/3* deletions that are not representative of the annual

average proportion. Subsequently, any recorded estimate may not be predictive of the number of cases that may be misdiagnosed due to *pfhrp2/3* deletions in the years between sampling intervals. For example, an overestimation of the annual average proportion of false-negative RDTs due to *pfhrp2/3* deletions could result in a switch to a less sensitive RDT, resulting in an increase in the number of malaria cases misdiagnosed if the annual average proportion of false-negative RDTs due to *pfhrp2/3* deletions is less than 5%. The alternative RDT may also be both more expensive and complicated to implement. Similarly, an underestimation of the annual average proportion of *P. falciparum* cases with false-negative HRP2 RDT results due to *pfhrp2/3* deletions would result in continued use of an overall less effective test and could provide *pfhrp2/3* deleted parasite populations an opportunity to expand.

In response to these concerns, I have extended my original methods (Watson et al. 2017) to characterise the impact of seasonal variations in transmission intensity on the proportion of false-negative RDTs due to *pfhrp2*-deleted parasites. The extended model predicts that more false-negative RDTs due to *pfhrp2* gene deletions are observed when monoclonal infections are more prevalent, with the highest proportion observed when sampling from younger children at the start of the rainy season. I continue to assess how samples collected within an 8-week interval can both over- and underestimate this proportion when compared to the annual average, which reflects the monitoring scheme recommended by the WHO for follow up studies if the outcomes of the original study are inconclusive. Lastly, I map the administrative regions in SSA with the greatest potential for estimates of the proportion of *P. falciparum* cases with false-negative HRP2 RDT results due *to pfhrp2* deletions to be not predictive of the annual average. In addition, I identify the optimum sampling intervals for each level one administrative region, which are most representative of the annual average.

2.2. Methods

2.2.1. Extensions to the *P. falciparum* transmission model

In my previous publication, I presented an extended version of an individual-based model of malaria transmission to characterise the key drivers of *pfhrp2* deletion selection, however it did not capture seasonality. To address this I incorporated seasonal variation in malaria transmission intensity through the inclusion of seasonal curves fitted to daily rainfall data available from the US Climate Prediction Center (2010). Rainfall data was available at a 10x10km spatial resolution from 2002 to 2009, with data missing for only 2 days. The data was subsequently aggregated to a series of 64 points per year, before Fourier analysis was conducted to capture the seasonal dynamics within this time period (Cairns et al. 2012). The first three frequencies of the resultant Fourier transformed data were used

to generate a normalised seasonal curve. This inclusion alters the rate at which new adult mosquitoes are born, with the differential equation governing the susceptible adult stage of the mosquito population now given by:

$$\frac{dS_M}{dt} = \theta(t)\mu_M M_v - \mu_M S_M - \Lambda_M S_M$$

where μ_M is the daily death rate of adult mosquitoes, M_v is the total mosquito population, i.e. $S_M + E_M + I_M$, Λ_M is the force of infection on the mosquito population and $\theta(t)$ is the normalised seasonal curve, with a period equal to 365 days.

2.2.2. Characterising the impact of seasonal transmission intensities upon the proportion of false-negative RDTs due to *pfhrp2* gene deletions

The impact of seasonality was examined by recording the proportion of clinical incidence that would have been misdiagnosed due to *pfhrp2* gene deletions across the year. This proportion was summarised at twelve 8-week intervals, i.e. January – March, February – April ... December – February. This proportion was recorded in both a high and low seasonality setting, characterised by a Markham Seasonality Index = 80% and 10% respectively (Cairns et al. 2015). These settings were examined at both a low and moderate transmission intensity (EIR = 1 and 10 respectively), with the starting proportion of *pfhrp2*-deleted parasites in the whole population set equal to 6% in agreement with previous observations of *pfhrp2* gene deletions in the DRC (Watson et al. 2017). The proportion of symptomatic cases seeking treatment was assumed to be 40% (f_T = 0.4). In all simulations, ten stochastic realisations of 100,000 individuals were simulated for 60 years to reach equilibrium first, before setting the frequency of *pfhrp2* deletions. Initially, I assumed there was no assumed fitness cost or selective advantage associated with *pfhrp2* gene deletion. This was modelled by assuming that individuals who are only infected with parasites with *pfhrp2* gene deletions will still be treated. This decision allowed us to control for selection within my investigation by ensuring that the changes observed in the observation of PfHRP2-negative clinical cases are only due to seasonal variation in transmission intensity, and not due to an increase in the frequency of *pfhrp2* gene deletions due to the selective advantage by evading diagnosis. As a result, when reporting the proportion of clinical cases that were misdiagnosed resulting from a false-negative PfHPR2-negative RDT I am reporting the proportion of cases that are infected with only *pfhrp2*-deleted parasites, i.e. individuals who would have been *pfhrp2*-negative and subsequently misdiagnosed. I also assume that 25% of individuals who are only infected with *pfhrp2*-deleted parasites will still be *pfhrp2*-positive due to the cross reactivity of PfHRP3 epitopes causing a positive PfHRP2-based RDT result (Baker et al. 2005).

Model predictions were subsequently compared to data collected from the Democratic Republic of Congo (DRC) as part of their 2013 – 2014 Demographic and Health Survey (DHS). In overview, 7,137 blood samples were collected from children under the age of 5 years old, which yielded 2,752 children diagnosed with *P. falciparum* infection by real-time PCR targeting the lactate dehydrogenase (*pfldh*) gene. The RDT barcodes for the 2,752 samples were identified and matched to the DHS survey to identify both the age of the children and the date of sample collection. The collection date was used to predict the mean clinical incidence from the previous 30 days for each sample. This was estimated using the deterministic implementation of my model fitted to the observed PCR prevalence of malaria from the DRC DHS 2013-2014 survey (Meshnick et al. 2015), incorporating the seasonality and treatment coverage for each province. Children who were younger than the median age in the 2,752 samples were grouped within a younger category. In addition, samples were classified as lower transmission if the clinical incidence of malaria in the month prior to sample collection was lower than the median clinical incidence. The counts of *pfhrp2*-negative samples within each group were subsequently compared using the Pearson chi-squared test with Rao-Scott corrections to account for the hierarchal survey design implemented within DHS surveys (Rao and Scott 1984). Pearson chi-squared tests were used in a similar analysis that was conducted using samples collected from the Gash-Barka and Debug regions in Eritrea between 2013 – 2014, for which the dates of sample collection were made available to me (Menegon et al. 2017).

Finally, the seasonal profiles for 598 first level administrative regions across sub-Saharan Africa were used to characterise the potential for estimates of the proportion of false-negative PfHRP2 RDTs due to *pfhrp2* gene deletions to be unrepresentative of the annual average. For each region, 100 simulation repetitions were conducted for 60 years to reach equilibrium first before fitting the frequency of *pfhrp2* gene deletions in each simulation such that the annual average proportion of false-negative RDT results due to *pfhrp2* deletions is equal to 5%. Each repetition was subsequently simulated for two further years, with 7,300 individuals seeking treatment sampled from each 8-week interval. This number approximates the recommended sample size within the WHO protocol for *pfhrp2* deletion prevalence at 5% ± 0.5%. For each sample the proportion of false-negative PfHRP2-based RDTs due to *pfhrp2* gene deletions was recorded. For each sample a binomial confidence interval was calculated and the resultant percentage of intervals that did not include the annual prevalence of 5% was calculated. For each region the number of 8-week intervals for which a premature decision to either swap from a PfHRP2-based RDT or continue using a PfHRP2-based RDT was made in more than 75% of simulations was recorded and mapped. The raw results of this analysis were subsequently used to create a database that details the optimum sampling intervals for estimating the annual proportion of false-negative RDT results due to *pfhrp2* deletions.

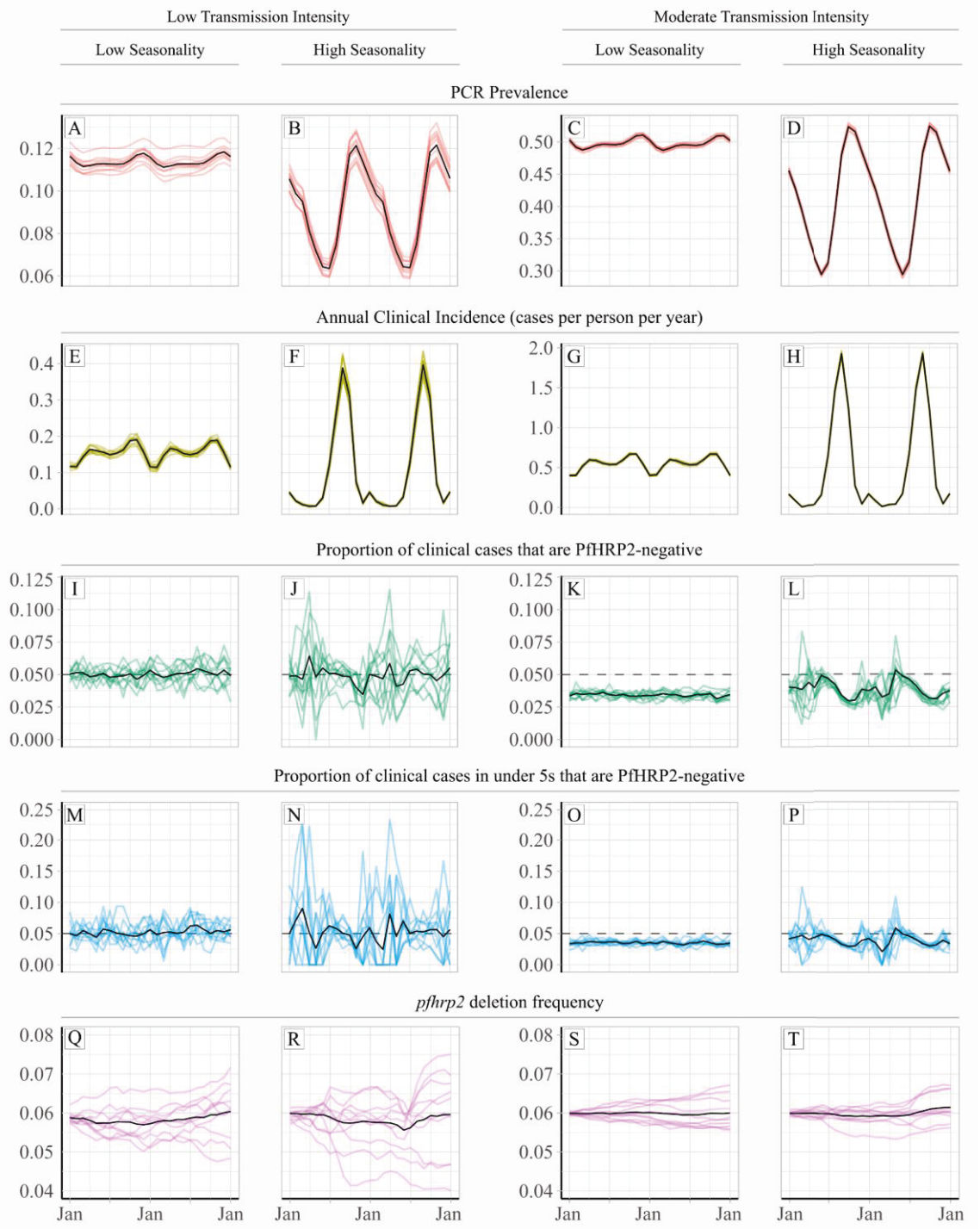


Figure 2.1 Relationship between seasonality, transmission intensity and proportion of clinical cases that are infected with only *pfhrp2*-deleted parasites. Graphs show in A – D and E - H the model predicted PCR prevalence and annual clinical incidence respectively at both a low and a moderate transmission intensity. In I – L and M - P the proportion of clinical cases only infected with pfhrp2-negative parasites is shown for both the whole population and in children under 5 years old respectively. Lastly, graphs Q - T show the population allele frequency of pfhrp2 gene deletions, which was set equal to 6% at the beginning of each simulation. 10 simulation realisations are shown in each graph, with the mean shown with by the black line. Lastly, the 5% threshold for switching RDT provided by the WHO is shown with the dashed horizontal line in plots I – P.

2.3. Results

Using the model, I first explored how the proportion of clinical cases only infected with *pfhrp2*-deleted parasites varies throughout a transmission season. I recorded the proportion of clinical cases that are PfHRP2-negative in four settings (a low and moderate transmission setting with both a low and highly seasonal transmission dynamic), which had a starting *pfhrp2* deletion frequency of 6%. 6% was chosen to reflect the previously estimated frequency of *pfhrp2* deletions prior to the introduction of RDTs in DRC (Watson et al. 2017). I initially assumed that the frequency of *pfhrp2* deletions was not increasing over time before considering scenarios in which the selective pressure for *pfhrp2* deletions causes an increase in the population frequency of *phrp2* deletions. This decision allowed for the impact of seasonality on the proportion of clinical cases that are *pfhrp2*-negative to be isolated, before allowing comparisons to scenarios in which the proportion of clinical cases that are *pfhrp2*-negative is increasing also due to changes in the population frequency of *phrp2* deletions.

My predictions suggest that the misdiagnosis of clinical cases due to *pfhrp2*-negative RDT results is heavily dependent on transmission intensity (Figure 2.1). For the same population frequency of *pfhrp2* gene deletions (Figure 2.1Q-T), the observed proportion of clinical cases that are *pfhrp2*-negative is predicted to be higher in lower transmission settings (Figure 2.1I-P).

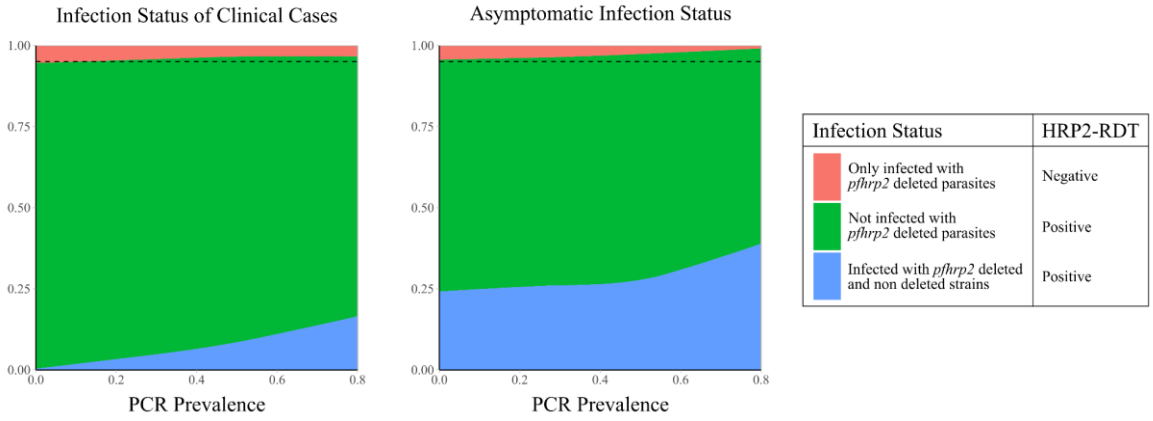


Figure 2.2 Model predicted relationship between clonality of infection in asymptomatic and clinical cases against prevalence of malaria. The graphs show the proportion of clinical cases of malaria and asymptomatic individuals that are infected with only pfhrp2-deleted parasites, not infected with pfhrp2-deleted parasites or polyclonally infected (infected with both pfhrp2-deleted parasites and wild type parasites). These proportions are shown with respect to increasing prevalence of malaria, identifying the increased proportion of individuals that are polyclonally infected at high transmission intensities. For both plots, the population frequency of pfhrp2 gene deletions in the population is 5%, which is shown by the dashed black line. Consequently, at very low PCR prevalence, when the majority of infections contain only one infecting strain and the proportion of polyclonal infections tends to zero, the observed prevalence of the deletions will tend to the population frequency of 5%.

The annual average proportion of clinical cases that are *pfhrp2*-negative was equal to 5% and 3.25% in the low and moderate transmission setting respectively. This observation is attributable to the lower rate of superinfection in low transmission settings. The lower rate of superinfection reduces the number of polyclonal infections and increases the chance that an individual is only infected with *pfhrp2*-negative parasites (Figure 2.2). When I considered scenarios with a selective advantage for *pfhrp2*-deletions (Figure 2.3), the population frequency of *pfhrp2* gene deletions increased over the two years observed (Figure 2.3Q-T) with a corresponding increase in the proportion of clinical cases that are *pfhrp2*-negative (Figure 2.3I-P).

An increased proportion of individuals only infected with *pfhrp2* gene deletions is predicted to occur at the beginning of the rainy season just before incidence starts to increase. During the rainy season, the observed proportion of cases expected to yield a false-negative RDT due to *pfhrp2*-deleted parasites (PfHRP2-negative) falls, with the lowest proportion observed after the end of the rainy season. These dynamics are more pronounced in highly seasonal transmission regions (Figure 2.1B, 1F, 1J, 1N, 1R and 1D, 1H, 1L, 1P, 1T). In the highly seasonal settings the observed proportion of clinical cases that are PfHRP2-negative is predicted to fluctuate above and below the 5% threshold for switching RDT provided by the WHO (Figure 2.1J, 1L, 1N and 1P). Smaller fluctuations are seen in less seasonal transmission regions (Figure 2.1A, 1E, 1I, 1M, 1Q and 1C, 1G, 1K, 1O, 1S), with no fluctuations in the observed proportion of clinical cases that are PfHRP2-negative occurring above 5% in the moderate transmission setting (Figure 2.1K and 1O). Similar patterns were observed in scenarios with an increasing frequency of *pfhrp2*-deletions, with fluctuations in the proportion of clinical cases that were PfHRP2-negative observed in the highly seasonal settings (Figure 2.3J, 3L, 3N and 3P). The highest proportion of cases expected to yield a false-negative RDT due to *pfhrp2*-deleted parasites was still observed at the beginning of the rainy season.

The specific 8-week interval during which samples are collected is predicted to impact the observed proportion of false-negative RDTs due to *pfhrp2* gene deletions (Figure 2.4). In a moderate transmission setting, a clear seasonal pattern is predicted (Figure 2.4C), with sampling at the beginning of the transmission seasons resulting in significant overestimation of the annual average proportion of false-negative RDTs. Subsequently, sampling at the end of the rainy season is predicted to yield estimates that are most representative of the annual average. In comparison, surveillance in regions with low seasonality is predicted to yield estimates representative of the annual average throughout the transmission season (Figure 2.4B and 2D). In all settings, using a sampling scheme spanning the entire transmission season produced estimates that accurately estimated the annual average. A moderate increase in the proportion of false-negative RDTs is also predicted when sampling younger individuals, with the same patterns also seen within asymptomatic individuals. This observation

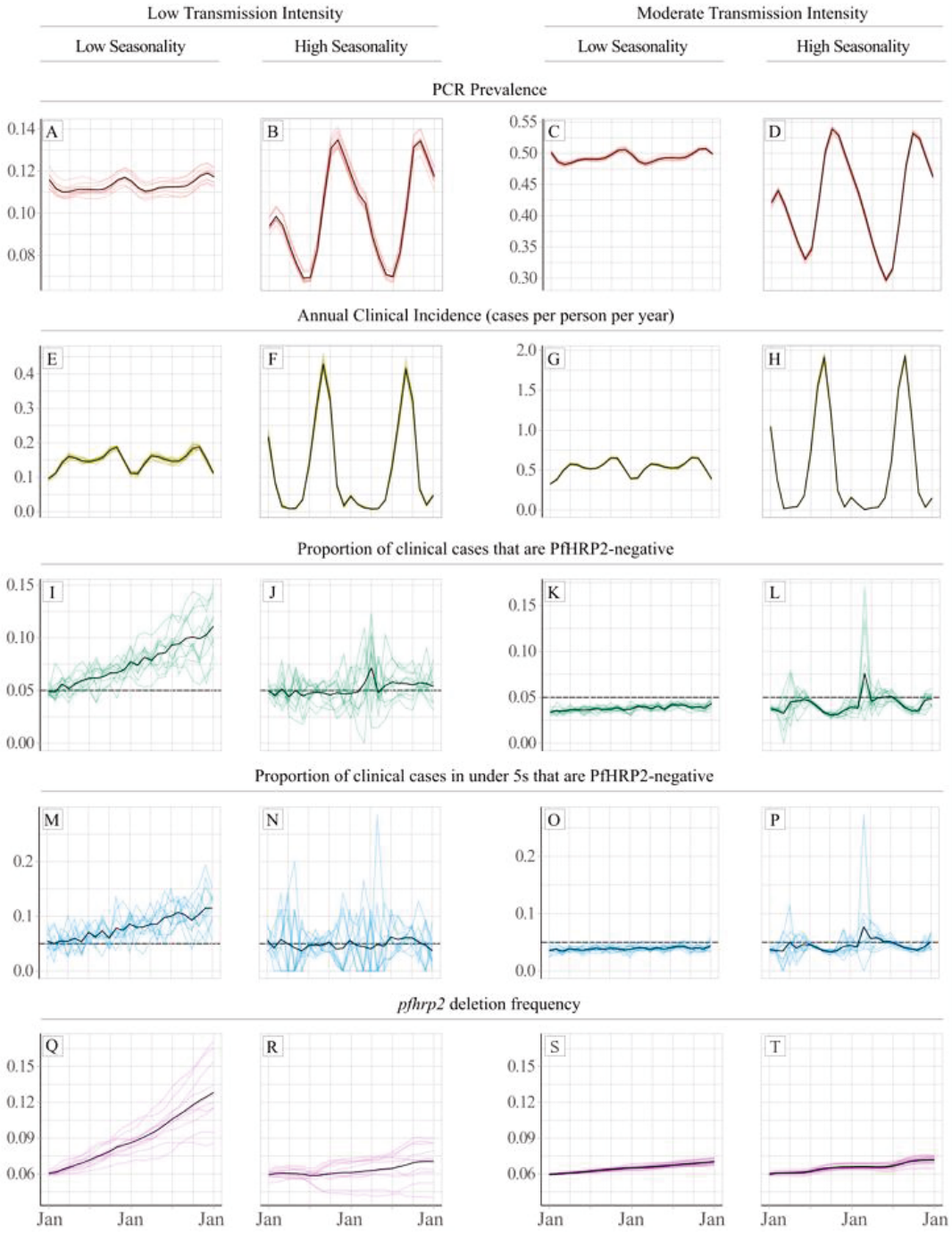


Figure 2.3 Impact of a selective advantage for *pfhrp2*-deleted parasites on the relationship between seasonality, transmission intensity and proportion of clinical cases that are infected with only pfhrp2-deleted parasites. Graphs show in A – D and E - H, the model predicted PCR prevalence and annual clinical incidence respectively at both a low and a moderate transmission intensity. In I – L and M - P the proportion of clinical cases only infected with *pfhrp2*-negative parasites is shown for both the whole population and in children under 5 years old, respectively. Lastly, graphs Q - T show the population allele frequency of pfhrp2 gene deletions, which was set equal to 6% at the beginning of each simulation before increasing due to selection. 10 simulation realisations are shown in each graph, with the mean shown with by the black line. Lastly, the 5% threshold for switching RDT provided by the WHO is shown with the dashed horizontal line in plots I – P.

reflects the increased probability that children younger than five years old yield symptoms after the first infection, due to their comparatively lower acquired clinical immunity. Similar seasonal dynamics were observed in the highly seasonal settings when I considered scenarios with a selective advantage for *pfhrp2*-deletions (Figure 2.5A, 5C).

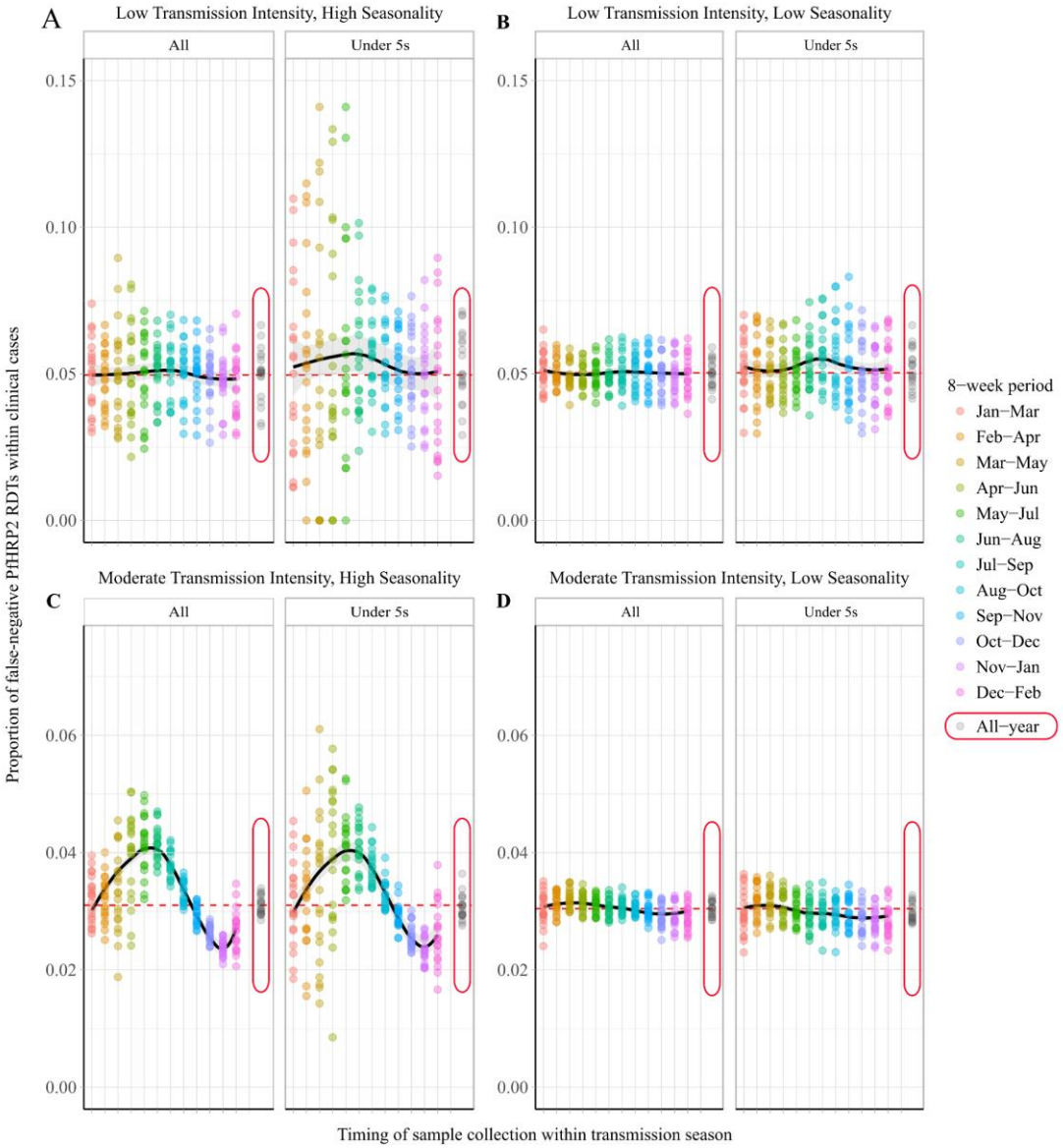


Figure 2.4 Proportion of false-negative PfHRP2 RDTs within clinical cases during 8-week intervals. Graphs show the proportion of clinical cases yielding false-negative PfHRP2 RDTs at 8-week intervals within a transmission season for both moderate (C, D) and low (A, B) transmission settings and high (A, C) and low (B, D) seasonality. In each panel the observed proportion *pfhrp2*-negative clinical cases is shown for the whole population and within children aged under 5 years old. Ten stochastic realisations are represented by the points in each plot, with the mean relationship throughout the transmission shown in black with a locally weighted scatterplot smoothing regression (loess). The annual average proportion of false-negative RDTs due to *pfhrp2* gene deletions is shown with the horizontal dashed red line, and a sampling scheme that occurs throughout the year, with samples collected proportionally to clinical incidence, is shown with grey points circled in red.

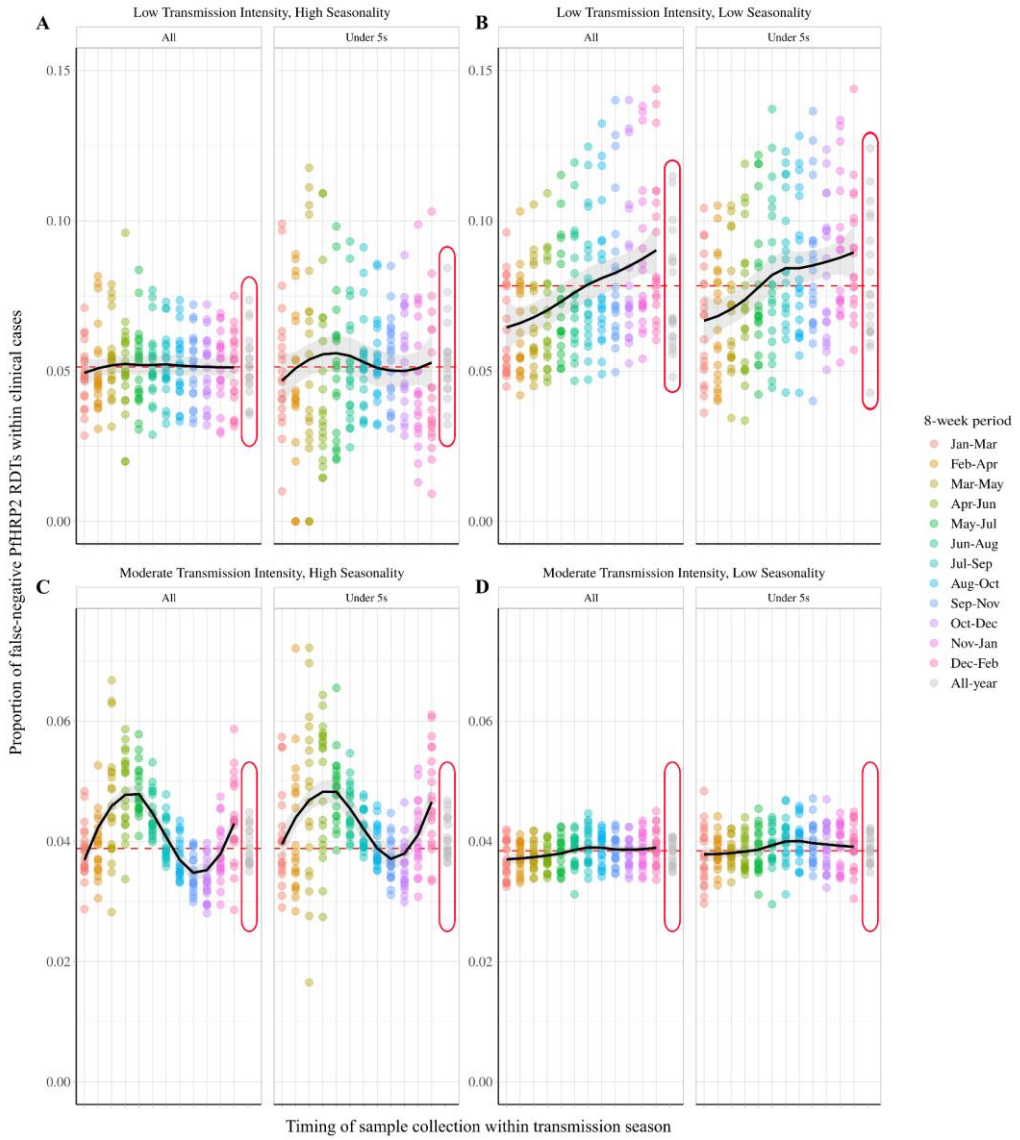


Figure 2.5 Impact of a selective advantage for pfhrp2-deleted parasites on the observed proportion of false-negative PfHRP2 RDTs within clinical cases during 8-week intervals. Graphs show the proportion of clinical cases yielding false-negative PfHRP2 RDTs at 8-week intervals within a transmission season for both moderate (C, D) and low (A, B) transmission settings and high (A, C) and low (B, D) seasonality. In each panel the observed proportion *pfhrp2*-negative clinical cases is shown for the whole population and within children aged under five years old. Ten stochastic realisations are represented by the points in each plot, with the mean relationship throughout the transmission shown in black with a locally weighted scatterplot smoothing regression (loess). The annual average proportion of false-negative RDTs due to pfhrp2 gene deletions is shown with the horizontal dashed red line, and a sampling scheme that occurs throughout the year, with samples collected proportionally to clinical incidence, is shown with grey points circled in red.

Using data from a national survey of *pfhrp2* gene deletions in the DRC, I found that the model-predicted outcomes above were similar to those observed in the field (Figure 2.6) (Parr et al. 2016). Among 2752 PCR-positive *P. falciparum* cases in the DRC, individuals were more likely to be infected with only *pfhrp2*-negative parasites if the clinical incidence in the month prior to sample collection was lower ($p = 4.1 \times 10^{-6}$), and if the individuals were younger ($p = 0.016$). These findings were maintained when comparing across age and transmission groups, with samples collected during

periods of lower transmission found to be more likely to be *pfhrp2*-negative in both older and younger age groups ($p = 6.6 \times 10^{-5}$ and 5.6×10^{-4} respectively). Lastly samples collected in younger individuals were more likely to be *pfhrp2*-negative in both lower and higher transmission groups when compared to older individuals (p = 0.06 and 0.06 respectively).

Lastly, I predicted and mapped the potential for estimates collected within 8-week intervals to be unrepresentative of the annual average proportion of false-negative RDTs due to *phrp2* gene deletions across 598 first administrative regions in SSA (Figure 2.7). I predict that 66 regions possess at least one 8-week interval for which a premature switch to a non PfHRP2-based RDT would have been made in more than 75% of simulations (Figure 2.7A) and 29 regions are predicted to possess at least one 8-week interval for which a premature decision to continue using PfHRP2-based RDTs would have been made in more than 75% of simulations (Figure 2.7B). Out of these 29 regions, 25 are also present within the formerly identified 66 regions. The data for each administrative region can be viewed online at https://shiny.dide.imperial.ac.uk/seasonal_hrp2/.

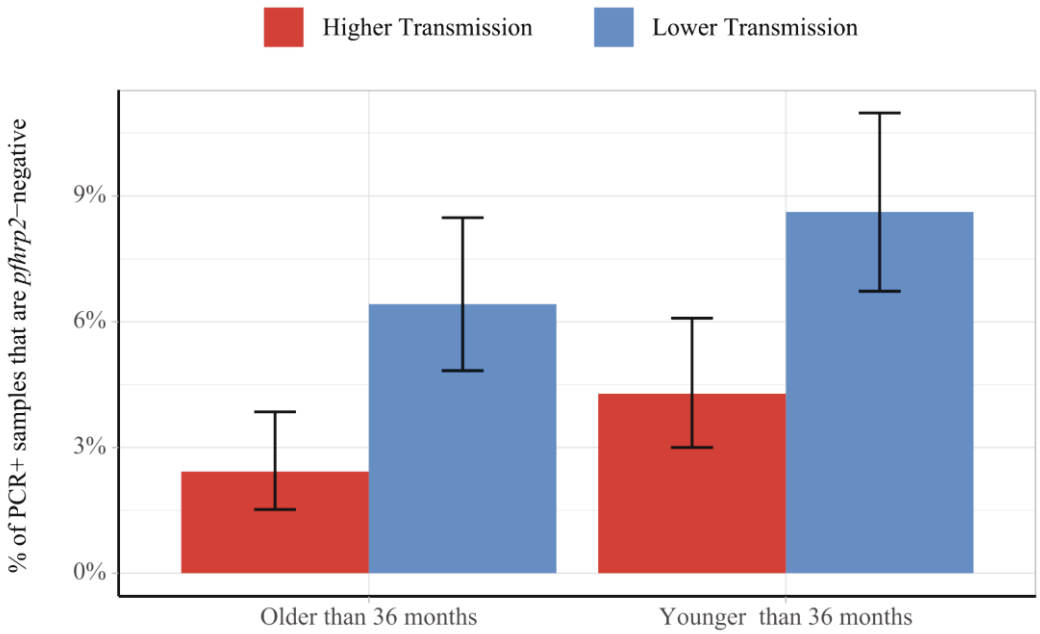


Figure 2.6 Impact of age and transmission intensity upon *pfhrp2* deletion in the Democratic Republic of the Congo, 2013-2014. Graphs show the percentage of PCR-positive *P. falciparum* samples taken from children under the age of 5 years from the 2013-2014 Demographic and Health Survey in Democratic Republic of Congo that are *pfhrp2*-negative. Children who are younger than the median age in the 2,752 samples are grouped within the younger category. In addition, samples are classified as lower transmission if the incidence of malaria in the month prior to sample collection is lower than the median clinical incidence. The 95% binomial confidence intervals are indicated with the vertical error bars.

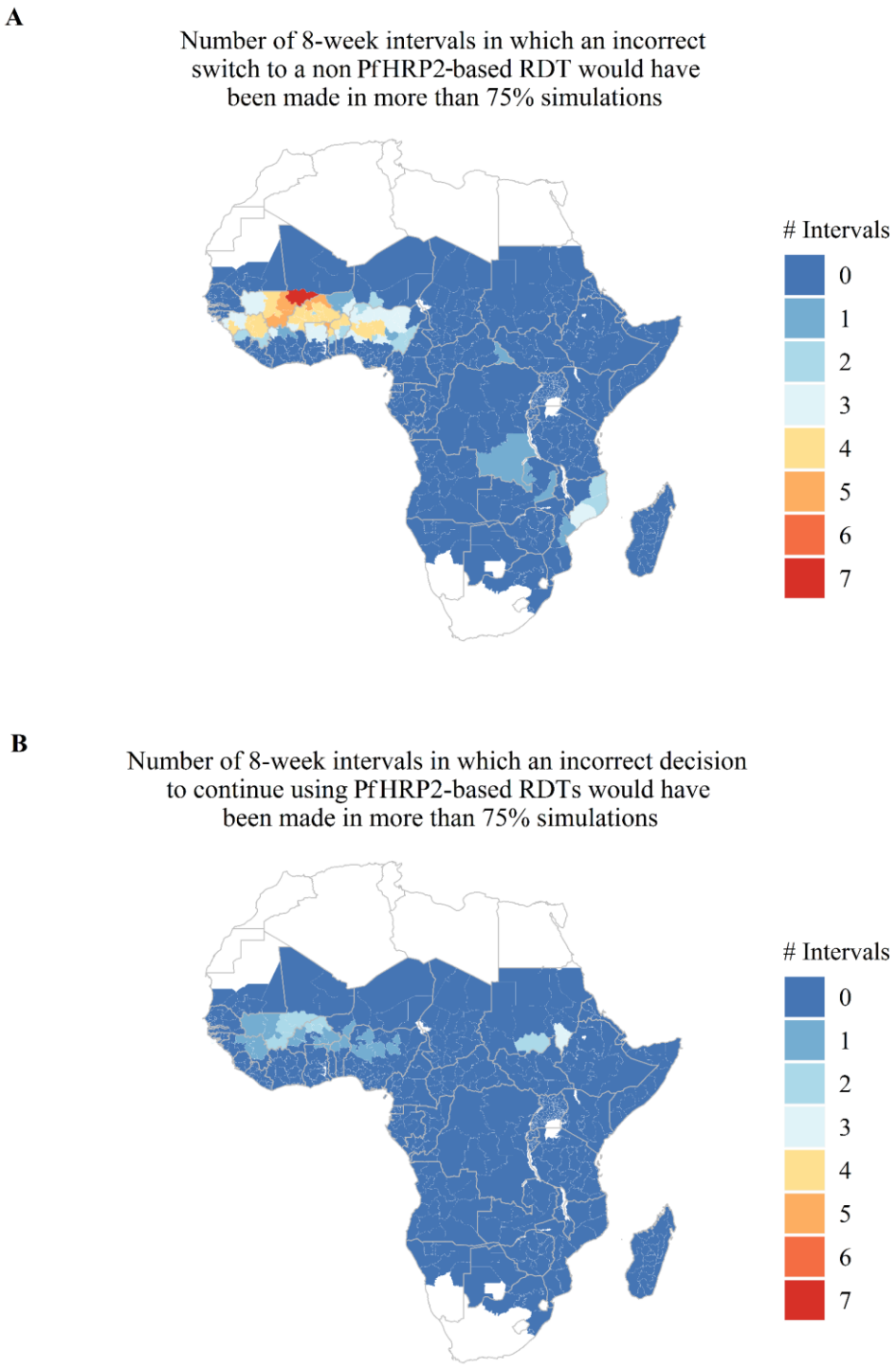


Figure 2.7 Predicted areas with the potential for collected estimates of the proportion of false-negative PfHRP2 RDTs due to *pfhrp2* deletions to be unrepresentative of the annual average. The maps show in **A)** the number of 8-week intervals at which an administrative region would prematurely swap to a non PfHRP2-based RDT due to overestimating the proportion of false-negative PfHRP2 RDTs due to *pfhrp2* gene deletions in more than 75% of simulations. In **B)** the opposing trend is shown, with the number of 8-week intervals at which an administrative region would prematurely continue to use PfHRP2-based RDTs due to underestimating the proportion of false-negative PfHRP2 RDTs due to *pfhrp2* gene deletions in more than 75% of simulations.

2.4. Discussion

This research characterises the potential for surveillance in highly seasonal areas within sub-Saharan Africa to produce estimates that fail to represent the annual average proportion of *P. falciparum* cases with false-negative HRP2 RDT results due to *pfhrp2* deletions. These findings highlight the impact of both the seasonal timing and the age of individuals sampled when estimating the proportion of false-negative RDTs due to *pfhrp2* deletions. Policy decisions based on the proportion of clinical cases presenting with false-negative RDTs due to *pfhrp2* gene deletions should thus be made with an awareness of the seasonal transmission dynamics of the region considered.

My modelling predicted that there would be increased observation of false-negative HRP2 RDT results after periods of lower transmission and within younger individuals. This prediction is consistent with a large, nationally representative survey of *pfhrp2*-negative samples among asymptomatic subjects in the DRC (Parr et al. 2016). These predictions are also in agreement with other observations from Dioro in the Ségou region of Mali, where in 2012 more than 80% of smear-positive individuals had false-negative RDTs when collected at the end of the dry season (Koita et al. 2013). The proportion of false-negative RDTs then rapidly decreased to 20% within 3-4 weeks after the start of the rainy season. It is, however, likely that a proportion of these false-negative RDTs were due to the increased observation of lower parasitaemia at lower transmission intensities such as at the end of the dry season (Okell et al. 2012). In addition, findings from Eritrea also support the model-predicted outcomes. Eritrea is a region with lower malaria prevalence compared to the Ségou region of Mali. The resultant decrease in transmission intensity is likely to result in an increased proportion of monoclonal infections throughout the transmission season. Consequently, I would predict less variability in the number of false-negative RDTs due to *pfhrp2* gene deletions at any given period within a transmission season. I also expect the observed prevalence of *pfhrp2* deletions to be more stochastic due to the lower effective population size of the parasite. Indeed, infections due to *pfhrp2*-deleted parasites identified in Eritrea between November 2013 and November 2014 were not more likely to have occurred after periods of lower transmission intensity ($p = 0.56$, $n = 144$, *pfhrp2* deletions at 9.7%) (Menegon et al. 2017).

Similar to the original publication (Watson et al. 2017), there are a number of modelling assumptions in this study. Firstly, there are modelling uncertainties when predicting the dynamics of false-negative RDTs due to *pfhrp2*-deleted parasites. To account for this uncertainty in this analysis, I have controlled for the drivers characterised in my earlier study by assuming there was no selective advantage associated with *pfhrp2*-deleted parasites and recording the number of individuals who would have been *pfhrp2*-negative and subsequently misdiagnosed. The absence of a selective advantage in this

way enabled the frequency of *pfhrp2* deletions to remain constant, which ensured that any observed dynamics in the estimates of false-negative RDTs due to *pfhrp2* deletions were due to the seasonality of transmission and not due to an increase in the population frequency of *pfhrp2* deletions. However, I am aware that there is likely a selective advantage for *pfhrp2* deleted parasites and subsequently I repeated the analyses with the selective advantage included. In these simulations I predicted a substantial increase in the frequency of *pfhrp2* gene deletions (Figure 2.3Q-T), however clear seasonal dynamics, with an increased proportion of false-negative RDTs due to *pfhrp2* deletions at the beginning of the transmission season, were still observed (Figure 2.5C). However, the observed dynamics were less clear in settings with the greatest increase in the frequency of *pfhrp2* deletions (Figure 2.5B).

Secondly, I assessed the potential for a region to yield unrepresentative estimates of the proportion of false-negative RDTs due to *pfhrp2* deletions through comparisons to the annual average proportion. This decision reflected firstly the monitoring period defined in the WHO technical guidance, with follow up studies recommended after two years if the 95% CI for the proportion of *P. falciparum* cases with false-negative HRP2 RDT results due to *pfhrp2/3* deletions is less than 5%, or one year if it does include 5%. It also reflected the modelling assumption that the population frequency of *pfhrp2* deletions is not increasing over time. However, in simulations in which a selective advantage to *pfhrp2* deleted parasites was included, a comparison to the annual average proportion is less suitable. For example, in Figure 2.5B, because I started the simulations in January the optimum sampling interval is simply the interval in the middle of the year, reflecting the constant increase in *pfhrp2* deleted parasites. In these scenarios it could be argued that the correct comparison would be to the average proportion of false-negative RDTs due to *pfhrp2/3* gene deletions in the year after sampling, which reflects how many cases could be misdiagnosed between sampling rounds. Unfortunately, this comparison is difficult without knowing how the proportion of false-negative RDTs due to *pfhrp2/3* gene deletions will change over time. However, I believe that it is more important to focus on the assumption that the strength of selection is negligible (see Figure 2.8). The rationale for this is that it is only in areas with a low selective pressure, for which the frequency of *pfhrp2/3* deletions is constant over time, that one could repeatedly make an incorrect decision with regards to whether to switch RDT (Figure 2.8A). In areas with a selective pressure it is still possible to incorrectly estimate the annual average for the following year, however the presence of the selective pressure is likely to cause any decision made to be simply premature as the frequency of *pfhrp2/3* deletions and subsequently false-negative PfHRP2 RDTs will increase over time (Figure 2.8B).

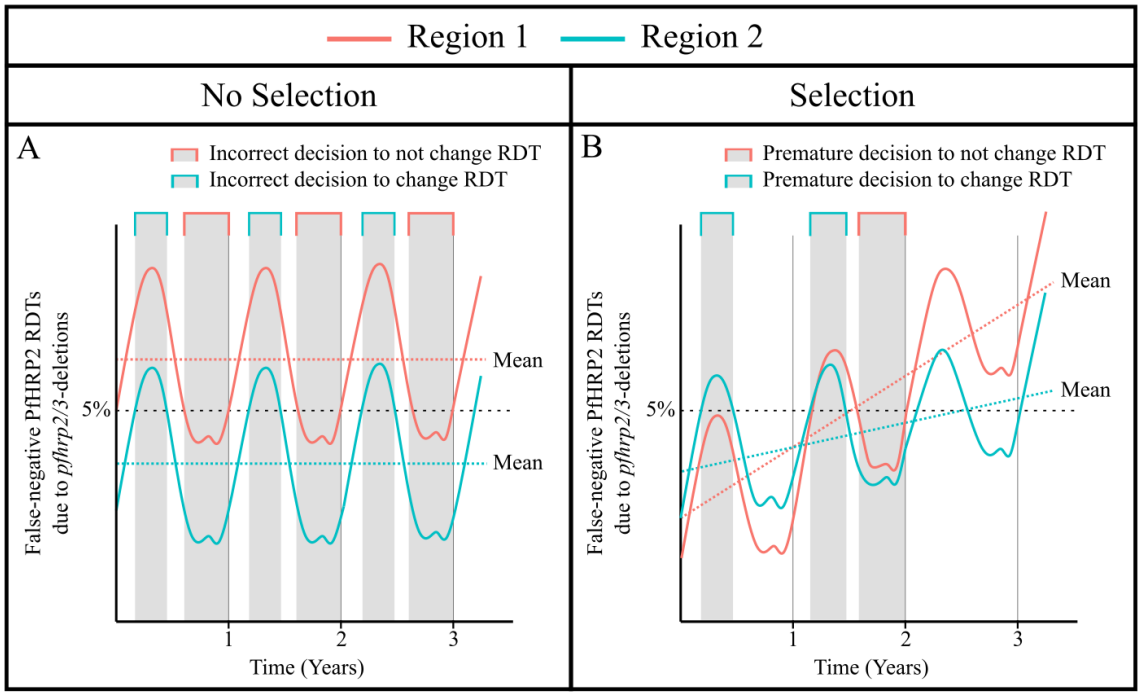


Figure 2.8 The impact of an assumed selective pressure for *pfhrp2/3*-deleted parasites on the decision to switch RDT. The graphs show two hypothetical scenarios with two different regions shown in red and blue for each region. In **A)** there are strong seasonal dynamics but no selective pressure. The absence of a selective pressure causes that the mean proportion of false-negative RDTs due to *pfhrp2/3* deletions over a one-year period to be constant and is shown with a horizontal dashed line. Consequently, there are time periods in which an incorrect decision to switch RDT could be made for the region in blue, and an incorrect decision to not switch RDT could be made for the region in red. In **B)** there are both seasonal dynamics and a selective pressure, which results in an increasing annual mean proportion of false-negative RDTs due to *pfhrp2/3* deletions over time. As in **A)** there are periods in which the observed proportion of false-negative RDTs due to *pfhrp2/3* deletions is both higher and lower than the rolling mean shown. However, decisions made in these periods are premature rather than definitively incorrect as the selection pressure would eventually cause the proportion to be greater than 5%.

Lastly, it is important to note again that the true strength of selection is unknown. The precise strength of selection is dependent on a number of factors such as the magnitude of any fitness costs associated with *pfhrp2* deletion, the degree to which microscopy based diagnosis is used, the level of non-adherence to RDT results, the treatment coverage and the prevalence of malaria in the region considered. Consequently, the results should not be interpreted as precise predictions of how unrepresentative future samples may be. They should instead be used to support surveillance efforts and to reinforce the need for longitudinal measures conducted at the same point within a transmission season. In addition, I recommend that if possible, sample collection in highly seasonal regions should not occur at the beginning of the transmission season, as this is predicted to lead to premature decisions to switch RDT irrespective of the strength of selection. It will, however, be possible after the samples have been collected to estimate the likely frequency of *pfhrp2* gene deletions by incorporating estimates of the multiplicity of infection within the sampled population. This frequency

could then be used to estimate how the proportion of false-negative RDT results due to *pfhrp2* deletions could increase in response to decreases in the prevalence of malaria.

2.5. Conclusion

In this chapter, I have extended my previous model of *pfhrp2* gene deletions to now include seasonal dynamics in transmission intensity. With these extensions, my model predicts that highly seasonal dynamics in malaria transmission intensity will cause seasonal dynamics in the observed proportion of false-negative RDT results due to *pfhrp2* gene deletions. The observed proportion of false-negative RDTs due to *pfhrp2* deletions is higher when monoclonal infections are more prevalent, with the highest prevalence observed when sampling at the start of the rainy season as individuals are less likely to already be infected. Similarly, the observed proportion of false-negative RDTs due to *pfhrp2* deletions is higher in younger individuals who have lower clinical immunity, as they are more likely to present with clinical symptoms after their first infection event. As the rainy season progresses, individuals are more likely to be superinfected and acquire wild type parasites, resulting in positive PfHRP2-based RDT results and a decrease in the observed proportion of false-negative RDTs due to *pfhrp2* deletions. In response to these dynamics, it may be sensible for national malaria control programmes conducting surveillance for *pfhrp2/3* deletions to choose a sampling interval towards the end of the transmission season, which is predicted to be most representative of the annual average proportion of false-negative RDTs due to *pfhrp2* deletions. To support surveillance efforts, I have published an online database detailing the optimum sampling interval as well as the fluctuations throughout the transmission season for each administrative region.

The *pfhrp2* model developed here is written in the programming language R (R Core Team 2019). The language is easier to quickly develop models such as this one, however, the simulation time required to run this model for 60 years in a high transmission setting with a population of 100,000 individuals is ~8 hours. In addition, the structure that was used when writing the model developed with modelling *pfhrp2* deletions and although it could be extended to track multiple strain types, the current model will be unsuitable if I want to model more involved parasite genetic traits, such as resistance or track identity by descent. Consequently, in the next chapter I will be developing a new model of malaria transmission. The model will still be based on the Imperial College transmission model, but it will be written more flexibly to accommodate parasite recombination, different parasite densities of multiple strains in one individual, vector based interventions and many other features.

Chapter 3. Distinguishing the role of superinfection and cotransmission upon *Plasmodium falciparum* complexities of infection

The number of genetically different strains within an infected individual is determined by three main processes; how quickly the body can clear an infection, how quickly an individual is superinfected (an infection occurring within someone who is already infected), and how many genetically different strains are passed on within an infection event. Both superinfections and cotransmission events, where more than one genetically identifiable strain is passed on during an inoculation event, drive the increase in new within host genetic diversity. However, the role of cotransmission events, specifically quantifying the contribution of both cotransmission and superinfection towards the parasite genetic diversity is poorly understood. In this chapter, I develop a new individual-based model of malaria transmission that will enable these processes to be estimated. In doing so I am moving away from my earlier dual-strain model of malaria transmission to one that enables both more nuanced descriptions of parasite genetics and the full lifecycle of the parasite to be explicitly modelled, enabling neutral genetic variation to be modelled. In addition, these changes will enable complex drug resistance profiles to be modelled more flexibly in future chapters, rather than the simple resistant/wild type framework that were used in chapter 2.

3.1. Introduction

The sexual life cycle of the *P. falciparum* parasite has been a topic of much debate (Mzilahowa et al. 2007), with the outcomes of the sexual diploid state of the parasite within the zygote being questioned. A number of researchers have previously concluded that there is no “effective” sexual stage, and that the parasite displays properties of a clonal organism due to substantial levels of selfing (Ayala 1998; Razakandrainibe et al. 2005). However, the current understanding is that a brief sexual stage occurs within the mosquito midgut and can result in the production of new genetic variants. This finding has been confirmed with the advance of direct DNA sequencing of oocysts extracted from the mosquito (Paul et al. 1995).

The number of potentially novel genetic variants produced, however, has been a topic of arguably more uncertainty. This has centred around whether the total number of genotypes present within a single oocyst can be at most 2 or 4:

> *"Our efforts to assemble and organize the known facts about P. falciparum meiotic recombination included close examination of the literature and consultation with a dozen prominent malariologists. This process revealed an unexpected gap in current biological knowledge, in the form of irreconcilable assumptions about whether at most one, two, or four genotypes might be represented within a single P. falciparum oocyst. Furthermore, the adherents of each set of assumptions seemed unaware that any such disagreement exists."* (McKenzie et al. 2001)

The reason for the author's considerable due diligence in assembling a conclusive theory on the outcome of the diploid stage is that it has significant impacts on the genetic structure of the parasite. If only one genotype was possible the speed at which the species could evolve would be substantially slower than if four genotypes were possible, and would have important implications into understanding how quickly the parasite may become resistant to new antimalarial drugs. The resultant consensus suggests that after fertilisation, the parasite is considered to be a diploid organism (2N). This is followed by an initial replication followed by an immediate two step meiotic division (Janse et al. 1986; Sinden 1991). This process can lead to up to 4 different genetic variants to be produced within a single oocyst if the male and female gamete were genetically different (Ranford-Cartwright et al. 1991).

The generation of new parasite genotypes within the mosquito can subsequently lead to multiple genetically distinct sporozoites being passed on during a successful bite. A few studies have measured the number of sporozoites in naturally infected mosquitoes. These estimates help to define the upper bound on the number of sporozoites that could be passed on and also how these limits may depend on transmission intensity. For example, a recent study conducted at a very low transmission site on the Thai–Myanmar reported a geometric mean of 57 sporozoites per mosquito (Pringle and Avery-Jones 1966). However, earlier studies conducted in Papua New Guinea and Africa reported higher sporozoite numbers, with geometric means >4000 (Pringle and Avery-Jones 1966; Gibson 1987; Kabiru et al. 1997). There are fewer studies attempting to quantify the number of sporozoites passed on by *An. gambiae* during an infectious bite, with one estimate suggesting a geometric mean of 10 sporozoites passed on (Beier et al. 1992). Importantly though, estimates of how many sporozoites successfully found a liver stage infection has not been empirically studied, although it is known to be more than one as shown by efforts to analyse the genetic relatedness within multiply infected individuals (Wong et al. 2017). Consequently, there is a need to estimate the percentage of sporozoites within a feed that survive to produce a liver stage infection and how this proportion contributes to within host genetic diversity.

Here, I detail the development of an extended version of the Imperial College Transmission model to examine these processes. The extended version incorporates both cotransmission of multiple

sporozoites from mosquitoes tracked at the individual level, enabling the full parasite life cycle to be modelled.

3.2. Methods

3.2.1. *P. falciparum* Transmission Model

An individual-based stochastic model with a fixed daily time step was developed to simulate the transmission dynamics of *P. falciparum*. Both the human and adult mosquito stages are modelled at an individual level, whereas parasites are modelled as discrete populations with each population relating to an infection event. The human transmission model is based upon previous modelling efforts (Griffin et al. 2010, 2014, 2015, 2016), which is described in its deterministic framework first, before detailing the human acquisition of immunity and the full set of equations detailing its stochastic implementation. The deterministic model described within the methods has been included as its equilibrium solution is used for model initialisation. Additionally, I developed a deterministic version of the earlier 2016 Griffin et al. model (Griffin et al. 2016) that incorporates interventions, which is used to indirectly incorporate the effects of intervention strategies as these are not modelled explicitly within the individual model.

I continue to describe the mosquito transmission model, which is again based on earlier modelling efforts (Griffin et al. 2010, 2014, 2015, 2016), before describing the stochastic equations detailing the new implementation of the adult mosquito stage at an individual-based level. Extensions detailing how the parasite populations are incorporated follow, by first describing the genetic barcode that each parasite population possesses. I then describe the within host parasite populations, which includes considerations surrounding the contribution of coinfection and superinfection towards the model's dynamics of within host multiplicities of infection, and how these relate to the probabilistic uptake of specific gametocyte strains by mosquitoes. This is followed by detailing the within-mosquito parasite populations, which explores the derivation of the distribution describing the model predicted oocyte intensities, and describes how recombination within the sexual stage is explicitly modelled.

3.2.1.1. Human transmission model

Individuals begin life susceptible to infection (state S) (Figure 3.1). At birth, individuals possess a level of maternal immunity that decays exponentially over the first 6 months. Each day individual *i* is probabilistically exposed to infectious bites governed by their individual force of infection (Λ_i). Λ_i is dependent on their pre-erythrocytic immunity, exposure to bites (dependent on both their age and their individual relative biting rate due to heterogeneous biting patterns in mosquitoes) and the size

of the infectious mosquito population. Infected individuals, after a latent period of 12 days (d_E), develop either clinical disease (state D) or asymptomatic infection (state A). This outcome is determined by their probability of acquiring clinical disease (ϕ_i), which is dependent on their clinical immunity. Individuals that develop disease have a fixed probability (f_T) of seeking treatment (state T). Treated individuals are assumed to always recover, i.e. fully-curative treatment, and then enter a protective state of prophylaxis (state P) at rate r_T, before returning to susceptible at rate r_S. Individuals that did not receive treatment recover to a state of asymptomatic infection at rate r_D. Asymptomatic individuals progress to a subpatent infection (state U) at rate r_A, before clearing infection and returning to susceptible at rate r_U. Additionally, superinfection is possible for all individuals in states D, A and U. Superinfected individuals who receive treatment will move to state T. Individuals who are superinfected but do not receive treatment in response to the superinfection will either develop clinical disease, thus moving to state D, or develop an asymptomatic infection and move to state A (except for individuals who were previously in state D, who will remain in state D).

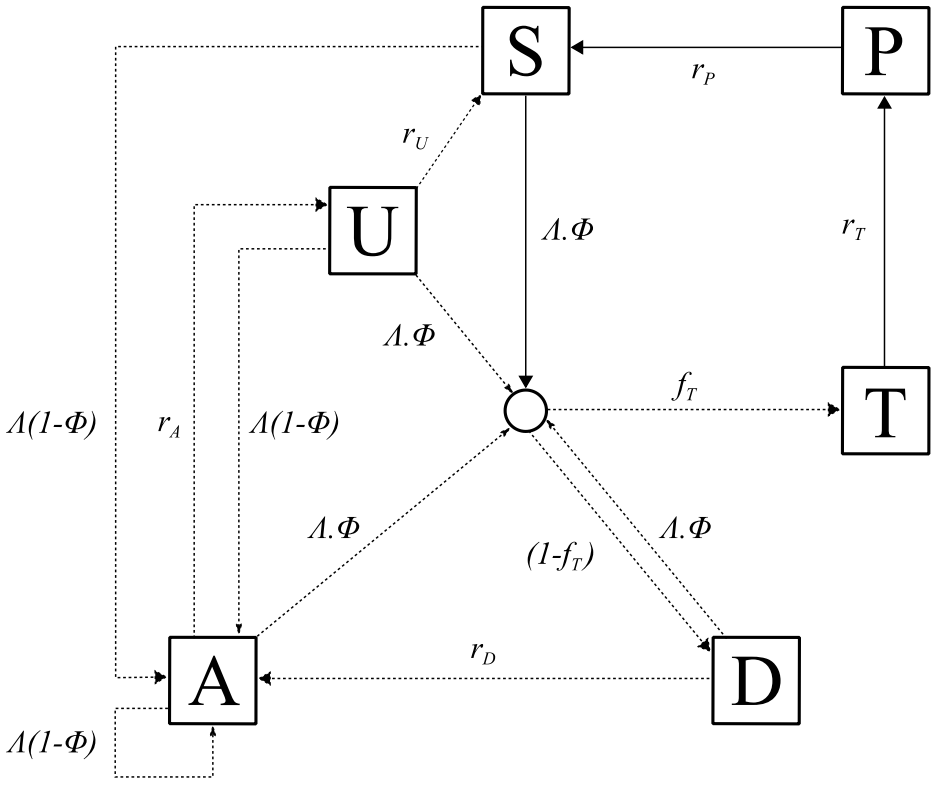


Figure 3.1 Transmission Model. Flow diagram for the human component of the transmission model, with dashed arrows indicating superinfection. S, susceptible; T, treated clinical disease; D, untreated clinical disease; P, prophylaxis; A, asymptomatic patent infection; U, asymptomatic sub-patent infection. All parameters are described and referenced within Table 1. Figure reproduced from Watson et al (CC BY 4.0) (Watson et al. 2017).

The movement between the human components of the transmission model is summarised with the following partial differential equations describing each compartment (t represents time and a represents age):

$$\frac{\partial S}{\partial t} + \frac{\partial S}{\partial a} = -\Lambda(t - d_E)S + \frac{P(t)}{d_P} + \frac{U(t)}{d_U}$$

$$\frac{\partial T}{\partial t} + \frac{\partial T}{\partial a} = \phi f_T \Lambda(t - d_E)(S(t) + D(t) + A(t) + U(t)) - \frac{T(t)}{d_T}$$

$$\frac{\partial D}{\partial t} + \frac{\partial D}{\partial a} = \phi(1 - f_T)\Lambda(t - d_E)(S(t) + A(t) + U(t)) - \frac{D(t)}{d_D}$$

$$\frac{\partial A}{\partial t} + \frac{\partial A}{\partial a} = (1 - \phi)\Lambda(t - d_E)(S(t) + U(t)) + \frac{D(t)}{d_D} - \phi\Lambda A(t) - \frac{A(t)}{d_A}$$

$$\frac{\partial U}{\partial t} + \frac{\partial U}{\partial a} = \frac{A(t)}{d_A} - \frac{U(t)}{d_U} - \Lambda(t - d_E)U(t)$$

$$\frac{\partial P}{\partial t} + \frac{\partial P}{\partial a} = \frac{T(t)}{d_T} - \frac{P(t)}{d_P}$$

When an individual enters a new infection state a waiting time is sampled from an exponential distribution for when the individual will move out of that infection state (except when individuals move into S). With the introduction of a fixed daily time-step, the day on which an individual transitions from state X to Y occurs is given by:

$$Day(X \rightarrow Y) \sim floor\big(Exp(\lambda)\big) + t_{now} + 1$$

where t_{now} is the current day, i.e. the day that the individual moved into state X, and λ is the transition rate. The set of state transitions for individuals and their associated transition rates are given in Table 3.1.

Table 3.1 Human Infection State Transitions

Process	Transition	Transition Rate
Progression of untreated disease to asymptomatic infection	D → A	$r_D = \frac{1}{d_D}$
Progression of asymptomatic infection to subpatent infection	A → U	$r_A = \frac{1}{d_A}$
Progression of subpatent infection to susceptible	U → S	$r_U = \frac{1}{d_U}$
Progression of treated disease to uninfected prophylactic period	T → P	$r_T = \frac{1}{d_T}$
Progression from uninfected prophylactic period to susceptible	P → S	$r_P = \frac{1}{d_P}$

We assume that each person has a unique biting rate, which is the product of their relative age dependent biting rate, ψ_i, given by:

$$\psi_i(a) = \frac{\sum_{i=1}^{n} \psi_i(a)}{n}\left(1 - \rho \exp^{\frac{a}{a_0}}\right)$$

and an assumed heterogeneity in biting patterns of mosquitoes, ζ_i, which persists throughout their lifetime and is drawn from a log-normal distribution with a mean of 1,

$$log(\zeta_i) \sim N\left(\frac{-\sigma^2}{2}, \sigma^2\right)$$

where $1 - \rho$ is the relative biting rate at birth when compared to adults and a_0 represents the time-scale at which the biting rate increases with age. The product of these biting rates is subsequently used to calculate the proportion of the whole population's bites that person *i* receives on a given day, π_i. Their daily entomological inoculation rate (EIR), ϵ_i, is thus calculated by multiplying by the number of infectious mosquitoes taking a blood meal from a human that day, which in turn yields their force of infection, which are given by:

$$\pi_i = \zeta_i \psi_i$$
$$\epsilon_i = I_{M.Feeding} \pi_i$$
$$\Lambda_i = \epsilon_i b_i$$

where $I_{M.Feeding}$ is the size of the feeding infectious mosquito population, and b_i is the probability of infection given an infectious mosquito bite.

The inclusion of individual mosquitoes results in the following stochastic implementation of infection. On any given day the number of infectious mosquitoes taking a blood meal from a human ($I_{M.Feeding}$) will result in the same number of infectious bites. These bites are allocated by sampling from the multinomial distribution using the conditional binomial method (Davis 1993), where sample weights are equal to π_i. Upon receiving a successful infectious bite, an individual will move to an untracked infection state, *I*, which leads to either clinical disease (D), treated clinical disease (T) or asymptomatic infection (A). This leads to the following transition rates related to infection below.

Table 3.2 Human Infection and Superinfection Transition Rates

Process	Transition	Transition Rate
Infection	S → I	$\Lambda_i(t - d_E)$
Super-infection from untreated clinical disease, asymptomatic infection or subpatent infection	D → I A → I U → I	$\Lambda_i(t - d_E)$

The probabilities of progressing from state I to D, T or U are determined by an individual's probability of clinical disease, ϕ_i, and the treatment coverage:

$$Prob(Clinical\ Disease) = \phi_i$$

$$Prob(Treated\ Clinical\ Disease \mid Clinical\ Disease) = f_T$$

The human population was assumed to have a maximum possible age of 100 years, with an average age of 21 years within the population yielding an approximately exponential age distribution typical of sub-Saharan countries. The day on which a human dies is thus allocated at birth by sampling from an exponential distribution with a mean equal to 21 years. When an individual dies, they are replaced with a new-born individual with the same individual biting rate due to heterogeneity in biting patterns.

3.2.1.2. Immunity and Detection Functions

We model 3 stages at which immunity may impact transmission, defined in the existing Griffin et al model (Griffin et al. 2016) as:

1. Pre-erythrocytic immunity, I_B; reduction in the probability of infection given an infectious mosquito bite.
2. Acquired and Maternal Clinical Immunity, I_{CA} and I_{CM} respectively; reduction in the probability of clinical disease given an infection due to the effects of blood stage immunity.
3. Detection immunity, I_D; reduction in the probability of detection and a reduction in the probability of onwards transmission.

Maternal clinical immunity is assumed to be at birth a proportion, P_M, of the acquired immunity of a 20 year-old and to decay at rate $\frac{1}{d_M}$. The remaining three types of immunity are described by the following partial differential equations, which describe how immunity increases due to exposure from zero at birth and decreases over time:

$$\frac{\partial I_B}{\partial t} + \frac{\partial I_B}{\partial a} = \frac{\epsilon}{\epsilon u_B + 1} - \frac{I_B}{d_B}$$

$$\frac{\partial I_{CA}}{\partial t} + \frac{\partial I_{CA}}{\partial a} = \frac{\Lambda}{\Lambda u_C + 1} - \frac{I_{CA}}{d_{CA}}$$

$$\frac{\partial I_D}{\partial t} + \frac{\partial I_D}{\partial a} = \frac{\Lambda}{\Lambda u_D + 1} - \frac{I_D}{d_{ID}}$$

where each u term represents the time during which immunity cannot be boosted further after a previous boost and each d term represents the duration of immunity.

The probabilities of infection, detection and clinical disease are subsequently created by transforming each immunity function by Hill functions. An individual's probability of infection, b_i, is given by:

$$b_i = b_0\left(b_1 + \frac{1-b_1}{1+\left(\frac{I_B}{I_{B0}}\right)^{\kappa_B}}\right)$$

where b_0 is the maximum probability due to no immunity, $b_0 b_1$ is the minimum probability and I_{B0} and κ_B are scale and shape parameters respectively.

An individual's probability of clinical disease, ϕ_i, is given by

$$\phi_i = \phi_0\left(\phi_1 + \frac{1-\phi_1}{1+\left(\frac{I_{CA}+I_{CM}}{I_{C0}}\right)^{\kappa_C}}\right)$$

where ϕ_0 is the maximum probability due to no immunity, $\phi_1\phi_0$ is the minimum probability and I_{C0} and κ_C are scale and shape parameters respectively.

An individual's probability of being detected by microscopy when asymptomatic, q_i, is given by

$$q_i = d_1 + \left(\frac{1-d_1}{1+\left(\frac{I_D}{I_{D0}}\right)^{\kappa_D} f_D}\right)$$

where d_1 is the minimum probability due to maximum immunity, and I_{D0} and κ_D are scale and shape parameters respectively. f_D is dependent only on an individual's age is given by

$$\frac{df_D}{da} = 1 - \frac{1-f_{D0}}{1+\left(\frac{a}{a_D}\right)^{\gamma_D}}$$

where f_{D0} represents the time-scale at which immunity changes with age, and a_D and γ_D are scale and shape parameters respectively.

The probability that an infected individual infects a mosquito upon being bitten is proportional to both their infectious state and their probability of detection, with a lower probability of detection assumed to correlate with a lower parasite density. Individuals who are in state D (clinically diseased), state U (sub-patent infection) and state T (receiving treatment) contribute to an onward infection within a mosquito with probabilities c_D, c_U and c_T. In state *A*, contribution to an onward infection within a mosquito occurs with probability c_A, and is given by $c_U + (c_D - c_U)q^{\gamma_I}$ where q is the probability of

being detected by microscopy when asymptomatic, and γ_I is a parameter that controls how quickly infectiousness falls within the asymptomatic state.

3.2.1.3. Human Stochastic Model Equations

Given the definitions above, the full stochastic individual-based human component of the model can be formally described by its Kolmogorov forward equations. As before, let i index individuals in the population. Then the state of individual i at time t is given by $\{j, k, t_k, l, t_l, m, t_m, a, t\}$, where a is age, j represents infection status (S, D, A, U, T or P), k is the level of infection-blocking immunity and t_k is the time at which infection blocking immunity was last boosted. Similarly, l and t_l denote the level and time of last boosting of clinical immunity, respectively, while m and t_m do likewise for parasite detection immunity. Let $\delta_{p,q}$ denote the Kronecker delta ($\delta_{p,q} = 1$ if $p = q$ and 0 otherwise) and $\delta(x)$ denote the Dirac delta function. Defining $P_i(j, k, t_k, l, t_l, m, t_m, a, t)$ as the probability density function for individual i being in state $\{j, k, t_k, l, t_l, m, t_m, a, t\}$ at time t, the time evolution of the system is governed by the following forward equation:

$$
\begin{aligned}
&\frac{\partial P_i(j,k,t_k,l,t_l,m,t_m,a,t)}{\partial t} + \frac{\partial P_i(j,k,t_k,l,t_l,m,t_m,a,t)}{\partial a} = \\
&\delta_{j,S}[r_P P_i(P,k,t_k,l,t_l,m,t_m,a,t) + r_U P_i(U,k,t_k,l,t_l,m,t_m,a,t)] \\
&+ \delta_{j,A}[r_D P_i(D,k,t_k,l,t_l,m,t_m,a,t)] \\
&+ \delta_{j,U}[r_A P_i(A,k,t_k,l,t_l,m,t_m,a,t)] \\
&+ \delta_{j,P}[r_T P_i(T,k,t_k,l,t_l,m,t_m,a,t)] \\
&+ (1-b_i)\epsilon_i(t-d_E)\left[\delta_{j,S} + \delta_{j,D} + \delta_{j,A} + \delta_{j,U}\right]\mathcal{O}_b \diamond P_i(j,k,t_k,l,t_l,m,t_m,a,t) \\
&+ b_i\epsilon_i(t-d_E)\left[\delta_{j,A}(1-\phi_i) + \delta_{j,D}\phi_i(1-f_T) + \delta_{j,T}\phi_i f_T\right]\mathcal{O}_b \diamond \mathcal{O}_c \diamond \mathcal{O}_d \diamond \sum_{j' \in \{S,A,U\}} P_i(j',k,t_k,l,t_l,m,t_m,a,t) \\
&+ b_i\epsilon_i(t-d_E)\,\mathcal{O}_b \diamond \mathcal{O}_c \diamond \mathcal{O}_d \diamond P_i(D,k,t_k,l,t_l,m,t_m,a,t) \\
&+ \left[r_B k\frac{\partial}{\partial k} + r_{CA} l\frac{\partial}{\partial l} + r_{ID} m\frac{\partial}{\partial m}\right] P_i(j,k,t_k,l,t_l,m,t_m,a,t) \\
&+ \mu\,\delta(a)\delta(t_k + T_{big})\delta(t_l + T_{big})\delta(t_m + T_{big})\delta_{j,S}\delta_{k,0}\delta_{l,0}\delta_{m,0}\sum_{j'} P_i(j',k,t_k,l,t_l,m,t_m,a,t) \\
&- \left[\mu + r_P\delta_{j,P} + r_U\delta_{j,U} + r_D\delta_{j,D} + r_A\delta_{j,A} + r_T\delta_{j,P} + h_i(t-d_E)\left[\delta_{j,S} + \delta_{j,D} + \delta_{j,A} + \delta_{j,U}\right]\right] P_i(j,k,t_k,l,t_l,m,t_m,a,t)
\end{aligned}
$$

Here $\mathcal{O}_b$, $\mathcal{O}_c$ and $\mathcal{O}_d$ are commutative integral operators with the following action on a density $(j, k, t_k, l, t_l, m, t_m, a, t)$:

$$
\mathcal{O}_b \diamond f = \delta(t - t_k)\int_0^\infty f(j,k-1,t-u_B-\tau,l,t_l,m,t_m,a,t)d\tau + \theta\left(\frac{t-t_k}{u_B}\right) f(j,k,t_k,l,t_l,m,t_m,a,t)
$$

$$
\mathcal{O}_c \diamond f = \delta(t - t_l)\int_0^\infty f(j,k,t_k,l-1,t-u_C-\tau,m,t_m,a,t)d\tau + \theta\left(\frac{t-t_l}{u_C}\right) f(j,k,t_k,l,t_l,m,t_m,a,t)
$$

$$\mathcal{O}_d \diamond f = \delta(t - t_m) \int_0^{\infty} f(j, k, t_k, l, t_l, m - 1, t - u_D - \tau, a, t) d\tau + \theta\left(\frac{t - t_m}{u_D}\right) f(j, k, t_k, l, t_l, m, t_m, a, t).$$

Finally, $\theta(x)$ is an indicator function such that $\theta(x) = 1$ if $x < 1$ and 0 otherwise.

For simulation, a discrete time approximation of this stochastic model was used, with a time-step of 1 day. For each individual k, l and m are set to zero at birth, while t_k, t_l and t_m are set to a large negative value $-T_{big}$ (to represent never having been exposed or infected so that their immunity will always be boosted upon their first exposure). Each immunity term increases by 1 for an individual whenever that individual receives an infectious bite (k), or is infected (l and m), if the previous boost to k, l and m occurred more than u_B, u_C and u_D days earlier, respectively. Immunity levels decay exponentially at rate r_B, r_{CA} and r_{ID}, where r_B, r_{CA} and r_{ID} are equal to $\frac{1}{d_B}$, $\frac{1}{d_{CA}}$ and $\frac{1}{d_{ID}}$ respectively.

3.2.1.4. Mosquito Population Dynamics

The adult stage of mosquito development was modelled individually and is similarly described in its deterministic framework before exploring its stochastic implementation. Adult mosquitoes will begin life susceptible to infection (S_M), and will seek a blood meal on the same day they are born and every 3 days after that until the mosquito dies. Each feeding day, mosquito *i* will be exposed to a force of infection, Λ_{Mi}, depending on the infection status and immunity of the human the mosquito is feeding on. The overall force of infection towards the mosquito population on a given day, Λ_M, is thus represented by the sum of the onward infection contributions from each infected human, delayed by d_g, delay due gametocytogenesis, which is given by:

$$\Lambda_M = \alpha_k Q_0 \left(\sum_{i=1}^{\Sigma_D} \pi_i c_D + \sum_{i=1}^{\Sigma_T} \pi_i c_T + \sum_{i=1}^{\Sigma_A} \pi_i c_A + \sum_{i=1}^{\Sigma_U} \pi_i c_U \right) (t - d_g)$$

where α_k is the daily rate at which a mosquito takes a blood meal, Q_0 is the proportion of bites that are on humans (anthropophagy) and d_g represents the delay from emergence of asexual blood-stage parasites to sexual gametocytes that contribute towards onward infectivity. Infected mosquitoes then pass through a latent infection stage (E_M) that will last 10 days representing the extrinsic incubation period for the parasite (d_{EM}), before becoming infectious to humans (I_M). Infectious mosquitoes remain infectious until they die. Whenever a mosquito dies, it is replaced with a new susceptible adult mosquito. Analogously to the human model, when a new adult mosquito emerges, the day on which it dies is drawn from an exponential distribution with a transition rate of μ_M= 0.132 days. The differential equations summarising the adult stage of mosquitoes are given by:

$$\frac{dS_M}{dt} = \mu_M M_v - \mu_M S_M - \Lambda_M S_M$$
$$\frac{dE_M}{dt} = \Lambda_M S_M - \mu_M E_M - \Lambda_M(t - d_{EM}) S_M(t - d_{EM}) exp^{-\mu_M d_{EM}}$$
$$\frac{dI_M}{dt} = \Lambda_M(t - d_{EM}) S_M(t - d_{EM}) exp^{-\mu_M d_{EM}} - \mu_M I_M$$

where μ_M is the daily death rate of adult mosquitoes, and M_v is the total mosquito population, i.e. $S_M + E_M + I_M$.

3.2.1.5. Mosquito Stochastic Model Equations

As with the human transmission model, the full stochastic individual-based mosquito component of the model can be formally described by its Kolmogorov forward equations. As before, let i denote each mosquito in the population, and j denote their infection status. Let $\delta_{p,q}$ denote the Kronecker delta function such that it equals 1 if $p = q$ and 0 otherwise. Defining $P_i(j,t)$ as the probability density function for mosquito i being in state $\{j,t\}$ at time t, the time evolution of the system is governed by the following forward equation:

$$\begin{aligned}\frac{\partial P_i(j,t)}{\partial t} = {} & \delta_{j,E_M}\left[\Lambda_{Mi}\left(P_i(S_M,t)\right)\right] + \delta_{j,I_M}\left[\Lambda_{Mi}(t - d_{EM})\left(P_i(S_M,t)\right)\right] \\ & + \delta_{j,S_M}\mu_M[P_i(S_M,t) + P_i(E_M,t) + P_i(I_M,t)] \\ & - P_i(j,t)\left[\mu_M + \Lambda_{Mi}\left[\delta_{j,S_M}\right] + \Lambda_{Mi}(t - d_{EM})\left[\delta_{j,S_M}\right]\right]\end{aligned}$$

3.2.1.6. Seasonality and Intervention Strategies

In simulations in which no seasonality is assumed, M_v remains constant throughout, i.e. whenever a mosquito dies it is always replaced. When seasonality is incorporated, the maximum value that M_v can be oscillates with a period of 365 days. This corresponds to a change in the birth rate of mosquitoes that reflects an assumed impact upon the seasonal carrying capacity of the environment as a result of rainfall patterns upon mosquito larval stage development. In these simulations, when a mosquito dies, it will only be replaced if the current total number of mosquitoes is less than the maximum value that M_v can be for that day. In simulations designed to replicate regional settings, a rainfall curve, $R(t)$, was estimated from rainfall data from 2002 to 2009 for the related first-administrative unit using the first three frequencies of the Fourier-transformed data (Garske et al. 2013). The seasonal total mosquito population size, $M_v(t)$, is thus given by:

$$M_v(t) = M_{v_0}\frac{R(t)}{\bar{R}}$$

Where $\bar{R}$ is the mean annual rainfall, and M_{v_0} represents the seasonal harmonic mean population size.

The computational constraints introduced by modelling individual mosquitoes and parasite population genetic dynamics necessitated modelling intervention strategies indirectly. This was handled by assuming that an introduction of intervention leads to a decrease in the average age of the mosquito population throughout the duration of the intervention due to an increased mortality rate. As a result, the average age reflects a new composite mortality rate due to both interventions and external causes. Similarly it leads to an increase in Q_0 to reflect mosquitoes that are repelled as a result of interventions but do not die. The daily rate of change to these parameters in response to ITN and IRS coverage is calculated using an equivalent deterministic version of the earlier model that included interventions (Griffin et al. 2016), before being introduced as a time-dependent variable within the stochastic model.

3.2.1.7. Importation Rate

The non-spatial, closed population nature of the model will result in the eventual fixation of a single genetic barcode. As such, when conducting simulations designed to replicate regional settings, an estimate of the importation rate was calculated, yielding to a daily probability that an infection is due to an imported case. The importation rate represents the sum of two different flows of infection into a regional setting:

1. Individuals who are infected outside the region while travelling to and from other areas
2. Visiting travellers from outside the region who infect mosquitoes within the admin unit

These two process are incorporated at the same stage within the model, whereby there is a temporally dependent daily probability that a generated recombinant genotype is due to an importation as follows:

$$Prob(Importation) = \delta_{imports}(t)$$

where $\delta_{imports}(t)$ is the population proportion of new infections resulting from importations on a given day. This parameter changes over time to reflect changes in regional seasonality (both within the region and neighbouring regions), and different rates of change in malaria prevalence across neighbouring regions (Cook et al. 2011). If the recombinant is due to an importation, then a random barcode is produced and passed on. This barcode will also be stored and associated with an oocyst within the mosquito considered if it was probabilistically determined to be due to the second flow of importation defined above, determined by the ratio of these two flows of infection. Predicted rates

of the two flows of infection above are calculated for each year between 2000 and 2015 using a fitted gravity model of human mobility (Marshall et al. 2018).

3.2.2. Parasite Dynamics

3.2.2.1. Parasite Genetic Barcode

Parasites are modelled as discrete populations resulting from an individual infection event associated with a mosquito or a human. Each asexual parasite is characterised by one genetic barcode, which contains information relating to 24-SNPs distributed across the parasite genome. These SNPs represent an increasingly used general SNP-based molecular barcode that has been used for the identification and tracking of *P. falciparum* clones (Daniels et al. 2008). Sexual stages of the parasite lifecycle within the mosquito are represented by both a female and male barcode, thus defining the range of recombinants that could be produced. The within human parasite dynamics and model considerations are discussed first before exploring the within mosquito parasite life cycle and associated modelling implications. A schematic overview of the modelled parasite lifecycle stages is shown in Figure 3.2.

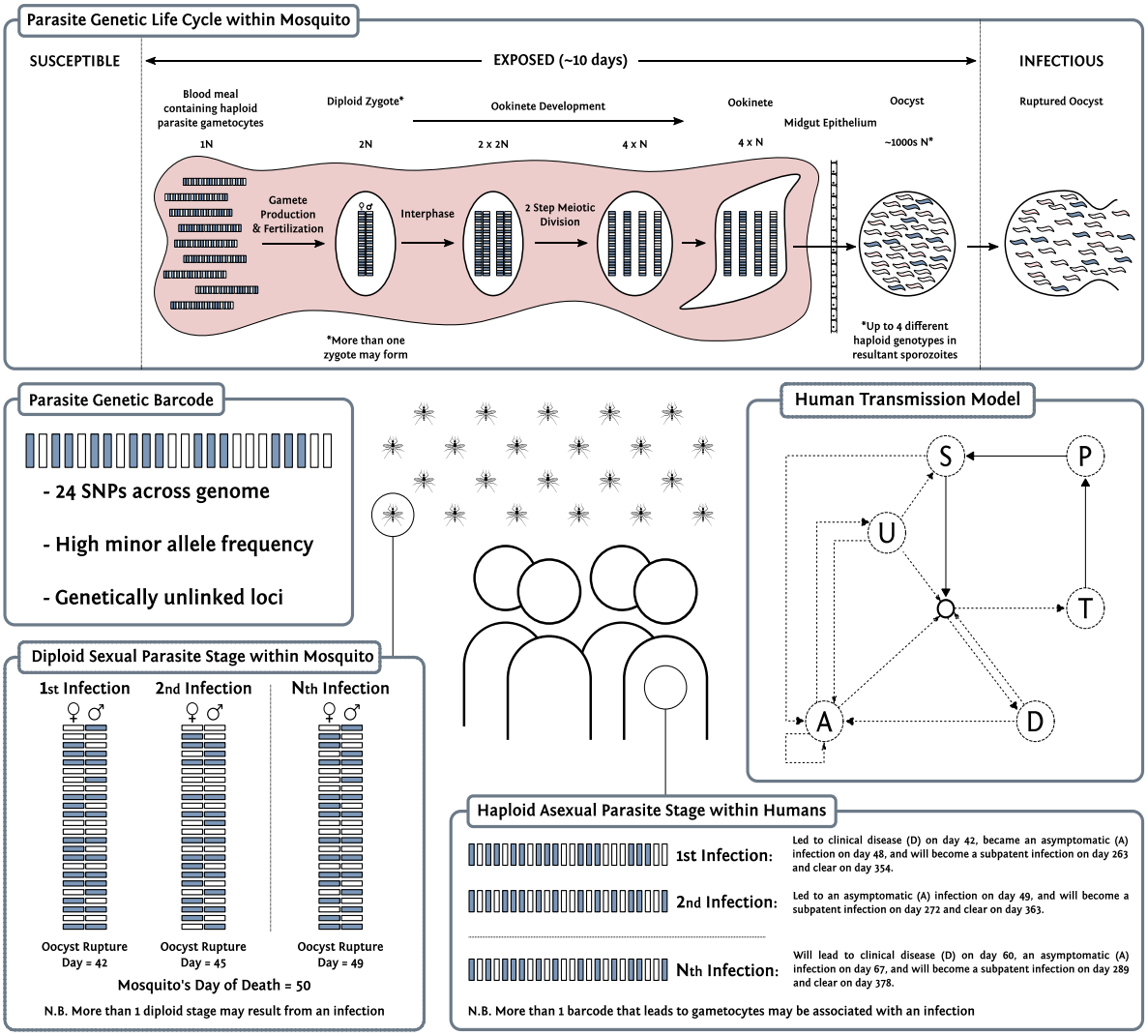


Figure 3.2 Parasite Dynamics within the transmission model. Individual mosquitoes are tracked, which allows for recombination to be modelled explicitly. Populations of parasite clones are tracked, and multiple oocysts are able to be formed from a feeding event, as well as multiple genetically distinct sporozoites onwardly transmitted. A barcode is associated with each parasite clone and can either represent biallelic single nucleotide polymorphisms, or unique identities that allow identity by descent to be calculated.

In simulations modelling identity by descent (IBD), I extend the barcode to consider 24 "identity-loci". An identity locus can take any integer value required, allowing true identities to be compared. In the SNP-loci barcode, each locus can only be 0 or 1, representing the minor and major allele for that barcode loci.

3.2.2.2. Within Human Parasite Dynamics

During a successful mosquito to human infection event, a number of asexual parasite barcodes may be introduced into the human. All introduced barcodes are assumed to be able to contribute to onward infectiousness from humans to mosquitoes after $d_E + d_g$ days, i.e. the time taken for the liver

stage infection and gametocytogenesis. If the individual's pre-erythrocytic immunity was boosted in the last u_B days no new parasite barcodes will be passed to the individual, otherwise more than one different asexual parasite barcode may be introduced during an infection event, representing cotransmission of genetically related parasites (assuming the mosquito was infected with more than one sporozoite genotype). The precise distribution describing the number of genotypes is unknown (Wong et al. 2017), but the mean number of sporozoites within an innoculation event is well characterised by a geometric distribution with mean equal to 10 (Pongtavornpinyo et al. 2009). The geometric mean will then be used to estimate the proportion of sporozoites that are successful, ξ, which yields the maximum number of successful sporozoites in an individual with no pre-erythrocytic immunity. If this number is less than 1, then a new total number of sporozoites is drawn until the maximum number of sporozoites after incorporating ξ is greater than 0. The observed number of successful sporozoites is then calculated by conducting Bernoulli trials for all but one of the successful sporozoites (as I assume one has to survive to found the infection) to see if they are successful, using the individual's probability of infection, b_i. In summary this can be written as:

$$Total_{spz} \sim Geom(p_{spz})$$

$$Max_{spz} = round(Total_{spz}.\xi)$$

$$Observed_{spz} = 1 + \sum_{1}^{Max_{spz}-1} bernoulli(b_i)$$

There is no assumed maximum number of parasites, with individuals assumed to clear strains on the day that they would have moved from a subpatent infection to susceptible for the strain considered, i.e. each acquired strain follows an assumed trajectory in parasitaemia representative of a normal infection cycle, with a mean duration of infectiousness equal to $d_A + d_U$. Acquired strains can thus move "infection state" independently of the human's infection state. For example, a given individual is infected on day 0 and develops an asymptomatic infection. The individual is scheduled to become subpatent on day 200, but they were bitten on day 150 and developed clinical symptoms and moved to state D. When this happens, the parasite density of the strain acquired on day 0 does not change and this strain will become a subpatent strain on day 200. After day 200, its probability of being onwardly transmitted is thus equal to c_U. After the parasite has moved to become a subpatent strain, the day at which the strain would have been cleared, i.e. the day the individual would have moved from state U to S if they had not been superinfected, is drawn and assigned to the parasite. On this drawn day the subpatent parasite strain is assumed to have been cleared. By tracking parasites in this way I am able to track the relative parasitemias of each acquired strain, enabling more accurate

sampling of within host parasite genetic diversity when passing on gametocytes to mosquitoes as well as enabling an equilibrium between clearing old strains and acquiring new strains, which represents the multiplicity of infection. This is shown in Figure 3.2, which also details the key features of the SNP barcode.

3.2.2.3. Within Mosquito parasite dynamics

When a mosquito is infected, the number of oocysts formed is sampled from a zero-truncated negative binomial distribution. The choice of a zero truncated negative binomial represents the increasingly identified zero-inflated negative binomial that describes the relationship between oocyst prevalence and mean oocysts per mosquito in SMFA studies (Churcher et al. 2012a; Stone et al. 2013, 2014). The related negative binomial distribution for the distribution of oocysts is given by

$$X_{oocysts} \sim NB(size_{oocysts}, shape_{oocysts})$$

where $X_{oocysts}$ represents the number of oocysts that will be formed, with mean equal to 2.5 and a shape equal to 1, which captures the mean and range of oocysts observed in natural *P. falciparum* infections (Churcher et al. 2013; Stone et al. 2013, 2014). For each oocyst formed, two barcodes are sampled from the infected host representing the female and male gametes that led to the oocysts formation. These two barcodes will result in up to 4 different potential genotypes (reflecting the immediate two step meiotic division that takes place after zygote formation) of sporozoite to be produced by the oocyst. When an infectious mosquito seeks a blood meal and leads to an onward infection, a value for $Observed_{spz}$ is sampled. The oocyst source for each onward infection within a coinfection is sampled from oocysts that have ruptured, i.e. the infection event that led to the oocyst occurred more than 10 days earlier. At this point recombination is simulated by randomly choosing either the male of female allele at each position in the barcode. The random sampling in this represents the assumed independent segregation events resulting from the absence of genetic linkage between barcode loci. Once a recombinant has been simulated it is stored and associated with the oocyst that generated it. If the same oocyst is chosen to lead to an additional infection, then the previously generated recombinant has a 25% chance of being onwardly transmitted and there is a 75% chance that a new recombinant is generated and subsequently saved. This process will continue until four recombinants have been simulated, at which point they each have a 25% chance of being onwardly transmitted if that oocyst is chosen to contribute a sporozoite to an onward infection. This process assumes that sporozoites will remain onwardly-transmissible for the remainder of the mosquito's life, with no effect upon their relative probability of being onwardly transmitted in relation to sporozoites that resulted from a more recently ruptured oocyst. It also assumes that each oocyst

contributes comparable numbers of sporozoites, such that the probability of a sporozoites originating from each oocyst is the same for all oocysts.

3.2.3. Model Parameter Values

All model parameters are shown in Table 3.3 with previously fitted model parameters sourced from Griffin et al. 2014 (Griffin et al. 2014), 2015 (Griffin et al. 2015) and 2016 (Griffin et al. 2016).

Table 3.3 Parameter estimates

Parameter	Symbol	Estimate
Human infection duration (days)		
Latent period	d_E	12
Patent infection	d_A	200
Clinical disease (treated)	d_T	5
Clinical disease (untreated)	d_D	5
Sub-patent infection	d_U	110
Prophylaxis following treatment	d_P	25
Treatment and Importation Parameters		
Probability of seeking treatment if clinically diseased	f_T	Variable
Importation Rate	$\delta_{imports}$	0.01
Infectiousness to mosquitoes		
Lag from parasites to infectious gametocytes	d_g	12 days
Untreated disease	c_D	0.0680 day^{-1}
Treated disease	c_T	0.0219 day^{-1}
Sub-patent infection	c_U	0.000620 day^{-1}
Parameter for infectiousness of state A	γ_1	1.824
Age and heterogeneity		
Age-dependent biting parameter	ρ	0.85
Age-dependent biting parameter	a_0	8 years
Daily mortality rate of humans	μ	0.000180
Variance of the log heterogeneity in biting rates	σ^2	1.67
Immunity reducing probability of infection		
Maximum probability due to no immunity	b_0	0.590
Maximum relative reduction due to immunity	b_1	0.5

Inverse of decay rate	d_B	10 years
Scale parameter	I_{B0}	43.879
Shape parameter	κ_B	2.155
Duration in which immunity is not boosted	u_B	7.199
Immunity reducing probability of clinical disease		
Maximum probability due to no immunity	ϕ_0	0.791
Maximum relative reduction due to immunity	ϕ_1	0.000737
Inverse of decay rate	d_{CA}	30 years
Scale parameter	I_{C0}	18.0237
Shape parameter	κ_C	2.370
Duration in which immunity is not boosted	u_C	6.0635
New-born immunity relative to mother's	P_M	0.774
Inverse of decay rate of maternal immunity	d_M	67.695
Immunity reducing probability of detection		
Minimum probability due to maximum immunity	d_1	0.161
Inverse of decay rate	d_{ID}	10 years
Scale parameter	I_{D0}	1.578
Shape parameter	κ_D	0.477
Duration in which immunity is not boosted	u_D	9.445
Scale parameter relating age to immunity	a_D	21.9 years
Time-scale at which immunity changes with age	f_{D0}	0.00706
Shape parameter relating age to immunity	γ_D	4.818
Mosquito Population Model		
Daily mortality of adults	μ_M	0.132
Daily biting rate	α_k	0.333
Anthropophagy	Q_0	0.92
Extrinsic incubation period	d_{EM}	10 days
Negative Binomial shape parameter for oocyst distribution	$shape_{oocysts}$	2.5
Negative Binomial size parameter for oocyst distribution	$size_{oocysts}$	1
Human Parasite Parameters		
Geometric probability of total sporozoites in an infectious bite	p_{spz}	1/10
Percentage of sporozoites successfully reaching blood-stage	ξ	20% (fitted)

3.2.4. Model Fitting

My extensions to the transmission model introduced a new parameter, ξ, which determines the percentage of the total sporozoites passed on within a feeding event that survive to yield a blood-stage infection and subsequently produce gametocytes. To fit this parameter I compared the model predicted relationship between the complexity of infection (COI) and age utilising previously SNP genotyped samples from five sites across Kenya (Omedo et al. 2017b) and Uganda (Chang et al. 2017), collected between 2008-2010 and 2012-2013 respectively. In brief, dried blood spots were collected and samples taken from individuals with evidence of asexual parasitemia by microscopy were selected for Sequenom SNP genotyping. Genotyping was conducted using the Sequenom MassARRAY iPLEX platform, yielding minor and major allele frequencies.

We applied *THE REAL McCOIL* proportional method to the SNP genotyped samples to estimate each individual's COI (Chang et al. 2017). *THE REAL McCOIL* is a Bayesian Markov chain Monte Carlo method for jointly estimating the complexity of infection of individuals from SNP genotyped samples as well as the population allele frequencies of the genotyped SNPs. I developed an extended version of the published implementation of *THE REAL McCOIL*, which uses dynamically-allocated memory containers to allow the processing of larger sample sizes and SNPs than the original implementation. The developed software, *McCOILR,* is available at https://github.com/OJWatson/McCOILR.

Samples were filtered first by excluding loci with more than 20% missing samples, followed by samples with more than 25% missing loci. We performed thirty repetitions of *THE REAL McCOIL* for each sample, with a burn-in period of 10^4 iterations followed by 10^6 sampling iterations. Sequencing measurement error was estimated along with COI and allele frequencies, the maximum observable COI was set equal to 25 and default priors were assigned for each parameter and I used standard methodology to confirm convergence between chains (Gelman and Rubin 1996). The observed relationship between COI and age was compared to the model predicted relationship for each administrative region studied. The model predicted relationship was created by conducting simulations calibrated to estimates of the administrative malaria prevalence from 2000 to 2015 (Bhatt et al. 2015), exploring values of ξ between 0.5% - 50%. For each region, 10 stochastic realisations of 100,000 individuals were simulated with a burn-in period of 50 years to ensure both an epidemiological and genetic equilibrium was reached by year 2000. For each of the five administrative regions of interest, I incorporate the historic scale up of insecticide treated nets and indoor residual spraying between 2000 and 2015, using data previously collated for the World Malaria Report (World Health Organization 2015d), and estimates for the coverage of treatment modelled using DHS and MICS survey data (Cohen et al. 2012b). Seasonality for each region was included by altering the total

number of mosquitoes using annually fluctuating seasonal curves fitted to daily rainfall data from 2002 to 2009 (Cairns et al. 2012). Lastly, I introduce rates of importation of infections that are calculated for each year between 2000 and 2015 using a fitted gravity model of human mobility (Marshall et al. 2018). These sources represent infections acquired from individuals travelling out of the region and returning with an infection, and also mosquitoes being infected by individuals travelling from outside into to the region of interest.

We calculated the "distance" between our model predictions and the observed data, $I(\xi)$, using the Kullback-Leibler (KL) divergence (Burnham et al. 2002). Using an individual's age and estimated COI, the distance between the observed and predicted distributions of COI for each age is given by:

$$I(\xi_i) \coloneqq I(pCOI_i(\xi), oCOI_i) = \sum_{COI=1}^{25} pCOI_i(\xi) ln\left(\frac{pCOI_i(\xi)}{oCOI_i}\right)$$

where $oCOI_i$ is the observed distribution of COI at age i and $pCOI_i(\xi)$ is one realisation of the model predicted distribution of COI at age i for a given frequency of successful sporozoites ξ (with only parasites that would have been detected by PCR being assumed to be detected by SNP genotyping). The total distance for a given value of ξ is subsequently given by:

$$\sum_{r}^{5}\left(\frac{\sum_{i}^{n_i} I(\xi_i)w_i}{\sum_{i}^{n_i} w_i}\right)_r$$

where w_i is the weight for age i (equal to the number of observations at age i) and n_i is the total number of unique sampled ages in administrative region r. This can be interpreted as the sum of the weighted KL divergence means within a region, with weights equal to the number of observations at each age. Each region thus contributes equally to the total distance, despite the difference in the number of individuals in each region.

Further model fit validation was conducted by incorporating a comparatively larger collection of estimates of the COI estimated using *msp2* genotyping, which is more commonly referred to as multiplicity of infection (MOI). *msp2* genotyping is known to underestimate COI in individuals with very high COIs, with COIs > 7 difficult to observe (Gupta et al. 2010). Consequently, to distinguish these estimates I refer to these as *msp2* COI. We compiled *P. falciparum* malaria *msp2* COI data where there were estimates of both the malaria prevalence and the *msp2* COI of study participants. This was conducted by updating a previous review (Karl et al. 2016), using the same search terms of "falciparum multiplicity infection prevalence msp2". Analogous relationships were predicted using the fitted model, with the model predicted *msp2* COI estimated by assuming that any observable COI from the

model greater than 7 results in an *msp2* COI of 7, which reflects the limits of resolution when using *msp2* genotyping (Gupta et al. 2010).

3.2.5. Contribution of superinfection and cotransmission to within host genetic diversity

The parameterised model was used to characterise the relative contribution of cotransmission events and superinfection towards within host parasite genetic diversity. Ten stochastic realisations of 100,000 individuals were simulated for 50 years at 15 different transmission intensities. The proportion of highly identical parasite strains (>50% of loci are IBD in pairwise comparisons between all unique parasite strains within an individual) within simulations was recorded in the last year of the simulations and used to estimate the proportion of within host genetic diversity that is due to cotransmission events rather than superinfection. Lastly, the simulations described will be used to compare the model predicted relationship between the mean IBD within samples and the proportion of samples that are multiply infected (COI > 1). This relationship has been recently explored using whole genome sequence data (Zhu et al. 2019) and subsequently provides another opportunity to assess the realism of the parameterised model.

3.3. Results

3.3.1. Complexity of Infection Data

First, I used *THE REAL McCOIL* (Chang et al. 2017) to estimate the COI from SNP genotyped samples collected previously from individuals with evidence of asexual parasitemia by microscopy from regions in Kenya and Uganda. These two datasets were selected as they recorded both the age of the sampled individuals and SNP intensities at a sufficiently large number of loci enabling the relationship between COI and age to be estimated. After excluding SNP loci with more than 20% missing data and subsequently removing samples with more than 25% missing SNP data from further analysis, the COI was estimated for 2419 samples from 95 primary schools in Western Kenya (1363 from Nyanza province and 1056 from Western province) and 584 samples from representative cross-sectional household surveys in three sub-counties in Uganda (462 from Nagongera in Tororo District, 74 from Kihihi in Kanungu District, and 48 Walukuba in Jinja District).

3.3.2. Model Fitting

The developed model was extended to include individual mosquitoes, enabling parasite populations and their genotypes to be tracked throughout the full lifecycle. This extension allows for the potential formation of multiple oocysts from an infectious event and multiple genetically distinct sporozoites to

be onwardly transmitted. The extensions require the parameterisation of the proportion of sporozoites from an infectious bite that survive to found a blood stage infection. This process will ultimately affect the level of new parasite genetic diversity introduced and may alter the observed COI. The model was parameterised using the earlier estimated relationships between COI and age in the five regions across Uganda and Kenya (Figure 3.3). I estimate that 20% of sporozoites onwardly transmitted within an infectious bite successfully progress to a blood-stage infection and produce gametocytes that may contribute to future mosquito infections. The model captures the observed peak in COI observed at age 7-8 (Figure 3.3a); however, the comparatively fewer samples at higher ages make it difficult to confirm that this is the true peak in COI. Additionally, this observed peak in COI also likely reflect the limits of detection, with more accurate model predictions occurring under the assumption that sub-patent parasite strains would not be detected (Figure 3.3b). Model fitting also showed that sensitivity to the percentage of sporozoites that survive is negligible between values of 15-20%, with the confidence intervals for the most likely parameter value of ζ (the % of sporozoites per infections surviving and eventually producing gametocytes) overlapping intervals for values of ζ ranging from 0.1 to 0.29 (Figure 3.3). Model fits were consistently better across the parameter range explored under the assumption that only PCR detectable infections contribute to the observed COI.

To further assess the fitted model, I wanted to incorporate estimates of COI based on *msp2* genotyping, which is more commonly measured. Estimating COI by *msp2* genotyping, however, does underestimate COI in individuals with high COI, with COIs > 7 difficult to resolve. I updated a previous literature review of paired estimates of *msp2* COI and parasite prevalence by PCR, which yielded 105 results. The fitted model predicts an increase in *msp2* COI with increasing malaria prevalence in agreement with the data collected within the literature search (Figure 3.3c). However, there are notably larger uncertainties in the recorded *msp2* COI at higher prevalence ranges in the studies found.

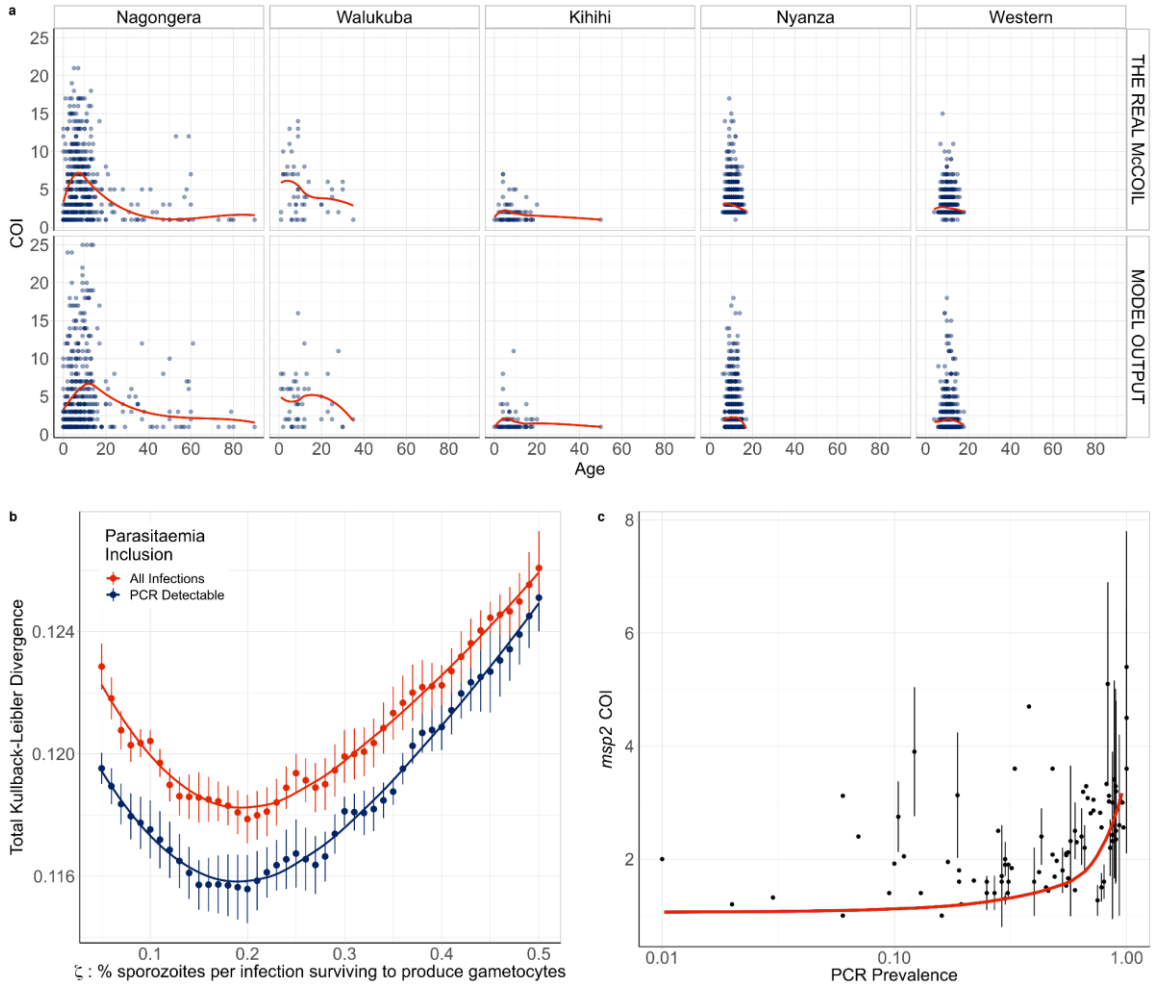


Figure 3.3 Relationship between complexity of infection and age. **a)** One realisation of the model predicted relationship between complexity of infection (COI) and age compared to the observed relationship estimated using *THE REAL McCOIL*. Each point represents an individual, with a local regression fit plotted in red. The relationship shown represents the selected best model fit, which estimates that 20% of sporozoites successfully progress to blood-stage infection in an individual with no immunity. In **b)** the results of the model fit are shown, with each point representing the mean Kullback-Leibler divergence and the whiskers representing the 95% confidence interval. Results of model fitting are shown for the assumption that sub-patent infections are either detected (red) or not detected (blue). In **c)** the model predicted relationship between COI measured by *msp2* genotyping and PCR prevalence is shown in red, with the point-ranges showing observed values of COI by *msp2* genotyping from the literature review.

3.3.3. Population genetic properties of the fitted transmission model

Using the fitted model, I first explored basic population genetic properties of the model to examine the relationships between population size, transmission intensity and genetic diversity within a closed population and no importation. This scenario enables the rate at which genetic diversity is lost due to genetic drift to be estimated. The rate of fixation was predicted to be substantially faster at lower transmission settings, reflecting the decreasing rate of recombination due to an increasing probability of infected individuals being monoclonally infected (Figure 3.4).

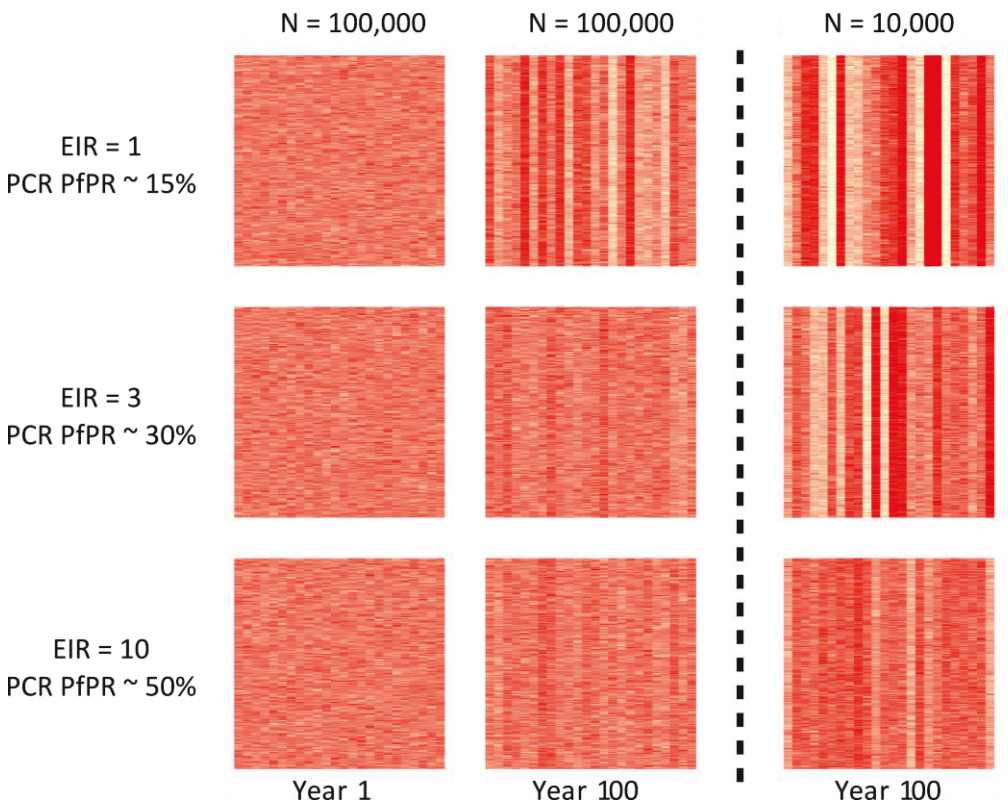


Figure 3.4 Genetic fixation under neutral selection. Each heatmap shows the human parasite population, with each row representing a unique parasite genetic barcode within a human, and each column representing one of the 24 SNPs of the barcode. The populations were initialised with an allele frequency of 50% at each barcode locus and simulated under the assumption that there is no importation of new genetic strains. Consequently, genetic drift causes genetic diversity to be lost over time, with the allele frequency of each barcode locus tending towards 0% (yellow) or 100% (red) after 100 years. The loss of diversity is greater at lower transmission intensities, as shown by increasing transmission intensities depicted in the rows of the plot, and at lower population sizes, as shown by the final column which was simulated using a population of 10,000 individuals rather than 100,000 individuals.

To further examine the impact of transmission intensity, I explored a scenario in which a fixed rate of 3% importation was assumed to yield a genetic equilibrium. A population of 10,000 individuals was simulated for 100 years with an EIR = 10 and a treatment coverage of 20%. After 100 years, the treatment coverage was increased by 1% for 35 years from 20% to 55% (Figure 3.5). Every 2 years 200 parasites were sampled and the frequency of each parasite genotype was calculated and plotted over time. After the increase in transmission intensity, a slight decrease was observed in the number of barcodes that occurred only once in the sample. During the same period, PCR prevalence decreased from 35% to 25%, suggesting that parasite genetic diversity is associated to malaria prevalence.

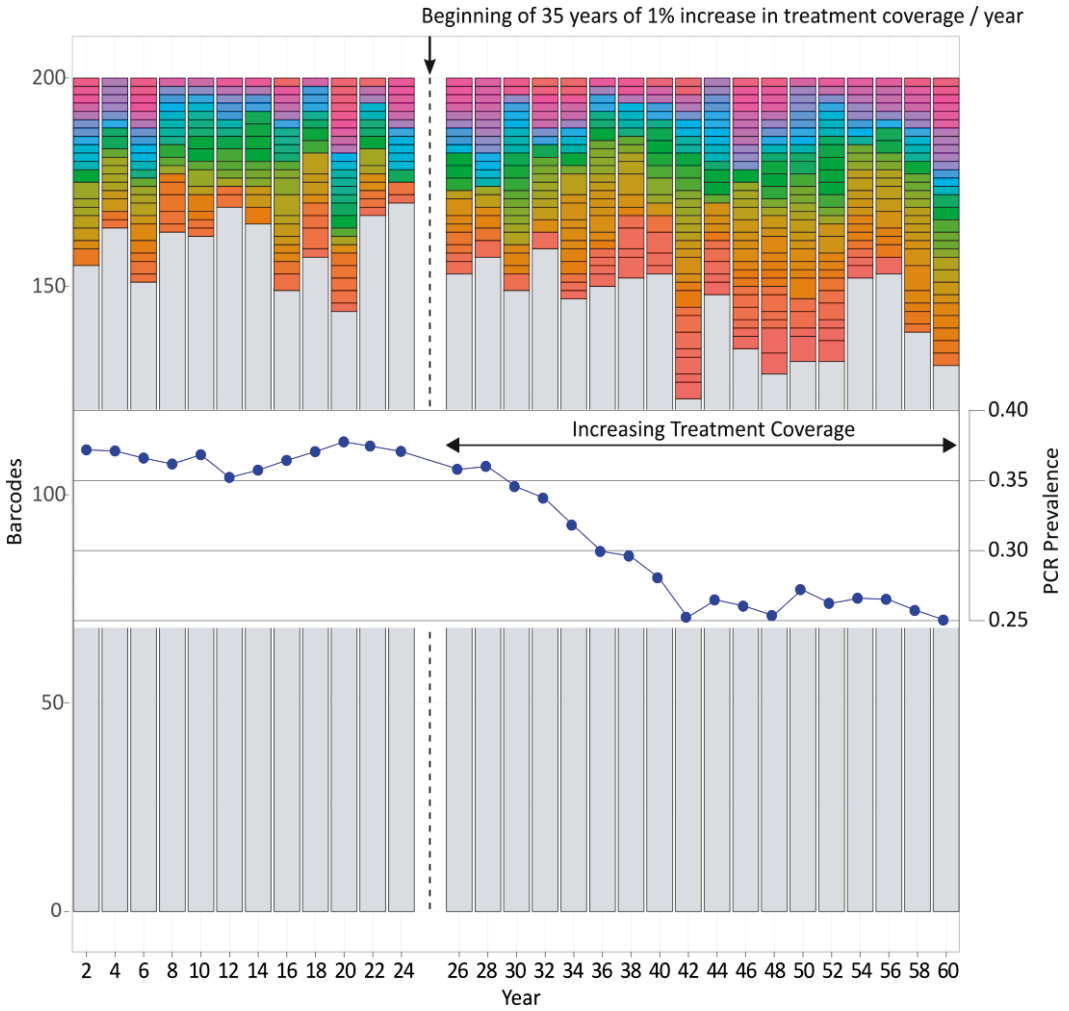


Figure 3.5 Decreasing parasite barcodes upon scale up of treatment. Each column in the graph shows 200 randomly sampled barcodes, with unique barcodes shown in grey and barcodes that appear more than once shown in various colours. Treatment coverage was scaled up from 20% to 55% from year 25 to 60, during which PCR prevalence of malaria decreases and a subtle decrease in the number of unique barcodes observed.

3.3.4. Contribution of cotransmission events to within host parasite diversity

Using the fitted model, I explored the relationship between the proportion of within host parasite strains that are highly related to other within host strains, with highly related parasite strains indicative of cotransmission events. The model predicted proportion of within host parasite diversity that is due to cotransmission events was shown to increase at lower transmission intensities (Figure 3.4a). I predict that at PCR prevalence less than 11.5%, more than 50% of strains within polygenomically infected individuals of all ages result from cotransmission events, rather than superinfection. This is based on the assumption that highly related parasites, which are more than 50% IBD with other parasites, have originated from a recent common ancestor, and as such reflects the proportion of within host genetic diversity that is due to cotransmission events rather than superinfection. I also predict this relationship is dependent on the age of individuals sampled, with

parasites within younger individuals more likely to be more highly related. This reflects the decreased chance that younger individuals will have been superinfected, which increases the likelihood that any polyclonal individuals are the result of a cotransmission event. In Figure 3.5b, the model predicted relationship between mean IBD in mixed infections and the fraction of mixed infections is shown. The model predicted relationship is comparable to estimates of IBD from whole genome sequence data collected from sites across Africa and Asia as part of the Pf3k study (Zhu et al. 2019). However, the model predicts significantly lower mean IBD in settings with a high fraction of mixed infections compared to the estimates based on the whole genome sequencing data, with samples from sites in Ghana, Malawi, Mali and the Democratic Republic of the Congo exhibiting higher mean IBD than predicted by the model.

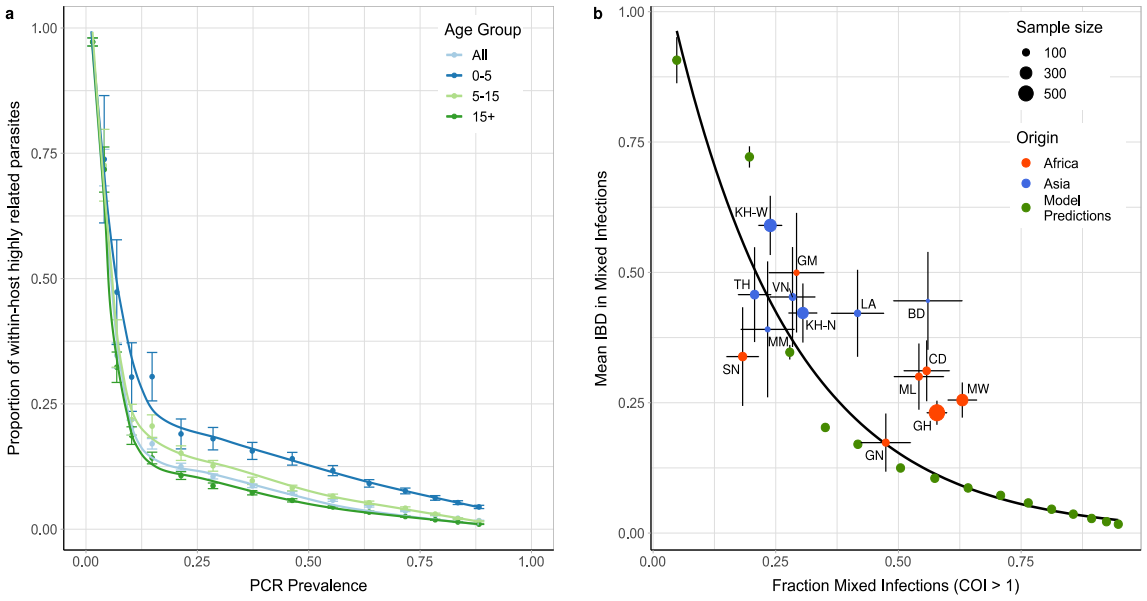


Figure 3.6 Contribution of superinfection and cotransmission to within host parasite relatedness. Model predicted relationship between the mean within host proportion of highly identical parasite strains (>50% of loci comparisons are identical by descent (IBD)) against PCR prevalence. The relationship is shown for all ages and for three age groups: 0-5 years, 5-15 years and 15+ years. 95% confidence intervals are shown with vertical error bars. In **b)** the mean IBD in mixed infections (COI > 1) is shown against the proportion of mixed infections. The model predicted relationship is shown with the green points and the black curve. The model estimates are compared to estimates of IBD from whole genome sequence data collected in sites across Africa and Asia, which were estimated previously in Zhu et al (Zhu et al. 2019). Populations are coloured by continent, with size reflecting sample size and error bars showing ±1 standard error of the mean. Abbreviations: SN-Senegal, GM-The Gambia, NG-Nigeria, GN-Guinea, CD-The Democratic Republic of Congo, ML-Mali, GH-Ghana, MW-Malawi, MM-Myanmar, TH-Thailand, VN-Vietnam, KH-Cambodia, LA-Laos, BD-Bangladesh.

3.4. Discussion

In this chapter I have extended a previously developed model of malaria transmission to include individual mosquitoes and discrete parasite populations. The percentage of sporozoites that are successful within an infectious bite was estimated to be 20% (95% CI 10%-29%), and was estimated by fitting the model to 3003 paired measures of the complexity of infection and age of individuals in 5 sites across Kenya and Uganda. The fitted model was used to initially estimate the proportion of the within host parasite genetic diversity that is the result of cotransmission events resulting in the acquisition of highly identical parasite strains, as opposed to strains acquired through superinfection events. I predict that for malaria prevalence greater than 11.5%, the majority of genetic variation within hosts is generated through superinfection events. To my knowledge this is the first attempt to characterise this relationship across the full transmission intensity spectrum seen within sub-Saharan Africa and represents a move towards standardising which genomic metrics should be used at different transmission ranges.

The main parameter fitted in the developed transmission model was the percentage of sporozoites from an infectious bite that survive to found an infection. The decision to include this parameter was prompted by an analysis of polyclonal infections in Senegal, which demonstrated that cotransmission of genetically related strains contributed towards within host genetic diversity (Wong et al. 2017). The relatedness within polyclonal infections was found to be significantly greater than would have been expected solely by random reassortment due to superinfection. However, the specific contribution of both coinfection and superinfection was unable to be quantified and was simply shown to be present to an unknown degree. In addition, how these two processes determine the population level of genetic diversity and complexity of infection (COI) was suggested as being likely dependent on the transmission intensity of the region of study.

In my model, I have assumed that the absolute number of sporozoites passed on during an inoculation event is independent of the transmission intensity. Consequently, this leads to superinfection contributing more to parasite diversity within hosts at higher transmission intensities. This observation has been previously hinted at by a study of kinship within multiple genotype infections in Malawi and Thailand, in which superinfection was shown to be the major driver of multiple infections in Malawi, whereas both mechanisms contributed comparably in Thailand due to the lower transmission intensity (Nkhoma et al. 2012). This relationship was shown with the fitted model in Figure 3.6a, with the proportion of highly related parasites within an infection decreasing at higher transmission intensities. Similarly, the model predicted an exponential relationship between mean IBD in mixed infections and the proportion of mixed infections (Figure 3.6b). The exponential relationship

was similar to findings in a recent study of IBD, which used whole genome sequence data to explore this relationship (Zhu et al. 2019). However, the model predicted significantly lower IBD at higher transmission settings (settings with a higher fraction of mixed infections) than observed in the data presented in Zhu et al. There are a number of reasons for this. Firstly, the whole genome sequence data was collected from individuals of unknown age as part of a convenience sample. If the samples were collected exclusively from younger individuals, the results in Figure 3.6a would suggest that the mean IBD would be higher than if the samples were collected across all ages. Secondly, in the study by Zhu et al., the estimated COI across all sites was less than 2, which is significantly lower than COI estimates from the sites in Kenya and Uganda in Figure 3.3. Given that some of the African study sites in Zhu et al. are in areas of high transmission intensity, it seems likely that the convenience sampling scheme used has selected for individuals with lower COIs. One explanation could be that the individuals chosen for sequencing receive treatment more regularly, which reduces the probability of parasite strains from superinfection events being present at the time of sampling. This could be due to their age, or due to their enrolment in the study that resulted in them being selected for sequencing. Ultimately, without this metadata it is very difficult to draw any conclusions about the validity of the model predictions in Figure 3.6b, although the broad similarity is encouraging. The importance of individual-level metadata will be further explored in chapter 4.

The developed model was also shown to be sufficient for recovering expected behaviour under simple population genetic models. For example, the rate at which genetic diversity is lost and the effective population size decreases was confirmed to be affected by both the transmission intensity, which promotes recombination events, and the human population size. Lastly, the impact of changing transmission intensity was briefly examined through the impact on the frequency of unique parasite barcodes after the scale up of treatment coverage. This brief assessment showed that parasite genetic variation is affected by changes in malaria prevalence and will be explored in greater detail in chapter 4.

This study has some important limitations. Firstly, we assumed there is only one parameter detailing the percentage of sporozoites that successfully progress to a blood-stage, which is the same for all study sites considered. This is likely a simplification, but our observation of 20% sporozoites surviving from an individual mosquito feed is comparable to Bejon et al's observation of 25% (14 sporozoites surviving from an assumed total of 55 sporozoites resulting from five mosquito bites) of sporozoites successfully progressing to blood-stage infection (Bejon et al. 2005). This estimate is also comparable with the two other attempts within the literature that look at estimating the number of successful sporozoites arising from an infectious mosquito bite. Both methods used similar approaches from vaccine efficacy studies, and estimated that 14 (Bejon et al. 2005) and 16 (White 2012) sporozoites on

average successfully progress to blood-stage infection from a total of 5 infectious bites. It is, however, higher than estimates based on transmission efficacy studies (Smith et al. 2010a). The model fitting, however, revealed that the sensitivity to this parameter was low, with the confidence intervals for a value of ζ equal to 0.20 overlapping intervals for values of ζ ranging from 0.1 to 0.29. This is highlighted when we re-examined the model predicted relationship between *msp2* COI and prevalence with these values, which showed only slight changes to the predicted COI (Appendix Figure 3.1). The fitted estimate was also based on model fits to the administrative mean prevalence as opposed to the recorded prevalence in the specific study sites. For example, the study site in Jinja District, Walukuba, was observed to have the lowest parasite prevalence of all three study sites in Uganda (Nankabirwa et al. 2015). If we had used this prevalence value as opposed to the administrative prevalence value, the parameterised model would have failed to predict the pattern of COI in Walukuba (Appendix Figure 3.2), which may suggest that this study site exhibits higher heterogeneity in the force of infection.

Secondly, although the developed model enables the simulation of parasite genetics, the model does not attempt to fully model the impact of different parasite variants on the acquisition of immunity. Instead the acquisition of immunity is still modelled through exposure to infectious bites rather than via exposure to specific parasite variants. Efforts to characterise and define the genetic diversity of *P. falciparum* have shown that the parasite is capable of maintaining considerable levels of genetic diversity. This is most noticeable among gene families such as the var genes (Verity et al. 2018), which are reflective of balancing selection maintaining phenotypic variation within the host and involved in moderation disease outcomes (Bopp et al. 2013). Developing mathematical models of infectious diseases that capture the full range and effects of pathogen genetic diversity, however, has been noted previously as one of three major challenges in the modelling of infectious diseases (Childs et al. 2015). Although I do not attempt to model immunity through exposure to different var repertoires, the model is sufficient to capture the level of within host parasite genetic diversity, as measured by the complexity of infection at different ages. In addition, my developed model addresses a number of other limitations and challenges posed in the review of modelling challenges presented by Childs et al (Childs et al. 2015), including:

- The generation of genetic diversity via recombination.
- Modelling frameworks enabling differential infectiousness of each parasite strain in multiply infected individuals.
- Altered clearance rates of parasite strains acquired from different infections.
- The role of superinfection on the duration and dynamics of an individual's relative infectiousness to mosquitoes.

The extension made in addressing the above challenges has introduced a number of modelling assumptions. Competition between parasite strains within a multiply infected individual is not modelled, which if included would also impact how the genetic variation within an infected individual may be passed over to the next human host (Conway et al. 1999; Gog et al. 2015). If between strain competition is sufficiently high to cause a substantial decrease in either the duration of an individual strain's infection or the rate at which new strains can be acquired, then the estimated frequency of sporozoites that survive is likely a conservative estimate. However, studies attempting to measure between strain competition are limited as a result of the complicated experimental set up. Most studies examine the complexity of infections using *msp1/2* genotyping and measure patient parasitaemia, which often shows that parasite density does not increase with increasing complexities of infection (Bell et al. 2006; Nkhoma et al. 2018a). These findings are often suggested to indicate within host competition and that the overall infectiousness of an infected individual does not change with regards to the number of genetically distinct strains. However, they do not give an indication of how the relative infectiousness of each strain is affected in mixed infections. For example, does the initial strain dominate through early monopolisation of within host resources? Or are there sufficient distinct niches within an individual that enable multiple parasite strains to exist in parallel but simply at lower parasite densities? Other individual factors are also shown to be of arguably more importance, such as patient age and temperature, with respect to onward infectiousness of an individual (Sondo et al. 2019). To measure the individual parasite densities of different strains, murine models of malaria dynamics have been used to some success for looking at mechanistic basis of within host competition (Sondo et al. 2019). However, these studies have often focussed on questions relating to competitive suppression and release with respect to antimalarial resistance, due to the profound effects that it would have on the speed at which resistance may spread (Huijben et al. 2011; Bushman et al. 2016, 2018). Consequently, at this stage it seemed not tractable to include within host competition that altered the relative infectiousness beyond the level expected based on the expected duration of infection of the individual parasite strain.

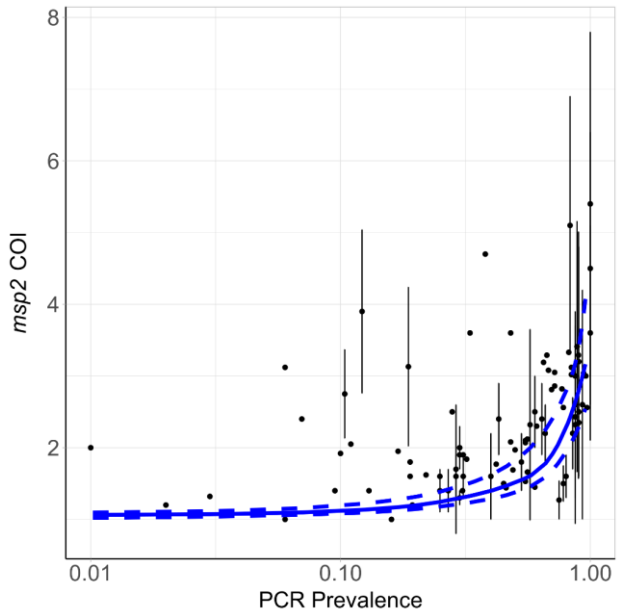
3.5. Conclusion

In this chapter I have developed a new individual-based model of the transmission of malaria for the simulation of parasite genetics. The model represents an extension of the version of the Imperial College transmission model previously developed for the simulation of *pfhrp2* gene deletions, which now allows parasites to be modelled more flexibly by tracking parasites using a genetic barcode. The extended model required the parameterisation of an additional parameter governing the percentage of sporozoites that survive from an infectious event. The parameter was fit through comparison to

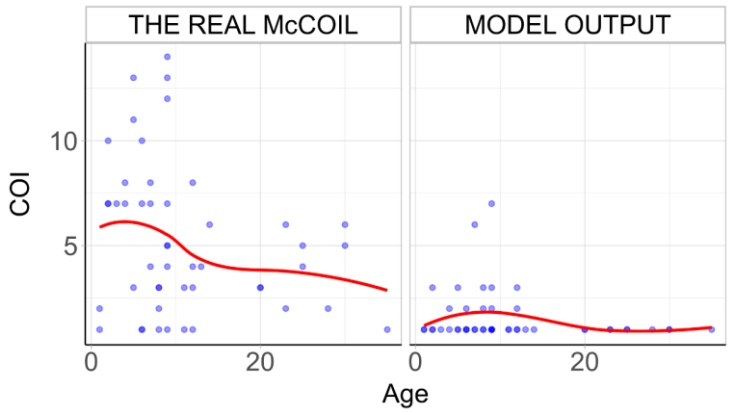
estimated relationships of the complexity of infection against age from five sites in Uganda and Kenya. Estimates of COI were produced using my developed wrapper for *THE REAL McCOIL* method, which now enables larger numbers of samples and loci to be handled. The fitted parameter was found to be equal to 20%, which is in line with previous estimates calculated from both vaccine efficacy studies and human challenge experiments. The fitted model was used to highlight that the model is able to reproduce expected population genetics results related to the loss of genetic diversity over time due to genetic drift as well as due to decreasing malaria prevalence resulting from increased treatment coverage. Lastly, the model was used to characterise the level of identity by descent between parasite strains within polyclonally infected individuals. Using this approach, I predicted that for malaria prevalence greater than 11.5% PCR PfPR, the majority of genetic variation within hosts is generated through superinfection events. My characterisation of the relationship between the level of highly related parasite strains within an individual and transmission intensity is the first to use a mathematical modelling approach to explore this dynamic. This has important implications in understanding how changes in transmission intensity will affect the level of inbreeding within the parasite population. These ideas will be developed further in chapter 4, in which multiple summary measures of parasite genetic diversity will be reviewed in an evaluation of how parasite genetics could be used as a tool for the surveillance of malaria.

3.6. Appendix

3.6.1. Appendix Figures



Appendix Figure 3.1 Model predicted relationship *msp2* COI and PCR prevalence. The blue solid line shows the relationship for the fitted value of ζ equal to 0.20. The dashed lines above and below this in blue show the relationship for values of ζ equal to 0.29 and 0.10 respectively. The point-ranges in black show the observed values of COI by msp2 genotyping from the literature review.



Appendix Figure 3.2 COI vs age using sub-county prevalence. The fitted model predicted relationship between COI and age for Walukuba, if the prevalence simulated was assumed to be equal to the prevalence within the sub-county surveyed, rather than the prevalence for the administrative region. Model fitting conducted in Figure 3.3a used the administrative region prevalence as estimated by the Malaria Atlas Project, which resulted in good agreement between COI and prevalence.

Chapter 4. Evaluating the performance of malaria genomics for inferring changes in transmission intensity using transmission modelling

In the previous chapter I detailed the development of a new mechanistic model of malaria transmission that enabled the simulation of malaria genetics. The model was used to explore the contribution of superinfection and cotransmission events towards within host parasite genetic diversity. In this chapter, I continue to use the same model to explore how a number of measures of parasite genetic diversity correlate to changes in transmission intensity and define their individual statistical power. Finally, I show how different metrics can be combined to predict the prevalence of malaria while highlighting the increased predictive accuracy when patient meta data is also recorded.

4.1. Introduction

Molecular tools are increasingly being used to understand the transmission histories and phylogenies of infectious pathogens (Hall et al. 2015). Using phylodynamic methods it is now possible to estimate the historic prevalence of infection directly from molecular data, even in organisms with relatively complex lifecycles (Volz et al. 2009). However, these tools typically rely on pathogens having an elevated mutation rate and not undergoing sexual recombination, which allows for the application of coalescent theory (Grenfell et al. 2004). Consequently, these techniques are yet to be adapted for the study of *P. falciparum* malaria, which is known to undergo frequent sexual recombination. In addition, malaria transmission between both the human and the mosquito hosts involves a series of population bottlenecks (Vaughan 2007; Churcher et al. 2010), which combined with the rapid sexual stage involving a single two-step meiotic division (Bennink et al. 2016), have marked effects on the population genetics of *P. falciparum* (McKenzie et al. 2001; Chang et al. 2013). This is extenuated by evidence of cotransmission of multiple clonally related parasites (Wong et al. 2017), which combined with host mediated immunity (Barry et al. 2007; Portugal et al. 2011) and density-dependent regulation of superinfection (Bruce et al. 2000; Pinkevych et al. 2013) result in a complicated network of processes driving the genetic diversity of the parasite population within an individual host.

Despite this substantial complexity, an increasingly nuanced understanding of the processes shaping parasite genetic diversity is appearing, with multiple genetic metrics proving promising for inferring transmission intensity (Daniels et al. 2013; Nkhoma et al. 2013). For example, measures of the multiplicity of *P. falciparum* infections have been shown to be useful for identifying hotspots of

malaria transmission (Bejon et al. 2010; Karl et al. 2016). The spatial connectivity of parasite populations has also been shown to be well predicted by pairwise measures of identity by descent (Omedo et al. 2017a; Taylor et al. 2017). More recently, it has been shown that malaria genotyping could be used to enhance epidemiological surveillance (Daniels et al. 2015), however, two main challenges have been identified before molecular tools could be used in an operational context. The first is that our understanding of the relationship between transmission intensity and within host parasite genetic diversity is incomplete. Combined models of both population genetics and malaria epidemiology would allow us to develop a more detailed view of both processes, yet these two approaches are largely explored separately. Recent efforts have been made to incorporate both modelling scales within one framework (Nguyen et al. 2015), with the concomitant modelling of resistance evolution both within and between hosts yielding important insights into the evolution of drug resistance (Legros and Bonhoeffer 2016). However, the realism of either the transmission process or the genetic evolutionary process has been limited in these models, with the representation of recombination and the parasite lifecycle within the mosquito often simplified. This makes the generalisability of using molecular tools for surveillance difficult. More realistic models are subsequently needed to better understand the relative contribution of superinfection and cotransmission of genetically related parasites towards the parasite genetic diversity observed within humans, which is known to depend on transmission intensity (Nkhoma et al. 2018b; Wong et al. 2018). The second challenge is to understand in what situations molecular tools will offer advantages over traditional surveillance, and how many samples are needed for reliable inference.

Here I use the mathematical transmission model developed in chapter 3 to address these challenges. Using the fitted model, I characterise how six measures of parasite genetic diversity respond to changes in transmission intensity. I continue by conducting a power analysis, assessing the ability of each metric to detect changes in transmission intensity as a function of the number of available samples. I conclude by building an ensemble statistical model, which demonstrates how routinely collected clinical genotype samples could be used for accurate prediction of malaria prevalence using as few as 200 SNP genotyped samples.

4.2. Methods

4.2.1. Impact of changes in transmission intensity on parasite genetic diversity

The effect of declines in transmission intensity on four measures of within host genetic diversity was explored. The four measures considered are:

- Mean COI - the number of genetically distinct parasite strains within an individual
- % Polygenomic - the percentage of polygenomic infections, i.e. COI >1
- % Unique - the percentage of unique parasite genotypes
- COU – coefficient of uniqueness, which is a newly defined metric given by:

$$COU = 1 - \frac{(\sum_i^n x_i^2) - \frac{1}{n}}{\left(1 - \frac{1}{n}\right)}; \quad 0 \leq COU \leq 1$$

where x_i is the frequency at which barcode i occurs within a sample of size n. COU = 0 when all barcodes within a sample are identical, and *COU* = 1 when all barcodes within a sample are unique.

Ten stochastic realisations of 100,000 individuals were simulated for 50 years with an initial parasite prevalence measured by PCR equal to ~70% and a fixed importation rate to ensure both a genetic and epidemiological equilibrium. Once at equilibrium, three differing levels of intervention scale-up (low, medium, high) were introduced that led to an absolute reduction in parasite prevalence from 70% to 45%, 20% and 5% respectively after 10 years. The scale-up of interventions resulted in an increase in the coverage of ITNs (maximum after 10 years: 30%, 60%, and 90%), IRS (maximum after 10 years: 20%, 40% and 60%) and treatment (maximum after 10 years: 15%, 30%, 45%). For all simulations, the monthly mean for each genetic marker was recorded for the whole population as well as within three age ranges (0-5 years old, 5-15 years old and over 15 years old), and within individuals who were asymptomatic or symptomatic at the time of sample collection.

An identical analysis was conducted at a lower starting prevalence, with maximum reductions in parasite prevalence by PCR from 35% to 20%, 2% and ~0% after 10 years, in order to assess the change in two measures of identity by descent (IBD), pIBD and iIBD (defined below). The population mean IBD (pIBD) I define as the mean number of loci in pairwise comparisons between samples that are identical across all loci in terms of their 24-locus identity barcode (focusing on genotypes that could be detected by microscopy only), i.e. it is the mean proportion of shared ancestry between samples. The individual mean IBD (iIBD) is the mean number of identical loci of the 24-locus identity barcode within individuals who are polygenomically infected. If all sampled individuals are monogenomic, then iIBD is set equal to 1.

4.2.2. Statistical power analysis of parasite genetic measures

To evaluate the utility of the considered measures of parasite genetic diversity, I conducted an analysis to characterise the predictive power of each metric for detecting changes in transmission intensity, and their sensitivities to the sample size chosen. In an analogous design to earlier simulations, I

measured sample mean measures of the COI, % Polygenomic, % Unique, COU, iIBD and pIBD at yearly intervals for the first five years after the initiation of the ten-year scale-up of interventions.

Sensitivity to the sample size of each metric was assessed by sequentially sampling subsets of the data and comparing the mean difference in metrics. Sample sizes between 10 and 600 individuals were explored, with 100 samples drawn from a stochastic realisation at years 0, 1, 2, 3, 4 and 5, and comparisons made between years 1-5 and year 0, i.e. 0-1, 0-2, ... 0-5. All samples were collected from individuals aged between 5-15 years old. One-tailed Monte Carlo p-values were generated for each subsample by 1000 permutations of the years that samples were collected from. The power of each metric was defined as the proportion of subsamples for which 95% of the permuted mean differences were greater or less than the observed mean difference, with the direction of the tail dependent on whether the metric is expected to decrease or increase respectively in response to a decrease in transmission intensity. The overall power for each metric was calculated as the mean power of ten stochastic realisations, and repeated at two different starting parasite prevalence by PCR (~45% and ~22.5%). Metrics based on comparisons of IBD were only assessed for the lowest starting parasite prevalence. The performance of each metric was also explored under the assumption that it was not possible to phase all genotypes within the samples collected, and that only the dominant genotype was able to be called.

4.2.3. Statistical modelling of the predictive performance of malaria genomics for surveillance

A statistical model was constructed to predict malaria prevalence using the genomic metrics explored thus far, with three different assumptions about the availability of patient metadata (no metadata, patient age only, and both patient age and symptomatic status of infection). To assess the utility of such a model for surveillance, samples of 200 individuals were taken from a range of simulations that span the transmission, seasonality and intervention coverage range seen in sub-Saharan Africa. I used the sampled mean measures of the genomic metrics discussed, and where available summaries of the age and clinical status of samples to create my model simulated datasets. 25% of simulated datasets were held back for out-of-sample testing.

I used an adapted workflow from previously published data science workflow for building an ensemble model for predicting malaria prevalence (Bhatt et al. 2016). Three different statistical models (gradient boosted trees, elastic net regression model and random forests) were fit to the model simulated data. The predictions of these level 1 models were subsequently used to train an ensemble model using a linear optimisation based on the root mean squared error (RMSE) of the level 0 models. When training both the level 1 models and the ensemble, K-fold cross validation sets were performed 25 times and

subsequently averaged to reduce any bias from the cross validation set chosen. The averaged cross validation results were used to assess the performance of the ensemble model on the testing dataset by comparing the RMSE, mean absolute error and the correlation under the different assumptions about the availability of patient metadata. Lastly, the trained ensemble model was used to predict the prevalence of malaria for the study sites considered within Uganda and Kenya.

4.3. Results

4.3.1. The impact of intervention strategies on parasite genetic diversity

Using the parameterised model, I first modelled how a reduction in transmission would affect four genetic metrics as the prevalence of malaria declined due to the scale up of interventions (Figure 4.1). The genetic metrics explored were: 1) the population mean complexity of infection (COI); 2) the percentage of samples that are polygenomic (COI > 1); 3) the percentage of unique parasite 24-SNP barcodes and 4) the coefficient of uniqueness (COU) (Figure 4.1).

The model was initiated at 70% PCR prevalence before three levels of intervention scale-up (low, medium and high) were simulated, reducing prevalence to ~45%, ~20% and ~5% respectively after ten years. I predict that all four metrics decline proportionally with declining malaria prevalence (Figure 4.1a). The model predicts that the specific relationship depends on the population chosen for genetic testing, with increasing COI predicted in older ages. The percentage of unique samples varied greatly depending on the on the sub-population sampled, reflecting difference in the absolute numbers of individuals that fall within each sub-population. Samples taken from individuals with asymptomatic malaria were predicted to have the highest COI and percentage of polygenomic samples. Across the scenarios simulated, metrics based on the complexity of infection (COI and % Polygenomic) showed a higher level of correlation with changes in the prevalence of malaria than measures based on the uniqueness of samples (COU and % Unique) (Table 4.1). In addition, samples collected only from patients with symptomatic malaria led to metrics that were the least correlated with reductions in prevalence, resulting from the decreased number of available samples. This effect was most noticeable when assessing the % of unique samples within clinical samples, which had a correlation coefficient of 0.24 with PCR prevalence (Table 4.1).

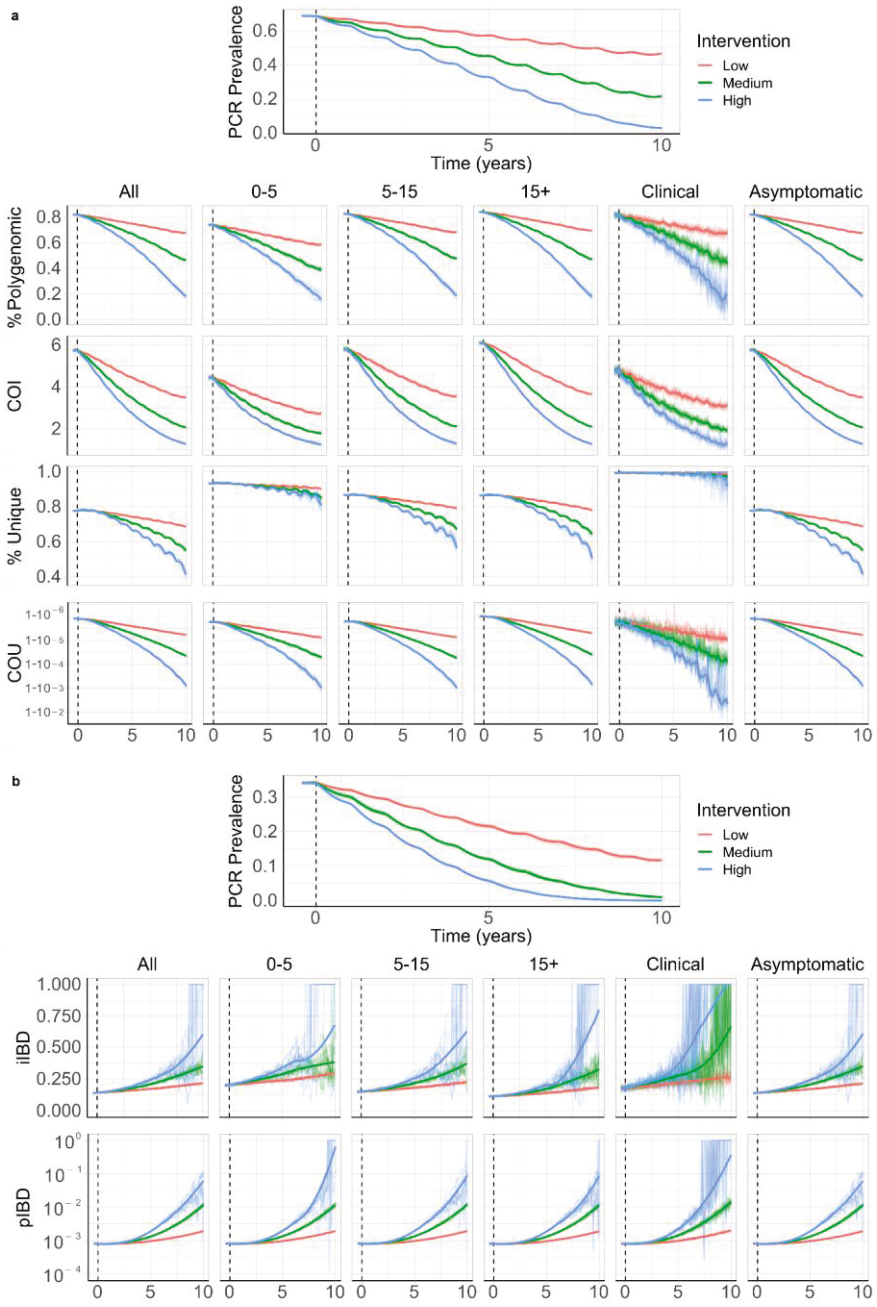


Figure 4.1 Age and sampling dependent impact of changes in transmission intensity upon genetic metrics of transmission intensity. In a) the top plot shows the change in PCR prevalence after the introduction of 3 different levels of intervention scale up, with both the 10 individual stochastic realisations and the mean local regression smoothed relationship shown. The following four rows show the population mean percentage of the population that are polygenomically infected, the complexity of infection (COI), the percentage of samples that are genotypically unique (% Unique) and the coefficient of uniqueness (COU) for the prevalence declines seen in the first row. The metrics are stratified into columns by the sampling scheme chosen. In b) the top plot shows the change in PCR prevalence, which reaches <1% in the highest intervention arm. The following rows show the within host identity by descent (iIBD) mean across the 24 identity loci considered, and the population mean pairwise measure of IBD (pIBD). In both the same sampling stratification is used as in a). In all plots the vertical dashed black line shows the time from which the scale up of interventions starts (Time = 0 years).

I also assessed measures of parasite genetic diversity based on comparisons of the number of loci that are identical-by-descent (IBD), which included the within host pairwise mean proportion of loci that are IBD (iIBD) and the population pairwise mean proportion of loci that are IBD (pIBD). I predict that both metrics increase in response to declines in prevalence, however, I predict that pIBD only increases substantially at PCR prevalences less than 15% (Figure 4.1b). Consequently, metrics based on IBD were explored at a lower starting prevalence of 35% PCR prevalence before the scale up of interventions. The shape of the increase in iIBD was predicted to be dependent on the population sampled (Figure 4.1b), with iIBD increasing quicker in symptomatic individuals. iIBD, however, becomes less informative as transmission intensity declines, with individuals less likely to be infected with multiple strains due to the lower rates of superinfection.

Table 4.1 Kendall rank correlation coefficients between genetic diversity metrics and parasite prevalence. Coefficients are bound between -1 and 1, with 1 indicating perfect ranked positive correlation and -1 indicating perfect ranked negative correlation.

Sampled	% Polygenomic	COI	% Unique	COU	iIBD	pIBD
All	0.97	0.96	0.83	0.93	-0.89	-0.86
0-5	0.96	0.96	0.73	0.93	-0.80	-0.86
5-15	0.97	0.96	0.83	0.93	-0.86	-0.86
15+	0.97	0.96	0.83	0.92	-0.84	-0.86
Clinical	0.87	0.91	0.24	0.75	-0.64	-0.85
Asymptomatic	0.97	0.96	0.83	0.93	-0.89	-0.86

4.3.2. Power Analysis

To evaluate the performance of each metric for detecting annual changes in the prevalence of malaria, I calculated the statistical power for each metric at different sample sizes, focussing on samples collected from children aged between 5-15 years old. I estimate that after 5 years of intervention scale up, corresponding to an absolute decrease in malaria prevalence by PCR of 20%, no more than 350 samples are required for each metric explored (except for iIBD) to detect the change in transmission intensity 80% of the time (Figure 4.2). The predictive power, however, declined across all metrics when the effect size, i.e. the decrease in prevalence, decreased. With 600 samples, each metric had less than 40% power to detect the decrease in prevalence after 1 year. The performance of each metric was additionally dependent on the starting prevalence, with metrics based on the uniqueness of samples (COU and % Unique) predicted to be more powerful at lower starting prevalences compared to higher prevalences (Figure 4.2b). Metrics based on measures of IBD were overall less powerful, with

the predictive power of iIBD being less than 80% across all years and sample sizes (Figure 4.2c). pIBD only exhibited a predictive power greater than 80% when detecting the largest change in prevalence between 22.5% and 8%, requiring over 225 samples.

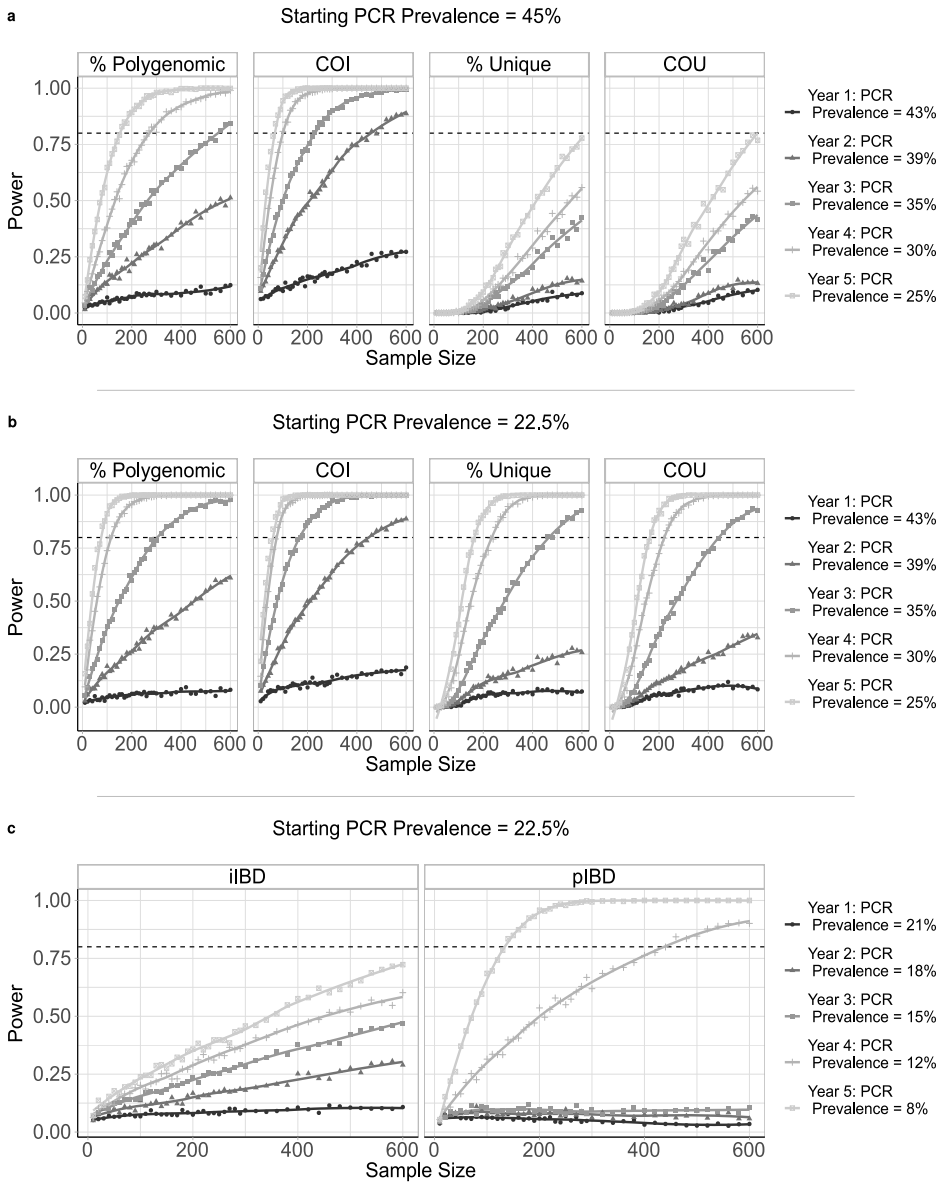


Figure 4.2 Predictive power of genetic metrics. The distribution of sample means of six metrics of parasite genetic diversity were compared for five years following the initiation of the scale up of intervention coverage. For each sample size, the power is defined as the proportion of 100 subsamples comparing year 0 and years 1-5 for which a significant difference in the mean was observed, estimated using one-tailed Monte Carlo p-values generated by 1000 permutations of the years' samples were collected in. In **a)** the metrics assessed are the percentage of samples that are polygenomic, the complexity of infection (COI), the percentage of barcodes within samples that are unique, and the coefficient of uniqueness (COU). The power of each metric was compared across five years in which a 20% absolute decrease in parasite prevalence from 45% was observed. The same information is shown in **b)**, but for a 14.5% absolute decrease in prevalence from 22.5% over 5 years. In **c)** the metrics considered are the mean within host identity by descent (iIBD) and the population mean pairwise measure of IBD (pIBD). In each plot 80% power is shown with the horizontal dashed line.

The power of COU, % Unique and pIBD were noticeably worse when it was assumed that samples from polygenomically infected individuals could not be phased (Appendix Figure 4.1). Under this assumption I assume that it is not possible to observe the genotype of each strain and consequently only the major haplotype within an individual is available, i.e. calling the most abundant allele at each locus of the barcode, which negates our ability to measure an individual's iIBD as we are only able to observe one parasite genotype per individual. Across the full range of malaria prevalence simulated, measures of COI and COU were consistently predicted to be the most powerful, with % unique samples and IBD metrics demonstrating increased power to detect changes in transmission in areas with lower baseline transmission intensities where I predict the genetic variation to be lower.

4.3.3. Statistical model for predicting transmission intensity

In order to translate the information I have characterised into an effective tool for assisting surveillance programs, a statistical model was created to predict malaria prevalence using genomic metrics derived from parasite SNP genotyping. Due to the difficulty in phasing high complexity infections, I assumed that all collected samples were unphased and as such I did not focus on metrics based on IBD when building our data set for training our statistical model.

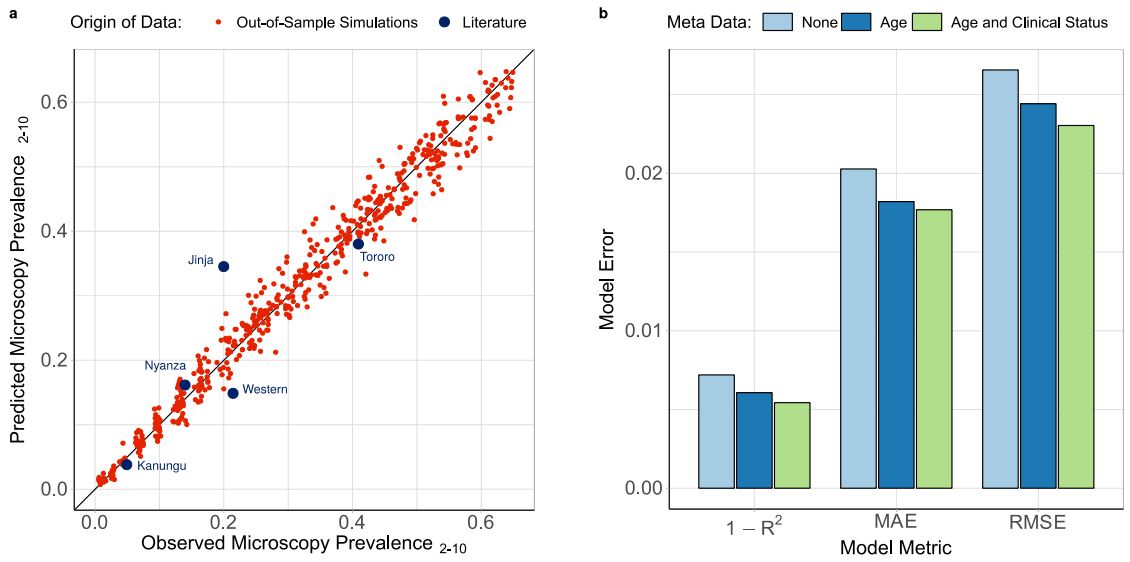


Figure 4.3 Ensemble statistical model predicted malaria prevalence vs observed prevalence. In **a)** the performance of the trained ensemble statistical model is shown, with the model predicted prevalence in red showing the predictions for the out-of-sample test dataset composed of model simulations held back from model fitting. The blue points show the predicted prevalence for the 5 administrative regions considered earlier. In **b)**, the performance of the ensemble model is shown under different assumptions about the availability of patient metadata within simulated data.

The fitted ensemble model performed well on out-of-sample simulation datasets, and was able to identify the underlying model behaviour used to generate the training dataset (Figure 4.3a). The best performing model provided accurate predictions of malaria prevalence when tested on SNP genotype data from the five administrative regions, with an observed mean absolute error equal to 0.055 for these five locations. The performance of the model was enhanced when sample metadata was available (Figure 4.3b), with the ensemble model trained and tested using data with no age or clinical status information consistently performing worse. Similar patterns were also observed when assessing the performance of each of the level 1 models in the ensemble model (Appendix Table 4.1). As in the power analysis, across the range of malaria transmission intensities assessed, measures of COI and COU were observed to be the most informative metrics and sample metadata was observed to contribute 28% towards the total model importance (Figure 4.4).

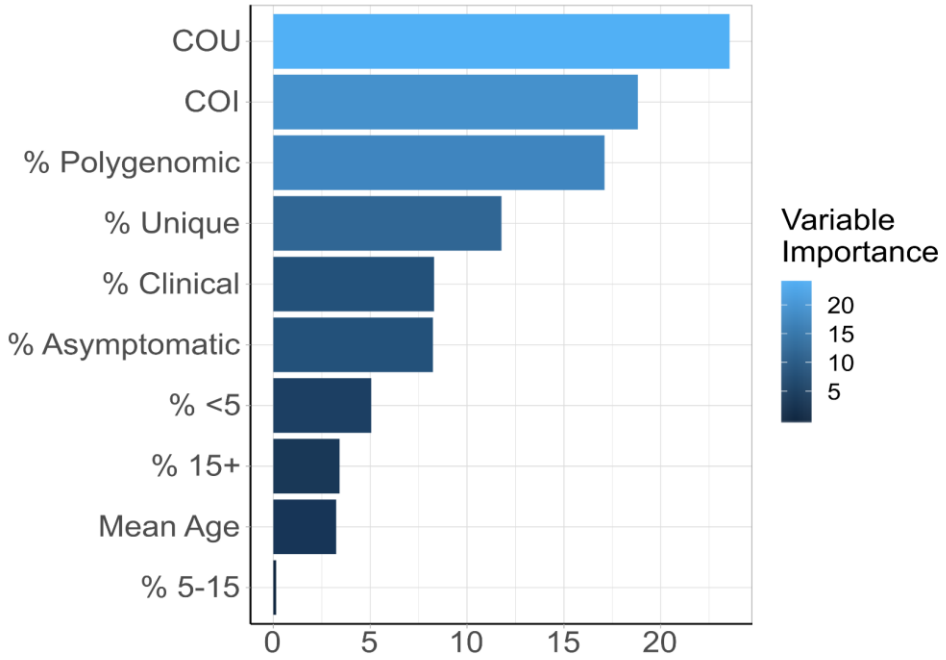


Figure 4.4 Importance of predictor variables within trained ensemble model. The newly defined measure, the coefficient of uniqueness (COU), was observed to be the most important metric, with the six metadata variables (age and clinical status) being the least important. The six metadata metrics contribute 28% of the total model importance, which shows why the inclusion of metadata yields better model predictions.

4.4. Discussion

The substantial reduction in the cost of generating genomic datasets over the last ten years and the establishment of scientific networks committed to generating and sharing genomic data has resulted in an abundance of sequenced *P. falciparum* genomes. This effort has resulted in the identification of loci associated with emerging drug resistance mechanisms (Cheeseman et al. 2012) and assisted in

developing putative novel drug targets (Ludin et al. 2012). Another potential use of malaria sequencing efforts is understanding how malaria genomes can be used to study transmission. Simple population genetics principles predict that in a closed population a reduction in transmission intensity will typically be accompanied by a reduction in parasite genetic diversity, resulting from reduced opportunities for outcrossing to occur within the sexual stages of the parasite's life cycle. However, there is as yet no consensus in the use of parasite genetics for inferring transmission intensity. There is a need to understand the contribution of superinfection and cotransmission towards the within host parasite genetic diversity, which is often highlighted within critiques of early attempts to utilise modelling approaches for transmission intensity inference (Greenhouse and Smith 2015).

Our newly defined measure of parasite diversity, the coefficient of uniqueness (COU), alongside COI were consistently powerful statistical tools for detecting changes in malaria prevalence. This is hardly surprising, as one should consider that the % unique samples and the % of polygenomic samples are simply the extreme cases of these metrics, and so I would expect them to contain less information. Additionally, the power analysis conducted was under the assumption that all samples that could be detected by PCR can be effectively phased. This is an overly ambitious assumption, and it is more correct to assess these metrics under the assumption that polygenomic samples cannot be phased (Appendix Figure 4.1). However, the increase in statistical power when we are able to phase samples should highlight a need within the research field for methods to compare unphased parasite samples, with the majority of samples at higher transmission intensities predicted to have a COI greater than 1.

In the absence of being able to phase polyclonal samples, however, the observed genomic metrics were still informative within the ensemble statistical model developed to translate parasite genetic information into estimates of malaria prevalence. For example, variable importance was observed for each predictor variable, however, COU and COI accounted for nearly half the variance explained. There is also a degree of compensation afforded between metrics, i.e. where one metric becomes less informative, another metric becomes more predictive. For example, at PCR PfPR less than 10%, COI and the % of samples that are polygenomic will become substantially less informative, whereas IBD measures will start being more informative. This is further demonstrated by only needing 200 samples within our statistical ensemble model to produce accurate predictions of the prevalence of malaria, with the addition of individual level metadata yielding further gains in model performance (Figure 4.4b). As more samples are added only modest improvements in model predictive performance are observed (Appendix Figure 4.2). The importance of meta data, specifically the age of individuals, is highlighted in the findings of the model predicted COI between age groups. In Figure 4.2, I compared the COI between asymptomatic and symptomatic individuals, in which I predicted across all ages that asymptomatic individuals have higher COI. However, this finding does not hold when I compare the

COI between symptomatic and asymptomatic individuals at different age groups and across different transmission intensities. For example, in the model fitting conducted in chapter 3, the model predicted that younger, symptomatic children in regions with lower transmission have higher COI than asymptomatic younger children (Appendix Figure 4.3). This finding is reversed, however, at higher transmission intensities reflecting the interaction between acquired clinical immunity and rates of superinfection.

This study has some important limitations. Firstly, from the model fitting conducted earlier, I assumed there is only one parameter detailing the percentage of sporozoites that successfully progress to a blood-stage, which is the same for all study sites considered. In chapter 3, I discussed how this is a likely oversimplification and that this parameter may also depend on the transmission intensity as well as spatial heterogeneity in the force of infection. I continued to discuss how if I conducted model fitting using the prevalence recorded in Walukuba as opposed to the administrative region's prevalence then the current model fit would not have captured the COI against age relationship. However, the fact that the model-predicted COI closely matches the observed data when using the administrative region's prevalence may suggest that parasite genetic metrics are more representative of the prevalence at larger spatial scales, which in turn may reflect human mobility between areas of differing transmission intensity and parasite genetic diversity. This point is worth highlighting as this feature could be of benefit from a surveillance point of view. For example, 200 samples collected in one area may be able to give accurate measures of malaria prevalence within a large area. This could be of particular utility in areas where community surveillance is not feasible, in which samples collected from symptomatic patients attending public health facilities could provide additional information in helping to translate clinical incidence into measures of parasite prevalence.

Secondly, I did not explicitly model the scale-up of vector based interventions, instead incorporating the effects of insecticide treated nets and indoor residual spraying through their impact on the average age of the mosquito population and the rate of anthrophagy. This assumption will cause each individual to experience the same relative reduction in molecular force of infection, i.e. the number of new *P. falciparum* clones acquired over time. Consequently, model predictions are likely to underestimate the variance in the reduction of within host parasite genetic diversity resulting from vector based interventions. This effect would lead to a decrease in the statistical power of the genetic metrics considered and subsequently the sample sizes presented within the power analysis are likely on the lower end of the sample sizes required for a given predictive power.

Thirdly, while the developed statistical model provided accurate estimates of malaria prevalence overall for the five regions, the prediction for Jinja was noticeably worse, which reflects the high COI

observed in that region given its comparatively low prevalence. While I was able to replicate the COI age relationship for this region during model parameterisation, this was largely due to the fact that the historic prevalence for the region was much higher. For this reason, the model predicts that individuals in the region will have higher acquired immunity and will subsequently be able to harbour more infections before developing a fever and potentially being treated and thus clearing infections. The developed statistical model, however, did not include any covariates for historic prevalence or genetic diversity. Subsequently, predictions made by this model largely reflect the mean diversity expected for a given prevalence and will suffer when making predictions for regions that have experienced a recent and large decline in prevalence. Consequently, as more genetic data is collected over time it should become increasingly feasible to extend the methods presented here to better handle rapid declines in prevalence and incorporate historic measures of genetic diversity. Lastly, in our model I have only considered neutral genetic markers that are unlinked. While these loci are informative for capturing standing genetic diversity, I have not considered how selective events may shape the genetic diversity. For example, if drug resistance were to spread quickly through an area it is likely that this would cause a decrease in genetic diversity in neighbouring regions (Imwong et al. 2017b). However, the precise impact that this will have on the metrics explored in this study will depend on both how quickly recombination will result in linkage disequilibrium decay and the strength of the selective sweep. Although these were not assessed in this paper, it would be possible to adapt our model to consider loci under selection and simulate how known factors that affect the speed of selection, such as transmission intensity, treatment rates and the metabolic costs associated with resistance, impact genetic metrics.

One remaining question, however, is that although this study has explored if parasite genetics could be a used for malaria surveillance, I have not explored whether they should. Any assessment of parasite genetics for surveillance must include an assessment of the costs of carrying out the analysis. For example, what are the cost comparisons for collecting 200 high quality blood samples to conduct SNP genotyping versus conducting a household survey in the same administrative region? Is the accuracy comparable to a household survey with this many samples? Additionally, how does the use of parasite genetics compare to using incidence data? Incidence data is more readily available, however, it is reliant on the accuracy of estimates of the population size. In situations where this is not possible, such as migratory populations and clinics with unknown health facility catchment areas, there may be a niche for parasite genetics. However, it is more likely that genetics will support current estimations based on converting incidence data to prevalence. In the study by Daniels et al. (Daniels et al. 2015), an increase in clinical incidence was observed in Senegal during a time of increased intervention scale up. In this instance, parasite genotyping was able to assess if transmission was

increasing or whether the increase in incidence was due to improved case reporting and access to treatment. Following studies evaluating the benefit of parasite genetics for surveillance will thus need to delineate both the circumstances in which genetics can support current surveillance methods, such as clinical incidence data, and map the regions in which accurate estimations of population sizes are not currently available and are unlikely to be possible in the future, such as displaced populations due to conflict. Additionally, any evaluation should also consider comparisons to alternative surveillance methods that may also be possible. A recent study exploring ante-natal care (ANC) clinics in Eastern Democratic Republic of Congo demonstrated that ANC prevalence could also be used for retrospective surveillance and evaluation of the impact of malaria interventions (Hellewell et al. 2018). This approach also has the benefit of being largely unbiased by population size due to the widespread and routine malaria testing conducted in ANC clinics regardless of symptoms.

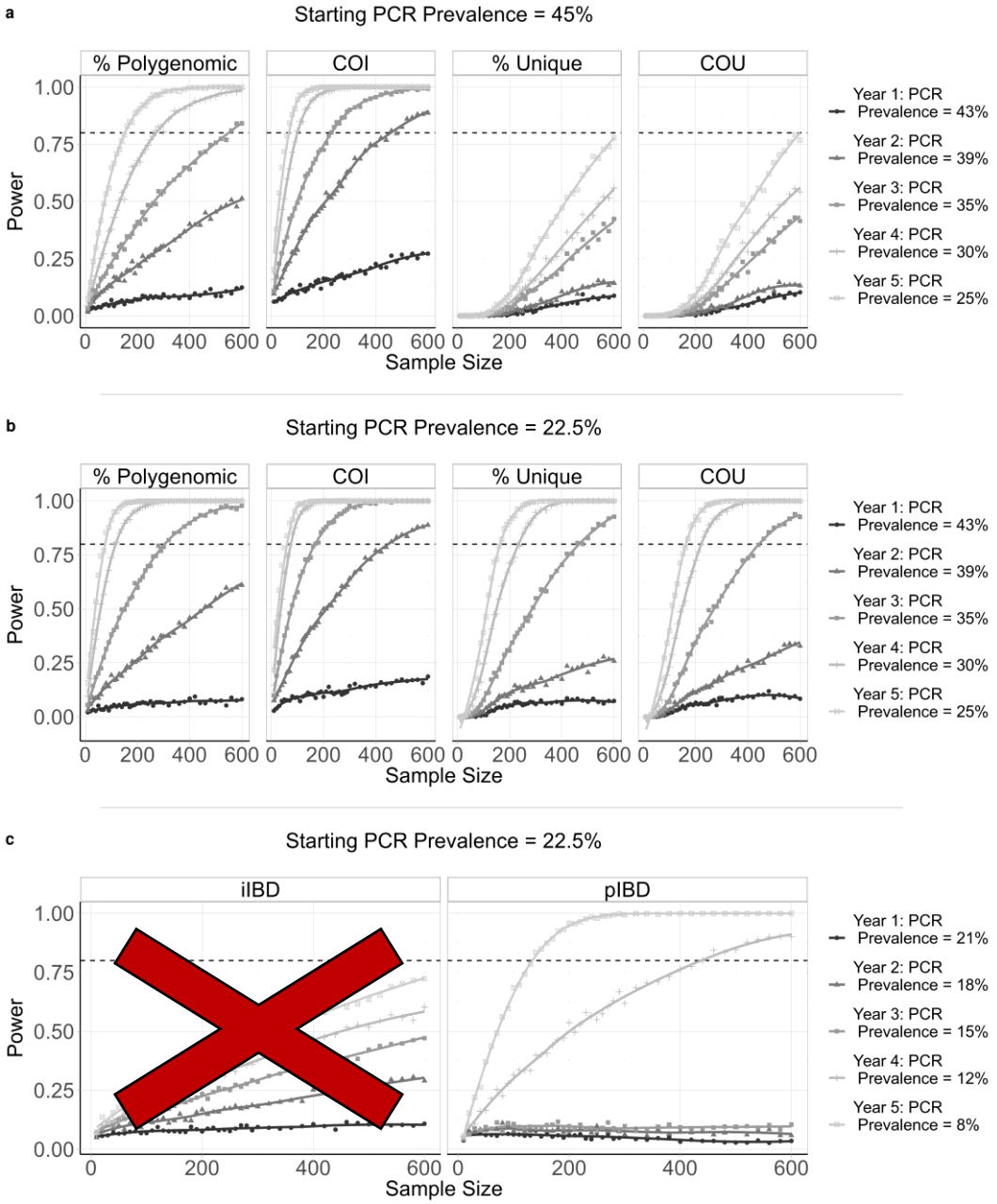
The 2018 world malaria report shows that the reductions in the global burden of malaria made since 2000 may be stalling, with 2 million more cases of malaria estimated in 2017 compared to 2016 (World Health Organization 2018c). These declines have necessitated the development of new tools to enhance current surveillance efforts. In this study, I have shown that that malaria genetic metrics could provide an additional toolkit for operational surveillance. In particular, a combination of metrics focussed on the complexity of infections, the frequency and uniqueness of genotyped barcodes and measures of identity by descent could be used for inferring the prevalence of malaria across the current range of malaria prevalence observed in SSA. It is important to highlight that there is still a need to understand the cost-effectiveness of these tools compared to current surveillance methods. In most endemic areas, routinely reported incidence data provides a temporally and spatially rich measure of malaria transmission. However, the methods developed here could complement measures of malaria incidence in areas where only clinical incidence data is available or areas in which the spatial coverage is poor. It is hoped that these findings, in particular the importance of sample metadata and quantifying the contribution of cotransmission and superinfection events have in shaping genetic diversity, can guide future efforts by the wider community for utilising malaria genotyping for epidemiological surveillance.

4.5. Conclusion

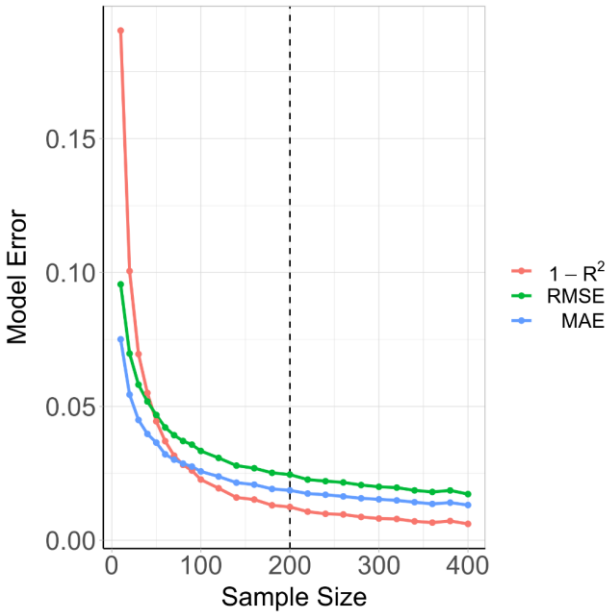
In this chapter I have used the model fitted in chapter 3 to predict how six measures of the parasite genetic diversity within individuals changes in response to transmission intensity. Once characterised, I built a statistical model that combines the information within each metric to predict the current prevalence of malaria. The increased prediction accuracy of the model when the age and clinical status of the sampled individuals is available is hoped to reinforce the need for patient metadata to be recorded and made available within all future attempts to use parasite genetics for surveillance. Although we have provided a use case for parasite genetics, I have explored in the discussion a number of reasons as to why genetics for surveillance may not be advised: cost effectiveness, limitations compared to alternative surveillance methods such as ANC surveillance and incomplete characterisation of the impact of varying spatial connection between populations. These are issues that still need to be addressed before recommendation about genetics for surveillance can be made. In the next two chapter, I will pivot how the developed model will be used in order to address questions regarding antimalarial resistance.

4.6. Appendix

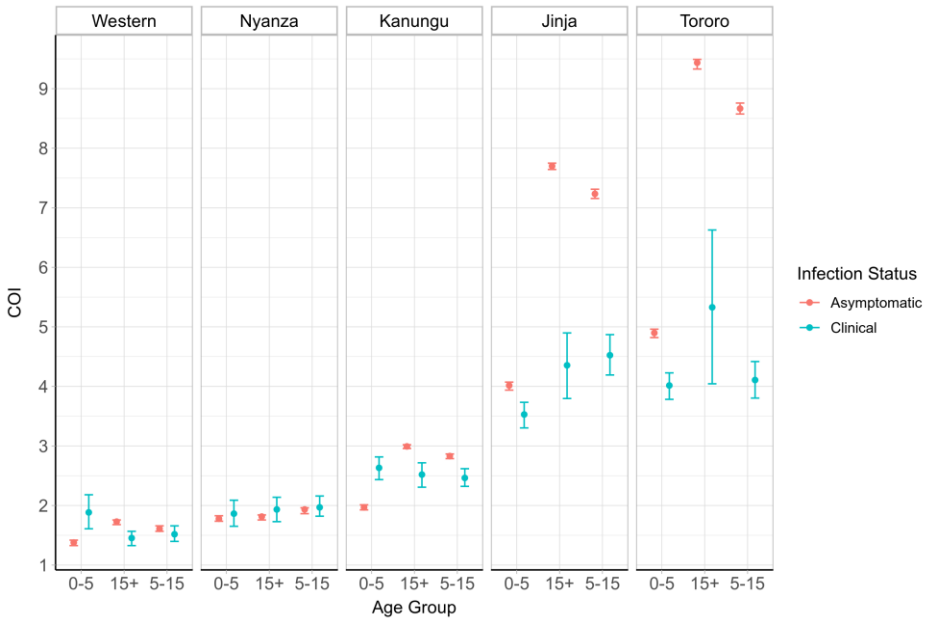
4.6.1. Appendix Figures



Appendix Figure 4.1 Predictive power of six metrics of parasite genetic diversity with respect to sample size under the assumptions that samples are unable to be phased. The same methods as those detailed in Figure 4.2 were used, with the only difference being that samples could not be phased and only the major haplotype could be called for an individual. iIBD is unable to be measured if samples cannot be phased and is subsequently crossed out. For pIBD, % Unique and COU it was assumed that the highest parasitaemia barcode was detected from each polygenomically infected individual. Lastly, there was no assumed difference in the ability to detect polygenomic samples or estimate the COI with unphased samples.



Appendix Figure 4.2 predictive performance of the ensemble model under different assumed sample sizes. Measures of the model error, root mean squared error (RMSE) and mean absolute error (MAE) as well as 1-R^2 are shown for sample sizes between 10 and 400. Model performance improves quickly over sample size ranges between 10 and 100, before slowing, with only very modest increases seen in model performance for sample sizes larger than 200.



Appendix Figure 4.3 Age and symptomatic status stratified COI from model fitting. Each plot shows the mean COI and 95% confidence interval for the study sites used in the model fitting. COI is stratified by age group and symptomatic status, showing that on the whole COI is higher in asymptomatic individuals, however, in lower transmission areas COI is higher in symptomatic young children.

4.6.2. Appendix Tables

Appendix Table 4.1 Statistical Model Performance

Meta Data	Model	RMSE*	MAE*	R^2*
None	Elastic Net	0.0276 (157.71%)	0.0225 (173.08%)	0.9935 (99.61%)
None	Gradient Boosted Trees	0.0214 (122.29%)	0.0159 (122.31%)	0.9961 (99.87%)
None	Random Forest	0.0211 (120.57%)	0.0151 (116.15%)	0.9962 (99.88%)
None	Weighted Mean Ensemble	0.0204 (116.57%)	0.0151 (116.15%)	0.9965 (99.91%)
Age	Elastic Net	0.0311 (177.71%)	0.0245 (188.46%)	0.9921 (99.47%)
Age	Gradient Boosted Trees	0.02 (114.29%)	0.0152 (116.92%)	0.9967 (99.93%)
Age	Random Forest	0.0197 (112.57%)	0.0143 (110%)	0.9968 (99.94%)
Age	Weighted Mean Ensemble	0.0195 (111.43%)	0.0144 (110.77%)	0.9969 (99.95%)
Age and Clinical Status	Elastic Net	0.0278 (158.86%)	0.0219 (168.46%)	0.9934 (99.6%)
Age and Clinical Status	Gradient Boosted Trees	0.0178 (101.71%)	0.0141 (108.46%)	0.9974 (100%)
Age and Clinical Status	Random Forest	0.0178 (101.71%)	0.013 (100%)	0.9973 (99.99%)
Age and Clinical Status	Weighted Mean Ensemble	0.0175 (100%)	0.013 (100%)	0.9974 (100%)

*** Absolute value (% relative to best performing model)**

4.6.3. Appendix Methods

4.6.3.1. Generalisation Methods

I chose to use 3 different statistical models (gradient boosted trees, elastic net regression model and random forests) to be used as level 0 models. I chose these 3 models due to their differences in their approaches, their consistent use for model fitting and consequent proven record of predictive accuracy and their ease of implementation. All models were applied using the R package *caretEnsemble* (Deane-Mayer and Knowles 2016), which flexibly allows the user to specify the desired models, how to control their fitting through cross-validation and how to combine them within an ensemble model.

4.6.3.1.1. Elastic net regression

Elastic net regressions are a form of penalised linear regression (Zou and Hastie 2005). The coefficients of the regression are found by minimising the prediction residuals and introducing two penalisation terms on the coefficients, given by:

$$f_M(x) = X^T\hat{\beta}$$
$$\hat{\beta} = argmin_{\beta}\|y - f_M(x)\|^2 + L_1\|\beta\| + L_2\|\beta\|^2$$

Where f(x) is our base linear regression and L_1 and L_2 are the two penalisation terms, which are the penalisation terms for both lasso and ridge regression respectively. The use of lasso regression enforces penalties both on the absolute size of the coefficient by L_2 regularisation and on the number of coefficients by L_1 regularisation. This is advantageous by helping to shrink the coefficients in the resultant linear model, which enables greater generalisation.

4.6.3.1.2. Random forests and Gradient Boosted Trees

Random forests and gradient boosted trees are two approaches that use an ensemble of regression trees (Breiman 1984). Trees are used to translate input variables into a prediction by partitioning the parameter space defined by our covariates, into *J* distinct regions in parameter space, Rj, which reflect the tips of the tree. Each region is assigned a constant, q_j that translates the region to a particular predictive value, $f(x)$. A tree is subsequently expressed as:

$$T(x;\theta) = \sum_{j=1}^{J} q_j \mathbb{I}\left(x \in R_j\right)$$

Where $\mathbb{I}$ is the indicator function, such that the subset x of the set of regions R_j is a function $\mathbb{I}_x: R_j \rightarrow \{0,1\}$, and θ are the coefficients of the decision tree.

Random forests combine the predictions of multiple randomly created decision trees via bagging (Hastie et al. 2009), in which bootstrapped samples of both the data and the covariates are used to average the ensemble of trees:

$$f_M(x) = \frac{1}{B}\sum_{b=1}^{B} T(x;\theta_b)$$

Where B is the total number of bootstrapped repetitions. Cross validation is used to find the optimal number of bootstrapped trees (Breiman 2001).

Gradient boosted trees iteratively update the regression tree by fitting to the residuals of the current tree. The resultant boosted tree model can be expressed as the addictive model of the decision trees, with the coefficients of each fitted tree found by using function gradient descent in which the negative gradient of the loss function in each tree model is used as the value of the residual in the boosting algorithm (Friedman 2011). This is given by:

$$f_M(x) = \sum_{m=1}^{M} T(x;\theta_m)$$

$$\hat{\theta}_m = argmin_{\theta_m} \sum_{i=1}^{M} L(y_i, f_{m-1}(x_i) + T(x_i;\theta_m))$$

For both random forests and the gradient boosted trees, extra model fitting parameters, such as the samples per tree node and the depth of the tree, were found using a grid search.

Chapter 5. Improving antimalarial cycling through increased diagnostic testing

In the last chapter I demonstrated how the transmission model developed in chapter 3 could be used to characterise how parasite genetics could be used for inferring malaria prevalence. The desire to include parasite genetics in surveillance is, however, arguably more invested in understanding how to detect antimalarial resistance and how best to slow its spread and prevent its emergence in Africa. The modelling framework developed in chapter 3 is well suited to answering questions related to antimalarial resistance by altering the parasite genetic barcode to instead model known loci associated with drug resistance. In this chapter I extend my transmission model to incorporate antimalarial resistance before using the extended model to review different drug strategies for preventing the spread of resistance. These include contrasting multiple first line therapies against sequential cycling of first line therapies before investigating how these strategies could be improved by reducing the treatment of non-malarial fevers.

5.1. Introduction

Artemisinin resistance has now emerged in the Greater Mekong Subregion (GMS) and is beginning to spread throughout the region. In 2012, mutations in the *P. falciparum* gene that encodes Kelch 13 (K13) were identified as the source of artemisinin resistance (Cheeseman et al. 2012). Efforts to sequence parasites in the GMS have revealed that clinically relevant parasites with K13 mutations are common in Asia (Ménard et al. 2016). However, parasites found to have K13 mutations in Africa have not been associated with a reduction in the efficacy of artemisinin and do not exhibit delayed parasite clearance (Wwarn and Group 2019). Genomic analysis of whole genome sequences from samples collected across Africa revealed that, although K13 mutations were present, they appeared to have arisen under neutral evolution (Amato et al. 2016). This suggests that although there is a large pool of K13 polymorphisms, which have been observed in longitudinal studies in both Rwanda (Tacoli et al. 2016) and Uganda (Ocan et al. 2016), they are not currently associated with artemisinin resistance, Consequently, the drug resistant phenotype observed in Asian K13 mutant parasites is likely reflective of additional resistance mechanisms occurring in the GMS (Fairhurst 2015).

The emergence of artemisinin resistance in GMS has been largely attributed to increased use of artemisinin-derivatives compared to Africa. Artemisinin-derivatives were adopted for first line treatments earlier in the GMS that Africa, with delays to ACTs being recommended as first line therapies due to some donors and international organisations being reluctant to support and fund

their comparatively higher cost (Attaran et al. 2004). In 2006, the WHO recommended ACTs as first line therapies, however, many countries in the GMS had been using artemisinin-derivatives prior to the 21st Century. Both the early adoption and the lack of availability and consensus on combination therapies likely increased the rates of artemisinin monotherapy being used in the GMS. For example, despite the national switch to artesunate-mefloquine (ASMQ) in 2000, one report revealed that 78% of all artemisinin delivered in Cambodia in 2002 was as a monotherapy (Yeung et al. 2008). Additionally, parasites in the GMS have now been shown to have acquired partner drug resistance, with copy number amplification of the *plasmepsin 2* and *plasmepsin 3* genes linked to piperaquine resistance (Witkowski et al. 2017; Amato et al. 2018). However, this was perhaps unsurprising, as piperaquine was used as a monotherapy within the GMS in the 1970s and 1980s in response to chloroquine resistance before itself being replaced due to observed piperaquine resistance (Davis et al. 2005). However, the emergence of the KEL1/PLA1 lineage in 2017, which is resistant to dihydroartemisinin-piperaquine (DHA-PPQ), shows that the parasite has been able to acquire multidrug resistance while offsetting any potential fitness costs associated with resistance and spread across the GMS (Imwong et al. 2017a).

The evidence is now clear that extensive use of antimalarials, in particular as monotherapies, will drive the emergence of drug resistance. However, what is not clear is the contribution of other likely drivers of resistance and how they differ depending on the stage of resistance development. Resistance development is often characterised by three distinct stages: emergence, spread and fixation (Figure 5.1). At each phase the processes driving the increase in resistance are likely to differ and so understanding which factors are important at each stage is key. For example, in efforts to prevent emergence the short half-life associated with artemisinin was suggested to reduce the window for selection due to the short drug elimination phase (Stepniewska and White 2008). Drugs with longer half-life, however, are more likely to enable parasites to be exposed to sub-optimal drug concentrations during reinfections and are consequently believed to be more susceptible to resistance (Hastings and Hodel 2014). Changes in drug concentrations resulting from increasing the number of doses have been identified as a possible factor in the probability of resistance emerging (Kay et al. 2015). Patient adherence to drug courses is cited as a cause of antibiotic resistance emergence (Wahl and Nowak 2000). The same is true for malaria, with poor compliance well-documented in malaria (Bruxvoort et al. 2014). This results in drugs falling below inhibitory concentrations, which reduces the therapeutic efficacy of antimalarials (Challenger et al. 2017) and may contribute to the emergence of resistance (White and Pongtavornpinyo 2003; Keoluangkhot et al. 2008). There are fewer studies, however, examining the factors associated with the spread of antimalarial resistance. In particular,

there is interest to see how drug strategies may be optimised to hinder the development of antimalarial resistance.

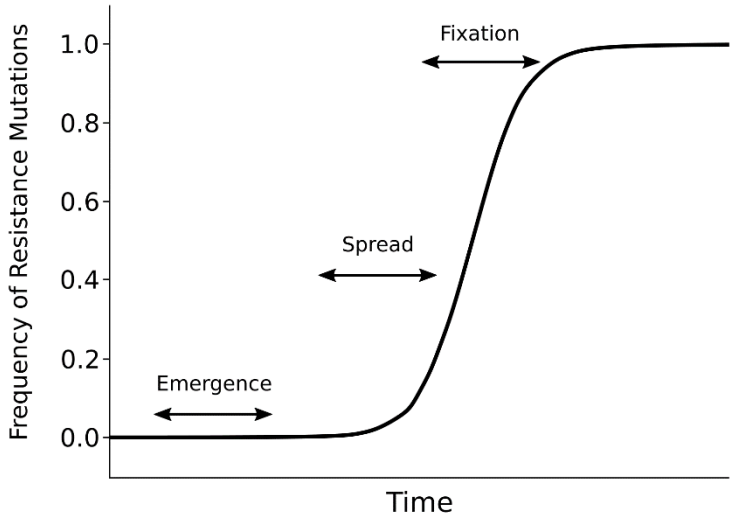


Figure 5.1 Stages of drug resistance development. Drug resistance development is often typified by three phases describing the speed at which resistance increases. Prior to emergence, mutations must first arise that confer drug resistance. Due to the rarity of these events and the low effective population size of mutant strains during emergence, this stage may vary in its length due to stochastic effects. After mutant strains have persisted through a number of transmission cycles, the spread of resistance occurs leading to the rapid increase in mutation frequency. Following spread, the increase in resistance mutations slows until they reach fixation.

One area of research is how drug distribution strategies may impact drug resistance. To answer these questions, mathematical models are often used to evaluate how different treatment strategies may impact the dynamics of resistance (Kouyos et al. 2014). However, the range of contributing factors means that there is no clear consensus, with models predicting very different evolutionary outcomes depending on the amount of drug used, the number of drugs available and the force of infection in the population (Zur Wiesch et al. 2011). One strategy that has been proposed for combatting antimalarial resistance is the use of multiple first line therapies (MFT), where different patients are allocated to different artemisinin combination therapies (ACTs). The use of multiple drugs with differing mechanisms of action is hoped to protect against the emergence of resistance, with some modelling studies suggesting MFT is advantageous to the current strategy of cycling through ACTs when treatment failures rise above 10% (Boni et al. 2008; Nguyen et al. 2015). However, alternative modelling studies predict little benefit to switching MFT, with any specific benefit dependent on a number of factors, including transmission intensity, the ratio of drugs used and the ability to maintain stocks of multiple drugs (Maude et al. 2009; Smith et al. 2010b; Antao and Hastings 2012). In addition, most modelling focus has been on the emergence of resistance with little attention on the spread of resistance, which is now established in most countries in the GMS. This lack of consensus resulted in

the technical expert group established by the WHO Malaria Policy Advisory Committee investigating drug resistance and containment to not recommend MFT in 2013 (World Health Organization 2013).

Another strategy, which many countries are also implementing, is to improve the current method of cycling first line therapies by reducing the presumptive treatment of fevers with antimalarials. Reducing overprescription may arguably be preferable for many control programmes compared to MFT, with increased costs associated with procuring and maintaining stock for multiple drugs. In addition, overprescription may represent a significant driver of resistance. In 2014, only 35.7% of self-reported fevers in sub-Saharan Africa were estimated to be accompanied by a malaria infection (Dalrymple et al. 2017). However, due to the historic burden of malaria in many countries many fevers are assumed to be due to malaria and are presumptively treated with antimalarials (Odaga et al. 2014). Presumptive treatment will lead to the overprescription of antimalarials, which can occur in scenarios where either RDTs are not available or the outcome of the RDT result is ignored (Umlauf 2017). This may then lead to sub-optimal drug concentrations in the population, which increases selection for partner drug resistance. Ensuring that all non-malarial fevers (NMFs) are only treated on the basis of a diagnostic test result is likely to reduce the selective pressure. The size of this reduction, however, is hypothesised to be dependent on transmission intensity (Figure 5.2), with higher transmission settings more likely to select for resistant parasites during reinfection events. Consequently, it is important to understand the transmission ranges at which reducing overprescription is beneficial in the long term because presumptive treatment of NMFs is known to lead to a reduction in clinical cases due to the period of prophylactic protection conferred. To answer this, I have extended the transmission model to investigate whether reducing overprescription of antimalarials could be as effective as MFT in slowing the spread of drug resistance. I have incorporated treatment with multiple ACTs and have adapted the malaria parasite barcode to model their resistance and fitness phenotypes. I scan across transmission intensity and treatment coverage gradients, evaluating the effectiveness of each strategy by comparing the mean number of treatment failures. I continue by characterising how both MFT and sequential cycling could be improved by reducing overprescription and increasing the number of ACTs available for use within MFT strategies.

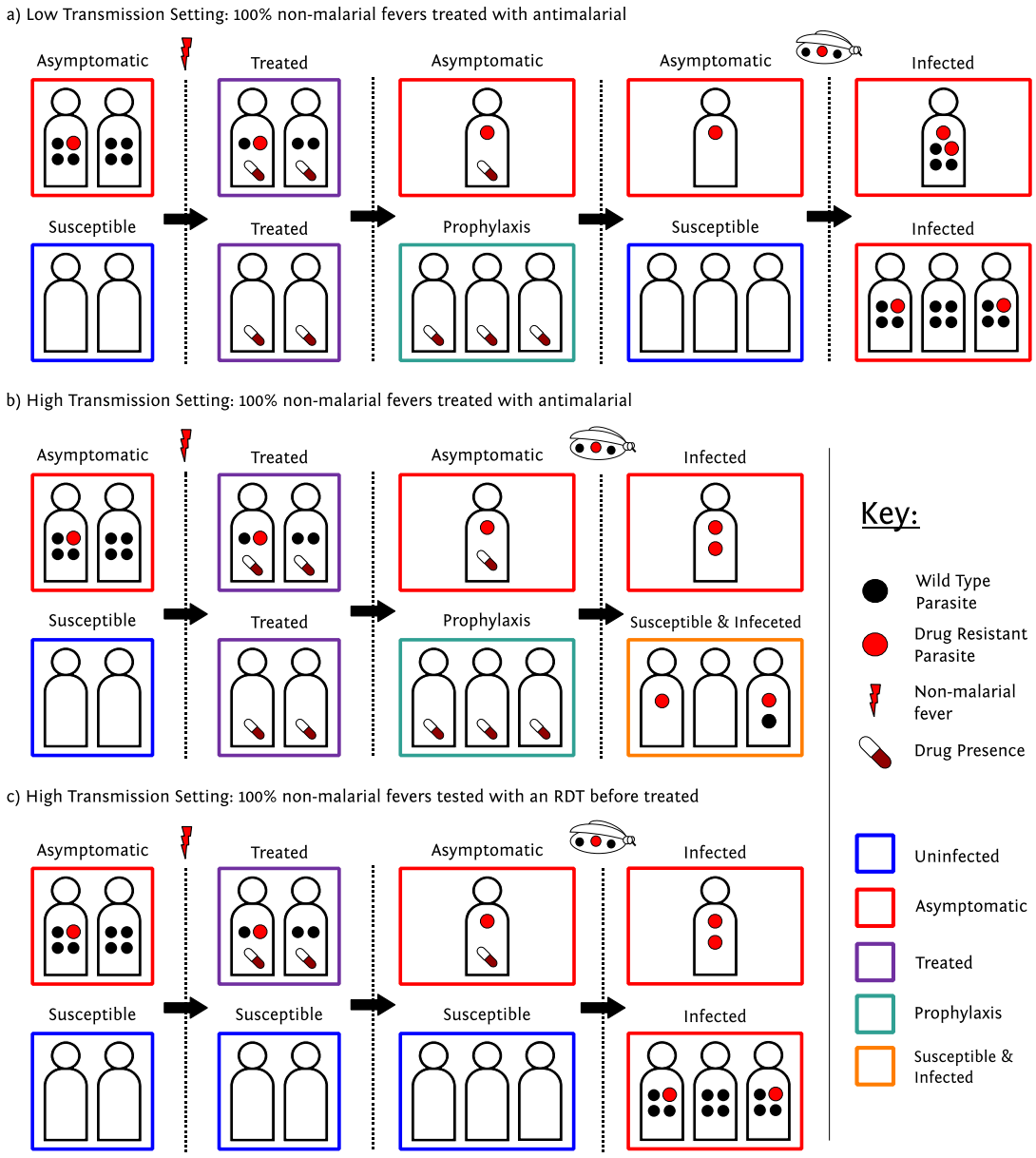


Figure 5.2 Hypothesis for increased benefit of reducing overprescription in high transmission settings. The diagram shows the outcomes for groups of asymptomatic individuals (red boxes) and susceptible individuals (blue boxes) after suffering a non-malarial fever (NMF). In a) 100% of NMFs will be treated regardless of infection state. Consequently, both groups are treated (purple boxes), with individuals infected with wild type parasites (black circles) successfully being treated and moving into a prophylactic state (green box). Resistant parasites (red circles) are less likely to be cleared and will lead to a continued asymptomatic infection. In a) the transmission intensity is low and so both groups of individuals will eliminate the partner drug before being reinfected. However, in b) the transmission intensity is assumed to be higher, which leads to reinfection occurring while individuals are still in a prophylactic state. Consequently, the partner drug is likely to prevent wild type parasites from being passed on, whereas resistant parasites are more likely to lead to an infection. This results in the population in b) having a higher frequency of resistant strains after reinfection compared to a). Lastly, in c) all NMFs are assumed to receive a diagnostic test. This decreases the probability that individuals will be in a state of prophylaxis during reinfection. Compared to setting b), this will lead to both an increased number of new infections as well as a decrease in the frequency of resistant parasites.

5.2. Methods

5.2.1. Transmission model extensions for modelling antimalarial resistance

5.2.1.1. Barcode Alterations

In order to model antimalarial resistance, I have adapted the parasite barcode for the simulation of resistance. In the new formulation, each position in the barcode represents either the absence (barcode position is equal to 0) or presence (barcode position is equal to 1) of a resistance mutation associated with resistance to a particular drug. For example, Table 5.1 shows how a barcode with three loci can be used to represent resistance to DHA-PPQ and ASMQ.

Table 5.1 Example barcode alterations to model antimalarial resistance. Example explores resistance to piperaquine (PPQ), artesunate (AS) and mefloquine (MQ).

Artemisinin Resistance	Piperaquine Resistance	Mefloquine Resistance	Phenotype
0	0	0	Wild type parasite. Fully susceptible.
1	0	0	Resistant to artemisinin
0	1	0	Resistant to PPQ
0	0	1	Resistant to MQ
1	1	0	Resistant to DHA-PPQ
1	0	1	Resistant to ASMQ
0	1	1	Resistant to PPQ and MQ
1	1	1	Multidrug resistant to DHA-PPQ and ASMQ

As before, barcodes are used to track populations of parasites that are introduced from an infectious mosquito bite. Previously, all barcode loci were assumed to be genetically unlinked and to consequently segregate independently during recombination. This assumption is still maintained when modelling resistance, with genes known to confer resistance to the five first line ACTs recommended by the WHO each occurring on different chromosomes.

Antimalarial resistance is assumed to introduce a fitness cost to resistant parasites compared to wild type parasites. Fitness costs are associated with each resistance locus and are assumed to be multiplicative. The fitness cost manifests as a reduction in parasite density that is assumed to reduce the probability of the resistant parasite being passed on to a mosquito. This can be expressed mathematically as follows. Let $b = [b_1, b_2, \dots, b_m]$ describe the vector of barcode loci for a barcode of length m. For example, the ASMQ resistant strain in Table 5.1 is represented by the vector $[1, 0, 1]$.

If v_j is the resistance cost associated with barcode locus j, then the comparative fitness cost due to resistance, r, for the given parasite is simply:

$$r = \prod_{i=j}^{m} v_j$$

In this study, the wild type allele at each locus Is assumed to have no fitness cost ($v_j = 1$) and thus $r = 1$ for the wild type parasite in Table 5.1.

When selecting which parasite strains contribute to an onward infection, the number of oocysts formed from the bite is first drawn from the negative binomial defined in chapter 3 before sampling parasites from the infected individual to contribute to the oocysts. The probability vector that describes the relative probability that a given strain will be chosen in an individual with n gametocytogenic parasite strains is now given by:

$$cr = [c_1 r_1, c_2 r_2, \dots c_n r_n]$$

r_i is the fitness cost associated with parasite strain i and c_i is the contribution of parasite i to onward infection, which will be either c_T, c_D, c_A or c_U depending on the infection status of strain i, denoted here as X_i. The probability that an infected individual infects a mosquito is still determined by the set of parameters determining the onward contribution to transmission, $\{c_T, c_D, c_A, c_U\}$, which are based on the infection status of the individual, denoted here as Y. However, if the strains responsible for the human's current infection state, i.e. all strains that match the human's infection state, are resistant then the probability of onward transmission is determined by the highest onward contribution of these strains, which is given by:

$$max\,\{cr_i : X_i = Y\}$$

In this way, fitness costs both affect the relative probability that a resistant strain is transmitted compared to a wild type strain in a mixed infection, while also reducing the probability that transmission occurs in individuals where the highest parasite density strain is resistant.

5.2.1.2. Clinical and Treatment Outcomes

In previous chapters, treatment was assumed to be 100% effective, i.e. all individuals who were treated would move to being treated, before recovering to a state of prophylaxis and eventually returning to being susceptible. With the addition of resistance, the treatment efficacy now varies and is determined both by the genotype of the parasite strains, the parasite density of each strain and the drug used to treat the infection.

Using the same representation of a parasite genotype as in Table 5.1, the efficacy of drug z can be expressed by the vector $ez = [ez_1, ez_2, \dots, ez_{2^n}]$ for simulations in which the number of loci being modelled is equal to n. ez_1 represents the efficacy of the drug against the wild type parasite, i.e. the barcode vector b represented by a vector of length n filled with zeros. The probability that any given strain is cleared by drug z can be expressed as $ez_{\beta(\boldsymbol{b})}$, where $\beta(b)$ is an adapted conversion of binary to decimal integers and is given by:

$$\beta(b) = \left(\sum_{j=1}^{m} b_j . 2^{j-1}\right) - 1$$

Importantly, ez reflects the probability that the drug will clear a parasite that has led to a symptomatic infection, i.e. the parasite strain is at a sufficiently high parasite density to trigger symptoms and seek treatment. Parasites from previous infections, however, are assumed to be at a lower parasite density and will be more likely to be cleared by the drug. In the model, the infection state of each strain, X_i, as well as the day the strain was acquired, t_0, and the time the strain will move out of the infection state, t_1, is tracked. This information is used to define the probability that strain i will recrudesce after treatment, $P(Recrudesce)_i$, which is given by:

$$P(Recrudesce)_i = \begin{cases} ez_{\beta(b_i)}\,, & X_i = Y \\ ez_{\beta(b_i)}\left(\dfrac{t_1 - t_c}{t_1 - t_0}\right), & X_i = A \\ 0\,, & X_i = U \end{cases}$$

where t_c is the current time. $P(Recrudesce)_i$ assumes that parasites below ~200p/µl (state U) will always be cleared regardless of the parasite phenotype. $P(Recrudesce)_i$ also assumes that the probability that an asymptomatic parasite above ~200p/µl (state A) will recrudesce is linearly related to the age of the infection and is at its highest when it first enters state A.

The probability that a treated infection will be successfully cleared by drug z, $P(Cleared)$, is equal to one minus the highest probability of a strain recrudescing, which is given by:

$$P(Cleared) = 1 - max\{P(Recrudesce)_i : 0 \leq i \leq n\}$$

This assumes that the multiplicity of infection does not directly affect the probability that an individual will be cleared, i.e. only one Bernoulli trial with probability $P(Cleared)$ is used to determine if all parasite strains were cleared.

If an individual fails treatment, it is assumed that they will recrudesce to yield a late parasitological failure (LPF) and move into state A after the prophylactic period of the drug has finished. Whether each parasite strain in a multiply infected individual recrudesces during a LPF is dependent on $P(Recrudesce)_i$. Bernoulli trials are conducted for each parasite strain except for one random strain for which $P(Recrudesce)_i = P(Cleared)$, which ensures that one of the most likely strains to recrudesce did actually recrudesce and cause a LPF.

If an individual successfully clears all parasites, they will either move directly into a state P or they will remain in the treated compartment for a longer duration resulting from slow parasite clearance (SPC). SPC is assumed to always occur if any of the pre-treatment strains that contributed to the clinical disease were resistant to any component of the drug. The duration of SPC was set equal to 10 days based on previous modelling studies estimating parasite clearance rates associated with SPC (Slater et al. 2016). During SPC it is assumed that all parasite strains not resistant to the drug given have cleared and will thus not contribute to onward infection during SPC.

Lastly, individuals in state P can only be reinfected by wild type parasites when they progress from state P to state S. In this study, the transition rate from state P -> S is now dependent on the ACT used, reflecting the different half-lives of available partner drugs. In addition, resistant strains are assumed to be able to infect individuals who are in state P if the strain is resistant to the partner drug used when treating the individual. The probability of reinfection in state P increases as the partner drug wanes and is modelled by assuming the reduction in chemoprophylaxis due to resistance halves the mean duration of prophylaxis, broadly in alignment with estimates of the protective duration from pooled analyses of the duration of chemoprophylaxis in regions with known resistance (Bretscher et al. 2019).

5.2.1.3. Non-malarial fevers

An age-dependent NMF rate was included to examine the impact of overprescription of antimalarials due to presumptive treatment of NMF. NMF is incorporated through the defined vector $nmf = [nmf_1, nmf_2, \dots, nmf_n]$ of age dependent NMF rates and the set $NMF_{intervals} = \{[0, age_1), [age_1, age_2), \dots, [age_n, 100)\}$ of right-open age intervals that divide the population into n age brackets with the population's maximum age set to 100 years. At the beginning of a simulation, the waiting time until the next NMF for each individual is drawn from an exponential distribution with rate nmf_i depending on their age. The outcome of a NMF depends on the individual's infection status, the assumed treatment coverage and the probability that treatment is based on the outcome of a

malaria diagnostic test, f_{test}. The probability that an individual will be treated after suffering a NMF, $P(NMF_T)$, is given by:

$$P(NMF_T) = \begin{cases} f_T(1 - f_{test}) , & Y \in \{S, U\} \\ f_T , & Y = D \\ f_T q_i , & Y = A \\ 0 , & Y \in \{T, P\} \end{cases}$$

Individuals currently treated or in prophylaxis are assumed above to not receive further antimalarial treatment when presenting with a NMF. Susceptible and subpatent individuals who seek treatment due to a NMF will only receive treatment if they first seek treatment and are then treated not based on a diagnostic test. Asymptomatic individuals will only be treated if they first seek treatment and if they would be detected using microscopy or RDT, q_i, which is defined in chapter 3. Lastly, diseased individuals who seek treatment due to a NMF will always receive treatment. After the NMF has been resolved a new day is drawn for the next NMF.

nmf and $NMF_{intervals}$ were sourced from a previous study on *pfhrp2* deletions, in which NMF was sourced from Demographic and Health Surveys in sub-Saharan Africa that surveyed whether individuals had been previously sick with a fever in the last 2 or 4 weeks and if and where they sought treatment for that fever (Watson et al. 2017). Briefly, surveys were subset based on the representativeness of the ages sampled before creating an age-bracketed annual rate of fever that led to treatment being sought. I used smaller age brackets at younger ages in order to capture the increased variability in fever at very young ages, before scaling the estimated rates by 35.7% to represent the likely NMF rate, as estimated from a large scale estimate across Africa (Dalrymple et al. 2017).

5.2.2. Drug policy strategies and parameter values

In a previous study by Nguyen et al., three different population-level treatment strategies were assessed in terms of how effectively they reduced the number of treatment failures and the emergence of artemisinin resistance (Nguyen et al. 2015). Building on this study, I have adapted the strategies they introduced to assess how effective each strategy is in slowing the continued spread of antimalarial resistance in a population. In addition, I introduce NMFs and examine how assumptions about the proportion of suspected malaria fevers that are tested before being treated impacts the performance of the strategies.

The first strategy is designed to reflect the current strategy in most countries in which drugs are sequentially cycled when the 60-day average treatment failure reaches the WHO-defined threshold

of 10% (World Health Organization 2010a). A one-year delay to switching first line ACT is included to reflect the time taken for national malaria control programmes to switch first line drugs. In this strategy I assume that 50% of suspected malaria fevers are treated on the basis of a diagnostic test, with the remaining 50% of suspected cases always treated. The second strategy is identical to the first strategy but with 100% of suspected malaria fevers treated on the basis of a diagnostic test. In both cycling strategies, if the last drug to be cycled reaches 10% treatment failure then the first drug used will be used next. The last strategy uses MFT, in which there is an equal probability that an individual is treated with any of the drugs available for MFT. In this strategy I assume that 50% of suspected malaria fevers are treated on the basis of a diagnostic test. The designed scenarios thus allow for the impact of increased diagnostic testing to be assessed when comparing strategy one vs two, while strategy three allows us to compare how effective MFT is vs sequential cycling in slowing the continued spread of resistance in a population.

I have assumed initially that three ACTs are available within the treatment strategies above: DHA-PPQ. ASMQ and AL which are cycled in this order. The order was chosen to reflect the emergence of artemisinin resistance in the GMS, with a number of countries recently switching from DHA-PPQ to ASMQ in response to rising DHA-PPQ treatment failure (Alonso 2017). For each ACT, resistance to the artemisinin component is encoded by the first barcode locus. Partner drug resistance is encoded with a different locus for each drug. One locus is thus assumed to be sufficient to confer resistance to each of the partner drugs, with a fitness cost of 0.99 assumed for each resistance locus as used in Nguyen et al. (Nguyen et al. 2015). 100 stochastic realisations of populations of 100,000 individuals were simulated for 40 years, with the first 20 years reflecting a period of "burn-in" ensuring an equilibrium is reached in simulations. After the burn in, the population is seeded with 10% artemisinin resistance and 1% resistance for each partner drug used. The frequency of resistance alleles and the percentage of treatment failures are recorded and summarised for the following 20 years and compared alongside any changes in the PCR prevalence of malaria. When comparing between settings with different treatment coverages, I assume that a clinical case not treated due to not seeking treatment in the first place $(1 - f_T)$ is counted as a treatment failure to allow fair comparisons between settings. Otherwise, polices with the lowest treatment coverage would appear optimal due to treating the fewest people and selecting for resistance the least.

Sensitivity analyses were conducted by scanning over a range of values for the transmission intensity (EIR: 0.25 - 10), treatment coverage (f_T: 0.5 – 0.9), the assumed percentage of NMF that are treated on the basis of a diagnostic test (f_{test}: 0.5 – 1.0), the number of ACTs available within MFT (2 – 5 first line therapies). The two additional ACTs to be included are artesunate-amodiaquine (ASAQ) and artesunate + sulfadoxine–pyrimethamine (AS+SP).

Parameters for drug efficacies were sourced from paired phenotype-genotype studies for which the proportion of treated cases that resulted in recrudescent events was measured amongst each resistant parasite genotype, i.e. wild type, artemisinin resistance, partner drug resistance and multidrug resistance. Where these studies are not available for an ACT, drug efficacies were first sourced from studies in which the drug was used in populations with assumed resistance to only one of the components. In the case of ASAQ and AS+SP, the decrease in efficacy from artemisinin resistance was set equal to the decrease for ASMQ, for which suitable phenotype-genotype studies were available. These measures were then combined to estimate the drug efficacy against multidrug resistant parasites by taking the squared product of the relevant monotherapy efficacies, which was found to reflect the conversion for ACTs in which complete phenotype-genotype studies were available. Consequently, modelling results are used primarily to answer questions related to the treatment strategies rather than performance of any one ACT. The efficacy and prophylactic properties assumed for each drug are shown in Table 5.2.

Table 5.2 Antimalarial efficacy and prophylactic properties

Drug	Duration of prophylaxis (days)	Efficacy*	Efficacy References
DHA-PPQ	25 (Okell et al. 2014)	0.985, 0.966, 0.786, 0.577	(Straimer et al. 2015)
ASMQ^{+}	25 (Okell et al. 2014)	0.983, 0.962, 0.743, 0.510	(Price et al. 2004)
AL^{+}	13 (Bretscher et al. 2019)	0.964, 0.951, 0.851, 0.653	(Price et al. 2006; Ndounga et al. 2015)
ASAQ^{+}	15 (Bretscher et al. 2019)	0.968, 0.947, 0.821, 0.604	(Mohamed et al. 2017)
AS+SP^{+}	25 (Watkins et al. 1997)	0.971, 0.950, 0.731, 0.482	(Alker et al. 2008)

* Probability of clearing a wild type, artemisinin resistant, partner drug resistant and multidrug resistant parasite

$^{+}$ Efficacy not based on complete phenotype-genotype study

5.3. Results

5.3.1. Impact of increasing diagnostic testing of non-malarial fevers on spread of resistance

In Figure 5.2, reducing overprescription was hypothesised to reduce the selective pressure for resistance in higher transmission settings, with selection increasingly occurring during reinfection of recently treated individuals. However, reducing presumptive treatment of NMFs will initially result in higher malaria prevalence due to fewer people being protected by waning partner drug concentrations. The extended transmission model was first used to characterise the trade-off between these two processes by quantifying the impact of increasing diagnostic testing of NMFs.

Figure 5.3 shows an example of these processes at two different transmission intensities, EIR = 1.5 and 4, which corresponds to a PCR PfPR of 15% and 25% respectively. In treatment strategies in which 100% of NMFs are treated on the outcome of diagnostic testing an absolute increase in PCR PfPR of ~2.5% is observed prior to the spread of resistance. In the lower transmission setting (Figure 5.3a), the change in treatment failure over time is comparable between both sequential cycling strategies, with the average useful therapeutic lifespan of DHA-PPQ equal to ~5 years, after which 10% treatment failure is observed. In the higher transmission setting the same therapeutic lifespan is observed when 100% of NMFs are tested (Figure 5.3b). However, when only 50% of NMFs are tested the therapeutic lifespan is approximately one year shorter, with 10% treatment failure occurring after four years on average. The increased speed of resistance spread results in the PCR PfPR after 20 years being higher in the 50% diagnostic testing cycling strategy. The MFT strategy is predicted to cause a constant increase in the frequency of resistance across all resistance loci. In the MFT strategy, the frequency of artemisinin resistance initially increases at a comparable rate to the cycling strategies. However, after 7.5 years (EIR = 4) and 10 years (EIR = 1.5) the frequency of artemisinin resistance in the MFT strategy increase quicker than the cycling strategies. Partner drug resistance increases at a constant rate in the MFT strategy, which is predicted by the different treatment efficacy associated with partner drug resistance as well as the different duration of post treatment prophylaxis. Treatment failure and PfPR initially increase more slowly in the MFT strategy, however, in both transmission settings the final prevalence of malaria after 20 years is highest in the MFT strategy.

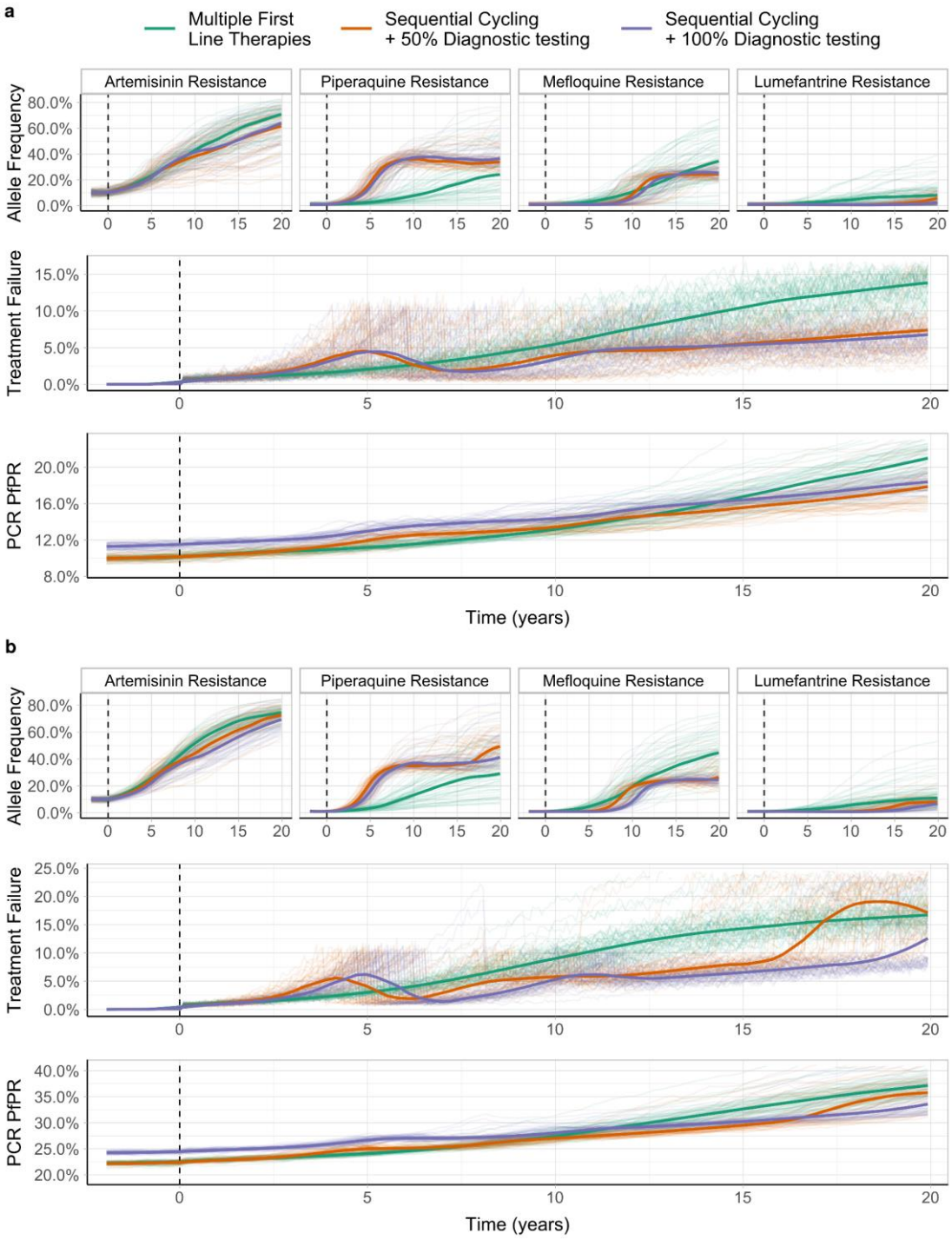


Figure 5.3 Increase in drug resistance mutations in response to multiple first line therapies (MFT) and sequential cycling. The graphs show the change in allele frequency of artemisinin resistance and the frequency of partner drug resistance for each partner drug at two different transmission intensities; a) EIR = 1.5 and b) EIR = 4 with a high assumed treatment coverage of 70%. The change in treatment failure and prevalence of malaria by PCR are shown below for each setting. 100 simulation repetitions are shown for each scenario with the median trend shown in bold. 50% of non-malarial fevers are assumed to be treated regardless of infection status in the cycling (red) and MFT (green) scenario, whereas one cycling scenario (purple) assumes 100% testing of non-malarial fevers. Resistance is assumed to not be under selection prior to year 0, shown with the vertical dashed line.

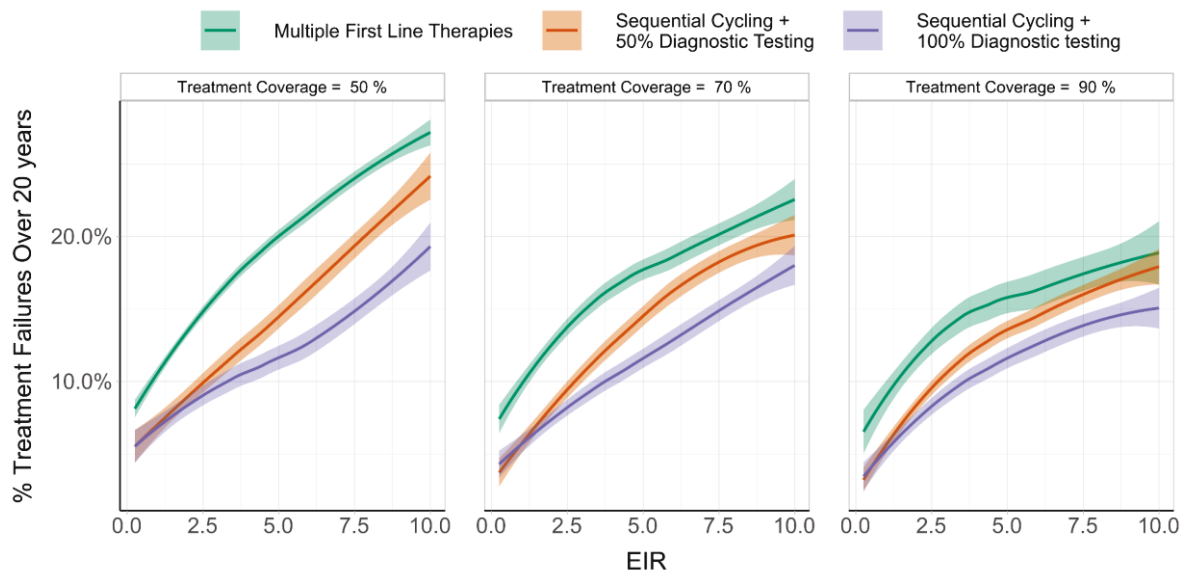


Figure 5.4 Effect of treatment coverage and transmission intensity on treatment failure due to resistance. The graphs show the average percentage of treatment failures over 20 years against EIR. 100 simulation repetitions were conducted for each strategy and the mean and 95% confidence interval for each treatment strategy shown. Each graph shows a different treatment coverage: 50%, 70% and 90% from left to right.

Using the same simulation set up we compared each treatment strategy by calculating the mean treatment failure over the 20-year study period. The percentage of treatment failures was predicted to be greatest in the MFT strategy across all transmission intensities and treatment coverages explored (Figure 5.4). In the lowest treatment coverage explored (f_T = 50%), treatment failure was significantly higher than both cycling strategies, with the 95% confidence interval for the percentage of treatment failures only overlapping with the 50% diagnostic testing strategy at the highest transmission intensity explored (EIR = 10). In the same setting, the sequential cycling strategies were not significantly different at lower transmission intensities (EIR < 3), however, above this EIR, treatment failure was significantly lower when all NMFs received diagnostic testing. Similar relationships were seen with increasing treatment coverage, although the difference in treatment failure between each strategy decreased. Lastly, the transmission intensity at which cycling with 100% diagnostic testing yields a significantly lower overall treatment failure rate compared to only 50% diagnostic testing occurs at lower EIRs with increasing treatment coverage.

In Figure 5.5 the speed of selection for artemisinin resistance is shown for each treatment strategy across the range of transmission intensities and treatment coverage explored earlier. Similar findings are observed with respect to the optimum treatment strategy, with 100% testing of NMFs consistently yielding the lowest frequency of artemisinin resistance. However, the differences between the treatment strategies are less pronounced than when examining treatment failures in Figure 5.4. In particular, with 50% assumed treatment coverage, MFT and cycling with 50% diagnostic testing are

not statistically different between an EIR 5 – 9, with MFT having a statistically lower frequency of artemisinin resistance when EIR = 10 (MFT 95% CI: 70.8%-72.9%, Cycling 95% CI: 73.5%-75.2%).

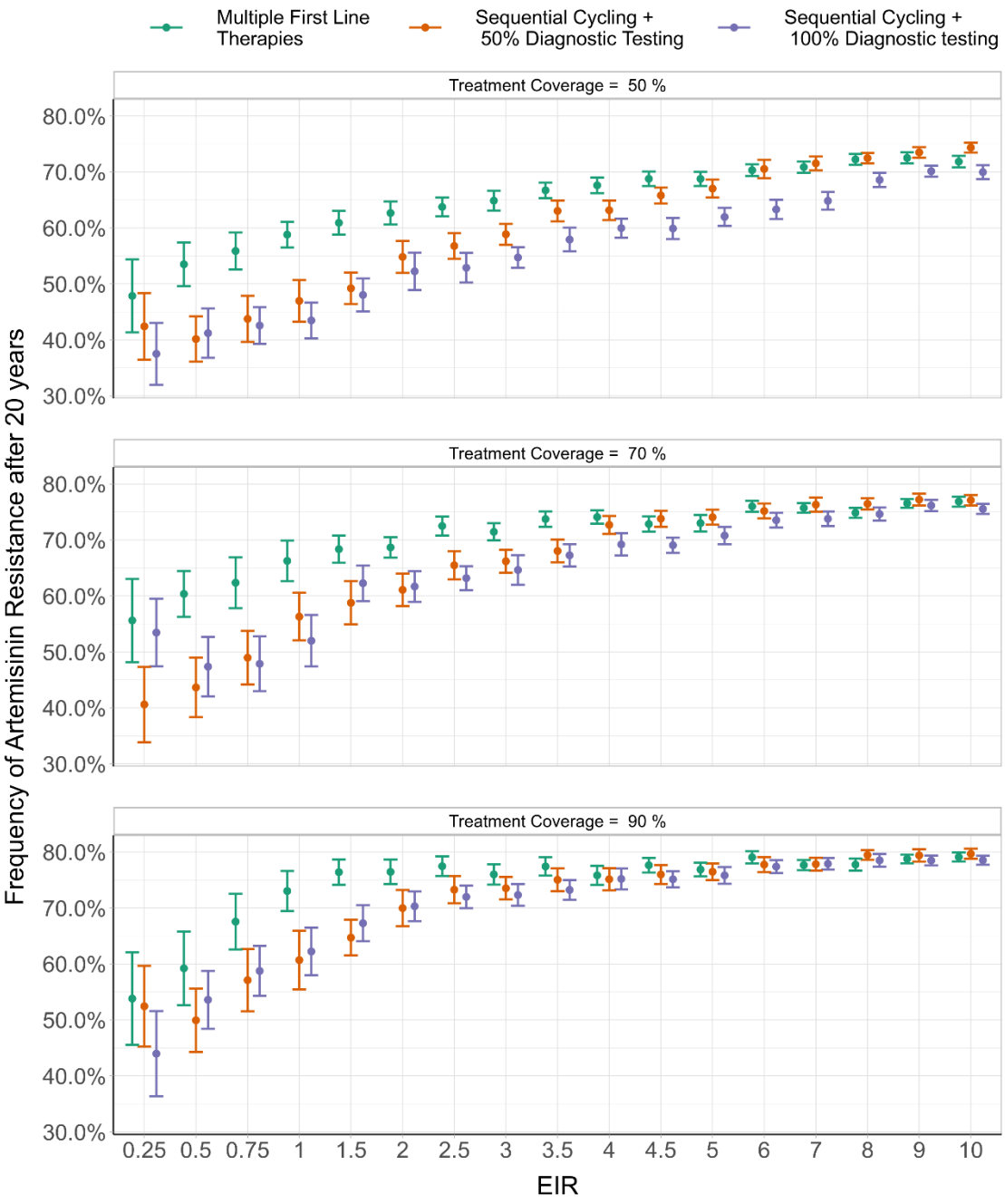


Figure 5.5 Speed of selection for artemisinin resistance under three treatment strategies. The graphs show the mean and 95% confidence interval for the maximum frequency of artemisinin resistance after 20 years from 100 simulation repetitions. Selection for artemisinin is predicted to be greatest in higher transmission settings. Each graph shows a different assumed treatment coverage, with selection increasing with increasing treatment coverage.

5.3.2. Improving current treatment strategies

To investigate how MFT could be improved, additional first line therapies were included in the MFT strategy up to a maximum of five ACTs available. Each ACT was assumed to be equally likely to be used, with 10% starting artemisinin and 1% starting partner drug resistance again assumed. As would be expected, for each additional ACT available the percentage of treatment failures over 20 years decreased (Figure 5.5). The rate at which treatment failure increased with increasing EIR was comparable across each MFT strategy above an EIR equal to 2.5. However, in scenarios with only two or three ACTs available the percentage of treatment failures increased quickest with respect to EIR for values of EIR less than 2.5. MFT was also predicted to outperform sequential cycling of three ACTs with 100% diagnostic testing of NMFs when four or more ACTs were available for MFT, although treatment failure was comparable between MFT with four first line therapies and the cycling strategy when treatment coverage was equal to 50%.

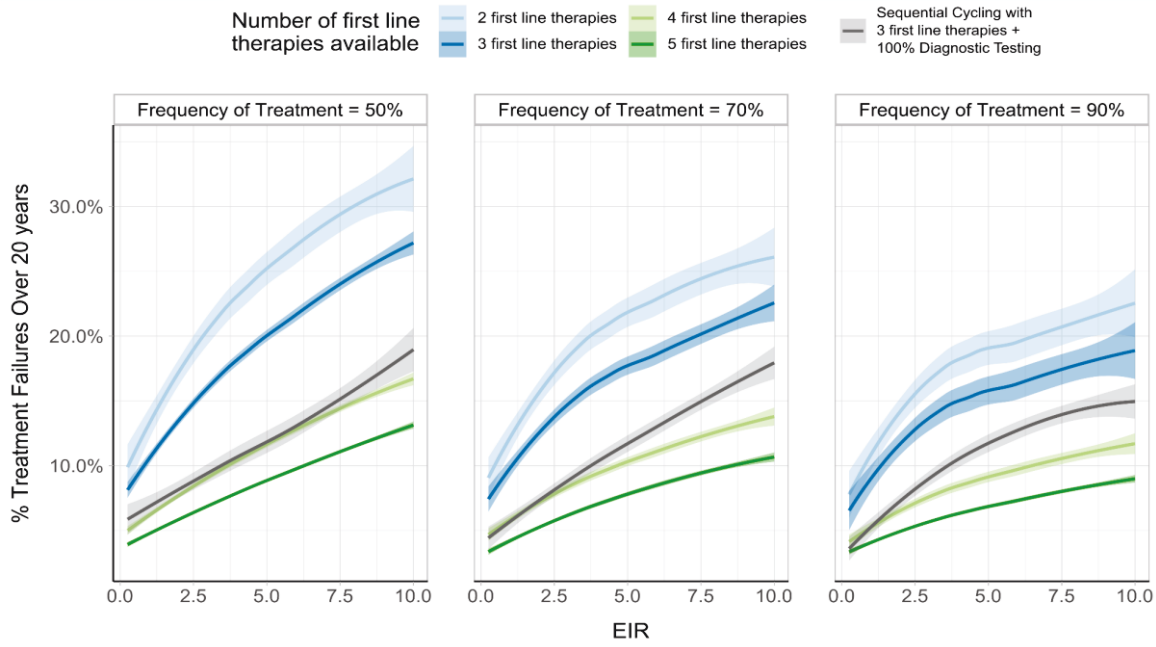


Figure 5.6 Benefit of increasing the number of first line therapies available within multiple first line therapies. Average treatment failures over a twenty-year period are shown for four different MFT strategies, which had two, three, four or five available first line therapies. In each strategy the same proportion of each drug was used and they are compared against a sequential cycling strategy using three drugs with 100% diagnostic testing of non-malarial fevers. 100 simulation repetitions were conducted for each strategy and the mean and 95% confidence interval for each treatment strategy shown.

In Figure 5.7, the benefit of increased diagnostic testing of NMFs is summarised across the transmission and treatment coverage ranges explored, with 50% - 100% NMF testing assessed. As demonstrated earlier, the benefit of increasing diagnostic testing is marginal in low transmission settings. The largest absolute decrease in mean treatment failure due to increased diagnostic testing

was equal to 6%, which reflects increasing NMF testing from 50% to 100% in the highest transmission settings explored with 50% treatment coverage. Smaller reductions in treatment failure due to increased diagnostic testing were predicted when treatment coverage was at its highest, with the maximum mean treatment failure equal to 18.4% when treatment coverage was equal to 90%.

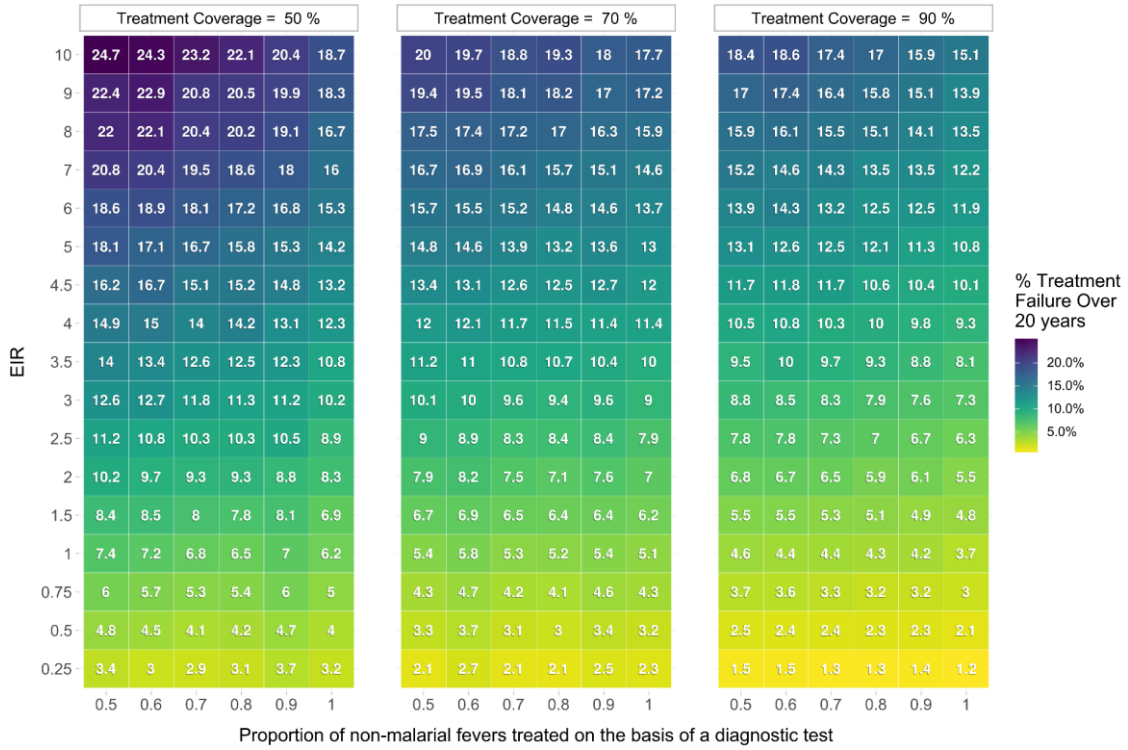


Figure 5.7 Impact of increasing the proportion of non-malarial fevers tested with an RDT within sequential cycling scenarios. Each heatmap square shows the mean percentage of treatment failures over 20 years across 100 simulation repetitions for a given EIR, treatment coverage and assumed proportion of non-malarial fevers tested with an RDT. Increasing the proportion of non-malarial fevers tested is shown to decrease the percentage of treatment failures over time.

5.4. Discussion

The results demonstrate a significant decrease in the continued spread of antimalarial resistance through increasing diagnostic testing of NMF at higher transmission intensities. By ensuring that all suspected cases of malaria receive a test before being treated, the amount of overprescription of antimalarials is reduced and the selection of resistant parasites during the elimination phase of partner drugs is decreased. This finding confirms the hypothesis that this effect occurs more frequently with increasing malaria prevalence due to the higher transmission intensity increasing the probability that infection events occur when partner drug concentrations are sufficiently high to prevent infection by wild type parasites but not resistant strains. This effect was shown to occur at values of EIR greater than 3, which corresponds to ~17.5 all-age prevalence by PCR.

My results predict that while MFT initially yields lower treatment failure rates, over a 20-year period MFT is predicted to lead to an increase in treatment failure. The difference between MFT and sequential cycling was greater in lower transmission settings with the lowest treatment coverage, with the differences between the treatment strategies increasingly non-significant at higher treatment coverages in higher malaria prevalence settings. MFT was predicted to be significantly improved with each additional ACT available to be used. Multiple drugs will force parasites to acquire multiple resistance mutations, which is both slower and introduces compounded fitness costs weakening the transmission efficacy of resistant parasites. In particular, the largest consecutive decrease in treatment failures was observed when increasing from three to four first line therapies available. With four available ACTs the mean treatment failure over the 20-year study period is predicted to be less than or equal to 15% across all transmission intensities and treatment coverages explored. MFT with four available drugs is predicted to have lower treatment failures over 20-years compared to cycling three drugs with 100% testing of NMFs at treatment coverages greater than 70%, with MFT requiring five first line ACTs to be significantly better than sequential cycling with three ACTs.

The study settings explored were based on a similar study by Nguyen et al., which concluded that MFT was a superior strategy compared to sequential cycling for tackling resistance (Nguyen et al. 2015). In their study, however, they primarily focussed on the emergence of resistance, i.e. how effectively can resistance arise and persist in a population that does not currently have resistance to the point where it is established, with “established” interpreted in their study as 1% population allele frequency. In this study we chose to continue these scenarios by considering the continued spread of resistance, which represents the next stage in the development of drug resistance (Mackinnon 2005). During the spread of resistance, as opposed to the emergence, I predict that the use of MFTs is not a superior strategy and shows at best a comparable ability to slow the spread of resistance.

The different conclusions made with regards to the benefit of MFT between the studies could be due to a number of reasons. Firstly, in my study I focussed on the continued spread of resistance as opposed to emergence. It is entirely plausible that cycling strategies and MFT strategies will have different benefits depending on the stage of resistance development. Secondly, there are a number of model differences between my study and the study by Nguyen et al. Most importantly is the presence of partial resistance in my model, whereas in their study resistance afforded complete lack of drug effect on the parasite. Unfortunately, this difference could not be incorporated in a sensitivity analysis in the current scenarios due to the focus on the spread rather than emergence of resistance. For example, the simulation is seeded with 10% artemisinin resistance, which would instantly yield 10% treatment failure under the assumption of complete resistance and so all cycling methods would be cycling every year by default. Future work will address this issues by exploring spread with starting

artemisinin resistance at 1%. I could, however, explore other modelling differences, such as the inclusion of cotransmission events in my model, which may affect selection of resistance. For example, consider the case when three parasite strains with different resistance mutations are introduced in the same bite. In an MFT strategy, whatever drug is chosen will encounter a parasite that is resistant to the drug, increasing the probability the individual will fail treatment. If only one parasite strain is introduced, then there is a one in three chance that the MFT chosen will encounter a resistant parasite. Additionally, cotransmission events will increase the chance that an individual is polyclonally infected, which increases the opportunity for recombination and may lead to faster build-up of multidrug resistant parasites. These assumptions were explored by adding two additional scenarios in which firstly no cotransmission occurred, i.e. only 1 strain is passed on, and secondly where cotransmission is tripled, i.e. three-times as many sporozoites are expected to survive from an infectious bite. These scenarios were explored at lower transmission settings (EIR = 0.25-2) focussing on prevalence settings similar to those explored in Nguyen et al. Assumed cotransmission levels were predicted to have no effect on mean treatment failures, regardless of the treatment strategy or transmission intensity. It is thus unlikely that the difference in performance of MFT vs cycling is due to this additional modelling complexity and is instead more likely to depend on the inclusion of partial resistance.

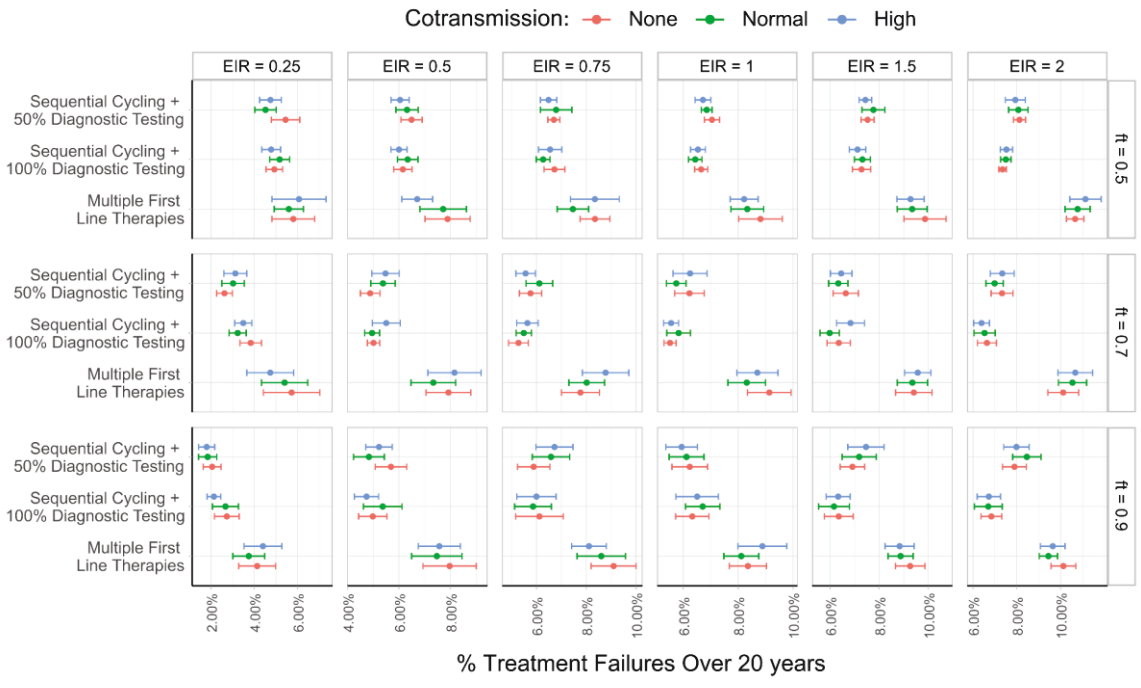


Figure 5.8 Effect of changing the level of cotransmission on treatment failures across different treatment strategies. The plots show the mean and 95% confidence interval for the average treatment failure over 20 years. A range of transmission intensities and treatment coverages are shown. Overall no consistent relationship was observed between the assumed number of sporozoites transmitted in an infectious bite and the performance of each of the treatment strategies explored.

There are a number of modelling assumptions that had to be introduced in order to conduct this analysis. The drug efficacies estimated for each of the ACTs used are overly simplified and should not be used to infer the specific performance of any one ACT. This is firstly because paired genotype-phenotype studies have not been conducted for each ACT and consequently the efficacies for all ACTs except DHA-PPQ were estimated using experimental set ups which are not ideal. Secondly, the representation of resistance to each ACT being encoded by two separate loci is overly simplified and does not capture possible compensatory mutations in the genetic background (Amato et al. 2016). For example, the *pfmdr1* gene has been linked to the efficacy of a number of antimalarials. This is dependent both on copy number variations (Borges et al. 2011), and specific SNPs in *pfmdr1* (Nwakanma et al. 2014). Additionally, both AL and ASAQ have been demonstrated to exert opposing selective pressures on *pfmdr1*. ASAQ selects for the YYY haplotype at positions 86, 184 and 1246 and AL selects for the NFD haplotype (Dokomajilar et al. 2006; Humphreys et al. 2007), which has been shown to be predicted by a country's first line ACT policy (Okell et al. 2018b). Although not modelled here, this relationship could be modelled in the future to explore whether the differential selection of opposing haplotypes could be harnessed to increase the therapeutic lifespan of ACTs. However, our current understanding of the drug efficacies of each ACT on all possible relevant parasite haplotypes is incomplete. Moving forwards, it will be necessary to conduct a systematic review of all phenotype-genotype studies available to construct more suitable values, which will likely require a PK-PD modelling approach to infer efficacies again where the data is not available. Although these intricacies were not included in my model, the framework developed is suitably flexible to extend the number of loci associated with resistance, with the model previously being used to easily track parasite barcodes of 576 loci when simulating parasite identities by descent.

The model also included other assumptions related to the role of pre-existing parasite strains on treatment outcomes. I assumed that the outcome of superinfection events is regardless of pre-existing strains from current asymptomatic infections did not alter treatment outcomes. This assumption is based on recent studies using malaria genotyping in longitudinal studies in Malawi that show that persistent asymptomatic infections do not alter the probability of new infections triggering a symptomatic infection (Buchwald et al. 2019), which builds on studies showing that pre-existing asymptomatic infections do not offer protection against new infections (Portugal et al. 2017). It was also assumed that the multiplicity of infection did not affect treatment outcome. However, alternate studies have shown that the multiplicity of infection is associated with subsequent risk of recrudescence in the DRC and Uganda (Dokomajilar et al. 2006; Humphreys et al. 2007). However, the increased number of strains in these infections are likely to be due to cotransmission events and not

due to low density asymptomatic infections. Subsequently, the assumption that the risk of recrudescence is independent of the number of parasite strains passed on may be incorrect. Lastly, my model did not model recrudescent infections due to treatment failure leading to new clinical cases and assumed that all recrudescent infections were asymptomatic. Although there is evidence that recrudescence can lead to a symptomatic infection (Mohamed et al. 2017), it is unclear how many of these cases are due to incomplete drug adherence, which has been shown to decrease treatment efficacy (Challenger et al. 2017).

The repeated emergence of antimalarial resistance has prompted policy makers to consider how drug treatment strategies can be improved to prevent resistance emergence. The continued spread of resistance, however, is equally pressing, with new reports since parameterising this model showing that DHA-PPQ treatment failure rates have increased in the last two years (Portugal et al. 2017). With no replacements in development for the artemisinin component in ACTs we need to manage the current spread of artemisinin resistance. Drug treatment strategies can be optimised for this, with the findings predicting that current strategies cycling first line therapies after 10% treatment failure has been reached can be further optimised by reducing the presumptive treatment of non-malarial fevers. Although there is additional prophylactic protection conferred by this practice, over a 20-year study period I predict that this will be offset by the increased spread of resistance in most transmission settings. Alternatively, the use of multiple first line therapies can be highly effective in slowing the spread of resistance. However, at least four ACTs need to be used to sufficiently increase the fitness cost barriers associated with resistance. Control programs are consequently recommended to assess if the logistics and costs of maintaining stocks of at least four ACTs are feasible. If not, increasing diagnostic testing is recommended as a strategy to slow the continued spread of resistance.

5.5. Conclusion

In this chapter I have extended the transmission model developed in chapter 3 to consider antimalarial resistance. These extensions introduced a number of additional modelling assumptions and highlighted the need for a greater understanding of the relationships between parasite resistance genotypes and phenotypes, as well as the impact of asymptomatic infections on treatment outcomes. Despite these challenges, the developed model was able to answer questions related to the optimum drug treatment strategies for slowing the spread of antimalarial resistance. The representation of resistant parasites with the barcode is also flexible enough to be extended to capture any epistatic relationships between resistance loci. Given the increasing number of genotype-phenotype studies being conducted as part of endeavours such as the Tracking Resistance to Artemisinin Collaboration, the model can hopefully be more completely parameterised and allow a more complete model of

parasite phenotypes to be included. In addition, the model incorporates a more nuanced characterisation of resistance phenotypes compared with previous modelling efforts that focussed on the emergence of drug resistance as opposed to its spread. Understanding the emergence of drug resistance is also highly important and will be explored in the next chapter by modelling the importation of resistant parasites into susceptible populations.

Chapter 6. Transmission of invasive artemisinin resistant parasites: a modelling study

The potential spread of artemisinin resistance out of the Greater Mekong Subregion (GMS) poses a threat to continued reductions in the burden of malaria in Africa. To date the multidrug resistant co-lineage of *P. falciparum,* KEL1/PLA1*,* has spread from Western Cambodia along the northern border with Thailand before spreading south along the Eastern border of Cambodia shared with Laos and Vietnam (Imwong et al. 2017a). The spread of a single dominant co-lineage, which expresses both artemisinin and piperaquine resistance, has prompted debate over whether the potential transnational spread of this lineage is a public health emergency of international concern (Alonso 2017). One area of research that has yet to be explored though is by what routes and how quickly could a multidrug resistant lineage invade other countries. Clearly, KEL1/PLA1 has been able to spread across multiple neighbouring countries, however, it was not found to have spread to Myanmar (Amato et al. 2018). This is primarily believed to be due to artemethur-lumefantrine being used primarily in Myanmar thus reducing the selective advantage of PLA1. However, dihydroartemisinin-piperaquine is increasingly used due to concerns of treatment failure associated with increased *mdr1* copy numbers*.* Alternatively, KEL1/PLA1 may have failed to spread to Myanmar because malaria parasites have evolved to be adapted to the local vector population (Molina-Cruz et al. 2015). Consequently, invasive parasite strains that are not adapted to the vector population may be less able evade the mosquito immune response and be transmitted through the mosquito.

In this chapter, I extend the resistance changes made to the model in chapter 5 to explore how parasite adaptation to the vector population impacts the speed at which an invasive parasite strain can spread within a new population. This work incorporates experimental data from collaborators in the Department of Life Sciences at Imperial College London, in which the fitness costs and resistance phenotype of artemisinin resistance has been examined using standard membrane feeding assays. The experimental methods included in sections 6.2.1 and 6.2.2 were not conducted by myself and were conducted by Kathrin Witmer, Farah Dahalan and other members of Jake Baum's research group. These sections are included to increase understanding of the experimental data and my involvement in the downstream statistical analysis. I use mixed effects models to estimate key parameters from the SMFA studies, before incorporating them within the transmission model. The model is used to characterise the impact on the speed with which invasive parasite lineages (i.e lineages that are not adapted to the vector population into which they are invading) could successfully spread within a non-resistant population.

6.1. Introduction

The development of artemisinin and its derivatives has radically changed the treatment of severe malaria, affording an effective treatment when previous first line treatments were failing due to resistance (Tu 2011). Artemisinin was first identified from a reference to the ability of qinghao, the Chinese name for the herb *Artemisia annua L.,* made in Ge Hong's *A Handbook of Prescriptions for Emergencies* (Tu 2011). The method of action of artemisinin is still uncertain, however the prominent theory is that parasite metabolism of haemoglobin causes iron-mediated activation of artemisinin that results in the drug being both highly reactive and rapidly consumed (Tilley et al. 2016). This mechanism of action causes artemisinin to have a very short half-life and therapeutic window, and consequently it is partnered with alternative antimalarial compounds that have a longer half-life and increase the probability of completely clearing an individual's parasites (Ashley and White 2005). Partnering artemisinin derivatives with a partner drug also decreases the probability of resistance emerging, as has been shown in the deployment of antituberculosis and antiretroviral drugs (Group 2004).

Despite efforts to slow the emergence of artemisinin resistance, the first documented cases of artemisinin resistance were made in Western Cambodia between 2008-2009 (Noedl et al. 2008; Das et al. 2009). Patients presented with delayed parasite clearance and increasing rates of 28-day parasitological failure, which suggested that new drug resistant parasite lineages were emerging. The Greater Mekong Subregion (GMS) is now the established origin of artemisinin resistance, with resistance to dihydroartemisinin-piperaquine (DHA-PPQ) spreading throughout the region (Hamilton et al. 2019). In addition, independent emergence of artemisinin resistance has now been identified in New Guinea, with identity by descent comparisons showing shared ancestry with parasites from Indonesian Papua as well as the gradual acquisition of a complex set of resistance variants (Miotto et al. 2019). The speed with which artemisinin resistance has emerged in South-East Asia has led to the concern that it may arrive in Africa, which could reverse recent declines in malaria-related mortality rates. Although there are still a number of therapeutically effective artemisinin combination therapies (ACTs), there is evidence that the spread of artemisinin resistance will lower the evolutionary threshold for further resistance mechanisms to evolve to currently efficacious partner drugs (Dondorp et al. 2010). Consequently, artemisinin resistance poses a threat by both reducing the number of effective frontline treatment and shortening the therapeutic lifespan of current frontline treatments.

Identification of the genetic determinants of artemisinin resistance occurred in 2012, which showed that the resistance phenotype is associated with polymorphisms in the propeller region of the Kelch 13 (K13) *P. falciparum* protein (Cheeseman et al. 2012). Since this discovery, numerous K13 single nucleotide mutations have been observed that exhibit a reduction in the rate at which parasites are

cleared by artemisinin. Four specific K13 mutations, C580Y, R539T, I543T, Y493H, have been shown directly to confer artemisinin resistance through the use of genome editing (Menard and Dondorp 2017). In addition, these mutations have increased at different rates within the GMS, suggesting that either the level of resistance conferred impacts the rate at which resistance may spread (Anderson et al. 2016), or that these specific mutations are aligned with different compensatory mechanisms in the genomic background that may offset any fitness cost associated with K13 mutations (Nair et al. 2018). The precise mechanism through which each K13 mutation confers artemisinin resistance is unknown, however, parasites with K13 mutations ($K13^{mut}$ parasites) exhibit an upregulation in the unfolded protein cell stress response (Mok et al. 2015). Consequently, it is likely that $K13^{mut}$ parasites may be able to better resist the impact of drug damage on cellular function.

The phenotype exhibited by drug resistant parasites is associated with a decreased clearance rate, which ultimately increases the probability that treatment failure occurs. Treatment failure can broadly be considered in three forms, each of which affords different degrees of transmission advantage to the resistant parasite. The first is that the uncleared parasite population increases in density triggering an additional symptomatic fever, often referred to as a late clinical treatment failure (Slater et al. 2016). The second occurs when uncleared parasites trigger a longer asymptomatic infection leading to a late parasitological failure. Lastly, resistant parasites that are successfully cleared by treatment are likely to have been cleared some days after a wild type parasite would have been cleared, which is referred to as slow parasite clearance. Although slow parasite clearance is not considered a treatment failure based on a 28-day follow up assessment of parasitaemia, each outcome results in an increased duration of infection for the resistant parasite, which increases the probability that a resistant parasite may be onwardly transmitted to a mosquito.

Artemisinin resistant parasites are also afforded increased transmission potential resulting from the effect of artemisinin on parasite gametocytes. Gametocytes mature over a 10-12 day period and are dependent on the size of the asexual parasite population (Bruce et al. 1990), with evidence that there is a delay from emergence of blood-stage parasites to onward infectivity (Ross et al. 2006a). Timing of sexual commitment is the result of a number of factors not limited to the balance between asexual parasite density (Smith et al. 2000), within host parasite genetic diversity (Bousema et al. 2008; Sowunmi et al. 2009) and human host factors such as anaemia and immunity (Sowunmi et al. 2009; Carter et al. 2013). In addition, the resultant sex ratio formed have been suggested to optimise the transmission strategy of the parasite (Paul et al. 2002), with gametocyte sex ratios typically female biased, but with increased male ratios occurring between different parasite clones, in the presence of competing clones (Ranford-Cartwright et al. 1993; Reece et al. 2008), throughout the course of an infection (Paul et al. 2000) and differing across geographic regions (Robert et al. 1996, 2003) and

timings within a transmission season (Sowunmi et al. 2008). There is evidence, however, that drug resistant strains, due to their comparatively lower parasite densities (Reece et al. 2010), commit to sexual stages earlier and exhibit increased male sex ratios (Bell et al. 2012). In addition, drug resistant parasites, in the presence of drugs, have been observed to have increased gametocyte carriage compared to wild type parasites (Lin et al. 2018). Consequently, artemisinin resistant strains may be more likely to be onwardly transmitted within a recently treated infection. In addition, artemisinin-derivatives have been shown to have transmission-blocking activity on K13 wild type ($K13^{WT}$) parasites through the selective targeting of exflagellation (Delves et al. 2012; Nair et al. 2018). $K13^{mut}$ field isolates that exhibit decreased clearance rates by dihydroartemisinin (DHA) have also shown resistance to the impact on exflagellation rates (Lozano et al. 2018). These findings, combined with the increased commitment to sexual stages and the proteomic evidence of the K13 protein in gametocytes (Lasonder et al. 2016), suggest that K13 mutations may have a role both in mature gametocyte stages and in increasing the probability of transmission to the mosquito.

P. falciparum parasites have also exhibited evolution to evade the immune system of mosquitoes, enabling successful infection and subsequent onwards transmission. There are more than 70 different *Anopheles* species worldwide that enable *P. falciparum* transmission (Sinka et al. 2012), which have resulted in the evolution of specific parasite-vector adaptations that afford different degrees of compatibility between the parasite and vector (Molina-cruz and Barillas-mury 2014). For example, a strain of *P. falciparum*, *Pf* NF54, originating from Africa has been shown to be ineffective at infecting *An. albimanus* (Baton and Ranford-cartwright 2012), which is an important malaria vector within central America (Sinka et al. 2010). Similarly, isolates of *P. falciparum* originating from South-East Asia have been shown to effectively infect *An. stephensi*, which is a major vector in India, however they are less effective at infecting *An. gambiae*. Recent evidence has shown that the mosquito immune system has selected for different *Pfs47* haplotypes of *P. falciparum*, each of which is adapted to the global distribution of *Anopheles* mosquitoes (Molina-Cruz et al. 2015). Importantly, the *Pfs47* protein is expressed on the surface of female gametocytes. This method of expression may enable a route by which male gametocytes of an invading *P. falciparum* strain, which are not adapted to the local vector population, can evade the immune system of the mosquito and fertilise a local female gametocyte. This situation may arise naturally if a human coinfected with both an invading $K13^{mut}$ strain and a local $K13^{WT}$ strain is treated with an ACT, with the artemisinin component sterilising $K13^{WT}$ male gametocytes more substantially than $K13^{mut}$ parasites.

Here, I fit statistical models to data comparing the transmission capabilities of one $K13^{mut}$ isolate versus a geographically-matched K13 wild type ($K13^{WT}$) isolate to estimate the impact that artemisinin has on the ability of the parasite to transmit to mosquitoes. Estimated parameters governing the relative

fitness costs associated with K13 mutations and the decreased sensitivity to artemisinin are incorporated into the transmission model developed in chapter 5, with extensions to incorporate the adaptation of parasites to different populations of mosquitoes. Through simulating a range of epidemiological settings I test whether artemisinin resistance confers an additional selective advantage in terms of local adaptation through the decreased impact of artemisinin on sexual stages of the parasite.

6.2. Methods

6.2.1. Standard Membrane Feeding Assays

Asexual blood stage parasites and gametocytes were cultured using a previously described method (Delves et al. 2016) with the following modifications. Asexual blood stage cultures were maintained in asexual culture medium (RPMI 1640 with 25mM HEPES (Life Technologies), 50 µg L^{-1} hypoxanthine (Sigma), 5% A+ human serum (Interstate Blood-Bank) and 5% AlbuMAX II (Life Technologies)). Gametocyte cultures were maintained in gametocyte culture medium (RPMI 1640 with 25mM HEPES (Life Technologies), 50 µg L^{-1} hypoxanthine (Sigma), 2 g L^{-1} sodium bicarbonate (Sigma), 5% A+ human serum (Interstate Blood-Bank) and 5% AlbuMAX II (Life Technologies)).

Gametocytes were induced and maintained as mentioned above. At day 14 post induction, gametocytes were spun down at 38°C and resuspended in 5ml of suspended animation buffer (SA) (Nijhout and Carter 1978). Gametocytes were purified using magnetic-activated cell sorting (MACS) before being resuspended in gametocyte medium including $25x10^6$ fresh red blood cells. Dihydroartemisinin (DHA) was added to ensure the desired end concentration within a 10ml gametocyte culture. DHA was again added after 24 hours to give a double-dosing regimen within a 48-hour period. After 48 hours, human serum and fresh blood were mixed with the parasite culture before being fed to adult *Anopheles stephensi* mosquitoes using a 3D printed feeder (Witmer et al. 2018).

6.2.2. Oocyst counts and size

On day 10, mosquitoes were dissected and midguts stained in 0.1% Mercurochrome before being inspected using a 10x magnification to count oocysts. To measure oocyst size, midguts of *Anopheles stephensi* fed on *P. falciparum*-infected blood were dissected and fixed with 4% formaldehyde, permeabilized with 0.1% Triton X-100 for one hour, blocked with 3% BSA for 30 minutes and stained with 1ug/ml in DAPI for 3 minutes. Midguts were washed with 1xPBS and mounted in Vectashield. Images were acquired on a Nikon Ti Eclipse inverted fluorescence microscope. Images of *P.*

falciparum-infected midguts were captured using Z-stack imaging and processed in ND processing using the Maximum Intensity Projection. Oocysts were automatically detected using Automated Spot Detection based on the intensity of the oocysts compared to midgut cells (NIS-Elements) and the prevalence and intensity of oocysts recorded.

6.2.3. Statistical Modelling of Oocyst Intensity and Prevalence

To assess the impact of artemisinin on the ability of each parasite line to form oocysts, I used generalised linear mixed effects models in order to incorporate data from different experimental replicates within the same modelling framework. These models have previously been used to model malaria transmission blocking interventions (Churcher et al. 2012b). We modelled either oocyst intensity or prevalence as the response with treatment (DHA concentration) included as a fixed effect and 0μM DHA represented by control groups treated with DMSO. The parasite line treated ($K13^{WT}$ or $K13^{mut}$) was included as a fixed effect to assess the differential impact of artemisinin on transmission success. The impact of treatment between experimental replicates was allowed to vary at random between replicates. A logistic regression (binomial error structure) was used to model the prevalence of mosquito infection, i.e. the presence or absence of oocysts, and a zero-inflated negative binomial distribution was used to model the intensity of infections, i.e. the numbers of mosquito oocysts (Blagborough et al. 2013). 95% confidence interval estimates were generated for the impact of drug concentration by bootstrapping methodology with 100,000 replicates.

6.2.4. Transmission modelling of male gametocyte sterilisation

Using the results from the statistical models, I extended the transmission model developed in chapter 5 to incorporate the effects of artemisinin on the development of male gametocytes and subsequent oocyst formation. Previously in the model, two parasite barcodes would be sampled from an infected individual if a feeding mosquito was infected by the individual. The barcodes are chosen based on their relative parasite densities, which we define as $c = \{c_1, c_2 \dots c_n\}$, where n is the total number of parasite strains in the individual and c_i is the infectiousness of parasite strain i to the mosquito. In accordance with the laboratory data, parasites that are resistant are assumed to have fitness costs that reduce their parasite density, which reduces their probability of being passed on to a mosquito. As in chapter 5, fitness costs are multiplicative with a fitness cost associated with each locus in the parasite barcode that confers a resistance phenotype. The resultant probability vector for each parasite strain being sampled is given by $cr = \{c_1 r_1, c_2 r_2 \dots c_n r_n\}$, where r_i is the fitness cost associated with parasite strain i.

To incorporate the effects of male gametocyte sterilisation, I adapted this process to consider the first barcode sampled from the infected individual to be the female gametocyte and the second to now be the male gametocyte. Female gametocytes are still sampled at random with probability cr, whereas male gametocytes are sampled using an adapted probability vector, cr_{male} given by:

$$cr_{male} = crs$$

Where $s = \{s_1, s_2 \dots s_n\}$ and s_i is the potential sterilising effect of artemisinin on strain i, which incorporates the effect of artemisinin on developing male gametocytes and is given by:

$$s_i = \begin{cases} 1, & if\ individual\ has\ not\ been\ treated\ with\ artemisinin\ in\ last\ 12\ days \\ 1, & if\ strain\ i\ is\ K13^{mut} and\ treated\ with\ artemisinin\ in\ last\ 12\ days \\ s, & if\ strain\ i\ is\ K13^{WT}\ and\ treated\ with\ artemisinin\ in\ last\ 12\ days \end{cases}$$

where s is the degree of sterilisation exerted by artemisinin. The 12-day duration reflects that artemisinin is known to target immature gametocytes in stages I-III due to their metabolic similarity to asexual parasite stages (Kumar and Zheng 1990; Lelièvre et al. 2012). In addition, artemisinin has been shown to sterilise male wild type gametocytes, which causes mature male gametocytes exposed to artemisinin to not undergo exflagellation in the mosquito and fail to form oocysts (Lozano et al. 2018). Consequently, an individual will only be infectious to mosquitoes after 12-days, with two days for artemisinin to clear and ten days for new mature gametocytes to be produced.

To explore the impact of male gametocyte sterilisation on the speed at which an invasive K13^{mut} parasite is able to invade a new population, I simulated a scenario of constant importation of a K13^{mut} parasites with *plasmepsin 2-3* duplications (referred to from here as K13^{mut}/PLA^{mut} parasite) into a region in which DHA-PPQ is the front-line ACT used. I assumed that 0.5% of new cases in the region were due to importations of double-mutant K13^{mut}/PLA^{mut} parasites.

To incorporate the results of the SMFA statistical modelling, the number of oocysts formed from a bite varied depending on the resistance profile of the parasite strains in the infecting human, i.e. the mean number of oocysts formed from an individual with only resistant parasites is different than for an individual with wild type parasites. Similarly, the mean number of oocysts formed from a bite on an individual treated with artemisinin in the last 10 days was estimated from the findings of the SMFA. Lastly, we used a range of values for s that include the mean estimated value from the SMFA as well as 0% sterilisation and 90% sterilisation as extreme cases to assess for the maximum predicted effect that could be associated with sterilisation of male K13^{WT} gametocytes by DHA. Lastly, it is assumed that K13^{mut} parasites commit to sexual stages 5 days earlier than K13^{WT} parasites, resulting from increased transmission advantage that has been observed for resistant parasite strains in the absence of drug pressure (Mockenhaupt et al. 2005).

I simulated populations of 100,000 individuals for 120 years. The first 40 years reflect a period of "burn-in" ensuring an equilibrium is reached in simulations, before starting 80 years in which invasive $K13^{mut}/PLA^{mut}$ parasites could be imported. I explore a range of different transmission intensities (EIR = 0.5, 0.75, 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10) and treatment coverages (f_T = 30%, 50%, 70%). The population allele frequency of each locus and the frequency of each parasite strain was tracked and the time at which the 60-day mean population frequency of the double mutant increased reached 2.5%, 5%, … 72.5%, 75% was recorded to give 30 measures of the speed of resistance invasion.

6.2.5. Transmission modelling of parasite adaptation to vector populations

To incorporate the role of parasite adaptations to local vector populations, I extend the parasite barcode to include an additional locus that confers the ability for ookinetes to successfully infect the local vector. Female gametocytes that do not possess this locus are assumed to be less capable of forming oocysts, with the probability that the ookinete will not develop into an oocysts given by p_{VA}. For example, I draw the number of oocysts that will form from a transmission event from a zero truncated negative binomial distribution. For each oocyst that would be formed, I sample a female gametocyte with probability cr and a male gametocyte with probability cr_{male}. If the female gametocyte does not possess the locus conferring local vector adaptation, the oocyst is assumed to have been formed with probability p_{VA}. I used three values of p_{VA} (0.1, 0.5, 1.0), which were chosen to explore both the most likely value suggested from studies of *Pfs47* (Molina-Cruz et al. 2015), as well as extreme values representing an absence of vector of adaptation ($p_{VA} = 1$) and severe adaptation ($p_{VA} = 0.1$).

The locus associated with local vector adaptation represents the gene *Pfs47*, which is 151kb away from the gene for K13 (Molina-Cruz et al. 2015). Consequently, we have adapted our recombination model such that the probability of a recombination event occurring is no longer independent between all loci in the parasite barcode. Using estimates of the recombination rate equal to 13.5kb/cM (Miles et al. 2016) and Haldane's mapping function (Haldane 1919) under the assumption of no interference we assume the probability of a recombination event happening between the locus encoding for artemisinin resistance and the locus encoding for local vector adaptation to be equal to 0.091 per meiosis event. For clarity we are assuming vector adaptation is encoded solely by the *Pfs47* locus and that the locus is biallelic, i.e. it is adapted to the vector population or it is not. To explore this assumption, we also model the case where the locus conferring vector adaptation is not under genetic linkage with K13. The same population settings as previously described are simulated, with three values of p_{VA} (0.1, 0.5, 1.0) chosen to explore the impact of vector adaptation, with p_{VA} = 1.0 acting as a control, i.e. there is no parasite adaptation to local vector populations.

6.3. Results

6.3.1. Transmission of field isolates with different K13 genotypes under DHA drug selection

Two parasite isolates, a $K13^{WT}$ (ANL1) and a $K13^{mut}$ isolate (APL5G) with evidence of high rates of mosquito infection were selected for further assessment via SMFA. The APL5G isolate possessed both the C580Y K13 mutation as well as amplification of the *plasmepsin* 2-3 gene cluster conferring piperaquine resistance. For a given gametocyte density, APL5G parasites yielded lower mean intensities and prevalence of oocysts in the absence of DHA (DMSO) compared to ANL1 parasites, potentially suggestive of a fitness cost (Figure 6.1). A significant decrease in both the oocyst intensity (incidence rate ratio = 0.73, 95% CI: 0.66-0.80) and prevalence (odds ratio = 0.46, 95% CI: 0.41-0.52) of mosquito infection was observed for ANL1 with increasing drug concentration. However, APL5G parasites were less affected by DHA, with a non-significant decrease in oocyst intensity observed (incidence rate ratio = 0.84, 95% CI: 0.65-1.07). Increasing concentrations of DHA were predicted to significantly reduce oocyst prevalence for APL5G (odds ratio = 0.72, 95% CI: 0.56-0.96), however, this effect was significantly less than the effect on ANL1 parasites (Table 6.1). The decreased impact of DHA on the ability of the $K13^{mut}$ isolate to infect mosquitoes indicates that that the resistance mutation may afford a transmission advantage in the presence of DHA, which offsets the fitness costs observed in the absence of drug.

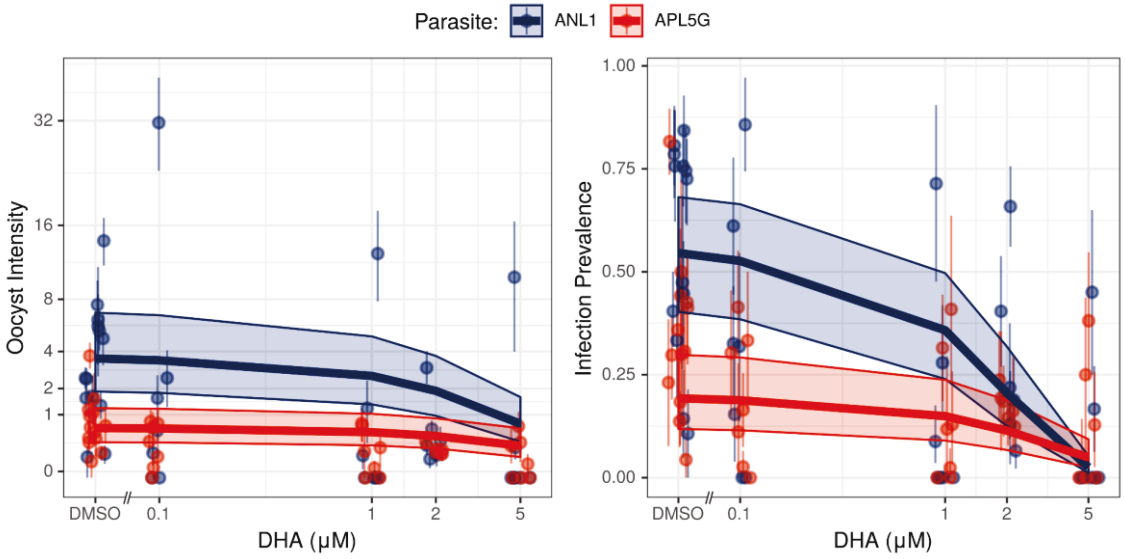


Figure 6.1 Modelling of standard membrane feeding data. Graphs show the oocyst intensity and infection prevalence of the $K13^{WT}$ isolate (ANL1) in blue and the $K13^{mut}$ isolate (APL5G) in red after exposure to DHA or DMSO. Points and whiskers on each plot show the mean and bootstrapped 95% CI for each replicate, with the predicted relationship and 95% CI shown with the trend line and shaded region. In the absence of DHA (DMSO), APL5G is predicted to produce significantly fewer oocysts and infections, whereas in the presence of DHA concentrations greater than 2μM DHA, the transmission potential of APL5G is comparable to ANL1.

Table 6.1 Effects of DHA on oocyst prevalence and intensity in different parasite lines

	Prevalence of Infection			Oocyst Intensity*		
	Odds Ratio	*95% CI*	*p*	*Incidence Rate Ratios*	*95% CI*	*p*
$K13^{mut}$ in the absence of DHA	0.20	0.16 – 0.25	**<0.001**	0.17	0.13 – 0.23	**<0.001**
DHA (μM): WT line	0.46	0.41 – 0.52	**<0.001**	0.73	0.66 – 0.80	**<0.001**
DHA (μM): $K13^{mut}$	0.72	0.56 – 0.96	**<0.001**	0.84	0.65 – 1.07	0.069
				* Negative binomial overdispersion parameter: 0.837		

6.3.2. Translating the impact of $K13^{mut}$ resistance in SMFA studies to transmission modelling

In order to translate the results of the SMFA into parameters for the transmission model, I used the results from the mixed-effect model to predict the prevalence of mosquito infection in the absence of artemisinin and in the presence of 2μM DHA. 2μM DHA was chosen to reflect the most likely peak dose concentration of DHA produced within the recommended dosage prescribed by the WHO for DHA-PPQ (World Health Organization 2015a). The WHO recommends doses of 4.0 mg/kg/day for DHA when combined with PPQ. Using the molar mass for DHA as 284.352 g/mol and a molarity of DHA equal to 2μM, the estimated peak dosage is 568.704 ng/ml, which is comparable to the median peak concentration observed in pharmacokinetic studies of DHA when prescribed as part of DHA-PPQ (Jamsen et al. 2011; Rijken et al. 2011; Chotsiri et al. 2017).

A significant reduction in the mean prevalence of infection in the absence of DHA is predicted between $K13^{WT}$ parasites (55.1%, 95% CI: 50.5% - 57.8%) and $K13^{mut}$ parasites (19.2%, 95% CI: 13.3% - 26.0%) (Figure 6.2). This finding suggests that the fitness cost associated with the $K13^{mut}$/PLA^{mut} genotype causes a 65.1% reduction in the onward probability of infection relative to wild type parasites in the absence of drug. This cost reflects the double mutant, which yields a comparative fitness of 79% for a single mutant under the assumption of a multiplicative fitness cost and an equal fitness cost associated with each resistance mutation. In addition, in the absence of DHA, a significant reduction in mean oocyst intensity is predicted in $K13^{mut}$ parasites (0.621, 95% CI: 0.32-1.22) compared to $K13^{WT}$ parasites (3.57, 95% CI: 1.87 – 6.85) (Figure 6.1a). Consequently, the mean number of oocysts that forms from an infection is described by a negative binomial with mean equal to 3.57 and dispersion parameter equal to 0.837 (Table 6.1). In onward infections originating from individuals only infected with $K13^{mut}$ parasites, the expected number of oocysts is first drawn from the described negative binomial distribution, before reducing the number of oocysts by 82.6% to a minimum of one oocyst. Oocyst

numbers are not reduced below one because the reduction in the probability of onwards infections is conferred by the earlier described fitness cost.

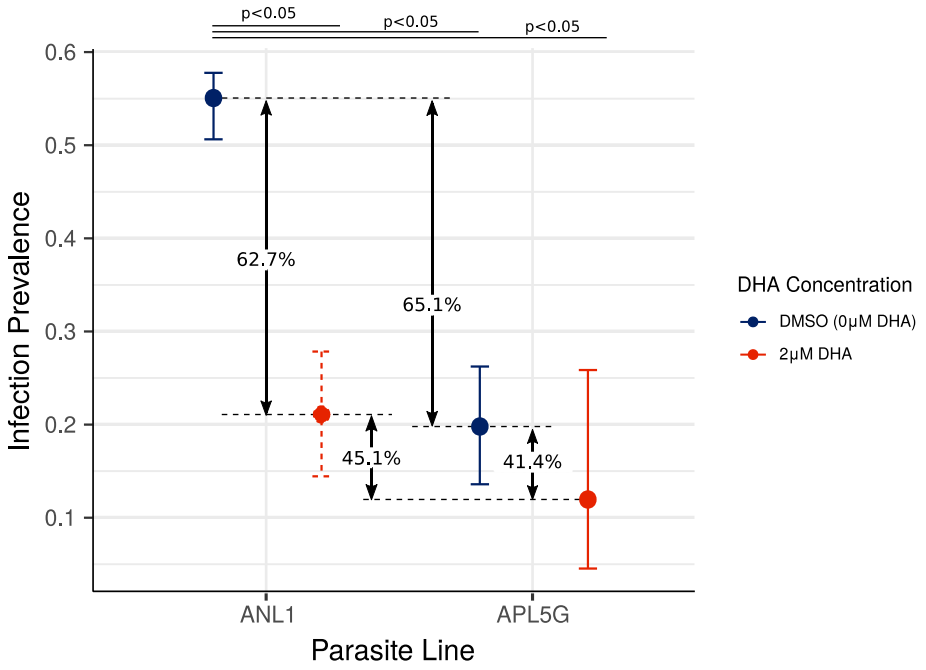


Figure 6.2 Fitness costs associated with DHA-PPQ resistance. The results of the standard membrane feeding assays allow for fitness costs associated with resistance to be estimated. The graph shows the predicted prevalence of infection of the $K13^{WT}$ stain (ANL1) and the $K13^{mut}$/PLA^{mut} strain (APL5G). The relative reduction in probability of infection between ANL1 and APL5G in the absence of DHA (DMSO) is 62.7% and reflects the comparative reduction in parasite transmission probability, which is the assumed fitness cost on transmission associated with dihydroartemisinin-piperaquine resistance. No significant difference was observed in infection prevalence in the presence of DHA.

In the presence of 2μM DHA, no significant difference was observed between $K13^{WT}$ parasites (20.6%, 95% CI: 27.7% - 14.3%) and $K13^{mut}$ parasites (12.1%, 95% CI: 4.1% - 25.4%). Consequently, we assume that in individuals who are treated with DHA-PPQ in the last 10 days, $K13^{WT}$ parasites are as equally likely as $K13^{mut}$ parasites to contribute male gametocytes to any oocysts that form from the bite and thus the chosen value for the sterilisation of male wild type gametocytes, s, must offset the fitness cost introduced above. $K13^{WT}$ parasites were significantly less likely to cause infection at 2μM DHA compared to the DMSO control, with a mean reduction of 65%, which is subsequently incorporated as our chosen value for s. The Imperial College transmission model is already parameterised with respect to the reduction in onward contribution to infection of individuals who are treated after developing symptoms compared to clinical cases not treated. Clinical cases of infection have a mean probability of causing onward infection equal to 0.068 in untreated cases and 0.0219 in treated cases, which is a 67.8% reduction. This value falls within the 95% confidence interval for the percentage reduction in $K13^{WT}$ parasites and consequently the onward probability of infection from a treated

clinical case, c_T, is kept equal to 0.0219. The only modification made is thus the decreased relative probability that in onward infections male gametocytes contribute to any oocysts formed. This effect is hypothesised to lead to an increased transmission advantage for K13^{mut} parasites in the presence of artemisinin, which is summarised in Figure 6.3.

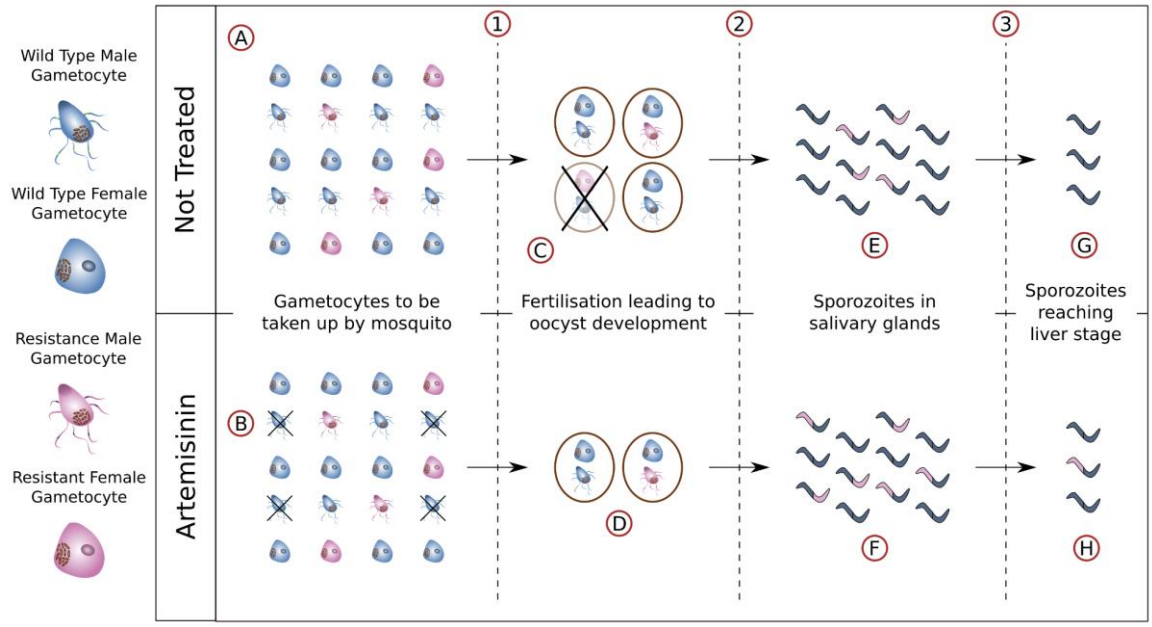


Figure 6.3 Hypothesis for population level advantage conferred to resistant parasites due to sterilisation of male wild type gametocytes by artemisinin. In (A) an individual is infected with a mix of wild type (K13^{WT}) and artemisinin resistant (K13^{mut}) parasites, and in (B) a similar individual has been treated with artemisinin, which has sterilised a number of K13^{WT} male gametocytes indicated with crosses. In both infections there are significantly more K13^{WT} parasites as well as a female biased sex ratio. **1)** Gametocytes are taken up by a mosquito bite. In (C), a number of ookinetes form that are predominately the result of fertilisation of two K13^{WT} gametes due to increased numbers of K13^{WT} gametocytes. Ookinetes formed from a female K13^{mut} gamete, however, are less likely to invade the mosquito and fail to form an oocyst indicated with the cross. In (D), fewer oocysts form due to the presence of artemisinin, however, the sterilisation of male K13^{WT} gametocytes increases the probability that an oocyst is formed due to the fertilisation of a female K13^{WT} gamete by a K13^{mut} male gamete. **2)** Oocysts burst releasing recombinant sporozoites. In (E), the majority of sporozoites are wild type, with only one-sixth of sporozoites produced being both K13^{mut} and adapted to the local mosquito population. In (F), due to the decreased number of oocysts and increased chance of K13^{mut} male gametes fertilising K13^{WT} female gametes, one-third of sporozoites are both K13^{mut} and adapted to the local mosquito population. **3)** Sporozoites are introduced by an infectious bite and a fraction survive to reach the liver stage. In (G), no K13^{mut} sporozoites have survived to yield an infection, whereas in (H), artemisinin resistance has successfully been introduced.

6.3.3. Modelled impact of male gametocyte sterilisation on speed of resistance invasion

Modelling the invasive dynamics of a K13^{mut}/PLA^{mut} parasite was simulated by assuming the constant importation of a fixed percentage of K13^{mut}/PLA^{mut} parasites into a human population with no resistant strains of malaria. An example of a single realisation of this process is shown in Figure 6.4, which

highlights how recombination between imported resistant parasites leads to the development of a vector adapted PLA^{mut} strain first, before recombining again with resistant parasites to cause the emergence and subsequent spread of a parasite strain that is both adapted to the local vector population and resistant to DHA-PPQ.

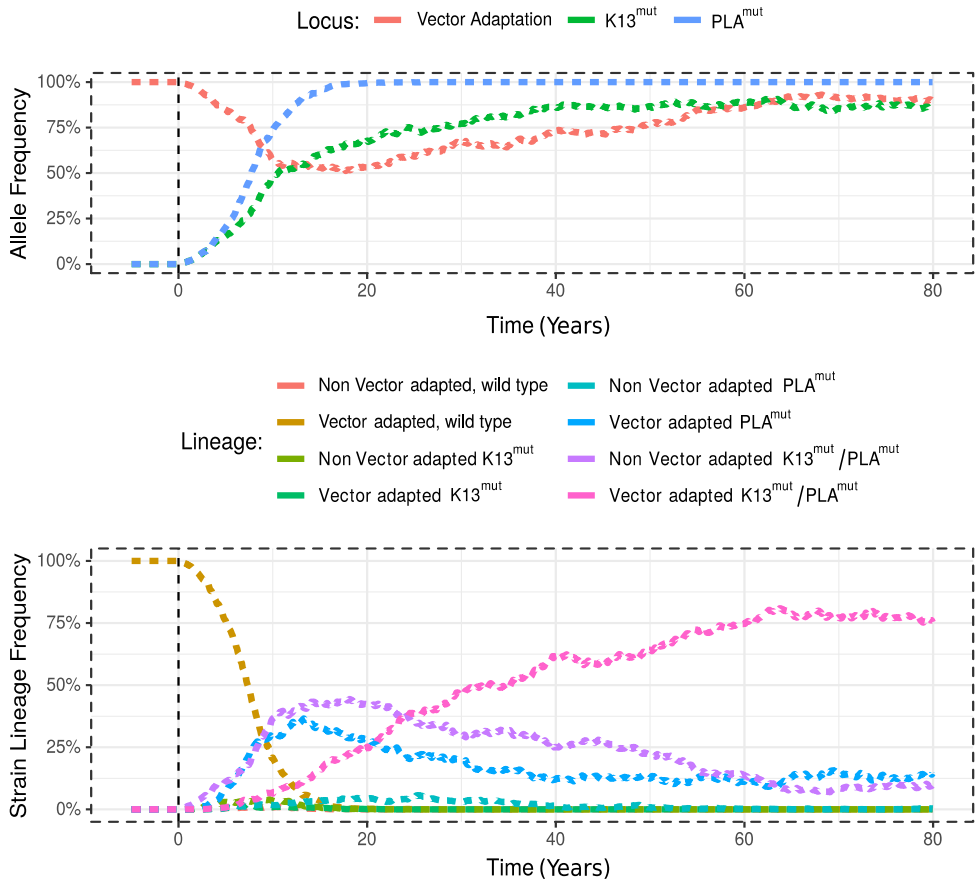


Figure 6.4 Example invasive simulation setting. In a) the frequency of each allele tracked in a simulation is presented, showing that prior to the initiation of importation there are no resistant parasites and all strains are vector adapted. After importation, vector adaptation frequency falls before increasing again. This can be explained by examining in b) the time series of the strain lineage frequency. Importation of K13^{mut}/PLA^{mut} parasites occurs first (in purple) before recombining with vector adapted parasites to form vector adapted PLA^{mut} parasites (in blue). The frequency of these parasites then subsequently falls due to recombination with invasive K13^{mut}/PLA^{mut} parasites to yield vector adapted K13^{mut}/PLA^{mut} parasites (in pink).

To assess for the impact of male gametocyte sterilisation, the mean rate of emergence was compared between 100 simulation realisations under different assumed levels of sterilisation, including the value predicted from Figure 6.2. One realisation of this is shown in Figure 6.5, which compares the impact of sterilisation for the following parameter set: EIR = 1, treatment coverage amongst symptomatic infections = 70%, relative vector adaptation of resistant female gametocytes = 100% and genetic linkage between vector adaptation locus and K13. In this scenario, vector adapted K13^{mut}/PLA^{mut} parasites spread faster in the absence of sterilisation. This difference, however, is slight

and only occurs after at least 30 years. When comparing the 95% confidence intervals of the time for the frequency of vector adapted K13^{mut}/PLA^{mut} parasites to reach 2.5%, 5% … 72.5%, 75% in each sterilisation level, 34 of the 90 comparisons are significantly different. Each of the 34 significant difference occurs due to increased speed of resistance spread in simulations with lower sterilisation levels (i.e. in 0% sterilisation vs both 65% and 90% sterilisation, and in 65% sterilisation vs 90% sterilisation).

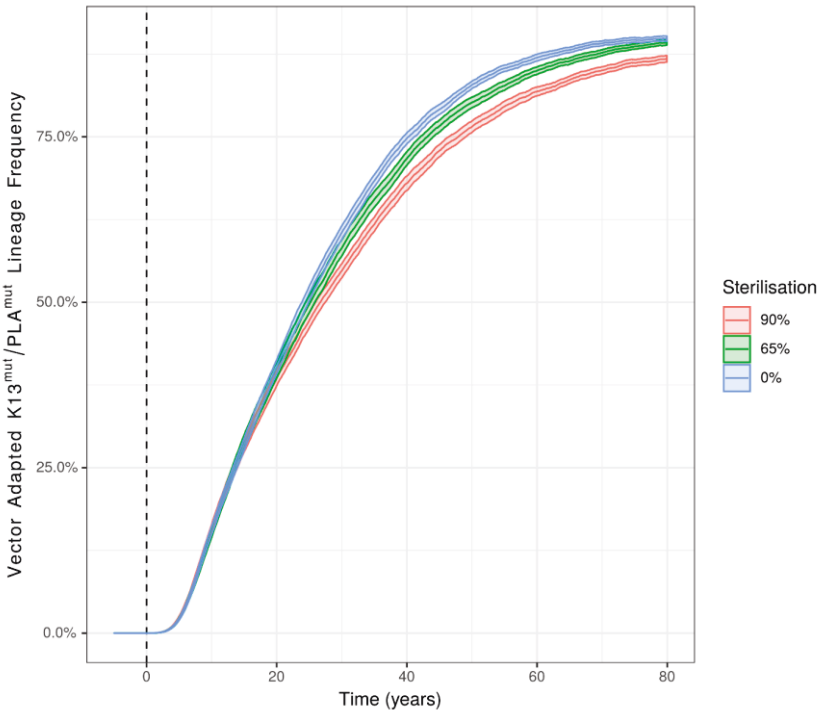


Figure 6.5 Summarised frequency of vector adapted K13^{mut}/PLA^{mut} parasites at different assumed degrees of male gametocyte sterilisation. The graph shows the mean and 95% confidence interval of a 100 simulation realisations for the invasion and spread of vector adapted K13^{mut}/PLA^{mut} parasites for one parameter setting (EIR = 1, treatment coverage = 70%, relative vector adaptation of resistant female gametocytes = 100%, genetic linkage between vector adaptation locus and K13) after the importation of K13^{mut}/PLA^{mut} parasites after 20 years, which is shown with the dashed line.

The same comparative approach was used for to assess for differences in the speed of resistance invasion across a range of transmission intensities, treatment coverages and assumed levels and linkage of parasite adaptation to local vectors. Overall, very little difference was observed when the assumed level of sterilisation changed (Figure 6.6). No differences were observed above an EIR = 3 or below an EIR = 1. Of the few differences observed, the majority occurred in the absence of parasite adaptation to local vector populations, with more differences observed when genetic linkage was assumed between the locus conferring vector adaptation and K13 (Figure 6.6b). Of the few significant differences observed overall, the speed at which resistant parasites invaded and spread was quicker in simulations with lower assumed sterilisation levels (398 significant comparisons) than in simulations with higher sterilisation levels (77 significant comparisons).

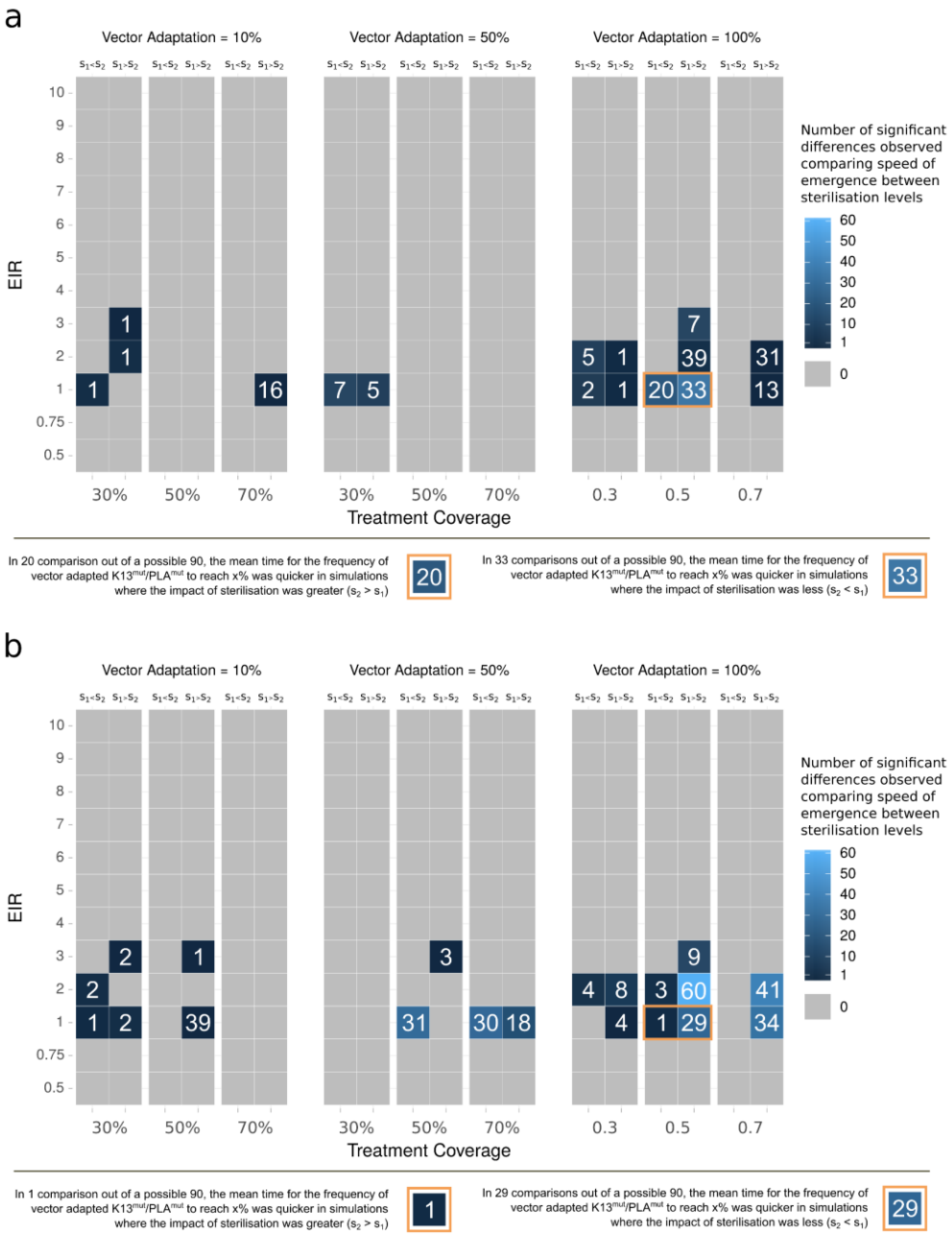


Figure 6.6 Impact of sterilisation, female gametocyte adaptation to local vectors, treatment coverage and transmission intensity on the speed at which K13^{mut}/PLA^{mut} parasites can invade a naïve population. Simulations for a given parameter set (EIR, treatment coverage and vector adaptation) were compared against the same parameter set at different levels of assumed male gametocyte sterilisation. Three sterilisation levels were explored (0%, 65% and 90%), with each sterilisation level compared against each other. The mean time taken for the 60-day mean frequency of vector adapted K13^{mut}/PLA^{mut} parasites to increase every 2.5% was recorded, resulting in a total of 90 comparisons (comparisons at 30 frequencies for 3 sterilisation levels). The number of comparisons for which a significant difference in the speed of invasion was observed is indicated in each plot tile. For each treatment coverage two columns are presented. The left column records the number of significant comparisons when the difference in emergence time is due to emergence occurring quicker in simulations with a higher sterilisation levels, indicated with $s_1 < s_2$. The opposite is shown in the right column, indicted with $s_1 > s_2$. The majority of simulations showed no impact of sterilisation on resistance invasion. Sterilisation had a marginally smaller impact when the locus conferring vector adaptation was assumed to be genetically unlinked, shown in **a)**, compared to when the locus is assumed to be *pfs47*, shown in **b)**. An explanation of the data shown in the plot is indicated for two parameter sets, which are highlighted in orange.

6.4. Discussion

The emergence of artemisinin resistance in the GMS poses a threat to malaria control, with many concerned that resistance will spread to Africa and reverse recent reductions in malaria burden. For example, chloroquine resistance first emerged in 1957 in Thailand before spreading to Africa (Payne 1987). By 1988, the vast majority of regions in sub-Saharan Africa had lost chloroquine as an efficacious antimalarial (D'Alessandro and Buttiëns 2001). However, as the use of chloroquine declined, numerous field studies observed a clear fitness cost associated with K76T chloroquine resistance, which led to a reduction in the allele frequency causing an expansion of the wild type parasite (Laufer et al. 2010; Nwakanma et al. 2014). These fitness costs, however, were unknown when chloroquine was spreading throughout Africa, with early mathematical models of chloroquine resistance noting the absence of key parameters such as the cost of resistance (Koella and Antia 2003). The cessation of chloroquine use confirmed a fitness cost, however, this strategy is not feasible with emerging artemisinin resistance with all currently recommended antimalarial treatments for uncomplicated malaria reliant on artemisinin derivatives.

In this study, SMFA were used primarily to assess for the differential impact of artemisinin on the transmission success of artemisinin resistant parasites. Using mixed effect generalised linear models I have shown a significant difference in the impact of DHA on the ability of $K13^{WT}$ and $K13^{mut}$ to infect mosquitoes. In the presence of concentrations of DHA reflective of normal DHA-PPQ dosage, the results demonstrate comparable infection prevalence between $K13^{WT}$ and $K13^{mut}$ isolates. In the absence of DHA, however, a significantly decreased probability of onwards transmission of $K13^{mut}$ parasites compared to $K13^{WT}$ parasites was observed, suggestive of a fitness cost associated with K13 mutations. Consequently, the results suggest that $K13^{mut}$ parasites have an increased transmission advantage in the presence of DHA, which offsets the fitness costs associated with resistance, through their reduced sterilisation of male gametocytes by artemisinin compared to wild type parasites.

The results of the SMFA study led to a hypothesis (Figure 6.3) that a selective advantage may be conferred by the reduced impact of DHA on the ability for $K13^{mut}$ to be onwardly transmitted compared to $K13^{WT}$ parasites. However, incorporating these findings within my adapted transmission model revealed that this effect does not cause a significant effect at the population level. Regardless of the assumed level of sterilisation, broadly similar speeds were observed for the rate at which invasive $K13^{mut}$/PLA^{mut} parasites were able to invade a previously susceptible population. This finding is similar to previous modelling studies that predicted that the addition of the gametocytocidal drug primaquine to treatment regimens affords no major gains in transmission reduction and offers little population-wide benefit.(Johnston et al. 2014)

No significant differences in the speed of resistance invasion were observed above an EIR = 3, with all settings leading to quick resistance invasion. Where differences were observed, resistance appeared to emerge slower when sterilisation was present. This may be due to sterilisation reducing the overall transmission intensity, which was shown in chapter 5 to cause resistance to spread quicker in higher transmission settings. In most parameter sets explored in Figure 6.5, for which any significant differences were observed, the differences occurred predominantly in one direction, i.e. more occurred in either lower or higher sterilisation levels. However, in the 14 parameter sets for which more than 10 significant comparisons were observed in total, four had comparable numbers (based on 95% binomial confidence interval from the counts of significant comparisons) in both directions (Table 6.2). The results presented, when viewed collectively, are suggestive of an absence of any population level selective advantage conferred by the increased transmission chance of $K13^{mut}$ parasites in the presence of DHA. Any parameter sets that exhibited significant differences due to invasion occurring quicker in simulations with greater sterilisation are more likely representative of a Type 1 error, resulting from the large number of simulations and parameter sets explored with a 100 simulation realisations per parameter set. This conclusion is also based on the lack of any clear directional pattern in the speed of resistance as sterilisation is assumed to increase.

Table 6.2 Simulation sets with non-consistent impacts of sterilisation on resistance invasion

Genetic Linkage	Vector Adaptation	Treatment Coverage	EIR	$s_1 < s_2$*+	$s_1 > s_2$
No	50%	30%	1	0.58 (7/12) (95% CI: 0.32-0.81)	0.42 (5/12) (95% CI: 0.19-0.68)
No	100%	50%	1	0.38 (20/53) (95% CI: 0.26-0.51)	0.62 (33/53) (95% CI: 0.49-0.74)
Yes	50%	70%	1	0.63 (30/48) (95% CI: 0.48-0.75)	0.37 (18/48) (95% CI: 0.25-0.52)
Yes	100%	30%	2	0.33 (4/12) (95% CI: 0.14-0.61)	0.67 (8/12) (95% CI: 0.39-0.86)

* Number of significant differences in which speed of invasion was quicker in simulations with greater sterilisation. The opposite finding (quicker in simulations with lower sterilisation) is shown with $s_1 > s_2$

+ 95% binomial confidence intervals based on the number of observations representing successes of a binomial covariate

A clear increase in the number of significant differences in speed of resistance invasion was observed when EIR was equal to 1 and 2. This finding likely reflects a transmission range in which the stochasticity in the chance of imported parasites causing onward infections is at its greatest while still ensuring eventual invasion. For example, at the lowest EIRs resistance failed to invade due to the low absolute number of resistant parasites being imported. An EIR equal to 0.75 in the simulated population of 100,000 individuals results in an average of 75,000 total infections in a year, which is 205 infections/day. With an importation rate of 0.5% we would predict that 1 infection on a given day is due to an invasive resistant parasite. Given the individual heterogeneity in biting rates, it is likely

that these rare imported events will lead to the resistant parasite being introduced into an individual with a high biting rate who is thus likely to have multiple parasites already. However, the increased biting rate of the individual also causes a higher acquired immunity and they are subsequently less likely to develop clinical symptoms and seek treatment allowing the transmission advantage conferred by $K13^{mut}$ to be realised. Conversely, if the imported parasite was introduced into a low immunity individual who would develop symptoms, there is less chance that this individual would have been either infected with wild type parasites (thus enabling the transmission advantage to be realised) or bitten while still infectious due to their lower individual biting rate. A similar explanation can be used to explain the absence of differences at high transmission intensities. In increased transmission intensities, individuals are predicted to have higher acquired immunity and subsequently more cases of malaria will result in asymptomatic infection. The duration of an asymptomatic infectious period is significantly longer than a treated infection, which leads to an increased chance of causing an onwards infection. Consequently, at higher transmission intensities the invasion of resistant parasites largely occurs via onward infections from asymptomatic individuals and thus any transmission advantage conferred in treated individuals is insignificant compared to the amount of infections originating from the asymptomatic reservoir.

There are a number of assumptions made in the transmission model, however, that may impact the findings of this study. Firstly, human immunity was assumed to be equally effective against both local parasites and the invasive resistant parasites. Immunity against severe malaria is well documented to be less geographically variant compared to immunity against uncomplicated malaria (Nielsen et al. 2004), with IgG mediated protection shown to be able to be transferred from sera collected in West-Africa adults to Thai individuals infected with *P. falciparum* (Sabchareon et al. 1991). An argument could be made that invasive parasites may lead to an increased chance of developing a fever and seeking treatment. This would be predicted to increase the opportunity for the transmission advantage of $K13^{mut}$ parasites to be realised. However, this would also likely decrease the rate at which resistance was able to invade if all imported cases were subject to treatment. This is because the results suggested in this study show that the majority of transmission in the population is due to asymptomatic infections. Consequently, if all imported strains triggered symptoms then there would likely be comparatively fewer asymptomatic infections with $K13^{mut}$/PLA^{mut} parasites compared to wild type parasites because DHA-PPQ is assumed to still achieve 28-day parasite clearance in 57.7% of infections (Witkowski et al. 2017). We have also assumed that the fitness cost associated with each mutant is the same, i.e. *plasmepsin* 2-3 gene amplifications have the same fitness cost as K13 mutations. If, however, fitness costs were more heavily associated with PLA^{mut} then we would likely

predict a slower rate of resistance invasion, reflecting the increased treatment failure associated with PLA^{mut} compared to $K13^{mut}$.

Secondly, the modelling assumes there is no inter-species competition between malaria parasites in mixed infections. This was not directly examined in the SMFA data, with mosquito infections produced by feeding on parasite cultures that were either $K13^{WT}$ or $K13^{mut}$, as opposed to a mixed culture. The presence of inter-strain competition may lead to a reduction in resistant parasite stages further reducing the chance that invasive parasite strains are able to recombine with local vector populations. Thirdly, we do not explicitly track parasite densities, with drug failure based on fixed probabilities dependent on the genotype of the parasite and the probability of onward infection of a given strain dependent on the strain's infection stage, human immunity and the results based on the SMFA data. This process does not thus include a complete PK-PD model, which would introduce an additional source of stochasticity resulting from the high variance observed in both drug absorption and parasite clearance rates (Johnston et al. 2014). This simplification may limit our modelling approach in its suitability to detect rare events that lead to resistance emerging earlier in particular settings. Lastly, the *An. stephensi* mosquitoes used in the SMFA were from a laboratory reared strain. The parasite strains chosen for further analysis by SMFA were the result of strains selected from an open-lab trial at 15 sites in 10 countries that were able to be successfully adapted to gametocyte cultures, which included the ANL1 $K13^{WT}$ and APL5G $K13^{mut}$ strains (Ashley et al. 2014). Although the strains were geographically-matched, the different affinities for each strain to be adapted in laboratory conditions and infect the chosen mosquito strain may be the cause of the observed fitness cost. The difference in infection prevalence in the absence of DHA may thus not be observed in natural infections or with a different vector. Consequently, follow up experiments that both explore the possibility of inter-species competition and the transmission capabilities of the two strains within a range a mosquito lines are warranted.

The results of the transmission modelling resulted in the rejection of the hypothesis suggested by the SMFA results that artemisinin resistance increases the speed with which invasive resistant lineages can invade susceptible populations. These results do, however, increase our understanding of the transmission phenotype associated with artemisinin resistance. For example, the results from the SMFA are to my knowledge the first to directly measure the fitness costs associated with artemisinin resistance at the transmission level, i.e. how fitness affects the ability for parasites to successfully infect mosquitoes. In contrast to this study, a number of studies have measured the in vitro fitness costs affecting asexual parasites with artemisinin resistance (Hott et al. 2015). These approaches often use competition assays or analysis of growth curves of naturally occurring $K13^{mut}$ isolates (Tirrell et al. 2019) or $K13^{mut}$ parasites resulting from genome editing to investigate the specific costs associated

with particular mutations (Nair et al. 2018). This can lead to comparative fitness ranks to be drawn between parasite strains and the use of CRISPR-Cas9 gene editing can be hugely informative for assessing if specific mutations, such as $K13^{mut}$, confer benefits that could explain the increased selection observed of this strain. These styles of studies, however, do not allow for any measured fitness cost to be suitable for inferring how these parasite strains differ in their transmission success.

This fitness cost is of significant importance and enables transmission modelling studies to characterise how this fitness cost alters the predicted speed at which resistance may spread. Previous modelling studies have highlighted how the specific predictions made are ultimately heavily dependent on this fitness costs, which has limited the use of transmission modelling in resistance forecasting (Koella and Antia 2003). However, we are able to address this limitation in our study and use the developed transmission model to answer other questions. For example, the simulations conducted can be used to confirm previous modelling predictions that resistance spreads quickest under increased treatment coverage (Figure 6.7). Consequently, although our original hypothesis was

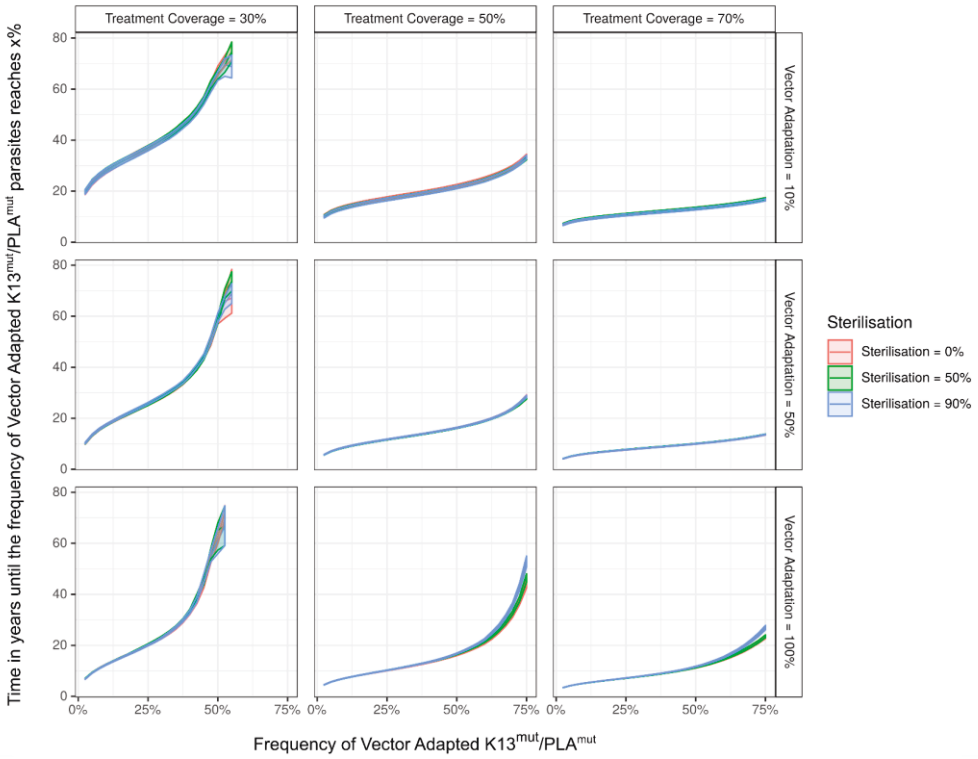


Figure 6.7 Speed of resistance invasion is faster under higher drug pressure. The graph shows the mean and 95% confidence interval of a 100 simulation realisations for the invasion and spread of vector adapted $K13^{mut}/PLA^{mut}$ parasites in a transmission setting within an EIR = 2 and no assumed genetic linkage between K13 and the loci responsible for vector adaptation. The speed at which vector adapted $K13^{mut}/PLA^{mut}$ parasites invade the local population is greatest in areas with higher treatment coverage. In this parameter set, the assumed level of male gametocyte sterilisation and the degree of adaptation of the parasite to the local vector population has only a marginal impact on the speed at which resistance invades, which is observed only in the highest treatment settings.

shown to not be correct from the transmission modelling, the incorporation of results from laboratory experiments within the transmission model has been shown to be a useful tool for testing how a selective advantage at the within host level may not always translate to one within the population.

6.5. Conclusion

In this chapter I have extended the transmission model again, building on the inclusion of resistance in chapter 5 to explore how artemisinin resistance may confer additional selective advantages at the transmission stages of the parasite's life cycle. This extended model was parameterised using results from standard membrane feeding assays conducted by collaborators in the Department of Life Sciences at Imperial College London. Their experimental set up explored the different oocyst intensity and prevalence of two parasite strains (a wild type and a DHA-PPQ resistant strain) in the presence and absence of DHA. I used mixed effect generalised linear models to account for clustering in their experimental replicates, which revealed a significant decrease in oocyst intensity and prevalence for the $K13^{mut}/PLA^{mut}$ parasite in the absence of DHA. This decrease was interpreted as the fitness cost associated with resistance, which was found to be offset by the significantly reduced impact of DHA on $K13^{mut}/PLA^{mut}$ oocyst prevalence compared to the wild type parasite. These findings were incorporated in the transmission model to test the hypothesis that the observed transmission advantage in the presence of DHA results in an increased selective advantage for artemisinin resistance. The transmission modelling largely rejected this hypothesis, however, with marginal impact resulting from male gametocyte sterilisation. This finding is believed to result from the substantially longer duration of infection associated with an asymptomatic infection compared to a treated infection. Consequently, any transmission advantage only occurs within a small time window compared to the more likely route for invasive resistant parasites to spread from longer asymptomatic infections. Despite this "negative" result, I have shown how the developed transmission model can be adapted to address nuanced questions related to the differential transmission of different parasite genetic variants.

Chapter 7. Discussion

The increasing interest in using malaria genetic approaches for operational surveillance will almost certainly lead to genetics being used for malaria control efforts. However, in what form this will take and how far away are we from that point is largely unknown and dependent on your definition of malaria genetics approaches. The identification of the molecular markers of antimalarial resistance has been used to validate clinical observations of treatment failures, which has driven decisions to switch first line antimalarials. These confirmatory uses of malaria genetics are a clear and undeniably valuable tool for assisting malaria control. However, a number of other uses have been suggested in a technical consultation proposed to the Malaria Policy Advisory Committee on the role of parasite genetics in optimising responses by national programmes (Noor and Ringwald 2018). Although malaria genotyping has been shown to be informative of malaria transmission in some settings (Daniels et al. 2015), the harder question is how useful will these approaches be and what else do we need to know before we can rule on the effective uses of parasite genetics.

Mathematical models of malaria transmission models have been used extensively in the last 10 years to evaluate the benefit and use cases for a number of malaria control interventions. Evaluating the role that parasite genetics could have should be no different. However, the complexity of the processes that shape parasite genetic diversity coupled with the recentness of interest in parasite genetics has led to transmission models of malaria genetics being largely unexplored. In this thesis, I have addressed this absence, building new models and applying those models to show the breadth of questions that can be answered using transmission models while also highlighting the areas that warrant further research. In this discussion I summarise the findings of my thesis before reviewing the limitations and future directions of my work on modelling parasite genetics.

7.1. Summary of findings

In chapter 2, I extended my previous model of *pfhp2* gene deletions to address the potential for biased estimates of the prevalence of *pfhrp2/3* deletions to be made when sampling at different times in a transmission season (Watson et al. 2019b). This work was undertaken to highlight that the proposed 8-week interval listed in the WHO guidance on *pfhrp2/3* surveillance to national malarial control programmes should be considered with an awareness of how seasonality in transmission will affect estimates. I predict that highly seasonal settings are more susceptible to systematic variations in the observed prevalence of *pfhrp2/3* deletions. This is due to the increased probability of an individual being only infected with *pfhrp2/3* deleted parasites at the beginning of a transmission season.

Estimates made during this time are likely to overestimate the percentage of cases that would be misdiagnosed throughout the transmission season. This may lead to a premature switch to a less sensitive diagnostic that causes more malaria cases to be misdiagnosed. To highlight this, I created an online database that explores the potential bias for each level one administrative region in sub-Saharan Africa to help guide control programmes implementing the WHO sampling scheme.

In chapter 3, I moved away from the model used for *pfhrp2* deletions. The model of *pfhrp2* deletions was purpose-built for modelling *pfhrp2* deletions and was consequently cumbersome for modelling more nuanced representations of parasite genetics. This chapter introduced an extended version of the Imperial College transmission model that is capable of simulating parasite genetics. Attention was placed on ensuring biological realism throughout and as such many factors omitted from previous modelling efforts were included. These included modelling mosquitoes at the individual level, which allows the full parasite life cycle to be tracked and recombination to be explicitly captured. Parasite densities were modelled indirectly by associating parasite strains with the predicted infection course for the human in the absence of future reinfections. The cotransmission of multiple parasite strains from one infectious bite was included and shown to be necessary when reproducing the relationship between age and the complexity of infection. Using these relationships, the percentage of sporozoites that survive from an infectious bite was estimated to be equal to 20%. With this parameter fitted, the model was used to show how simple population genetic predictions can be made using the model. Lastly, I characterised the contribution of both superinfection and cotransmission events towards parasite genetic diversity within hosts, estimating that in settings with than 11.5% malaria prevalence by PCR, the majority of genetic variation within hosts is generated through superinfection events.

The parameterised model developed in chapter 3 was used to explore how parasite genetics could be used for surveillance purposes. The genetic barcode used to model parasite genotypes allows sufficient neutral genetic variation to be simulated, which allows estimates of multiple genetic metrics to be produced from the model. The relationship between six genetic metrics and malaria prevalence was characterised and the statistical power of each metric for detecting changes in malaria prevalence in response to the scale up of malaria prevalence estimated. To demonstrate how parasite genetics could be operationally used, I built an ensemble statistical model that uses summary parasite genetic metrics to predict current malaria prevalence. This exercise was conducted to show how the differences in the metrics could be combined to accurately predict malaria prevalence, as well as to highlight the importance in ensuring that sample metadata, such as the age and clinical status of the individuals contributing parasite samples, is made available as standard. The exercise is hoped to also delineate how parasite genetics could be used and to guide continued efforts in harnessing parasite genetics for inferring transmission intensity and evaluation intervention campaigns.

In chapter 5, I adapted the transmission model to simulate antimalarial resistance. By changing the information represented in the parasite genetic barcode, the resistance phenotype of parasite strains can be evaluated and used to infer treatment outcomes. The model was used to explore whether current drug treatment strategies can be improved to prevent the continued spread of resistance. I focussed on the spread of resistance, as opposed to the earlier stage of resistance emergence, to explore whether the same conclusions about drug treatment strategies are made. Three drug treatment strategies were explored, comparing multiple first line therapies against the current strategy of cycling first line antimalarials when they begin to fail. I predicted that multiple first line therapies are not superior to cycling at slowing the spread of resistance. Additionally, cycling of antimalarials can be improved by reducing the overprescription of antimalarials through the presumptive treatment of non-malarial fevers. The benefit of increasing diagnostic testing of non-malarial fevers is characterised, predicting a significant benefit conferred when increasing diagnostic testing in settings above an EIR greater than 3 (PCR prevalence ~ 17.5%). Lastly, I demonstrate how multiple first line therapies can be improved by increasing the number of available artemisinin combination therapies. This creates a significant evolutionary barrier to the parasite due to the difficulties in acquiring resistance to multiple drugs and the compounded fitness costs associated with possessing mutations in multiple genes.

In the final results chapter, I continued modelling antimalarial resistance but focussed on the mechanisms by which resistant parasites may invade new populations. This work was in collaboration with a parasitology group at Imperial College investigating the impact of sterilisation of male parasite gametocytes by artemisinin. I used mixed-effects generalised linear models to show that artemisinin significantly reduces the ability of wild type parasites to form oocysts and infect mosquitoes. Artemisinin was found to have a significantly smaller impact on the ability of resistant parasites to form oocysts, which offsets the fitness costs associated with resistance in the absence of artemisinin. These results were incorporated into an extended version of the transmission model developed in chapter 5 to characterise if the transmission advantage conferred to resistant parasites in the presence of artemisinin confers a population level selective advantage. I predict that the reduced impact of artemisinin on the transmission of artemisinin resistant parasites does not confer a significant population-level transmission advantage. This finding was robust to changes in transmission intensity, treatment coverage, the degree of sterilisation by artemisinin and assumptions about local parasites being adapted to the local vector population. The parameter estimates for the fitness costs associated with resistance, however, are an invaluable asset to resistance modelling efforts and can be used in the future for more accurate predictions from resistance modelling studies.

7.2. Limitations and Future Directions

The majority of results presented in the thesis are based on mathematical modelling of the transmission dynamics of malaria parasites. Consequently, the findings need to be viewed in the context of the assumptions and limitations of the model. This critique will also be made in light of the available data for parameterisation efforts, highlighting where the modelling could be developed in the future.

The original model of *pfhrp2* gene deletions has been used to characterise the drivers of selection for *pfhrp2* deletions and was able to reproduce how the frequency of deletions in the Democratic Republic of Congo varies across transmission settings and within a transmission season. However, one area of weakness is the incorporation of *pfhrp3* deletions. In the original study, the role of PfHRP3 epitopes in triggering a positive PfRHP2-based RDT result was included by assuming that false-negative HRP2-based RDTs still occur in 25% of HRP2-/HRP3+, which was based on earlier estimates of the impact of parasite density on HRP3 cross reactivity (Baker et al. 2005). In the last 2 years, the literature surrounding the cross reactivity of HRP3 has developed, suggesting that the 25% chosen may be too low, with evidence that HRP2-/HRP3+ infections >1000 parasites/µL are more likely to yield a positive RDT (Beshir et al. 2017). Moving forward, it seems increasingly likely that future studies will require the frequency of *pfhrp3* gene deletions to be also modelled. Sensible parasite densities will also need to be included to reflect the chance that an individual infected with HRP2-/HRP3+ parasites will yield a positive RDT, with clinical cases caused by HRP2-/HRP3+ parasites almost certain to be detected given the parasite densities usually associated with symptomatic infection being above 1000 parasites/µL (Kitchen 1949). Non-malarial fevers, however, were the largest driver of *pfhrp2* deletions in my earlier study. A recent systematic review of parasite densities would suggest that many asymptomatic infections would fall under 1000 parasites/µL (Slater et al. 2019). Consequently, I would predict that the treatment seeking rates in Africa would cause individuals who are HRP2-/HRP3+ to have only a modest effect on slowing selection of HRP2-. However, it will likely still ensure that symptomatic malaria cases are not misdiagnosed, assuming they have presented at clinic prior to their parasite density falling below 1000 parasites/µL. The model developed in chapter 2 will likely be insufficient to accurately capture all the dynamics above and any future *pfhrp2/3* deletion studies I conduct will likely adapt the model developed in chapters 3 to 6.

Although the transmission model developed in chapter 3 involved considerable extensions to enable parasite genetics to be modelled, there were very few new parameters introduced. This was entirely by design. Considerable effort has been undertaken by researchers at Imperial College to parameterise the model (Griffin et al. 2010, 2014, 2015). Although computational running time of my

model is sufficiently fast, the additional complexities from including individual mosquitoes and parasite populations does hinder computational performance, which would make conducting a full parameter fitting exercise costly. However, it was something that was considered numerous times, most noticeably in reference to the immunity component, ID. The Imperial College transmission model does not explicitly model parasite density, but models how ID alters the probability that an asymptomatic infection will be detected by microscopy. One reason for this is that there was no need to do so, with the model parametrised to malaria prevalence data using microscopy detection as well as clinical incidence data. However, in my model I wanted to include a better representation of the likely parasite densities. This was something I wanted to capture as it shapes many of the results presented in this thesis. For example, in chapter 3, I estimated the percentage of sporozoites that survive to found liver stage infections. This parameter modifies the mean COI by allowing more parasite strains to be introduced. However, translating my model estimated COI to the COI that would be estimated from an equivalent individual in the field is something I am still largely unsure of. For example, below what threshold will a parasite strain not contribute to recorded SNP frequencies from Sequenom genotyping? Will all minor clones be detected, or will some be sequestered at the time of sampling? Will some minor clones be attributed to sequencing error? From my model I am able to observe the true COI, however, it is unclear how this translates to observed COI from methods like *THE REAL McCOIL*. In my model fitting I made different assumptions about whether subpatent strains would be detected given the improved detection sensitivities of molecular methods compared to microscopy. These were then validated by bringing in data from *msp2* based estimates of COI to give increased confidence to the estimated parameter. However, given that I track the time at which a parasite strain is acquired, whether it triggered a symptomatic infection and at what time it will become subpatent I feel the model could be extended to further capture the realism of within host parasite densities.

Within host parasite densities were a common area of uncertainty and it still remains unclear to what extent inter-strain dynamics occur. The model assumes that individuals in state D are unable to be superinfected, due to the short lived interferon response from the current infection preventing immediate superinfection events (Recker et al. 2011). However, how parasites associated with asymptomatic infections both moderate new infections and impact within host parasite densities is largely unknown. Studies showing how asymptomatic infections do not affect the probability of being clinically infected conclude that it is more likely that the specific var gene expression of previous infections is more likely to determine if a new infection will trigger a symptomatic infection (Buchwald et al. 2019). The obvious extension to include this would involve using a section of the parasite barcode to track var type expression and include parameters detailing the rate of antigenic switching, which is

the approach used in the Intellectual Ventures model (Eckhoff 2011b) and other models investigating how antigenic variation is related to transmission intensity (Holding et al. 2018). However, this is a very different model of immunity acquisition to the purely exposure driven immunity that the Imperial College transmission uses and would require immunity to be rewritten. Given the model is still able to capture the relationships between COI and age in chapter 3 this does not seem necessary. Instead, variant exposure could be tracked within an individual, which would describe the parasite var types that have resulted in the current overall acquired immunity and used to infer each parasite strain's density. The inclusion of explicit parasite densities based on antigenic variants would also open the possibility for the model to be extended to include PK-PD modelling when determining clinical outcomes and related parasite density profiles during recrudescent infections. This is similar to a model variant used in OpenMalaria when considering within host parasite dynamics (OpenMalaria #892bd5). In any future work related to antimalarial resistance it would be good to include this approach, at the very least as a sensitivity analysis, to explore how more accurate PK-PD modelling changes conclusions made about the optimum treatment strategies. This would directly tie into the planned systematic review of studies exploring the treatment efficacy of each ACT on known parasite resistant genotypes, or at least in areas with equivalent studies that have measured resistance allele frequencies.

7.3. Implications

The results presented in the thesis have implications for the analysis and surveillance of parasite genetic variants. Most immediately, the results in chapter 2 were conducted in collaboration with members from the WHO. The findings of the study have now been considered in recommendations made to national malaria control programmes that were conveyed during a WHO organised workshop for African malaria control programs hosted in Zambia earlier this year. As more studies estimating the prevalence of *pfhrp2/3* deletions are conducted in the near future, I will be able to evaluate the predictions and update the recommendations made accordingly.

In chapter 3, I estimate the percentage of sporozoites that survive from an infectious bite, which adds to growing evidence that cotransmission of multiple parasite strains from individual bites occurs. Although there has long been evidence that multiple sporozoites are introduced in a bite, studies to assess the genetic relatedness of the sporozoites have only been conducted recently, with one of the first studies looking at polygenomic infections in Senegal (Wong et al. 2017). One reason this finding was suggested so recently is the difficulty in analysing polygenomic infections and being able to phase these samples to identify the individual haplotypes present. In the last 2 years, however, the development of statistical methods that are now capable of estimating COI (Chang et al. 2017),

identity by descent within samples (Schaffner et al. 2018) and phasing polyclonal samples (Zhu et al. 2019) suggests that our understanding of the factors shaping parasite genetic diversity is increasing. These tools will be of particular importance if parasite genetics are to ever be used for operational surveillance and inferring transmission intensity. This is demonstrated in the results of chapter 4, which show the added information that can be gained when polyclonal samples are assumed to be able to be phased, rather than having to rely on calling the major haplotype ore relying solely on monoclonal samples.

The results in chapters 5 have implications for the control of antimalarial resistance. For example, I predict that ensuring diagnostic testing is made available for all non-malarial fevers could lead to up to a 25% decrease in relative treatment failure over a 20-year period. Additionally, I predict that the use of multiple first line therapies does not confer significant benefits over the current strategy of cycling first line therapies for slowing the continued spread of antimalarial resistance. This is the opposite finding to a previous study that compared the two strategies for preventing the emergence of resistance. Emergence represents a different stage of resistance development and as such my results suggest that policy decisions should be made in light of the current resistance profile of a country, i.e. preventing emergence or limiting the spread of existing resistance. I will also be continuing these studies to examine the emergence of resistance in order to give a more complete overview of how treatment strategies should be optimised.

Finally, the results in chapter 6 provide estimates of the fitness costs associated with resistance to dihydroartemisinin-piperaquine. Estimates of fitness costs are a rarity and consequently these results, along with the estimated difference in the impact of artemisinin on wild type and resistant parasite transmission, will be of use to the wider resistance community. These effects, however, were predicted to not impact the selection for resistance at a population level. This exercise highlights the importance of considering experimental results in the wider epidemiological context of an infectious pathogen. Additionally, this chapter provides a demonstration of how malaria transmission models that incorporate parasite genetics can be adapted to answer more nuanced questions about malaria epidemiology.

7.4. Conclusion

The proposed technical consultation to the World Health Organisation on the role of parasite genetics in malaria surveillance will be reported on at one of the next meetings of the Malaria Policy Advisory Committee. The stage is certainly set for malaria parasite genetics - but in what capacity? One area I am confident that parasite genetics will be needed is to address emerging threats of resistance. For example, the guidance proposed by the WHO to survey for the prevalence of *pfhrp2/3* deletions is a necessary study to understand its current spread and guide RDT research and development accordingly. Similarly, phenotype-genotype studies have helped identify molecular markers of emerging resistance, which allow us to understand its spread and predict the next best course of action. In this thesis, I have shown how mathematical models that incorporate parasite genetics are a useful tool for helping to guide these efforts, which have yielded direct impact to current malaria responses. However, I think conclusions about the role that genetics has to play in inferring malaria prevalence will be less clear as there are still many unknowns in the processes shaping neutral genetic variation. The results in this thesis are consequently hoped to contribute to our understanding of these processes and to characterise the areas that warrant further study. Despite these challenges, considerable methodological and statistical methods have been developed in the last two years that have substantially aided our understanding of the genetic diversity of *P. falciparum*. I would thus not be surprised if the outcomes of the consultation recommend parasite genetics to be used for more than just the study of resistance. If they do, I hope this thesis helps demonstrate the role that transmission modelling of parasite genetics should have in these recommendations.

References

Abdallah JF, Okoth SA, Fontecha G a, Torres REM, Banegas EI, Matute ML, et al. Prevalence of pfhrp2 and pfhrp3 gene deletions in Puerto Lempira, Honduras. Malar J. 2015;14(1):19.

Akinyi S, Hayden T, Gamboa D, Torres K, Bendezu J, Abdallah JF, et al. Multiple genetic origins of histidine-rich protein 2 gene deletion in Plasmodium falciparum parasites from Peru. Sci Rep. 2013 Sep 30;3(2797):1–8.

Alker A, Kazadi W, Kutelemeni A, Bloland P, Tshefu A, Meshnick S. Dhfr and dhps genotype and sulfadoxine–pyrimethamine treatment failure in children with falciparum malaria in the Democratic Republic of Congo. Trop Med Int Health. 2008 Dec 1;13:1384–91.

Almelli T, Ndam NT, Ezimegnon S, Alao MJ, Ahouansou C, Sagbo G, et al. Cytoadherence phenotype of Plasmodium falciparum- infected erythrocytes is associated with specific pfemp-1 expression in parasites from children with cerebral malaria. Malar J. 2014;13(1):333.

Alonso P. Antimalarial drug resistance in the Greater Mekong Subregion: How concerned should we be? [Internet]. 2017 [cited 2017 Oct 4]. Available from: http://www.who.int/malaria/media/drug-resistance-greater-mekong-qa/en/

Alonso PL, Eubank S, Ghani A, Hay SI, Sinden R, Smith D, et al. A Research Agenda for Malaria Eradication: Modeling. Plos Med. 2011;8(1).

Alphey L, Benedict M, Bellini R, Clark GG, Dame DA, Service MW, et al. Sterile-Insect Methods for Control of Mosquito-Borne Diseases: An Analysis. Vector-Borne Zoonotic Dis. 2010;10(3):295–311.

Aly ASI, Vaughan AM, Kappe SHI. Malaria Parasite Development in the Mosquito and Infection of the Mammalian Host. Annu Rev Microbiol. 2009;63(1):195–221.

Amato R, Miotto O, Woodrow CJ, Almagro-Garcia J, Sinha I, Campino S, et al. Genomic epidemiology of artemisinin resistant malaria. Elife. 2016;5(MARCH2016):1–29.

Amato R, Pearson RD, Almagro-Garcia J, Amaratunga C, Lim P, Suon S, et al. Origins of the current outbreak of multidrug-resistant malaria in southeast Asia: a retrospective genetic study. Lancet Infect Dis. 2018;18(3):337–45.

Amino R, Thiberge S, Martin B, Celli S, Shorte S, Frischknecht F, et al. Quantitative imaging of Plasmodium transmission from mosquito to mammal. Nat Med. 2006 Feb;12(2):220–4.

Anderson RM, May RM. Infectious diseases of humans: dynamics and control. Infect Dis humans Dyn Control. 1991;vii + 757 pp.

Anderson TJC, Nair S, McDew-White M, Cheeseman IH, Nkhoma S, Bilgic F, et al. Population parameters underlying an ongoing soft sweep in southeast Asian malaria parasites. Mol Biol Evol. 2016;34(1):msw228.

Anderson TJC, Roper C. The origins and spread of antimalarial drug resistance: Lessons for policy makers. Acta Trop. 2005;94(3 SPEC. ISS.):269–80.

Angulo I, Fresno M. Cytokines in the pathogenesis of and protection against malaria. Vol. 9, Clinical and Diagnostic Laboratory Immunology. 2002. p. 1145–52.

Ansari HR, Templeton TJ, Subudhi AK, Ramaprasad A, Tang J, Lu F, et al. Genome-scale comparison of expanded gene families in Plasmodium ovale wallikeri and Plasmodium ovale curtisi with Plasmodium malariae and with other Plasmodium species. Int J Parasitol. 2016;46(11):685–96.

Antao T, Hastings I. Policy options for deploying anti-malarial drugs in endemic countries: a population genetics approach. Malar J. 2012;11:422.

Antao T, Hastings IM. Environmental, pharmacological and genetic influences on the spread of drug-resistant malaria. Proc Biol Sci. 2011;278(1712):1705–12.

Aponte JJ, Schellenberg D, Egan A, Breckenridge A, Carneiro I, Critchley J, et al. Efficacy and safety of intermittent preventive treatment with sulfadoxine-pyrimethamine for malaria in African infants: a pooled analysis of six randomised, placebo-controlled trials. Lancet. 2009;374(9700):1533–42.

Aron JL. Mathematical modeling of immunity to malaria. Math Biosci. 1988;90(1–2):385–96.

Ashley EA, Dhorda M, Fairhurst RM, Amaratunga C, Lim P, Suon S, et al. Spread of Artemisinin Resistance in *Plasmodium falciparum* Malaria. N Engl J Med. 2014;371(5):411–23.

Ashley EA, Phyo AP. Drugs in Development for Malaria. Drugs. 2018 Jun 1;78(9):861–79.

Ashley EA, White NJ. Artemisinin-based combinations. Curr Opin Infect Dis. 2005;18:531–6.

Atiele H, Zouh G, Yaw A, Mwanzo I, Githeko A, Yan G. Insecticide-treated net (ITN) ownership, usage, and malaria transmission in the highlands of western Kenya. Parasites and Vectors. 2011;4(113):1–10.

Attaran A, Barnes KI, Curtis C, d'Alessandro U, Fanello CI, Galinski MR, et al. WHO, the Global Fund, and medical malpractice in malaria treatment. Lancet. 2004;363(9404):237–40.

Aurrecoechea C, Brestelli J, Brunk BP, Dommer J, Fischer S, Gajria B, et al. PlasmoDB: A functional genomic database for malaria parasites. Nucleic Acids Res. 2009;37(SUPPL. 1):539–43.

Ayala FJ. Is sex better? Parasites say "no". Proc Natl Acad Sci U S A. 1998;95(7):3346–8.

Bachmann A, Scholz JAM, Janßen M, Klinkert M-Q, Tannich E, Bruchhaus I, et al. A comparative study of the localization and membrane topology of members of the RIFIN, STEVOR and PfMC-2TM protein families in Plasmodium falciparum-infected erythrocytes. Malar J. 2015;14(1):274.

Baker J, McCarthy J, Gatton M, Kyle DE, Belizario V, Luchavez J, et al. Genetic diversity of Plasmodium falciparum histidine-rich protein 2 (PfHRP2) and its effect on the performance of PfHRP2-based rapid diagnostic tests. J Infect Dis. 2005;192:870–7.

Barry AE, Leliwa-Sytek A, Tavul L, Imrie H, Migot-Nabias F, Brown SM, et al. Population genomics of the immune evasion (var) genes of Plasmodium falciparum. PLoS Pathog. 2007;3(3):1–9.

Baton LA, Ranford-cartwright LC. Ookinete destruction within the mosquito midgut lumen explains Anopheles albimanus refractoriness to Plasmodium falciparum (3D7A) oocyst infection. Int J Parasitol. 2012;42(3):249–58.

Battle KE, Lucas TCD, Nguyen M, Howes RE, Nandi AK, Twohig KA, et al. Mapping the global endemicity and clinical burden of Plasmodium vivax, 2000–17: a spatial and temporal modelling study. Lancet. 2019;394(10195):332–43.

Beardmore RE, Peña-Miller R, Gori F, Iredell J, Barlow M. Antibiotic cycling and antibiotic mixing: Which one best mitigates antibiotic resistance? Mol Biol Evol. 2017;34(4):802–17.

Beier JC, Beier MS, Vaughan JA, Pumpuni CB, Davis JR, Noden BH. Sporozoite transmission by Anopheles freeborni and Anopheles gambiae experimentally infected with Plasmodium falciparum. J Am Mosq Control Assoc. 1992 Dec;8(4):404–8.

Bejon P, Andrews L, Andersen RF, Dunachie S, Webster D, Walther M, et al. Calculation of Liver-to-

Blood Inocula, Parasite Growth Rates, and Preerythrocytic Vaccine Efficacy, from Serial Quantitative Polymerase Chain Reaction Studies of Volunteers Challenged with Malaria Sporozoites. J Infect Dis. 2005;191(4):619–26.

Bejon P, Williams TN, Liljander A, Noor AM, Wambua J, Ogada E, et al. Stable and Unstable Malaria Hotspots in Longitudinal Cohort Studies in Kenya. Smith TA, editor. PLoS Med. 2010 Jul 6;7(7):e1000304.

Bell AS, Huijben S, Paaijmans KP, Sim DG, Chan BHK, Nelson WA, et al. Enhanced transmission of drug-resistant parasites to mosquitoes following drug treatment in rodent malaria. PLoS One. 2012;7(6).

Bell AS, de Roode JC, Sim D, Read AF. Within-Host Competition in Genetically Diverse Malaria Infections: Parasite Virulence and Competitive Success. Evolution (N Y). 2006;60(7):1358.

Bennink S, Kiesow MJ, Pradel G. The development of malaria parasites in the mosquito midgut. Cell Microbiol. 2016;18(7):905–18.

Bergstrom CT, Lo M, Lipsitch M. Ecological theory suggests that antimicrobial cycling will not reduce antimicrobial resistance in hospitals. Proc Natl Acad Sci U S A. 2004;101(36):13285–90.

Berhane A, Anderson K, Mihreteab S, Gresty K, Rogier E, Mohamed S, et al. Major Threat to Malaria Control Programs by *Plasmodium falciparum* Lacking Histidine-Rich Protein 2, Eritrea. Emerg Infect Dis. 2018;24(3):462–70.

Beshir KB, Sepúlveda N, Bharmal J, Robinson A, Mwanguzi J, Busula AO, et al. Plasmodium falciparum parasites with histidine-rich protein 2 (pfhrp2) and pfhrp3 gene deletions in two endemic regions of Kenya. Sci Rep. 2017;7(1):1–10.

Bhatt S, Cameron E, Flaxman SR, Weiss DJ, Smith DL, Gething PW. Improved prediction accuracy for disease risk mapping using Gaussian Process stacked generalisation. 2016;1–33.

Bhatt S, Weiss DJ, Cameron E, Bisanzio D, Mappin B, Dalrymple U, et al. The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. Nature. 2015;526(7572):207–11.

Blagborough a M, Churcher TS, Upton LM, Ghani a C, Gething PW, Sinden RE. Transmission-blocking interventions eliminate malaria from laboratory populations. Nat Commun. 2013;4(1812):1812.

Böhme U, Otto TD, Sanders M, Newbold CI, Berriman M. Progression of the canonical reference malaria parasite genome from 2002–2019 [version 1; peer review: 2 approved, 1 approved with reservations]. Wellcome Open Res. 2019;4:1–22.

Boni MF, Smith DL, Laxminarayan R. Benefits of using multiple first-line therapies against malaria. Proc Natl Acad Sci. 2008;105(37):14216–21.

Bopp SER, Manary MJ, Bright AT, Johnston GL, Dharia N V., Luna FL, et al. Mitotic Evolution of Plasmodium falciparum Shows a Stable Core Genome but Recombination in Antigen Families. PLoS Genet. 2013;9(2):1–15.

Borges S, Cravo P, Creasey A, Fawcett R, Modrzynska K, Rodrigues L, et al. Genomewide scan reveals amplification of mdr1 as a common denominator of resistance to mefloquine, lumefantrine, and artemisinin in Plasmodium chabaudi malaria parasites. Antimicrob Agents Chemother. 2011 Oct;55(10):4858–65.

Bousema J, Drakeley C, Mens P, Arens T, Houben R, Omar S, et al. Increased Plasmodium falciparum gametocyte production in mixed infections with P. malariae. Am J Trop Med Hyg. 2008;78(3):442–8.

Boyd MF, Kitchen SF. The Duration of the Intrinsic Incubation Period in Falciparum Malaria in Relation to Certain Factors Affecting the Parasites 1. Am J Trop Med Hyg. 1937 Nov 1;s1-17(6):845–8.

Bradley J, Stone W, Da DF, Morlais I, Dicko A, Cohuet A, et al. Predicting the likelihood and intensity of mosquito infection from sex specific plasmodium falciparum gametocyte density. Elife. 2018;7:1–13.

Brady OJ, Slater HC, Pemberton-Ross P, Wenger E, Maude RJ, Ghani AC, et al. Role of mass drug administration in elimination of Plasmodium falciparum malaria: a consensus modelling study. Lancet Glob Heal. 2017;5(7):e680–7.

Breiman L. Classification and regression trees. Chapman & Hall/CRC; 1984.

Breiman L. Random forests. Mach Learn. 2001;45(1):5–32.

Bretscher M, Dahal P, Griffin J, Stepniewska K, Ghani A, Guerin P. A pooled analysis of the duration of chemoprophylaxis against malaria after treatment with artesunate-amodiaquine and artemether-lumefantrine. medRxiv. 2019;

Briët OJ, Chitnis N. Effects of changing mosquito host searching behaviour on the cost effectiveness of a mass distribution of long-lasting, insecticidal nets: A modelling study. Malar J. 2013;12(1):1–11.

Briët OJT, Penny MA, Hardy D, Awolola TS, Van Bortel W, Corbel V, et al. Effects of pyrethroid resistance on the cost effectiveness of a mass distribution of long-lasting insecticidal nets: A modelling study. Malar J. 2013;12(1):1–12.

Bruce MC, Alano P, Suthie S, Carter R. Commitment of the malaria parasite plasmodium falciparum to sexual and asexual development. Parasitology. 1990;100(2):191–200.

Bruce MC, Donnelly CA, Alpers MP, Galinski MR, Barnwell JW, Walliker, David, et al. Cross-Species Interactions Between Malaria Parasites in Humans. Science (80-). 2000;287(5454):845–8.

Bruxvoort K, Goodman C, Patrick Kachur S, Schellenberg D. How patients take malaria treatment: A systematic review of the literature on adherence to antimalarial drugs. PLoS One. 2014;9(1).

Buchwald AG, Sixpence A, Chimenya M, Damson M, Sorkin JD, Wilson ML, et al. Clinical implications of asymptomatic plasmodium falciparum infections in Malawi. Clin Infect Dis. 2019 Jan 1;68(1):106–12.

Bull PC, Marsh K. The role of antibodies to Plasmodium falciparum-infected-erythrocyte surface antigens in naturally acquired immunity to malaria. Vol. 10, Trends in Microbiology. Elsevier Ltd; 2002. p. 55–8.

Burkot TR. Non-Random Host Selection by Anopheline Mosquitoes. Parasitol Today. 1988;4(6):156–62.

Burnham KP, Anderson DR, Burnham KP. Model selection and multimodel inference : a practical information-theoretic approach. 2nd ed. Springer-Verlag; 2002.

Burt A. Site-specific selfish genes as tools for the control and genetic engineering of natural populations. Proc R Soc B Biol Sci. 2003;270(1518):921–8.

Bushman M, Antia R, Udhayakumar V, de Roode JC. Within-host competition can delay evolution of drug resistance in malaria. PLoS Biol. 2018;16(8):1–25.

Bushman M, Morton L, Duah N, Quashie N, Abuaku B, Koram KA, et al. Within-host competition and drug resistance in the human malaria parasite *Plasmodium falciparum*. Proc R Soc B Biol Sci.

2016;283(1826):20153038.

Cairns M, Ghani A, Okell L, Gosling R, Carneiro I, Anto F, et al. Modelling the protective efficacy of alternative delivery schedules for intermittent preventive treatment of malaria in infants and children. PLoS One. 2011;6(4).

Cairns M, Roca-Feltrer A, Garske T, Wilson AL, Diallo D, Milligan PJ, et al. Estimating the potential public health impact of seasonal malaria chemoprevention in African children. Nat Commun. 2012;3:1–9.

Cairns ME, Walker PGT, Okell LC, Griffin JT, Garske T, Asante KP, et al. Seasonality in malaria transmission: Implications for case-management with long-acting artemisinin combination therapy in sub-Saharan Africa. Malar J. 2015;14(1):1–13.

Calderaro A, Piccolo G, Gorrini C, Rossi S, Montecchini S, Dell'Anna ML, et al. Accurate identification of the six human Plasmodium spp. causing imported malaria, including Plasmodium ovale wallikeri and Plasmodium knowlesi. Malar J. 2013 Sep 13;12:321.

Cameron E, Battle KE, Bhatt S, Weiss DJ, Bisanzio D, Mappin B, et al. Defining the relationship between infection prevalence and clinical incidence of Plasmodium falciparum malaria. Nat Commun. 2015;6:1–10.

Carlton JM, Angiuoli S V., Suh BB, Kooij TW, Pertea M, Silva JC, et al. Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii. Nature. 2002;419(6906):512–9.

Carneiro IA, Smith T, Lusingu JPA, Malima R, Utzinger J, Drakeley CJ. Modeling the relationship between the population prevalence of Plasmodium falciparum malaria and anemia. Am J Trop Med Hyg. 2006;75(2):82–9.

Carrara VI, Sirilak S, Thonglairuam J, Rojanawatsirivet C, Proux S, Gilbos V, et al. Deployment of early diagnosis and mefloquine-artesunate treatment of falciparum malaria in Thailand: The Tak Malaria Initiative. PLoS Med. 2006;3(6):0856–64.

Carrel M, Patel J, Taylor SM, Janko M, Mwandagalirwa MK, Tshefu AK, et al. The geography of malaria genetics in the Democratic Republic of Congo: A complex and fragmented landscape. Soc Sci Med. 2015 May;133:233–41.

Carter LM, Kafsack BFC, Llinás M, Mideo N, Pollitt LC, Reece SE. Stress and sex in malaria parasites: Why does commitment vary? Evol Med Public Heal. 2013;2013(1):135–47.

de Carvalho LP, Sandri TL, José Tenório de Melo E, Fendel R, Kremsner PG, Mordmüller B, et al. Ivermectin Impairs the Development of Sexual and Asexual Stages of Plasmodium falciparumIn Vitro. Antimicrob Agents Chemother. 2019 Aug 1;63(8):e00085-19.

Centers for Disease Control and Prevention [Public Domain]. Malaria Lifecycle [Internet]. [cited 2019 Sep 7]. Available from: https://www.cdc.gov/malaria/about/biology/

Cerqueira GC, Cheeseman IH, Schaffner SF, Nair S, McDew-White M, Phyo AP, et al. Longitudinal genomic surveillance of Plasmodium falciparum malaria parasites reveals complex genomic architecture of emerging artemisinin resistance. Genome Biol. 2017;18(1):1–13.

Challenger JD, Bruxvoort K, Ghani AC, Okell LC. Assessing the impact of imperfect adherence to artemether-lumefantrine on malaria treatment outcomes using within-host modelling. Nat Commun. 2017;8(1).

Chang H-H, Childs LM, Buckee CO. Variation in infection length and superinfection enhance selection efficiency in the human malaria parasite. Sci Rep. 2016 May 19;6(April):26370.

Chang H-H, Moss EL, Park DJ, Ndiaye D, Mboup S, Volkman SK, et al. Malaria life cycle intensifies both natural selection and random genetic drift. Proc Natl Acad Sci. 2013;110(50):20129–34.

Chang H-H, Wesolowski A, Sinha I, Jacob CG, Mahmud A, Uddin D, et al. Mapping imported malaria in Bangladesh using parasite genetic and human mobility data. Ferguson NM, Wenger EA, Brady O, editors. Elife. 2019;8:e43481.

Chang H-H, Worby CJ, Yeka A, Nankabirwa J, Kamya MR, Staedke SG, et al. THE REAL McCOIL: A method for the concurrent estimation of the complexity of infection and SNP allele frequency for malaria parasites. PLOS Comput Biol. 2017;13(1):e1005348.

Charchuk R, Paul MKJ, Claude KM, Houston S, Hawkes MT. Burden of malaria is higher among children in an internal displacement camp compared to a neighbouring village in the Democratic Republic of the Congo. Malar J. 2016;15(1):431.

Cheeseman IH, Miller BA, Nair S, Nkhoma S, Tan A, Tan JC, et al. A Major Genome Region Underlying Artemisinin Resistance in Malaria. Science (80-). 2012;336(April):79–82.

Chenet SM, Taylor JE, Blair S, Zuluaga L, Escalante AA. Longitudinal analysis of Plasmodium falciparum genetic variation in Turbo, Colombia: implications for malaria control and elimination. Malar J. 2015;14(1):363.

Cheng Q, Gatton ML, Barnwell J, Chiodini P, McCarthy J, Bell D, et al. Plasmodium falciparum parasites lacking histidine-rich protein 2 and 3: a review and recommendations for accurate reporting. Malar J. 2014;13(1):283.

Childs LM, Abuelezam NN, Dye C, Gupta S, Murray MB, Williams BG, et al. Modelling challenges in context: Lessons from malaria, HIV, and tuberculosis. Epidemics. 2015;10:102–7.

Chiodini PL, Bowers K, Jorgensen P, Barnwell JW, Grady KK, Luchavez J, et al. The heat stability of Plasmodium lactate dehydrogenase-based and histidine-rich protein 2-based malaria rapid diagnostic tests. Trans R Soc Trop Med Hyg. 2007 Apr 1;101(4):331–7.

Chitnis N, Schapira A, Smith T, Steketee R. Comparing the effectiveness of malaria vector-control interventions through a mathematical model. Am J Trop Med Hyg. 2010;83(2):230–40.

Chotsiri P, Wattanakul T, Hoglund RM, Hanboonkunupakarn B, Pukrittayakamee S, Blessborn D, et al. Population pharmacokinetics and electrocardiographic effects of dihydroartemisinin–piperaquine in healthy volunteers. Br J Clin Pharmacol. 2017;83(12):2752–66.

Churcher TS, Blagborough AM, Delves M, Ramakrishnan C, Kapulu MC, Williams AR, et al. Measuring the blockade of malaria transmission - An analysis of the Standard Membrane Feeding Assay. Int J Parasitol. 2012a;42(11):1037–44.

Churcher TS, Blagborough AM, Delves M, Ramakrishnan C, Kapulu MC, Williams AR, et al. Measuring the blockade of malaria transmission - An analysis of the Standard Membrane Feeding Assay. Int J Parasitol. 2012b;42(11):1037–44.

Churcher TS, Bousema T, Walker M, Drakeley C, Schneider P, Ouédraogo AL, et al. Predicting mosquito infection from Plasmodium falciparum gametocyte density and estimating the reservoir of infection. Elife. 2013;2013(2):1–12.

Churcher TS, Dawes EJ, Sinden RE, Christophides GK, Koella JC, Basanez MG, et al. Population biology of malaria within the mosquito: density-dependent processes and potential implications for transmission-blocking interventions. Malar J. 2010;9(1):311.

Cohen JM, Moonen B, Snow RW, Smith DL. How absolute is zero? An evaluation of historical and current definitions of malaria elimination. Malar J. 2010;9(1):1–13.

Cohen JM, Smith DL, Cotter C, Ward A, Yamey G, Sabot OJ, et al. Malaria resurgence: a systematic review and assessment of its causes. Malar J. 2012a;11(1):122.

Cohen JM, Woolsey AM, Sabot OJ, Gething PW, Tatem AJ, Moonen B. Optimizing Investments in Malaria Treatment and Diagnosis. Science (80-). 2012b;338(6107):612–4.

Conway DJ, Roper C, Oduola a M, Arnot DE, Kremsner PG, Grobusch MP, et al. High recombination rate in natural populations of Plasmodium falciparum. Proc Natl Acad Sci U S A. 1999;96(8):4506–11.

Cook J, Kleinschmidt I, Schwabe C, Nseng G, Bousema T, Corran PH, et al. Serological markers suggest heterogeneity of effectiveness of malaria control interventions on Bioko Island, Equatorial Guinea. PLoS One. 2011;6(9):1–9.

Cooper DJ, Rajahram GS, William T, Jelip J, Mohammad R, Benedict J, et al. Plasmodium knowlesi Malaria in Sabah, Malaysia, 2015–2017: Ongoing Increase in Incidence Despite Near-elimination of the Human-only Plasmodium Species. Clin Infect Dis. 2019;(Xx):1–7.

Cox FEG. History of the discovery of the malaria parasites and their vectors. Parasit Vectors. 2010;3.

Craig MH, Snow RW, le Sueur D. A climate-based distribution model of malaria transmission in sub-Saharan Africa. Parasitol Today. 1999;15(3):105–11.

Crowell V, Briët OJ, Hardy D, Chitnis N, Maire N, Pasquale A Di, et al. Modelling the cost-effectiveness of mass screening and treatment for reducing Plasmodium falciparum malaria burden. Malar J. 2013;12(1):1–14.

D'Alessandro U, Buttiëns H. History and importance of antimalarial drug resistance. Trop Med Int Heal. 2001;6(11):845–8.

Dalrymple U, Cameron E, Bhatt S, Weiss DJ, Gupta S, Gething PW. Quantifying the contribution of plasmodium falciparum malaria to febrile illness amongst african children. Elife. 2017;6:1–17.

Daniels R, Chang HH, Séne PD, Park DC, Neafsey DE, Schaffner SF, et al. Genetic Surveillance Detects Both Clonal and Epidemic Transmission of Malaria following Enhanced Intervention in Senegal. PLoS One. 2013;8(4):4–10.

Daniels R, Volkman SK, Milner DA, Mahesh N, Neafsey DE, Park DJ, et al. A general SNP-based molecular barcode for Plasmodium falciparum identification and tracking. Malar J. 2008;7:223.

Daniels RF, Schaffner SF, Wenger EA, Proctor JL, Chang H-H, Wong W, et al. Modeling malaria genomics reveals transmission decline and rebound in Senegal. Proc Natl Acad Sci U S A. 2015;112(22):7067–72.

Das D, Phyo AP, Tarning J, Ph D, Lwin KM, Ariey F, et al. Artemisinin Resistance in Plasmodium falciparum Malaria. N Engl J Med. 2009;361(5):455–67.

Davis CS. The computer generation of multinomial random variates. Comput Stat Data Anal. 1993;16(2):205–17.

Davis TME, Hung T-Y, Sim I-K, Karunajeewa HA, Ilett KF. Piperaquine: a resurgent antimalarial drug. Drugs. 2005;65(1):75–87.

Day T, Read AF. Does high-dose antimicrobial chemotherapy prevent the evolution of resistance? PLoS Comput Biol. 2016;12(1):1–20.

Deane-Mayer ZA, Knowles JE. caretEnsemble: Ensembles of Caret Models. R package version 2.0.0. 2016.

Delves M, Plouffe D, Scheurer C, Meister S, Wittlin S, Winzeler EA, et al. The activities of current antimalarial drugs on the life cycle stages of plasmodium: A comparative study with human and rodent parasites. PLoS Med. 2012;9(2).

Delves MJ, Straschil U, Ruecker A, Miguel-Blanco C, Marques S, Baum J, et al. Routine in vitro culture of P. falciparum gametocytes to evaluate novel transmission-blocking interventions. Nat Protoc. 2016;11(9):1968–80.

Demographic and Health Surveys Program. MIS overview [Internet]. 2018 [cited 2018 Feb 19]. Available from: https://dhsprogram.com/what-we-do/survey-types/mis.cfm

Devine GJ, Perea EZ, Killeen GF, Stancil JD, Clark SJ, Morrison AC. Using adult mosquitoes to transfer insecticides to Aedes aegypti larval habitats. Proc Natl Acad Sci U S A. 2009;106(28):11530–4.

Dietz K, Raddatz G, Molineaux L. Mathematical model of the first wave of Plasmodium falciparum asexual parasitemia in non-immune and vaccinated individuals. Am J Trop Med Hyg. 2006;75(2):46–55.

Dietz K, Wernsdorfer WH, McGregor I. Mathematical models for transmission and control of malaria. Malaria Princ Pract Malariol Vol 2. 1988;1091–133.

Ding XC, Ubben D, Wells TNC. A framework for assessing the risk of resistance for anti-malarials in development. Malar J. 2012;11(1):292.

Dokomajilar C, Nsobya SL, Greenhouse B, Rosenthal PJ, Dorsey G. Selection of Plasmodium falciparum pfmdr1 alleles following therapy with artemether-lumefantrine in an area of Uganda where malaria is highly endemic. Antimicrob Agents Chemother. 2006 May;50(5):1893–5.

Dondorp AM, Yeung S, White L, Nguon C, Day NPJ, Socheat D, et al. Artemisinin resistance: Current status and scenarios for containment. Nat Rev Microbiol. 2010;8(4):272–80.

Donnelly MJ, McCall PJ, Lengeler C, Bates I, D'Alessandro U, Barnish G, et al. Malaria and urbanization in sub-Saharan Africa. Malar J. 2005;4.

Doolan DL, Dobaño C, Baird JK. Acquired immunity to Malaria. Vol. 22, Clinical Microbiology Reviews. 2009. p. 13–36.

Durrhelm DN, Becker PJ, Billinghurst K, Brink A. Diagnostic disagreement--the lessons learnt from malaria diagnosis in Mpumalanga. South African Med J. 1997 May;87(5):609–11.

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(1):69–77.

Eckhoff PA. A malaria transmission-directed model of mosquito life cycle and ecology. Malar J. 2011a;10(1):303.

Eckhoff PA. A malaria transmission-directed model of mosquito life cycle and ecology. Malar J. 2011b;10(303).

Eckhoff PA. Malaria parasite diversity and transmission intensity affect development of parasitological immunity in a mathematical model. Malar J. 2012;11:1–14.

Eckhoff PA, Wenger EA, Godfray HCJ, Burt A. Impact of mosquito gene drive on malaria elimination in a computational model with explicit spatial and temporal dynamics. Proc Natl Acad Sci U S A. 2016 Dec 27;114(2):E255-264.

Enserink M. Is setting a deadline for eradicating malaria a good idea? Scientists are divided [Internet]. Science. 2019 [cited 2019 Nov 12]. Available from:

https://www.sciencemag.org/news/2019/08/setting-deadline-eradicating-malaria-good-idea-scientists-are-divided

Eyles DE, Young MD. The duration of untreated or inadequately treated Plasmodium falciparum infections in the human host. Journal Natl Malar Soc. 1951 Dec;10(4):327–36.

Fairhurst RM. Understanding artemisinin-resistant malaria: what a difference a year makes. Curr Opin Infect Dis 2015. 2015;28(5):417–25.

Feachem RGA, Chen I, Akbari O, Bertozzi-Villa A, Bhatt S, Binka F, et al. Malaria eradication within a generation: ambitious, achievable, and necessary. Lancet. 2019 Sep 11;394(10203):1056–112.

Ferreira MU, da Silva Nunes M, Wunderlich G. Antigenic diversity and immune evasion by malaria parasites. Clin Diagn Lab Immunol. 2004 Nov;11(6):987–95.

Fidock DA, Nomura T, Talley AK, Cooper RA, Dzekunov SM, Ferdig MT, et al. Mutations in the P. falciparum digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. Mol Cell. 2000 Oct;6(4):861–71.

Foy BD, Alout H, Seaman JA, Rao S, Magalhaes T, Wade M, et al. Efficacy and risk of harms of repeat ivermectin mass drug administrations for control of malaria (RIMDAMAL): a cluster-randomised trial. Lancet. 2019;393(10180):1517–26.

Frech C, Chen N. Variant surface antigens of malaria parasites: Functional and evolutionary insights from comparative gene family classification and analysis. BMC Genomics. 2013;14(1):1–23.

Friedman JH. Greedy function approximation: A gradient boosting machine. Statistics (Ber). 2011;29(5):1189–232.

Galinsky K, Valim C, Salmier A, de Thoisy B, Musset L, Legrand E, et al. COIL: a methodology for evaluating malarial complexity of infection using likelihood from single nucleotide polymorphism data. Malar J. 2015;14(1):4.

Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, et al. Genome sequence of the human malaria parasite Plasmodium falciparum. Nature. 2002;419(6906):498–511.

Garske T, Ferguson NM, Ghani AC. Estimating Air Temperature and Its Influence on Malaria Transmission across Africa. PLoS One. 2013;8(2).

Gatton ML, Dunn J, Chaudhry A, Ciketic S, Cunningham J, Cheng Q. Implications of parasites lacking plasmodium falciparum histidine-rich protein 2 on Malaria morbidity and control when rapid diagnostic tests are used for diagnosis. J Infect Dis. 2017;215(7):1156–66.

Gatton ML, Hogarth W, Saul A. Time of treatment influences the appearance of drug-resistant parasites in Plasmodium falciparum infections. Parasitology. 2003/02/06. 2001;123(6):537–46.

Gelman A, Rubin DB. Markov chain Monte Carlo methods in biostatistics. Stat Methods Med Res. 1996 Dec 2;5(4):339–55.

Gerald N, Mahajan B, Kumar S. Mitosis in the human malaria parasite plasmodium falciparum. Eukaryot Cell. 2011;10(4):474–82.

Gerardin J, Eckhoff P, Wenger EA. Mass campaigns with antimalarial drugs: a modelling comparison of artemether-lumefantrine and DHA-piperaquine with and without primaquine as tools for malaria control and elimination. BMC Infect Dis. 2015a;15(1):144.

Gerardin J, Ouédraogo AL, Mccarthy KA, Eckhoff PA, Wenger EA. Characterization of the infectious reservoir of malaria with an agent-based model calibrated to age-stratified parasite densities and

infectiousness. Malar J. 2015b;14(231).

Gething PW, Van Boeckel TP, Smith DL, Guerra CA, Patil AP, Snow RW, et al. Modelling the global constraints of temperature on transmission of Plasmodium falciparum and P. vivax. Parasit Vectors. 2011a;4.

Gething PW, Patil AP, Smith DL, Guerra C a., Elyazar IRF, Johnston GL, et al. A new world malaria map: Plasmodium falciparum endemicity in 2010. Malar J. 2011b;10(1):378.

Gibson FD. The efficiency of sporozoite transmission in the human malarias , Plasmodium falciparum and P . vivax *. Bull World Heal Org. 1987;65(3):375–80.

Gimnig JE, Walker ED, Otieno P, Kosgei J, Olang G, Ombok M, et al. Incidence of malaria among mosquito collectors conducting human landing catches in western kenya. Am J Trop Med Hyg. 2013;88(2):301–8.

Gnanguenon V, Azondekon R, Oke-Agbo F, Beach R, Akogbeto M. Durability assessment results suggest a serviceable life of two, rather than three, years for the current long-lasting insecticidal (mosquito) net (LLIN) intervention in Benin. BMC Infect Dis. 2014;14(1):1–10.

Gog JR, Pellis L, Wood JLN, McLean AR, Arinaminpathy N, Lloyd-Smith JO. Seven challenges in modeling pathogen dynamics within-host and across scales. Epidemics. 2015;10:45–8.

Gosling RD, Okell L, Mosha J, Chandramohan D. The role of antimalarial treatment in the elimination of malaria. Clin Microbiol Infect. 2011;17(11):1617–23.

Govella NJ, Chaki PP, Geissbuhler Y, Kannady K, Okumu F, Charlwood JD, et al. A new tent trap for sampling exophagic and endophagic members of the Anopheles gambiae complex. Malar J. 2009;8(1):1–12.

Grandesso F, Nabasumba C, Nyehangane D, Page AL, Bastard M, De Smet M, et al. Performance and time to become negative after treatment of three malaria rapid diagnostic tests in low and high malaria transmission settings. Malar J. 2016;15(1):1–12.

Greenhouse B, Smith DL. Malaria genotyping for epidemiologic surveillance. Proc Natl Acad Sci. 2015;112(22):6782–3.

Greenwood B. Review: Intermittent preventive treatment - a new approach to the prevention of malaria in children in areas with seasonal malaria transmission. Trop Med Int Heal. 2006;11(7):983–91.

Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, Mumford JA, et al. Unifying the epidemiological and evolutionary dynamics of pathogens. Science. 2004;303(5656):327–32.

Griffin JT, Bhatt S, Sinka ME, Gething PW, Lynch M, Patouillard E, et al. Potential for reduction of burden and local elimination of malaria by reducing Plasmodium falciparum malaria transmission: a mathematical modelling study. Lancet Infect Dis. 2016;3099(15):1–8.

Griffin JT, Ferguson NM, Ghani AC. Estimates of the changing age-burden of Plasmodium falciparum malaria disease in sub-Saharan Africa. Nat Commun. 2014;5(3136).

Griffin JT, Hollingsworth TD, Okell LC, Churcher TS, White M, Hinsley W, et al. Reducing Plasmodium falciparum Malaria Transmission in Africa: A Model-Based Evaluation of Intervention Strategies. Krishna S, editor. PLoS Med. 2010 Aug 10;7(8):e1000324.

Griffin JT, Hollingsworth TD, Reyburn H, Drakeley CJ, Riley EM, Ghani AC. Gradual acquisition of immunity to severe malaria with increasing exposure. Proc R Soc B Biol Sci. 2015;282:20142657.

Group IAS. Artesunate combinations for treatment of malaria: Meta-analysis. Lancet. 2004;363(9402):9–17.

Guerra CA, Gikandi PW, Tatem AJ, Noor AM, Smith DL, Hay SI, et al. The limits and intensity of Plasmodium falciparum transmission: implications for malaria control and elimination worldwide. PLoS Med. 2008 Mar;5(2):e38.

Guerra CA, Snow RW, Hay SI. Mapping the global extent of malaria in 2005. Trends Parasitol. 2006 Aug;22(8):353–8.

Guilbride DL, Guilbride PDL, Gawlinski P. Malaria's deadly secret: a skin stage. Trends Parasitol. 2012 Apr;28(4):142–50.

Gupta S, Swinton J, Anderson RM. Theoretical Studies of the Effects of Heterogeneity in the Parasite Population on the Transmission Dynamics of Malaria. Proc R Soc London Ser B-Biological Sci. 1994;256(1347):231–8.

Gupta V, Dorsey G, Hubbard AE, Rosenthal PJ, Greenhouse B. Gel versus capillary electrophoresis genotyping for categorizing treatment outcomes in two anti-malarial trials in Uganda. Malar J. 2010;9(1):1–8.

Haldane JBS. The combination of linkage values and calculation of distance between the loci of linked factors. J Genet. 1919;8:299–309.

Hall M, Woolhouse M, Rambaut A. Epidemic Reconstruction in a Phylogenetics Framework: Transmission Trees as Partitions of the Node Set. PLoS Comput Biol. 2015;11(12):e1004613.

Hamilton WL, Amato R, van der Pluijm RW, Jacob CG, Quang HH, Thuy-Nhien NT, et al. Evolution and expansion of multidrug-resistant malaria in southeast Asia: a genomic epidemiology study. Lancet Infect Dis. 2019;3099(19):10–5.

Hasibeder G, Dye C. Population dynamics of mosquito-borne disease: Persistence in a completely heterogeneous environment. Theor Popul Biol. 1988;33(1):31–53.

Hastie T, Tibshirani R, Friedmna J. The Elements of Statistical Learning. 2nd ed. Springer; 2009.

Hastings IM. A model for the origins and spread of drug-resistant malaria. Parasitology. 1997 Aug;115(Pt 2):133–41.

Hastings IM, Hodel EM. Pharmacological considerations in the design of anti-malarial drug combination therapies - Is matching half-lives enough? Malar J. 2014;13(1):1–15.

Hastings IM, Hodel EM, Kay K. Quantifying the pharmacology of antimalarial drug combination therapy. Sci Rep. 2016;6.

Hay SI, Guerra CA, Gething PW, Patil AP, Tatem AJ, Noor AM, et al. A world malaria map: plasmodium falciparum endemicity in 2007. PLoS Med. 2009;6(3):0286–302.

Hay SI, Sinka ME, Okara RM, Kabaria CW, Mbithi PM, Tago CC, et al. Developing Global Maps of the Dominant Anopheles Vectors of Human Malaria. Plos Med. 2010;7(2).

Hay SI, Smith DL, Snow RW. Measuring malaria endemicity from intense to interrupted transmission. Lancet Infect Dis. 2008/04/02. 2008 Jun;8(6):369–78.

Hellewell J, Walker P, Ghani A, Rao B, Churcher TS. Using ante-natal clinic prevalence data to monitor temporal changes in malaria incidence in a humanitarian setting in the Democratic Republic of Congo. Malar J. 2018;17(1):1–11.

Hemingway J, Shretta R, Wells TNC, Bell D, Djimdé AA, Achee N, et al. Tools and Strategies for Malaria

Control and Elimination: What Do We Need to Achieve a Grand Convergence in Malaria? PLoS Biol. 2016;14(3):1–14.

Hoffman SL, Bancroft WH, Gottlieb M, James SL, Burroughs EC, Stephenson JR, et al. Funding for malaria genome sequencing. Nature. 1997;387(6634):647.

Hofmann N, Mwingira F, Shekalaghe S, Robinson LJ, Mueller I, Felger I. Ultra-Sensitive Detection of Plasmodium falciparum by Amplification of Multi-Copy Subtelomeric Targets. PLoS Med. 2015;12(3).

Holding T, Valletta JJ, Recker M. Multiscale immune selection and the transmission-diversity feedback in antigenically diverse pathogen systems. Am Nat. 2018 Dec 1;192(6):E189–201.

Holtz TH, Kachur SP, Roberts JM, Marum LH, Mkandala C, Chizani N, et al. Use of antenatal care services and intermittent preventive treatment for malaria among pregnant women in Blantyre District, Malawi. Trop Med Int Heal. 2004;9(1):77–82.

Hott A, Tucker MS, Casandra D, Sparks K, Kyle DE. Fitness of artemisinin-resistant Plasmodium falciparum in vitro. J Antimicrob Chemother. 2015 Oct 1;70(10):2787–96.

Houwen B. Blood film preparation and staining procedures. Clin Lab Med. 2002 Mar 1;22(1):1–14.

Howard S, Omumbo J, Nevill C, Some E, Donnelly C, Snow R. Evidence for a mass community effect of insecticide-treated bednets on the incidence of malaria on the Kenyan coast. Vol. 94, Transactions of the Royal Society of Tropical Medicine and Hygiene. London :; 2000. p. 357–60.

Huijben S, Sim DG, Nelson WA, Read AF. The fitness of drug-resistant malaria parasites in a rodent model: Multiplicity of infection. J Evol Biol. 2011;24(11):2410–22.

Humphreys GS, Merinopoulos I, Ahmed J, Whitty CJM, Mutabingwa TK, Sutherland CJ, et al. Amodiaquine and artemether-lumefantrine select distinct alleles of the Plasmodium falciparum mdr1 gene in Tanzanian children treated for uncomplicated malaria. Antimicrob Agents Chemother. 2007 Mar;51(3):991–7.

Imwong M, Hien TT, Thuy-Nhien NT, Dondorp AM, White NJ. Spread of a single multidrug resistant malaria parasite lineage (PfPailin) to Vietnam. Lancet Infect Dis. 2017a;17(10):1022–3.

Imwong M, Suwannasin K, Kunasol C, Sutawong K, Mayxay M, Rekol H, et al. The spread of artemisinin-resistant Plasmodium falciparum in the Greater Mekong Subregion: a molecular epidemiology observational study. Lancet Infect Dis. 2017b;17(5):491–7.

Iqbal J, Khalid N, Hira PR. Comparison of two commercial assays with expert microscopy for confirmation of symptomatically diagnosed malaria. J Clin Microbiol. 2002;40(12):4675–8.

Jakubowski A, Stearns SC, Kruk ME, Angeles G, Thirumurthy H. The US President's Malaria Initiative and under-5 child mortality in sub-Saharan Africa: A difference-in-differences analysis. PLoS Med. 2017;14(6):1–20.

Jamsen KM, Duffull SB, Tarning J, Lindegardh N, White N, Simpson JA. Optimal designs for population pharmacokinetic studies of oral artesunate in patients with uncomplicated falciparum malaria. Malar J. 2011;10(1):181.

Janse CJ, Van der Klooster PFJ, Van der Kaay HJ, Van der Ploeg M, Overdulve JP. Rapid repeated DNA replication during microgametogenesis and DNA synthesis in young zygotes of Plasmodium berghei. Trans R Soc Trop Med Hyg. 1986 Jan 1;0(1):154–7.

Jiang H, Li N, Gopalan V, Zilversmit MM, Varma S, Nagarajan V, et al. High recombination rates and hotspots in a Plasmodium falciparum genetic cross. Genome Biol. 2011;12(4):R33.

Johnston GL, Gething PW, Hay SI, Smith DL, Fidock DA. Modeling Within-Host Effects of Drugs on Plasmodium falciparum Transmission and Prospects for Malaria Elimination. PLoS Comput Biol. 2014;10(1).

Johnston SP, Pieniazek NJ, Xayavong M V., Slemenda SB, Wilkins PP, Da Silva AJ. PCR as a confirmatory technique for laboratory diagnosis of malaria. J Clin Microbiol. 2006;44(3):1087–9.

Kabiru EW, Mbogo CM, Muiruri SK, Ouma JH, Githure JI, Beier JC. Sporozoite loads of naturally infected Anopheles in Kilifi District, Kenya. J Am Mosq Control Assoc. 1997 Sep;13(3):259–62.

Kamya MR, Arinaitwe E, Wanzira H, Katureebe A, Barusya C, Kigozi SP, et al. Malaria transmission, infection, and disease at three sites with varied transmission intensity in Uganda: Implications for malaria control. Am J Trop Med Hyg. 2015;92(5):903–12.

Karl S, White MT, Milne GJ, Gurarie D, Hay SI, Barry AE, et al. Spatial effects on the multiplicity of Plasmodium falciparum infections. PLoS One. 2016;11(10):1–20.

Kay K, Hodel EM, Hastings IM. Altering antimalarial drug regimens may dramatically enhance and restore drug effectiveness. Antimicrob Agents Chemother. 2015;59(10):6419–27.

Ken-Dror G, Hastings IM. Markov chain Monte Carlo and expectation maximization approaches for estimation of haplotype frequencies for multiply infected human blood samples. Malar J. 2016;15(1):430.

Keoluangkhot V, Green MD, Nyadong L, Fernández FM, Mayxay M, Newton PN. Impaired clinical response in a patient with uncomplicated falciparum malaria who received poor-quality and underdosed intramuscular artemether. Am J Trop Med Hyg. 2008;78(4):552–5.

Kilian A, Byamukama W, Pigeon O, Gimnig J, Atieli F, Koekemoer L, et al. Evidence for a useful life of more than three years for a polyester-based long-lasting insecticidal mosquito net in Western Uganda. Malar J. 2011;10.

Killeen GF, Fillinger U, Knols BGJ. Advantages of larval control for African malaria vectors: low mobility and behavioural responsiveness of immature mosquito stages allow high effective coverage. Malar J. 2002;1:8.

Killeen GF, Ross A, Smith T. Infectiousness of malaria-endemic human populations to vectors. Am J Trop Med Hyg. 2006;75(2):38–45.

Kim Y, Escalante AA, Schneider KA. A population genetic model for the initial spread of partially resistant malaria parasites under anti-malarial combination therapy and weak intrahost competition. PLoS One. 2014;9(7):19–25.

Kingman JFC. Origins of the Coalescent: 1974–1982. Genetics. 2000 Dec 1;156(4):1461–3.

Kirchner S, Power BJ, Waters AP. Recent advances in malaria genomics and epigenomics. Genome Med. 2016;8(1):1–17.

Kitchen SF. Symptomatology: general considerations and falciparum malaria. In: Boyd MF, editor. Malariology. Philadelphia: WB Saunders. London: WB Saunders; 1949. p. 996–1017.

Klein EY, Smith DL, Laxminarayan R, Levin S. Superinfection and the evolution of resistance to antimalarial drugs. Proc Biol Sci. 2012;279(1743):3834–42.

Knols BGJ, deJong R, Takken W. Differential attractiveness of isolated humans to mosquitoes in Tanzania. Trans R Soc Trop Med Hyg. 1995;89(6):604–6.

Koella JC, Antia R. Epidemiological models for the spread of anti-malarial resistance. Malar J. 2003;2:3.

Koenderink JB, Kavishe RA, Rijpma SR, Russel FGM. The ABCs of multidrug resistance in malaria. Trends Parasitol. 2010 Sep 1;26(9):440–6.

Koita OA, Ndiaye J-L, Nwakanma D, Sangare L, Ndiaye D, Joof F, et al. Seasonal changes in the frequency of false negative rapid diagnostic tests based on histidine rich protein 2 (HRP2). Am J Trop Med Hyg. 2013;89(Supplement 5):1.

Kouyos RD, Metcalf CJE, Birger R, Klein EY, zur Wiesch PA, Ankomah P, et al. The path of least resistance: aggressive or moderate treatment? Proc R Soc London B Biol Sci. 2014;281(1794):20140566.

Kuhnert D, Stadler T, Vaughan TG, Drummond AJ. Phylodynamics with Migration: A Computational Framework to Quantify Population Structure from Genomic Data. Mol Biol Evol. 2016;33(8):2102–16.

Kumar N, Zheng H. Stage-specific gametocytocidal effect in vitro of the antimalaria drug qinghaosu on Plasmodium falciparum. Parasitol Res. 1990;76(3):214–8.

Kwiatkowski D. Malaria genomics: Tracking a diverse and evolving parasite population. Int Health. 2015;7(2):82–4.

Lamikanra AA, Brown D, Potocnik A, Casals-Pascual C, Langhorne J, Roberts DJ. Malarial anemia: of mice and men. Blood. 2007 Jul 1;110(1):18–28.

Langhi DM, Bordin JO. Duffy blood group and malaria. Hematology. 2006;11(5–6):389–98.

Langhorne J, Ndungu FM, Sponaas AM, Marsh K. Immunity to malaria: more questions than answers. Nat Immunol. 2008;9(7):725–32.

Lasonder E, Rijpma SR, Van Schaijk BCLL, Hoeijmakers WAMM, Kensche PR, Gresnigt MS, et al. Integrated transcriptomic and proteomic analyses of P. falciparum gametocytes: molecular insight into sex-specific processes and translational repression. Nucleic Acids Res. 2016;44(13):6087–101.

Laufer MK, Takala-Harrison S, Dzinjalamala FK, Stine OC, Taylor TE, Plowe CV. Return of Chloroquine-Susceptible Falciparum Malaria in Malawi Was a Reexpansion of Diverse Susceptible Parasites. J Infect Dis. 2010 Sep;202(5):801–8.

Lawson DJ, Hellenthal G, Myers S, Falush D. Inference of Population Structure using Dense Haplotype Data. 2012;8(1):11–7.

Legros M, Bonhoeffer S. A combined within-host and between-hosts modelling framework for the evolution of resistance to antimalarial drugs. J R Soc Interface. 2016;13(117):20160148.

Lelièvre J, Almela MJ, Lozano S, Miguel C, Franco V, Leroy D, et al. Activity of clinically relevant antimalarial drugs on plasmodium falciparum mature gametocytes in an atp bioluminescence “transmission blocking” assay. PLoS One. 2012 Apr 13;7(4).

Lengeler C. Insecticide-treated bed nets and curtains for preventing malaria. Cochrane Database Syst Rev. 2004;(2):CD000363.

Lenhart A, Orelus N, Maskill R, Alexander N, Streit T, McCall PJ. Insecticide-treated bednets to control dengue vectors: Preliminary evidence from a controlled trial in Haiti. Trop Med Int Heal. 2008;13(1):56–67.

Lin JT, Patel JC, Levitz L, Wojnarski M, Chaorattanakawee S, Gosi P, et al. Gametocyte Carriage, Antimalarial Use, and Drug Resistance in Cambodia, 2008–2014. Am J Trop Med Hyg. 2018 Nov 7;99(5):1145–9.

Lindsay SW, Snow RW. The trouble with eaves - House entrey by vectors of malaria. Trans R Soc Trop Med Hyg. 1988;82(4):645–6.

Llor C, Bjerrum L. Antimicrobial resistance: Risk associated with antibiotic overuse and initiatives to reduce the problem. Ther Adv Drug Saf. 2014;5(6):229–41.

Lozano S, Gamallo P, González-Cortés C, Presa Matilla JL, Fairhurst RM, Herreros E, et al. Gametocytes from K13 propeller mutant plasmodium falciparum clinical isolates demonstrate reduced susceptibility to dihydroartemisinin in the male gamete exflagellation inhibition assay. Antimicrob Agents Chemother. 2018 Dec 1;62(12):e01426-18.

Ludin P, Woodcroft B, Ralph SA, Mäser P. In silico prediction of antimalarial drug target candidates. Int J Parasitol Drugs Drug Resist. 2012;2:191–9.

Maccallum WG. On The Flagellated Form Of The Malarial Parasite. Lancet. 1897 Jun 5;150(3872):1240–1.

MacDonald G. The analysis of infection rates in diseases in which superinfection occurs. Trop Dis Bull. 1950;(47):907–915.

MacDonald G. The analysis of equilibrium in malaria. Trop Dis Bull. 1952a Sep;49(9):813–29.

MacDonald G. The Analysis of the Sporozoite Rate. Trop Dis Bull. 1952b;49(6):569-586 pp.

MacDonald G. Epidemiological basis of malaria control. Bull World Heal Org. 1956;(15):613–626.

MacDonald G. The Epidemiology and Control of Malaria. London, Oxford Univ. Pr.; 1957.

Mackinnon MJ. Drug resistance models for malaria. Acta Trop. 2005;94(3 SPEC. ISS.):207–17.

Maire N, Aponte JJ, Ross A, Thompson R, Alonso P, Utzinger J, et al. Modeling a field trial of the RTS,S/AS02A malaria vaccine. Am J Trop Med Hyg. 2006a;75(2):104–10.

Maire N, Smith T, Ross A, Owusu-Agyei S, Dietz K, Molineaux L. A model for natural immunity to asexual blood stages of Plasmodium falciparum malaria in endemic areas. Am J Trop Med Hyg. 2006b;75(2):19–31.

Maire N, Tediosi F, Ross A, Smith T. Predictions of the epidemiologic impact of introducing a pre-erythrocytic vaccine into the expanded program on immunization in sub-Saharan Africa. Am J Trop Med Hyg. 2006c;75(2):111–8.

Malaria Policy Advisory Committee. Meeting report of the WHO Evidence Review Group on mass drug administration for malaria. 2019.

malERA Consultative Group on Vaccines. A research agenda for malaria eradication: vaccines. PLoS Med. 2011 Jan 25;8(1):e1000398.

Mappin B, Dalrymple U, Cameron E, Bhatt S, Weiss DJ, Gething PW. Standardizing Plasmodium falciparum infection prevalence measured via microscopy versus rapid diagnostic test. Malar J. 2015;14:460.

Marshall JM, White MT, Ghani AC, Schlein Y, Muller GC, Beier JC. Quantifying the mosquito ' s sweet tooth : modelling the effectiveness of attractive toxic sugar baits (ATSB) for malaria vector control. Malar J. 2013;12(291):1–13.

Marshall JM, Wu SL, C HMS, Kiware SS, Ouédraogo AL, Touré MB, et al. Mathematical models of human mobility of relevance to malaria transmission in Africa. Nat Sci Reports. 2018;(May):1–27.

Maude RJ, Pontavornpinyo W, Saralamba S, Aguas R, Yeung S, Dondorp AM, et al. The last man

standing is the most resistant: eliminating artemisinin-resistant malaria in Cambodia. Malar J. 2009;8:31.

Maude RJ, Socheat D, Nguon C, Saroth P, Dara P, Li G, et al. Optimising strategies for Plasmodium falciparum Malaria elimination in Cambodia: Primaquine, mass drug administration and Artemisinin resistance. PLoS One. 2012;7(5).

Maxwell CA, Msuya E, Sudi M, Njunwa KJ, Carneiro IA, Curtis CF. Effect of community-wide use of insecticide-treated nets for 3-4 years on malarial morbidity in Tanzania. Trop Med Int Heal. 2002;7(12):1003–8.

McCarthy KA, Wenger EA, Huynh GH, Eckhoff PA. Calibration of an intrahost malaria model and parameter ensemble evaluation of a pre-erythrocytic vaccine. Malar J. 2015;14(1):1–10.

McKenzie FE, Ferreira MU, Baird JK, Snounou G, Bossert WH. Meiotic recombination, cross-reactivity, and persistence in Plasmodium falciparum. Evolution (N Y). 2001;55(7):1299–307.

McKenzie FE, Sirichaisinthop J, Miller RS, Gasser RA, Wongsrichanalai C. Dependence of malaria detection and species diagnosis by microscopy on parasite density. Am J Trop Med Hyg. 2003;69(4):372–6.

Medley GF, Sinden RE, Fleck S, Billingsley PF, Tirawanchai N, Rodriguez MH. Heterogeneity in patterns of malarial oocyst infections in the mosquito vector. Parasitology. 1993;106(Pt 5):441–9.

Meis J, Rijntjes PJM, Verhave JP, Ponnudurai T, Hollingdale MR, Smith JE, et al. Fine-structure of the malaria parasite Plasmodium falciparum in human hepatocytes in vitro. Cell Tissue Res. 1986;244(2):345–50.

Le Menach A, Takala S, McKenzie FE, Perisse A, Harris A, Flahault A, et al. An elaborated feeding cycle model for reductions in vectorial capacity of night-biting mosquitoes by insecticide-treated nets. Malar J. 2007;6:10.

Menard D, Dondorp A. Antimalarial drug resistance: a threat to malaria elimination. Cold Spring Harb Perspect Med. 2017 Jul;7(7).

Ménard D, Khim N, Beghain J, Adegnika AA, Shafiul-Alam M, Amodu O, et al. A Worldwide Map of Plasmodium falciparum K13-Propeller Polymorphisms. N Engl J Med. 2016;374(25):2453–64.

Menegon M, L'Episcopia M, Nurahmed AM, Talha AA, Nour BYM, Severini C. Identification of Plasmodium falciparum isolates lacking histidine-rich protein 2 and 3 in Eritrea. Infect Genet Evol. 2017;55(March):131–4.

Meshnick S, Janko M, Doctor, Stephanie Anderson O, Thwai K, Levitz L, Emch M. Demographic and health survey (DRC-DHSII) 2013-2014 Supplemental malaria report. 2015;

Miles A, Iqbal Z, Vauterin P, Pearson R, Campino S, Theron M, et al. Indels, structural variation, and recombination drive genomic diversity in Plasmodium falciparum. Genome Res. 2016;26:1288–99.

Miller JL, Sack BK, Baldwin M, Vaughan AM, Kappe SHI. Interferon-Mediated Innate Immune Responses against Malaria Parasite Liver Stages. Cell Rep. 2014 Apr 24;7(2):436–47.

Miotto O, Almagro-Garcia J, Manske M, Macinnis B, Campino S, Rockett KA, et al. Multiple populations of artemisinin-resistant Plasmodium falciparum in Cambodia. Nat Genet. 2013;45(6):648–55.

Miotto O, Amato R, Ashley EA, MacInnis B, Almagro-Garcia J, Amaratunga C, et al. Genetic architecture of artemisinin-resistant Plasmodium falciparum. Nat Genet. 2015;47(3):226–34.

Miotto O, Sekihara M, Tachibana S-I, Yamauchi M, Pearson RD, Amato R, et al. Emergence of artemisinin-resistant Plasmodium falciparum with kelch13 C580Y mutations on the island of New Guinea. bioRxiv. 2019 Jan;621813.

Miura K. Progress and prospects for blood-stage malaria vaccines. Vol. 15, Expert Review of Vaccines. Taylor and Francis Ltd; 2016. p. 765–81.

Mockenhaupt FP, Teun Bousema J, Eggelte TA, Schreiber J, Ehrhardt S, Wassilew N, et al. Plasmodium falciparum dhfr but not dhps mutations associated with sulphadoxine-pyrimethamine treatment failure and gametocyte carriage in northern Ghana. Trop Med Int Health. 2005 Sep;10(9):901–8.

Mogeni P, Omedo I, Nyundo C, Kamau A, Noor A, Bejon P, et al. Effect of transmission intensity on hotspots and micro-epidemiology of malaria in sub-Saharan Africa. BMC Med. 2017;15(1):1–11.

Mohamed AO, Abdel Hamid MM, Mohamed OS, Elkando NS, Suliman A, Adam MA, et al. Efficacies of DHA-PPQ and AS/SP in patients with uncomplicated Plasmodium falciparum malaria in an area of an unstable seasonal transmission in Sudan. Malar J. 2017;16(1):10–5.

Mok S, Ashley EA, Ferreira PE, Zhu L, Lin Z, Yeo T, et al. Population transcriptomics of human malaria parasites reveals the mechanism of artemisinin resistance. Science (80-). 2015;347(6220):431–5.

Molina-cruz A, Barillas-mury C. The remarkable journey of adaptation of the Plasmodium falciparum malaria parasite to New World anopheline mosquitoes. Mem Inst Oswaldo Cruz. 2014;109(5):662–7.

Molina-Cruz A, Canepa GE, Kamath N, Pavlovic N V., Mu J, Ramphul UN, et al. Plasmodium evasion of mosquito immunity and global malaria transmission: The lock-and-key theory . Proc Natl Acad Sci. 2015;112(49):15178–83.

Molina-Cruz A, Garver LS, Alabaster A, Bangiolo L, Haile A, Winikor J, et al. The Human Malaria Parasite Pfs47 Gene Mediates Evasion of the Mosquito Immune System. Science (80-). 2013 May 24;340(6135):984–7.

Molineaux L. The Garki Project. Geneva: World Health Organisation; 1980.

Molineaux L, Gramiccia G. The Garki Project. Research on the epidemiology and control of malaria in the Sudan Savanna of West Africa. 1980;

Mu J, Myers RA, Jiang H, Liu S, Ricklefs S, Waisberg M, et al. Plasmodium falciparum genome-wide scans for positive selection, recombination hot spots and resistance to antimalarial drugs. Nat Genet. 2010;42(3):268–71.

Mueller GC, Beier JC, Traore SF, Toure MB, Traore MM, Bah S, et al. Successful field trial of attractive toxic sugar bait (ATSB) plant-spraying methods against malaria vectors in the Anopheles gambiae complex in Mali, West Africa. Malar J. 2010;9.

Müller GC, Junnila A, Traore MM, Revay EE, Traore SF, Doumbia S, et al. A novel window entry/exit trap for the study of endophilic behavior of mosquitoes. Acta Trop. 2017;167:137–41.

Murillo Solano C, Akinyi Okoth S, Abdallah JF, Pava Z, Dorado E, Incardona S, et al. Deletion of Plasmodium falciparum Histidine-Rich Protein 2 (pfhrp2) and Histidine-Rich Protein 3 (pfhrp3) Genes in Colombian Parasites. PLoS One. 2015;10(7):e0131576.

Murray CJL, Rosenfeld LC, Lim SS, Andrews KG, Foreman KJ, Haring D, et al. Global malaria mortality between 1980 and 2010: A systematic analysis. Lancet. 2012 Feb 4;379(9814):413–31.

Murray CK, Gasser RA, Magill AJ, Miller RS. Update on rapid diagnostic testing for malaria. Clin

Microbiol Rev. 2008;21(1):97–110.

Mzilahowa T, McCall PJ, Hastings IM. “Sexual” population structure and genetics of the malaria agent P. falciparum. PLoS One. 2007;2(7).

Nabarro D. Roll Back Malaria. Parassitologia. 1999 Sep;41(1–3):501–4.

Nair S, Li X, Arya GA, McDew-White M, Ferrari M, Nosten F, et al. Fitness Costs and the Rapid Spread of kelch13-C580Y Substitutions Conferring Artemisinin Resistance. Antimicrob Agents Chemother. 2018;62(9):e00605-18.

Nair S, Williams JT, Brockman A, Paiphun L, Mayxay M, Newton PN, et al. A selective sweep driven by pyrimethamine treatment in Southeast Asian malaria parasites. Mol Biol Evol. 2003;20(9):1526–36.

Nájera JA, González-Silva M, Alonso PL. Some lessons for the future from the global malaria eradication programme (1955-1969). PLoS Med. 2011;8(1).

Nankabirwa JI, Yeka A, Arinaitwe E, Kigozi R, Drakeley C, Kamya MR, et al. Estimating malaria parasite prevalence from community surveys in Uganda: a comparison of microscopy, rapid diagnostic tests and polymerase chain reaction. Malar J. 2015;14(1):528.

Ndounga M, Pembe Issamou Mayengue, Casimiro PN, Koukouikila-Koussounda F, Bitemo M, Diassivy Matondo B, et al. Artesunate-amodiaquine versus artemether-lumefantrine for the treatment of acute uncomplicated malaria in Congolese children under 10 years old living in a suburban area: a randomized study. Malar J. 2015;14(1):1–11.

Neafsey DE, Galinsky K, Jiang RHY, Young L, Sykes SM, Saif S, et al. The malaria parasite Plasmodium vivax exhibits greater genetic diversity than Plasmodium falciparum. Nat Genet. 2012 Aug 5;44:1046.

Nguyen TD, Olliaro P, Dondorp AM, Baird JK, Lam HM, Farrar J, et al. Optimum population-level use of artemisinin combination therapies: A modelling study. Lancet Glob Heal. 2015;3(12):e758–66.

Nielsen MA, Vestergaard LS, Lusingu J, Giha HA, Grevstad B, Goka BQ, et al. Geographical and Temporal Conservation of Antibody Recognition of. Society. 2004;76(2):3531–5.

Nijhout MM, Carter R. Gamete development in malaria parasites: bicarbonate-dependent stimulation by pH in vitro. Parasitology. 1978 Feb;76(1):39–53.

Nkhoma SC, Banda RL, Khoswe S, Dzoole-Mwale TJ, Ward SA. Intra-host dynamics of co-infecting parasite genotypes in asymptomatic malaria patients. Infect Genet Evol. 2018a;65(April):414–24.

Nkhoma SC, Nair S, Al-Saai S, Ashley E, McGready R, Phyo AP, et al. Population genetic correlates of declining transmission in a human pathogen. Mol Ecol. 2013;22(2):273–85.

Nkhoma SC, Nair S, Cheeseman IH, Rohr-Allegrini C, Singlam S, Nosten F, et al. Close kinship within multiple-genotype malaria parasite infections. Proc Biol Sci. 2012;279(1738):2589–98.

Nkhoma SC, Trevino SG, Gorena KM, Nair S, Khoswe S, Jett C, et al. Resolving within-host malaria parasite diversity using single-cell sequencing. bioRxiv. 2018b;:391268.

Noedl H, Se Y, Schaecher K, Smith BL, Socheat D, Fukuda MM. Evidence of Artemisinin-Resistant Malaria in Western Cambodia. N Engl J Med. 2008;359(24):2619–20.

Noor AM, Ringwald P. Proposed technical consultation on the role of parasite genetics in malaria surveillance to optimize response by national programmes [Internet]. 2018 [cited 2019 Sep 9].

Available from: https://www.who.int/malaria/mpac/mpac-october2018-session8-parasite-genetics-surveillance.pdf

Nussenzweig RS, Vanderberg J, Most H, Orton C. Protective immunity produced by the injection of x-irradiated sporozoites of plasmodium berghei. Nature. 1967 Oct 14;216(5111):160–2.

Nwakanma DC, Duffy CW, Amambua-Ngwa A, Oriero EC, Bojang KA, Pinder M, et al. Changes in malaria parasite drug resistance in an endemic population over a 25-year period with resulting genomic evidence of selection. J Infect Dis. 2014;209(7):1126–35.

O'Meara WP, Smith DL, McKenzie FE. Potential impact of intermittent preventive treatment (IPT) on spread of drug-resistant malaria. PLoS Med. 2006;3(5):633–42.

Ocan M, Bwanga F, Okeng A, Katabazi F, Kigozi E, Kyobe S, et al. Prevalence of K13-propeller gene polymorphisms among Plasmodium falciparum parasites isolated from adult symptomatic patients in northern Uganda. BMC Infect Dis. 2016;16(1):1–9.

Ochola LB, Vounatsou P, Smith T, Mabaso MLH, Newton C. The reliability of diagnostic techniques in the diagnosis and management of malaria in the absence of a gold standard. Lancet Infect Dis. 2006 Sep 1;6(9):582–8.

Odaga J, Sinclair D, Lokong JA, Donegan S, Hopkins H, Garner P. Rapid diagnostic tests versus clinical diagnosis for managing people with fever in malaria endemic settings. Cochrane Database Syst Rev. 2014;2014(4).

Okell L, Reiter L, Ebbe L, Baraka V, Bisanzio D, Watson O, et al. Emerging implications of policies on malaria treatment: genetic changes in the Pfmdr-1 gene affecting susceptibility to artemether-lumefantrine and artesunate-amodiaquine in Africa. BMJ Glob Heal. 2018a;3(5):e000999.

Okell LC, Bousema T, Griffin JT, Ouédraogo AL, Ghani AC, Drakeley CJ. Factors determining the occurrence of submicroscopic malaria infections and their relevance for control. Nat Commun. 2012;3:1237.

Okell LC, Cairns M, Griffin JT, Ferguson NM, Tarning J, Jagoe G, et al. Contrasting benefits of different artemisinin combination therapies as first-line malaria treatments using model-based cost-effectiveness analysis. Nat Commun. 2014;5:5606.

Okell LC, Drakeley CJ, Ghani AC, Bousema T, Sutherland CJ. Reduction of transmission from malaria patients by artemisinin combination therapies: a pooled analysis of six randomized trials. Malar J. 2008;7.

Okell LC, Ghani AC, Lyons E, Drakeley CJ. Submicroscopic Infection in Plasmodium falciparum–Endemic Populations: A Systematic Review and Meta-Analysis. J Infect Dis. 2009;200:1509–17.

Okell LC, Griffin JT, Kleinschmidt I, Hollingsworth TD, Churcher TS, White MJ, et al. The potential contribution of mass treatment to the control of plasmodium falciparum malaria. PLoS One. 2011;6(5).

Okell LC, Reiter LM, Ebbe Lene Sandøand Baraka V, Bisanzio D, Watson OJ, Bennett A, et al. Emerging implications of policies on malaria treatment: genetic changes in the Pfmdr-1 gene affecting susceptibility to artemether--lumefantrine and artesunate--amodiaquine in Africa. BMJ Glob Heal. 2018b;3(5):e000999.

Okumu FO, Moore SJ. Combining indoor residual spraying and insecticide-treated nets for malaria control in Africa: a review of possible outcomes and an outline of suggestions for the future. Malar J. 2011;10(1):208.

Omedo I, Mogeni P, Bousema T, Rockett K, Amambua-Ngwa A, Oyier I, et al. Micro-epidemiological

structuring of Plasmodium falciparum parasite populations in regions with varying transmission intensities in Africa. Wellcome Open Res. 2017a;2(May):10.

Omedo I, Mogeni P, Rockett K, Kamau A, Hubbart C, Jeffreys A, et al. Geographic-genetic analysis of Plasmodium falciparum parasite populations from surveys of primary school children in Western Kenya. Wellcome Open Res. 2017b;2(May):1–25.

OpenMalaria #892bd5. Models of natural history of malaria in humans [Internet]. Available from: https://github.com/SwissTPH/openmalaria/wiki/ModelWithinHost#asexual-infection-models

Organization WH. Malaria Rapid Diagnostic Test Performance. Summary results of WHO product testing of malaria RDTs: round 1-7 (2008–2016). 2017.

Parr JB, Verity R, Doctor SM, Janko M, Carey-Ewend K, Turman BJ, et al. Pfhrp2-deleted Plasmodium falciparum parasites in the Democratic Republic of Congo: A national cross-sectional survey. J Infect Dis. 2016 Nov 14;216(1):36–44.

Patel JC, Taylor SM, Juliao PC, Parobek CM, Janko M, Gonzalez LD, et al. Genetic evidence of importation of drug-resistant Plasmodium falciparum to Guatemala from the Democratic Republic of the Congo. Emerg Infect Dis. 2014;20(6):932–40.

Paul REL, Brey PT, Robert V. Plasmodium sex determination and transmission to mosquitoes. Trends Parasitol. 2002;18(1):32–8.

Paul REL, Coulson TN, Raibaud A, Brey PT. Sex Determination in Malaria Parasites. Science (80-). 2000;287(5450):128–31.

Paul REL, Packer MJ, Walmsley M, Lagog M, Ranford-Cartwright LC, Paru R, et al. Mating Patterns in Malaria Parasite Populations of Papua New Guinea. Science (80-). 1995;269(5231):1709–11.

Payne D. Spread of chloroquine resistance in Plasmodium falciparum. Parasitol Today. 1987;3(8):241–6.

Payne D. Use and limitations of light microscopy for diagnosing malaria at the primary health care level. Bull World Health Organ. 1988;66(5):621–6.

Penny MA, Maire N, Bever C a, Pemberton-Ross P, Briët OJT, Smith DL, et al. Distribution of malaria exposure in endemic countries in Africa considering country levels of effective treatment. Malar J. 2015;14(1):384.

Penny MA, Maire N, Studer A, Schapira A, Smith TA. What Should Vaccine Developers Ask? Simulation of the Effectiveness of Malaria Vaccines. PLoS One. 2008;3(9):Article No.: e3193.

Penny MA, Verity R, Bever CA, Sauboin C, Galactionova K, Flasche S, et al. Public health impact and cost-effectiveness of the RTS,S/AS01 malaria vaccine: A systematic comparison of predictions from four mathematical models. Lancet. 2016;387(10016):367–75.

Peterson DS, Walliker D, Wellems TE. Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase confers resistance to pyrimethamine in falciparum malaria. Proc Natl Acad Sci U S A. 1988 Dec;85(23):9114–8.

Pinkevych M, Petravic J, Chelimo K, Vulule J, Kazura JW, Moormann AM, et al. Density-dependent blood stage Plasmodium falciparum suppresses malaria super-infection in a malaria holoendemic population. Am J Trop Med Hyg. 2013;89(5):850–6.

Plucinski MM, Dimbu PR, Fortes F, Abdulla S, Ahmed S, Gutman J, et al. Posttreatment HRP2 Clearance in Patients with Uncomplicated Plasmodium falciparum Malaria. J Infect Dis. 2018a;217(5):685–92.

Plucinski MM, Herman C, Jones S, Dimbu R, Fortes F, Ljolje D, et al. Screening for Pfhrp2/3-Deleted Plasmodium falciparum, Non-falciparum, and Low-Density Malaria Infections by a Multiplex Antigen Assay. J Infect Dis. 2018b;(September).

Plucinski MM, McElroy PD, Dimbu PR, Fortes F, Nace D, Halsey ES, et al. Clearance dynamics of lactate dehydrogenase and aldolase following antimalarial treatment for Plasmodium falciparum infection. Parasites and Vectors. 2019;12(1):1–6.

Pluess B, Tanser FC, Lengeler C, Sharp BL. Indoor residual spraying for preventing malaria. Cochrane Database Syst Rev. 2010;(4):CD006657.

Pongtavornpinyo W, Hastings IM, Dondorp A, White LJ, Maude RJ, Saralamba S, et al. Probability of emergence of antimalarial resistance in different stages of the parasite life cycle. Evol Appl. 2009;2(1):52–61.

Portugal S, Carret C, Recker M, Armitage AE, Gonçalves L a, Epiphanio S, et al. Host-mediated regulation of superinfection in malaria. Nat Med. 2011;17(6):732–7.

Portugal S, Tran TM, Ongoiba A, Bathily A, Li S, Doumbo S, et al. Treatment of chronic asymptomatic plasmodium falciparum infection does not increase the risk of clinical malaria upon reinfection. Clin Infect Dis. 2017;64(5):645–53.

Price RN, Tjitra E, Guerra CA, Yeung S, White NJ, Anstey NM. Vivax malaria: Neglected and not benign. Am J Trop Med Hyg. 2007;77(6):79–87.

Price RN, Uhlemann A-C, van Vugt M, Brockman A, Hutagalung R, Nair S, et al. Molecular and Pharmacological Determinants of the Therapeutic Response to Artemether-Lumefantrine in Multidrug-Resistant Plasmodium falciparum Malaria. Clin Infect Dis. 2006;42(11):1570–7.

Price RN, Uhlemann AC, Brockman A, McGready R, Ashley E, Phaipun L, et al. Mefloquine resistance in Plasmodium falciparum and increased pfmdr1 gene copy number. Lancet. 2004 Jul 31;364(9432):438–47.

Pringle G, Avery-Jones S. An assessment of the sporozoite inoculation rate as a measure of malaria transmission in the Ubembe area of North-east Tanzania. J Trop Med Hyg. 1966 Jun;69(6):132–9.

Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000 Jun;155(2):945–59.

Protopopoff N, Mosha JF, Lukole E, Charlwood JD, Wright A, Mwalimu CD, et al. Effectiveness of a long-lasting piperonyl butoxide-treated insecticidal net and indoor residual spray interventions, separately and together, against malaria transmitted by pyrethroid-resistant mosquitoes: a cluster, randomised controlled, two-by-two fact. Lancet. 2018;391(10130):1577–88.

R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2019.

Ranford-Cartwright LC, Balfe P, Carter R, Walliker D. Genetic hybrids of Plasmodium falciparum identified by amplification of genomic DNA from single oocysts. Mol Biochem Parasitol. 1991;49(2):239–43.

Ranford-Cartwright LC, Balfe P, Carter R, Walliker D. Frequency of cross-fertilization in the human malaria parasite Plasmodium falciparum. Parasitology. 1993;107(01):11.

Rao JNK, Scott AJ. On Chi-Squared Tests for Multiway Contingency Tables with Cell Proportions Estimated from Survey Data. Ann Stat. 1984;12(1):46–60.

Rao VB, Schellenberg D, Ghani AC. Overcoming health systems barriers to successful malaria

treatment. Trends Parasitol. 2013;29(4):164–80.

Rasmussen DA, Volz EM, Koelle K. Phylodynamic Inference for Structured Epidemiological Models. PLoS Comput Biol. 2014;10(4).

Razakandrainibe FG, Durand P, Koella JC, De Meeus T, Rousset F, Ayala FJ, et al. "Clonal" population structure of the malaria agent Plasmodium falciparum in high-infection regions. Proc Natl Acad Sci. 2005;102(48):17388–93.

Recker M, Buckee CO, Serazin A, Kyes S, Pinches R, Christodoulou Z, et al. Antigenic variation in Plasmodium falciparum malaria involves a highly structured switching pattern. PLoS Pathog. 2011;7(3).

Reece SE, Ali E, Schneider P, Babiker HA. Stress, drugs and the evolution of reproductive restraint in malaria parasites. Proc R Soc B Biol Sci. 2010;277(1697):3123–9.

Reece SE, Drew DR, Gardner A. Sex ratio adjustment and kin discrimination in malaria parasites. Nature. 2008;453(7195):609–14.

Reiker T, Chitnis N, Smith T. Modelling reactive case detection strategies for interrupting transmission of Plasmodium falciparum malaria. Malar J. 2019;18(1):1–13.

Rickman LS, Jones TR, Long GW, Paparello S, Schneider I, Paul CF, et al. Plasmodium falciparum infected anopheles-stephensi inconsistently transmit malaria to humans. Am J Trop Med Hyg. 1990;43(5):441–5.

Rijken MJ, McGready R, Phyo AP, Lindegardh N, Tarning J, Laochan N, et al. Pharmacokinetics of dihydroartemisinin and piperaquine in pregnant and nonpregnant women with uncomplicated falciparum malaria. Antimicrob Agents Chemother. 2011;55(12):5500–6.

Robert V, Read AF, Essong J, Tchuinkam T, Mulder B, Verhave J-P, et al. Effect of gametocyte sex ratio on infectivity of Plasmodium falciparum to Anopheles gambiae. Trans R Soc Trop Med Hyg. 1996;90(6):621–4.

Robert V, Sokhna CS, Rogier C, Ariey F, Trape J-F. Sex ratio of Plasmodium falciparum gametocytes in inhabitants of Dielmo, Senegal. Parasitology. 2003;127(1):1–8.

Roberts L, Enserink M. Malaria - Did they really say ... eradication? Science (80-). 2007;318:1544–5.

Rodriguez-Barraquer I, Arinaitwe E, Jagannathan P, Boyle MJ, Tappero J, Muhindo M, et al. Quantifying heterogeneous malaria exposure and clinical protection in a cohort of Ugandan children. J Infect Dis. 2016;214.

Roper C, Pearce R, Bredenkamp B, Gumede J, Drakeley C, Mosha F, et al. Antifolate antimalarial resistance in southeast Africa: A population-based analysis. Lancet. 2003;361(9364):1174–81.

Roper C, Pearce R, Nair S, Sharp B, Nosten F, Anderson T. Intercontinental Spread of Pyrimethamine-Resistant Malaria. Science (80-). 2004;305(5687):1124.

Ross A, Killeen G, Smith T. Relationships between host infectivity to mosquitoes and asexual parasite density in Plasmodium falciparum. Am J Trop Med Hyg. 2006a;75(Suppl 2):32–7.

Ross A, Maire N, Molineaux L, Smith T. An epidemiologic model of severe morbidity and mortality caused by Plasmodium falciparum. Am J Trop Med Hyg. 2006b;75(2):63–73.

Ross A, Smith T. The effect of malaria transmission intensity on neonatal mortality in endemic areas. Am J Trop Med Hyg. 2006;75(2):74–81.

Ross R. Prevention of Malaria in Mauritius. WATERLOW SONS Ltd. 1908;

Rutledge GG, Böhme U, Sanders M, Reid AJ, Cotton JA, Maiga-Ascofare O, et al. Plasmodium malariae and P. ovale genomes provide insights into malaria parasite evolution. Nature. 2017;542(7639):101–4.

Sabchareon A, Burnouf T, Ouattara D, Attanath P, Bouharoun-Tayoun H, Chantavanich P, et al. Parasitologic and clinical human response to immunoglobulin administration in falciparum malaria. Am J Trop Med Hyg. 1991;45(3):297–308.

Saralamba S, Pan-Ngum W, Maude RJ, Lee SJ, Tarning J, Lindegardh N, et al. Intrahost modeling of artemisinin resistance in Plasmodium falciparum. Proc Natl Acad Sci U S A. 2011;108(1):397–402.

Schaffner SF, Taylor AR, Wong W, Wirth DF, Neafsey DE. HmmIBD: Software to infer pairwise identity by descent between haploid genotypes. Malar J. 2018;17(1):10–3.

Schwenk R, Asher L V, Chalom I, Lanar D, Sun P, White K, et al. Opsonization by antigen-specific antibodies as a mechanism of protective immunity induced by Plasmodium falciparum circumsporozoite protein-based vaccine. Parasite Immunol. 2003 Jan;25(1):17–25.

Schwenk RJ, Richie TL. Protective immunity to pre-erythrocytic stage malaria. Trends Parasitol. 2011;27(7):306–14.

Sherrard-Smith E, Griffin JT, Winskill P, Corbel V, Pennetier C, Djénontin A, et al. Systematic review of indoor residual spray efficacy and effectiveness against Plasmodium falciparum in Africa. Nat Commun. 2018;9(1).

Sidjanski S, Vanderberg JP. Delayed migration of Plasmodium sporozoites from the mosquito bite site to the blood. Am J Trop Med Hyg. 1997 Oct;57(4):426–9.

Simpson J a, Zaloumis S, DeLivera AM, Price RN, McCaw JM. Making the Most of Clinical Data: Reviewing the Role of Pharmacokinetic-Pharmacodynamic Models of Anti-malarial Drugs. AAPS J. 2014;16(5):962–74.

Simpson JA, Watkins ER, Price RICN, Aarons L, Kyle DE, White NJ. Mefloquine Pharmacokinetic-Pharmacodynamic Models : Implications for Dosing and Resistance. Antimicrob Agents Chemother. 2000;44(12):3414–24.

Sinden R, Gilles HM. The malaria parasites. In: Warrell DA, Gilles HM, editors. Essential Malariology. 4th ed. New York: Oxford University Press; 2002.

Sinden RE. Mitosis and meiosis in malarial parasites. Acta Leiden. 1991;60(1):19–27.

Singh N, Bharti PK, Singh MP, Mishra S, Shukla MM, Sharma RK, et al. Comparative Evaluation of Bivalent Malaria Rapid Diagnostic Tests versus Traditional Methods in Field with Special Reference to Heat Stability Testing in Central India. PLoS One. 2013;8(3):1–9.

Sinka ME, Bangs MJ, Manguin S, Rubio-palis Y, Chareonviriyaphap T, Coetzee M, et al. A global map of dominant malaria vectors. Parasit Vectors. 2012;5(69):1–11.

Sinka ME, Rubio-palis Y, Manguin S, Patil AP, Temperley WH, Gething PW, et al. The dominant Anopheles vectors of human malaria in the Americas: occurrence data , distribution maps and bionomic précis. Parasit Vectors. 2010;3(72):1–26.

Slater HC, Griffin JT, Ghani AC, Okell LC. Assessing the potential impact of artemisinin and partner drug resistance in sub-Saharan Africa. Malar J. 2016;15(10).

Slater HC, Okell LC, Ghani AC. Mathematical Modelling to Guide Drug Development for Malaria Elimination. Trends Parasitol. 2017;33(3):175–84.

Slater HC, Ross A, Felger I, Hofmann NE, Robinson L, Cook J, et al. The temporal dynamics and infectiousness of subpatent Plasmodium falciparum infections in relation to parasite density. Nat Commun. 2019;10(1).

Slater HC, Ross A, Ouédraogo AL, White LJ, Nguon C, Walker PGT, et al. Assessing the impact of next-generation rapid diagnostic tests on Plasmodium falciparum malaria elimination strategies. Nature. 2015a;528(7580):S94-101.

Slater HC, Ross A, Ouédraogo AL, White LJ, Nguon C, Walker PGT, et al. Assessing the impact of next-generation rapid diagnostic tests on Plasmodium falciparum malaria elimination strategies. Nature. 2015b;528(December).

Slater HC, Walker PGT, Bousema T, Okell LC, Ghani AC. The potential impact of adding ivermectin to a mass treatment intervention to reduce malaria transmission: A modelling study. J Infect Dis. 2014;210(12):1972–80.

Smit MR, Ochomo EO, Aljayyoussi G, Kwambai TK, Abong'o BO, Chen T, et al. Safety and mosquitocidal efficacy of high-dose ivermectin when co-administered with dihydroartemisinin-piperaquine in Kenyan adults with uncomplicated malaria (IVERMAL): a randomised, double-blind, placebo-controlled trial. Lancet Infect Dis. 2018 Jun 1;18(6):615–26.

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. Chitnis CE, editor. PLoS Pathog. 2012a Apr 5;8(4):e1002588.

Smith DL, Drakeley CJ, Chiyaka C, Hay SI. A quantitative analysis of transmission efficiency versus intensity for malaria. Nat Commun. 2010a;1(8):108.

Smith DL, Dushoff J, McKenzie FE. The risk of a mosquito-borne infection in a heterogeneous environment. Plos Biol. 2004;2(11):1957–64.

Smith DL, Klein EY, McKenzie FE, Laxminarayan R. Prospective strategies to delay the evolution of anti-malarial drug resistance: weighing the uncertainty. Malar J. 2010b;9:217.

Smith RC, Barillas-Mury C. Plasmodium Oocysts: Overlooked Targets of Mosquito Immunity. Trends Parasitol. 2016;32(12):979–90.

Smith T, Killeen GF, Maire N, Ross A, Molineaux L, Tediosi F, et al. Mathematical modeling of the impact of malaria vaccines on the clinical epidemiology and natural history of Plasmodium falciparum malaria: Overview. Am J Trop Med Hyg. 2006a;75(2 Suppl):1–10.

Smith T, Ross A, maire N, Chitnis N, Studer A, Hardy D, et al. Ensemble modeling of the likely public health impact of a pre-erythrocytic malaria vaccine. Plos Med. 2012b;9(e1001157).

Smith T, Ross A, Maire N, Rogier C, Trape JF, Molineaux L. An epidemiologic model of the incidence of acute illness in Plasmodium falciparum malaria. Am J Trop Med Hyg. 2006b;75(2):56–62.

Smith TG, Lourenco P, Carter R, Walliker D, Ranford-Cartwright LC. Commitment to sexual differentiation in the human malaria parasite, Plasmodium falciparum. Parasitology. 2000;121(2):127–33.

Sondo P, Derra K, Lefevre T, Diallo-Nakanabo S, Tarnagda Z, Zampa O, et al. Genetically diverse Plasmodium falciparum infections, within-host competition and symptomatic malaria in humans. Sci Rep. 2019;9(1):1–9.

Sowunmi A, Balogun ST, Gbotosho GO, Happi CT. Plasmodium falciparum gametocyte sex ratios in children with acute, symptomatic, uncomplicated infections treated with amodiaquine. Malar J. 2008 Sep;7:169.

Sowunmi A, Gbotosho GO, Happi CT, Folarin OA, Balogun ST. Population structure of Plasmodium falciparum gametocyte sex ratios in malarious children in an endemic area. Parasitol Int. 2009;58(4):438–43.

Stepniewska K, White NJ. Pharmacokinetic determinants of the window of selection for antimalarial drug resistance. Antimicrob Agents Chemother. 2008;52(5):1589–96.

Stevenson MM, Riley EM. Innate immunity to malaria. Nat Rev Immunol. 2004;4(3):169–80.

Stone W, Sawa P, Lanke K, Rijpma S, Oriango R, Nyaurah M, et al. A Molecular Assay to Quantify Male and Female Plasmodium falciparum Gametocytes: Results From 2 Randomized Controlled Trials Using Primaquine for Gametocyte Clearance. J Infect Dis. 2017 Aug;216(4):457–67.

Stone WJR, Churcher TS, Graumans W, Van Gemert GJ, Vos MW, Lanke KHW, et al. A scalable assessment of Plasmodium falciparum transmission in the standard membrane-feeding assay, using transgenic parasites expressing green fluorescent protein-luciferase. J Infect Dis. 2014;210(9):1456–63.

Stone WJR, Eldering M, van Gemert G-J, Lanke KHW, Grignard L, van de Vegte-Bolmer MG, et al. The relevance and applicability of oocyst prevalence as a read-out for mosquito feeding assays. Sci Rep. 2013;3:3418.

Straimer J, Gnädig NF, Witkowski B, Amaratunga C, Duru V, Ramadani AP, et al. K13-propeller mutations confer artemisinin resistance in Plasmodium falciparum clinical isolates. Science (80-). 2015;347(6220):428–31.

Strode C, Donegan S, Garner P, Enayati AA, Hemingway J. The Impact of Pyrethroid Resistance on the Efficacy of Insecticide-Treated Bed Nets against African Anopheline Mosquitoes: Systematic Review and Meta-Analysis. PLoS Med. 2014;11(3).

Sutherland CJ, Tanomsing N, Nolder D, Oguike M, Jennison C, Pukrittayakamee S, et al. Two Nonrecombining Sympatric Forms of the Human Malaria Parasite Plasmodium ovale Occur Globally . J Infect Dis. 2010;201(10):1544–50.

Tacoli C, Gai PP, Bayingana C, Sifft K, Geus D, Ndoli J, et al. Artemisinin resistance-associated K13 polymorphisms of plasmodium falciparum in Southern Rwanda, 2010-2015. Am J Trop Med Hyg. 2016;95(5):1090–3.

Takala-Harrison S, Clark TG, Jacob CG, Cummings MP, Miotto O, Dondorpe AM, et al. Genetic loci associated with delayed clearance of plasmodium falciparum following artemisinin. Proc Natl Acad Sci U S A. 2013;110(1):240–5.

Tangpukdee N, Duangdee C, Wilairatana P, Krudsood S. Malaria diagnosis: A brief review. Korean J Parasitol. 2009;47(2):93–102.

Tarning J, Kloprogge F, Dhorda M, Jullien V, Nosten F, White NJ, et al. Pharmacokinetic properties of artemether, dihydroartemisinin, lumefantrine, and quinine in pregnant women with uncomplicated plasmodium falciparum malaria in uganda. Antimicrob Agents Chemother. 2013;57(10):5096–103.

Tatem AJ, Jia P, Ordanovich D, Falkner M, Huang Z, Howes R, et al. The geography of imported malaria to non-endemic countries: a meta-analysis of nationally reported statistics. Lancet Infect Dis. 2017;17(1):98–107.

Taylor AR, Schaffner SF, Cerqueira GC, Nkhoma SC, Anderson TJC, Sriprawat K, et al. Quantifying connectivity between local Plasmodium falciparum malaria parasite populations using identity by descent. PLoS Genet. 2017;1–20.

Tediosi F, Hutton G, Maire N, Smith TA, Ross A, Tanner M. Predicting the cost-effectiveness of introducing a pre-erythrocytic malaria vaccine into the expanded program on immunization in Tanzania. Am J Trop Med Hyg. 2006;75(2):131–43.

Teklehaimanot HD, Lipsitch M, Teklehaimanot A, Schwartz J. Weather-based prediction of Plasmodium falciparum malaria in epidemic-prone regions of Ethiopia I. Patterns of lagged weather effects reflect biological mechanisms. Malar J. 2004;3(41):1–11.

Tessema S, Wesolowski A, Chen A, Murphy M, Wilheim J, Mupiri AR, et al. Using parasite genetic and human mobility data to infer local and cross-border malaria connectivity in Southern Africa. Elife. 2019;8:1–20.

Thomas SM, Ndir O, Dieng T, Mboup S, Wypij D, Maguire JH, et al. Analysis of pfcrt point mutations and chloroquine susceptibility in isolates of Plasmodium falciparum. Mol Biochem Parasitol. 2002;66:474–80.

Tilley L, Straimer J, Gnädig NF, Ralph SA, Fidock DA. Artemisinin Action and Resistance in Plasmodium falciparum. Vol. 32, Trends in Parasitology. Elsevier Current Trends; 2016. p. 682–96.

Tiono AB, Ouédraogo A, Ogutu B, Diarra A, Coulibaly S, Gansané A, et al. A controlled, parallel, cluster-randomized trial of community-wide screening and treatment of asymptomatic carriers of Plasmodium falciparum in Burkina Faso. Malar J. 2013;12:1–11.

Tirrell AR, Vendrely KM, Checkley LA, Davis SZ, McDew-White M, Cheeseman IH, et al. Pairwise growth competitions identify relative fitness relationships among artemisinin resistant Plasmodium falciparum field isolates. Malar J. 2019 Aug 28;18(1):295.

Trampuz A, Jereb M, Muzlovic I, Prabhu RM. Clinical review: Severe malaria. Vol. 7, Critical Care. 2003. p. 315–23.

Tu Y. The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. Nat Med. 2011;17(10):1217–20.

Tusting LS, Bisanzio D, Alabaster G, Cameron E, Cibulskis R, Davies M, et al. Mapping changes in housing in sub-Saharan Africa from 2000 to 2015. Nature. 2019;568(7752):391–4.

Umlauf R. Precarity and Preparedness: Non-Adherence as Institutional Work in Diagnosing and Treating Malaria in Uganda. Med Anthropol. 2017 Jul;36(5):449–63.

Vaughan JA. Population dynamics of Plasmodium sporogony. Trends Parasitol. 2007;23(2):63–70.

Verity R, Hathaway NJ, Waltmann A, Doctor SM, Watson OJ, Patel JC, et al. Plasmodium falciparum genetic variation of var2csa in the Democratic Republic of the Congo. Malar J. 2018;17(46).

Volkman SK, Neafsey DE, Schaffner SF, Park DJ, Wirth DF. Harnessing genomics and genome biology to understand malaria biology. Nat Rev Genet. 2012;13(5):315–28.

Volz EM. Complex population dynamics and the coalescent under neutrality. Genetics. 2012;190(1):187–201.

Volz EM, Kosakovsky Pond SL, Ward MJ, Leigh Brown AJ, Frost SDW. Phylodynamics of Infectious Disease Epidemics. Genetics. 2009;183(4):1421–30.

Wagner JC, Platt RJ, Goldfless SJ, Zhang F, Niles JC. Efficient CRISPR-Cas9–mediated genome editing in Plasmodium falciparum. Nat Methods. 2014 Aug 10;11(9):915.

Wahl LM, Nowak MA. Adherence and drug resistance: predictions for therapy outcome. Proceedings Biol Sci. 2000 Apr;267(1445):835–43.

Walker K, Lynch M. Contributions of Anopheles larval control to malaria suppression in tropical Africa: review of achievements and potential. Med Vet Entomol. 2007;21(1):2–21.

Walker PGT, Floyd J, ter Kuile F, Cairns M. Estimated impact on birth weight of scaling up intermittent preventive treatment of malaria in pregnancy given sulphadoxine-pyrimethamine resistance in Africa: A mathematical model. PLoS Med. 2017;14(2):1–19.

Walker PGT, Griffin JT, Ferguson NM, Ghani AC, Alonso P, Tanner M, et al. Estimating the most efficient allocation of interventions to achieve reductions in Plasmodium falciparum malaria burden and transmission in Africa: a modelling study. Lancet Glob Heal. 2016 Jul;4(7):e474-84.

Walker PGT, White MT, Griffin JT, Reynolds A, Ferguson NM, Ghani AC. Malaria morbidity and mortality in Ebola-affected countries caused by decreased health-care capacity, and the potential effect of mitigation strategies: A modelling analysis. Lancet Infect Dis. 2015;15(7):825–32.

Wang P, Lee CS, Bayoumi R, Djimde A, Doumbo O, Swedberg G, et al. Resistance to antifolates in Plasmodium falciparum monitored by sequence analysis of dihydropteroate synthetase and dihydrofolate reductase alleles in a large number of field samples of diverse origins. Mol Biochem Parasitol. 1997 Nov;89(2):161–77.

Watkins WM, Mberu EK, Winstanley PA, Plowe CV. The efficacy of antifolate antimalarial combinations in Africa: a predictive model based on pharmacodynamic and pharmacokinetic analyses. Parasitol Today. 1997 Dec;13(12):459–64.

Watkins WM, Sibley CH, Hastings IANM. The search for effective and sustainable treatments for Plasmodium falciparum malaria in Africa: a model of the selection of resistance by antifolate drugs and their combinations. Am J Trop Med Hyg. 2005;72(2):163–73.

Watson OJ, Slater HC, Verity R, Parr JB, Mwandagalirwa MK, Tshefu A, et al. Modelling the drivers of the spread of Plasmodium falciparum hrp2 gene deletions in sub-Saharan Africa. Elife. 2017 Aug 24;6:e25008.

Watson OJ, Sumner KM, Janko M, Goel V, Winskill P, Slater HC, et al. False-negative malaria rapid diagnostic test results and their impact on community-based malaria surveys in sub-Saharan Africa. BMJ Glob Heal. 2019a;4(e001582).

Watson OJ, Verity R, Ghani AC, Garske T, Cunningham J, Tshefu A, et al. Impact of seasonal variations in Plasmodium falciparum malaria transmission on the surveillance of pfhrp2 gene deletions. Elife. 2019b May 2;8(e40339).

Weiss DJ, Lucas TCD, Nguyen M, Nandi AK, Bisanzio D, Battle KE, et al. Mapping the global prevalence, incidence, and mortality of Plasmodium falciparum, 2000–17: a spatial and temporal modelling study. Lancet. 2019;394(10195):322–31.

Wellems TE, Walker-Jonah A, Panton LJ. Genetic mapping of the chloroquine-resistance locus on Plasmodium falciparum chromosome 7. Proc Natl Acad Sci U S A. 1991 Apr 15;88(8):3382–6.

Wenger EA, Eckhoff PA. A mathematical model of the impact of present and future malaria vaccines. Malar J. 2013;12(1):1–13.

Wesolowski A, Taylor AR, Chang HH, Verity R, Tessema S, Bailey JA, et al. Mapping malaria by combining parasite genomic and epidemiologic data. BMC Med. 2018;16(1):1–8.

White M. Models for measuring and predicting malaria vaccine efficacy. Imperial College London; 2012.

White MT, Conteh L, Cibulskis R, Ghani AC. Costs and cost-effectiveness of malaria control

interventions - a systematic review. Malar J. 2011;10(1):337.

White NJ, Pongtavornpinyo W. The de novo selection of drug-resistant malaria parasites. Proc R Soc B Biol Sci. 2003;270(1514):545–54.

Zur Wiesch PA, Kouyos R, Engelstädter J, Regoes RR, Bonhoeffer S. Population biological principles of drug-resistance evolution in infectious diseases. Lancet Infect Dis. 2011;11(3):236–47.

Winskill P, Walker PG, Griffin JT, Ghani AC. Modelling the cost-effectiveness of introducing the RTS,S malaria vaccine relative to scaling up other malaria interventions in sub-Saharan Africa. BMJ Glob Heal. 2017;2(1):e000090.

Witkowski B, Duru V, Khim N, Ross LS, Saintpierre B, Beghain J, et al. A surrogate marker of piperaquine-resistant Plasmodium falciparum malaria: a phenotype–genotype association study. Lancet Infect Dis. 2017 Feb 1;17(2):174–83.

Witmer K, Sherrard-Smith E, Straschil U, Tunnicliff M, Baum J, Delves M. An inexpensive open source 3D-printed membrane feeder for human malaria transmission studies. Malar J. 2018 Dec;17(282).

Wong J, Bayoh N, Olang G, Killeen GF, Hamel MJ, Vulule JM, et al. Standardizing operational vector sampling techniques for measuring malaria transmission intensity: Evaluation of six mosquito collection methods in western Kenya. Malar J. 2013;12(1):1–11.

Wong W, Griggs AD, Daniels RF, Schaffner SF, Ndiaye D, Bei AK, et al. Genetic relatedness analysis reveals the cotransmission of genetically related Plasmodium falciparum parasites in Thiès, Senegal. Genome Med. 2017;9(1):5.

Wong W, Wenger EA, Hartl DL, Wirth DF. Modeling the genetic relatedness of Plasmodium falciparum parasites following meiotic recombination and cotransmission. Pascual M, editor. PLOS Comput Biol. 2018 Jan 9;14(1):e1005923.

Wongsrichanalai C, Barcus MJ, Muth S, Sutamihardja A, Wernsdorfer WH. A review of malaria diagnostic tools: microscopy and rapid diagnostic test (RDT). Am J Trop Med Hyg. 2007 Dec;77(6 Suppl):119–27.

World Health Organization. Algeria and Argentina certified malaria-free by WHO [Internet]. [cited 2019 Sep 22]. Available from: https://www.who.int/news-room/detail/22-05-2019-algeria-and-argentina-certified-malaria-free-by-who

World Health Organization. World Malaria Report 2008. 2008.

World Health Organization. Guidelines for the treatment of malaria, 2nd edition. 2010a;.

World Health Organization. WHO Policy recommendation on Intermittent Preventive Treatment during infancy for Plasmodium falciparum malaria control in Africa Contra-indications [Internet]. 2010b [cited 2017 Oct 9]. Available from: http://www.who.int/malaria/news/WHO_policy_recommendation_IPTi_032010.pdf

World Health Organization. Interim position statement- the role of larviciding for malaria-control in sub-Saharan africa. Glob Malar Progr. 2012a;(April):1–21.

World Health Organization. Updated WHO policy recommendation: intermittent preventive treatment of malaria in pregnancy using sulfadoxine-pyrimethamine (IPTp-SP) [Internet]. 2012b. Available from: http://www.who.int/malaria/publications/atoz/who_iptp_sp_policy_recommendation/en/

World Health Organization. WHO Policy Recommendation: Seasonal Malaria Chemoprevention (SMC)

for Plasmodium falciparum malaria control in highly seasonal transmission areas of the Sahel sub-region in Africa Background [Internet]. 2012c [cited 2017 Oct 9]. Available from: http://www.who.int/malaria/publications/atoz/who_smc_policy_recommendation/en/

World Health Organization. Minutes of the Drug Resistance and Containment Technical Expert Group. 2013;(April):1–43.

World Health Organization. Guidelines for the treatment of malaria. Third Edition. 2015a.

World Health Organization. Malaria Rapid Diagnostic Test Performance. Results of WHO product testing of malaria RDTs: round 6 (2014–2015). 2015b;.

World Health Organization. Recommendations on the role of mass drug administration, mass screening and treatment, and focal screening and treatment for malaria [Internet]. 2015c [cited 2017 Oct 9]. Available from: http://www.who.int/malaria/publications/atoz/role-of-mda-for-malaria/en/

World Health Organization. World Malaria Report 2015. 2015d.

World Health Organization. Malaria Threats Map [Internet]. 2018a [cited 2018 May 1]. Available from: http://www.who.int/malaria/maps/threats-about/en/

World Health Organization. Protocol for estimating the prevalence of pfhrp2/pfhrp3 gene deletions among symptomatic falciparum patients with false-negative RDT results [Internet]. 2018b [cited 2018 Mar 16]. Available from: http://www.who.int/malaria/publications/atoz/hrp2-deletion-protocol/en/

World Health Organization. World Malaria Report. 2018c.

World Health Organization. World Malaria Report. World Malaria Report. 2018d.

World Health Organization, Global Malaria Programme. A Framework for Malaria Elimination. WHO Press, World Health Organization. 2017.

WorldWide Antimalarial Research Network (WWARN). WWARN Explorer [Internet]. Available from: http://www.wwarn.org/tracking-resistance/wwarn-explorer

Wu L, van den Hoogen LL, Slater H, Walker PGT, Ghani AC, Drakeley CJ, et al. Comparison of diagnostics for the detection of asymptomatic Plasmodium falciparum infections to inform control and elimination strategies. Nature. 2015;528(7580):S86–93.

WWARN DP Study Group. The Effect of Dosing Regimens on the Antimalarial Efficacy of Dihydroartemisinin-Piperaquine: A Pooled Analysis of Individual Patient Data. PLoS Med. 2013;10(12):1–17.

Wwarn K, Group GS. Association of mutations in the Plasmodium falciparum Kelch13 gene (Pf3D7_1343700) with parasite clearance rates after artemisinin-based treatments-a WWARN individual patient data meta-analysis. BMC Med. 2019;17(1):1.

Yeung S, Van Damme W, Socheat D, White NJ, Mills A. Access to artemisinin combination therapy for malaria in remote areas of Cambodia. Malar J. 2008;7:1–14.

Zhu L, Marshall JM, Qualls WA, Schlein Y, McManus JW, Arheart KL, et al. Modelling optimum use of attractive toxic sugar bait stations for effective malaria vector control in Africa. Malar J. 2015;14(1):492.

Zhu SJ, Hendry JA, Almagro-Garcia J, Pearson RD, Amato R, Miles A, et al. The origins and relatedness structure of mixed infections vary with local prevalence of P. falciparum malaria. Elife. 2019;8:1–

41.

Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc Ser B Stat Methodol. 2005;67(5):768.

National Weather Service Climate Prediction Center [Internet]. 2010 [cited 2010 Jul 1]. Available from: http://www.cpc.noaa.gov/products/international/

WHO | Malaria Vaccine Implementation Programme (MVIP) - Programme Advisory Group. WHO. 2018;

Appendix

eLIFE elifesciences.org

RESEARCH ADVANCE

Impact of seasonal variations in *Plasmodium falciparum* malaria transmission on the surveillance of *pfhrp2* gene deletions

Oliver John Watson[1]*, Robert Verity[1], Azra C Ghani[1], Tini Garske[1], Jane Cunningham[2], Antoinette Tshefu[3], Melchior K Mwandagalirwa[3,4], Steven R Meshnick[4,5], Jonathan B Parr[5], Hannah C Slater[1]

[1]MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom; [2]Global Malaria Programme, World Health Organization, Geneva, Switzerland; [3]School of Public Health, University of Kinshasa, Kinshasa, Democratic Republic of the Congo; [4]Department of Epidemiology, Gillings School for Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, United States; [5]Division of Infectious Diseases, Department of Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, United States

Abstract Ten countries have reported *pfhrp2/pfhrp3* gene deletions since the first observation of *pfhrp2*-deleted parasites in 2012. In a previous study (Watson et al., 2017), we characterised the drivers selecting for *pfhrp2/3* deletions and mapped the regions in Africa with the greatest selection pressure. In February 2018, the World Health Organization issued guidance on investigating suspected false-negative rapid diagnostic tests (RDTs) due to *pfhrp2/3* deletions. However, no guidance is provided regarding the timing of investigations. Failure to consider seasonal variation could cause premature decisions to switch to alternative RDTs. In response, we have extended our methods and predict that the prevalence of false-negative RDTs due to *pfhrp2/3* deletions is highest when sampling from younger individuals during the beginning of the rainy season. We conclude by producing a map of the regions impacted by seasonal fluctuations in *pfhrp2/3* deletions and a database identifying optimum sampling intervals to support malaria control programmes.

DOI: https://doi.org/10.7554/eLife.40339.001

***For correspondence:** o.watson15@imperial.ac.uk

Competing interests: The authors declare that no competing interests exist.

Funding: See page 12

Received: 25 July 2018
Accepted: 29 April 2019
Published: 02 May 2019

Reviewing editor: Ben Cooper, Mahidol Oxford Tropical Medicine Research Unit, Thailand

Introduction

Diagnostic testing of suspected malaria cases has more than doubled in the last 15 years, with 75% of suspected cases seeking treatment from the public health sector receiving a diagnostic test in 2017 (***World Health Organization, 2018a***). Much of this progress reflects the increased distribution of rapid diagnostic tests (RDTs), with the most commonly used RDTs targeting the *P. falciparum* protein HRP2 (PfHRP2). In 2014, a review of published reports of *pfhrp2/3* deletions was conducted and included a critical assessment of the comprehensiveness of the diagnostic investigation. (***Cheng et al., 2014***). The findings of this review highlighted a need for a harmonized approach to investigating and confirming or excluding *pfhrp2/3* deletions and called for further studies to determine the prevalence and impact of *pfhrp2/3* gene deletions. Since that review, false-negative RDT results due to *pfhrp2/3* gene deletions have been reported in 10 countries in sub-Saharan Africa (SSA) (***World Health Organization, 2018b***). The frequency of *pfhrp2/3* deletions varies across SSA,

with the highest burden observed in Eritrea where 80.8% of samples from Ghindae Hospital were both *pfhrp2*-negative and *pfhrp3*-negative in 2016 (**Berhane et al., 2018**).

Mathematical modelling has predicted that the continued use of only PfHRP2 RDTs will quickly select for parasites without the *pfhrp2* gene (**Gatton et al., 2017**). This selection pressure occurs due to the misdiagnosis of infections caused by parasites lacking the *pfhrp2* gene, which will subsequently contribute more towards onwards transmission than wild-type parasites that are correctly diagnosed due to the expression of *pfhrp2*. In 2017, we conducted an analysis of the drivers of *pfhrp2* gene deletion selection, identifying the administrative regions in SSA with the greatest potential for selecting for *pfhrp2-deleted* parasites (**Watson et al., 2017**). The regions identified were areas with both a low prevalence of malaria and a high frequency of people seeking treatment and being treated on the basis of PfHRP2-based RDT diagnosis. The precise strength of selection, however, is not known, with other factors such as the rate of non-malarial fevers and non-adherence to RDT outcomes likely to impact the number of misdiagnosed cases receiving treatment.

In February 2018, the World Health Organization (WHO) issued guidance for national malaria control programmes on how to investigate suspected false-negative RDTs with an emphasis on *pfhrp2/3* gene deletions. (**World Health Organization, 2018c**). The primary study outcome to be calculated in the guidance is as follows:

$$\begin{array}{c}\text{Proportion of } P.\,falciparum \text{ cases}\\ \text{with false}-\text{negative HRP2 RDT}\\ \text{results due to } pfhrp2/3 \text{ deletions}\end{array} = \frac{\begin{array}{c}\#\text{ of confirmed } falciparum \text{ patients with } pfhrp2/3\\ \text{gene deletions and HRP2 RDT negative results}\end{array}}{\begin{array}{c}\#\text{ of confirmed } P.\,falciparum \text{ cases}\\ (\text{by either RDT or microscopy})\end{array}}$$

The guidance recommends that a national change to non PfHRP2-based RDTs should be made if the estimated proportion of *P. falciparum* cases with false-negative HRP2 RDT results due to *pfhrp2/3* deletions is above 5%. If the estimated proportion is less than 5% the country is recommended to establish a monitoring scheme whereby the study is repeated in two years if the 95% confidence interval does not include 5%, or one year if it does include 5%. The 5% threshold approximates the point at which the number of cases missed due to false-negative PfHRP2-based RDTs caused by *pfhrp2/3* deletions may become greater than the number of cases that would be missed due to the decreased sensitivity of non PfHRP2-based RDTs. The guidance also specifies a sampling scheme to be used when estimating the prevalence of *pfhrp2/3* gene deletions. Samples are to be collected from at least 10 health facilities per province to be tested, with sampling focussed on symptomatic *P. falciparum* patients presenting at the health facilities. All samplings are to be ideally completed within an 8-week period.

The 8-week interval permits for a rapid turnaround and allows for efficient investigations and policy responses. However, the timing of the 8-week interval chosen within a transmission season is important. The chosen interval could lead to estimates of the proportion of *P. falciparum* cases with false-negative HRP2 RDT results due to *pfhrp2/3* deletions that are not representative of the annual average proportion. Subsequently, any recorded estimate may not be predictive of the number of cases that may be misdiagnosed due to *pfhrp2/3* deletions in the years between sampling intervals. For example, an overestimation of the annual average proportion of false-negative RDTs due to *pfhrp2/3* deletions could result in a switch to a less sensitive RDT, resulting in an increase in the number of malaria cases misdiagnosed if the annual average proportion of false-negative RDTs due to *pfhrp2/3* deletions is less than 5%. The alternative RDT may also be both more expensive and complicated to implement. Similarly, an underestimation of the annual average proportion of *P. falciparum* cases with false-negative HRP2 RDT results due to *pfhrp2/3* deletions would result in continued use of an overall less effective test and could provide *pfhrp2/3* deleted parasite populations an opportunity to expand.

In response to these concerns, we extended our original methods (**Watson et al., 2017**) to characterise the impact of seasonal variations in transmission intensity on the proportion of false-negative RDTs due to *pfhrp2*-deleted parasites. We present an extended version of our previous model, which predicts that more false-negative RDTs due to *pfhrp2* gene deletions are observed when monoclonal infections are more prevalent, with the highest proportion observed when sampling

from younger children at the start of the rainy season. We continue to assess how samples collected within an 8-week interval can both over- and underestimate this proportion when compared to the annual average, which reflects the monitoring scheme recommended by the WHO for follow up studies if the outcomes of the original study are inconclusive. Lastly, we map the administrative regions in SSA with the greatest potential for estimates of the proportion of *P. falciparum* cases with false-negative HRP2 RDT results due *to pfhrp2* deletions to be not predictive of the annual average. In addition, we identify the optimum sampling intervals for each level one administrative region, which are most representative of the annual average.

Results

Using our model, we first explored how the proportion of clinical cases only infected with *pfhrp2*-deleted parasites varies throughout a transmission season. We recorded the proportion of clinical cases that are PfHRP2-negative in four settings (a low and moderate transmission setting with both a low and highly seasonal transmission dynamic), which had a starting *pfhrp2* deletion frequency of 6%. 6% was chosen to reflect our previously estimated frequency of *pfhrp2* deletions prior to the introduction of RDTs in the Democratic Republic of the Congo (DRC) (***Watson et al., 2017***). We initially assumed that the frequency of *pfhrp2* deletions was not increasing over time before considering scenarios in which the selective pressure for *pfhrp2* deletions causes an increase in the population frequency of *phrp2* deletions. This decision allowed for the impact of seasonality on the proportion of clinical cases that are *pfhrp2*-negative to be isolated, before allowing comparisons to scenarios in which the proportion of clinical cases that are *pfhrp2*-negative is increasing also due to changes in the population frequency of *phrp2* deletions.

Our predictions suggest that the misdiagnosis of clinical cases due to *pfhrp2*-negative RDT results is heavily dependent on transmission intensity (***Figure 1***). For the same population frequency of *pfhrp2* gene deletions (***Figure 1Q–T***), the observed proportion of clinical cases that are *pfhrp2*-negative is predicted to be higher in lower transmission settings (***Figure 1I–P***). The annual average proportion of clinical cases that are *pfhrp2*-negative was equal to 5% and 3.25% in the low and moderate transmission setting, respectively. This observation is attributable to the lower rate of superinfection in low transmission settings. The lower rate of superinfection reduces the number of polyclonal infections and increases the chance that an individual is only infected with *pfhrp2*-negative parasites (***Figure 1—figure supplement 1***). When we considered scenarios with a selective advantage for *pfhrp2*-deletions (***Figure 1—figure supplement 2***), the population frequency of *pfhrp2* gene deletions increased over the two years observed (***Figure 1—figure supplement 2Q–T***) with a corresponding increase in the proportion of clinical cases that are *pfhrp2*-negative (***Figure 1—figure supplement 2I–P***).

An increased proportion of individuals only infected with *pfhrp2* gene deletions is predicted to occur at the beginning of the rainy season just before incidence starts to increase. During the rainy season, the observed proportion of cases expected to yield a false-negative RDT due to *pfhrp2*-deleted parasites (PfHRP2-negative) falls, with the lowest proportion observed after the end of the rainy season. These dynamics are more pronounced in highly seasonal transmission regions (***Figure 1B, F, J, N, R, D, H, L, P and T***). In the highly seasonal settings, the observed proportion of clinical cases that are PfHRP2-negative is predicted to fluctuate above and below the 5% threshold for switching RDT provided by the WHO (***Figure 1J, L, N and P***). Smaller fluctuations are seen in less seasonal transmission regions (***Figure 1A, E, I, M, Q, C, G, K, O and S***), with no fluctuations in the observed proportion of clinical cases that are PfHRP2-negative occurring above 5% in the moderate transmission setting (***Figure 1K and O***). Similar patterns were observed in scenarios with an increasing frequency of *pfhrp2*-deletions, with fluctuations in the proportion of clinical cases that were PfHRP2-negative observed in the highly seasonal settings (***Figure 1—figure supplement 2J, L, N and P***). The highest proportion of cases expected to yield a false-negative RDT due to *pfhrp2*-deleted parasites was still observed at the beginning of the rainy season.

The specific 8-week interval during which samples are collected is predicted to impact the observed proportion of false-negative RDTs due to *pfhrp2* gene deletions (***Figure 2***). In a moderate transmission setting, a clear seasonal pattern is predicted (***Figure 2C***), with sampling at the beginning of the transmission seasons resulting in significant overestimation of the annual average proportion of false-negative RDTs. Subsequently, sampling at the end of the rainy season is predicted to

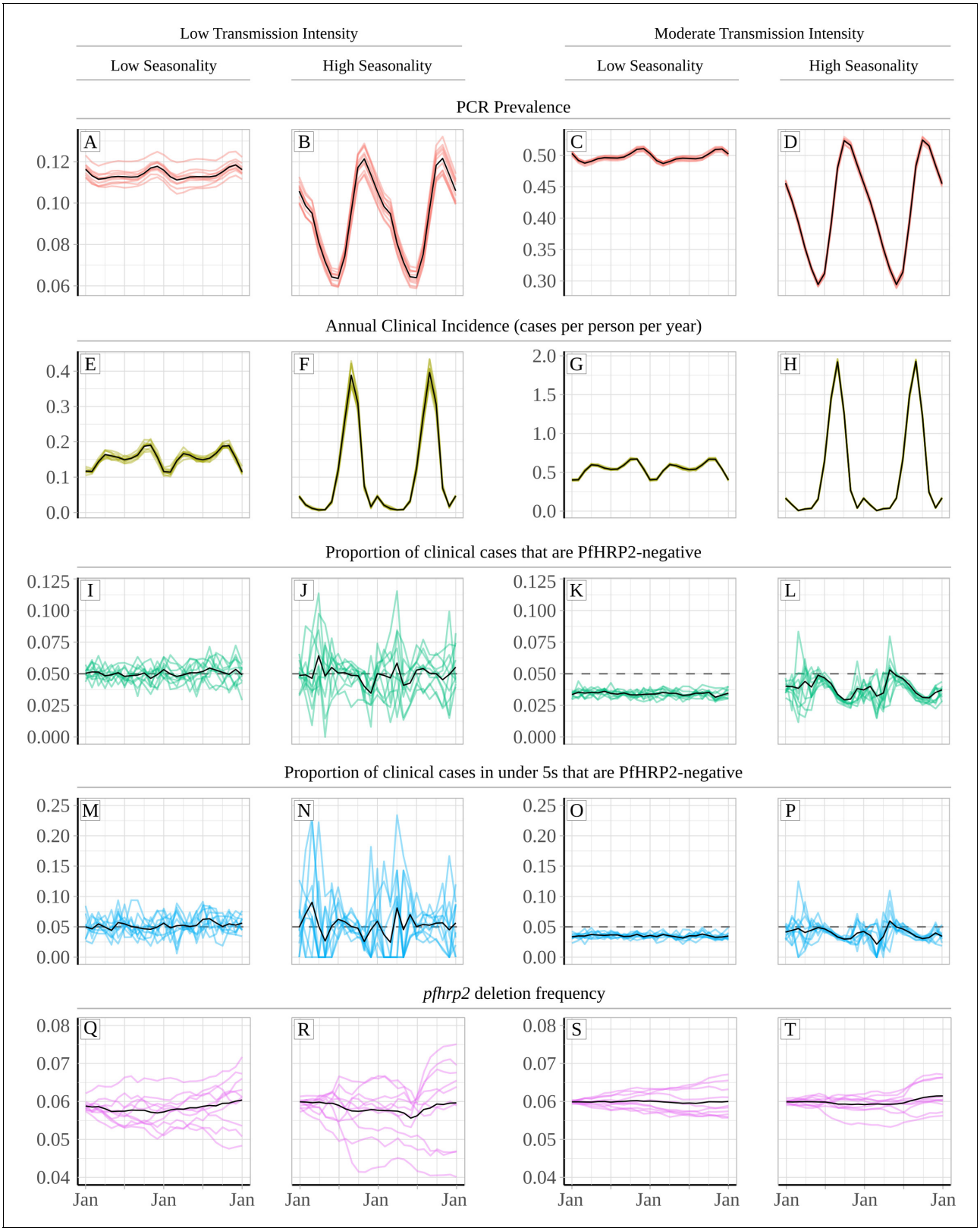


Figure 1. Relationship between seasonality, transmission intensity and proportion of clinical cases that are infected with only *pfhrp2*-deleted parasites. Graphs show in (**A – D**) and (**E - H**) the model predicted PCR prevalence and annual clinical incidence respectively at both a low and a moderate transmission intensity. In (**I – L**) and (**M - P**) the proportion of clinical cases only infected with *pfhrp2*-negative parasites is shown for both the whole population and in children under 5 years old, respectively. Lastly, graphs (**Q - T**) show the population allele frequency of *pfhrp2* gene deletions, which

Figure 1 continued on next page

Figure 1 continued

was set equal to 6% at the beginning of each simulation. 10 simulation realisations are shown in each graph, with the mean shown with by the black line. Lastly, the 5% threshold for switching RDT provided by the WHO is shown with the dashed horizontal line in plots (**I – P**).

DOI: https://doi.org/10.7554/eLife.40339.002

The following figure supplements are available for figure 1:

Figure supplement 1. Model predicted relationship between clonality of infection in asymptomatic and clinical cases against prevalence of malaria.

DOI: https://doi.org/10.7554/eLife.40339.003

Figure supplement 2. Impact of a selective advantage for *pfhrp2*-deleted parasites on the relationship between seasonality, transmission intensity and proportion of clinical cases that are infected with only *pfhrp2*-deleted parasites.

DOI: https://doi.org/10.7554/eLife.40339.004

yield estimates that are most representative of the annual average. In comparison, surveillance in regions with low seasonality is predicted to yield estimates representative of the annual average throughout the transmission season (***Figure 2B and D***). In all settings, using a sampling scheme spanning the entire transmission season produced estimates that accurately estimated the annual average. A moderate increase in the proportion of false-negative RDTs is also predicted when sampling younger individuals, with the same patterns also seen within asymptomatic individuals. This observation reflects the increased probability that children younger than 5 years old yield symptoms after the first infection, due to their comparatively lower acquired clinical immunity. Similar seasonal dynamics were observed in the highly seasonal settings when we considered scenarios with a selective advantage for *pfhrp2*-deletions (***Figure 2—figure supplement 1A and C***).

Using data from a national survey of *pfhrp2* gene deletions in the DRC, we found that the model-predicted outcomes above were similar to those observed in the field (***Figure 3***) (***Parr et al., 2017***). Among 2752 PCR-positive *P. falciparum* cases in the DRC, individuals were more likely to be infected with only *pfhrp2*-negative parasites if the clinical incidence in the month prior to sample collection was lower ($p=4.1\times10^{-6}$), and if the individuals were younger (p=0.016). These findings were maintained when comparing across age and transmission groups, with samples collected during periods of lower transmission found to be more likely to be *pfhrp2*-negative in both older and younger age groups ($p=6.6\times10^{-5}$ and 5.6×10^{-4}, respectively). Samples collected in younger individuals were more likely to be *pfhrp2*-negative in both lower and higher transmission groups when compared to older individuals (p=0.06 and 0.06, respectively).

Lastly, we predicted and mapped the potential for estimates collected within 8-week intervals to be unrepresentative of the annual average proportion of false-negative RDTs due to *phrp2* gene deletions across 598 first administrative regions in SSA (***Figure 4***). We predict that 66 regions possess at least one 8-week interval for which a premature switch to a non PfHRP2-based RDT would have been made in more than 75% of simulations (***Figure 4A***) and 29 regions are predicted to possess at least one 8-week interval for which a premature decision to continue using PfHRP2-based RDTs would have been made in more than 75% of simulations (***Figure 4B***). Out of these 29 regions, 25 are also present within the formerly identified 66 regions. The data for each administrative region can be viewed online at the following interactive database https://shiny.dide.imperial.ac.uk/seasonal_hrp2/.

Discussion

This research characterises the potential for surveillance in highly seasonal areas within sub-Saharan Africa to produce estimates that fail to represent the annual average proportion of *P. falciparum* cases with false-negative HRP2 RDT results due to *pfhrp2* deletions. These findings highlight the impact of both the seasonal timing and the age of individuals sampled when estimating the proportion of false-negative RDTs due to *pfhrp2* deletions. Policy decisions based on the proportion of clinical cases presenting with false-negative RDTs due to *pfhrp2* gene deletions should thus be made with an awareness of the seasonal transmission dynamics of the region considered.

Our modelling predicted that there would be increased observation of false-negative HRP2 RDT results after periods of lower transmission and within younger individuals. This prediction is consistent with a large, nationally representative survey of *pfhrp2*-negative samples among asymptomatic subjects in the DRC (***Parr et al., 2017***). These predictions are also in agreement with other

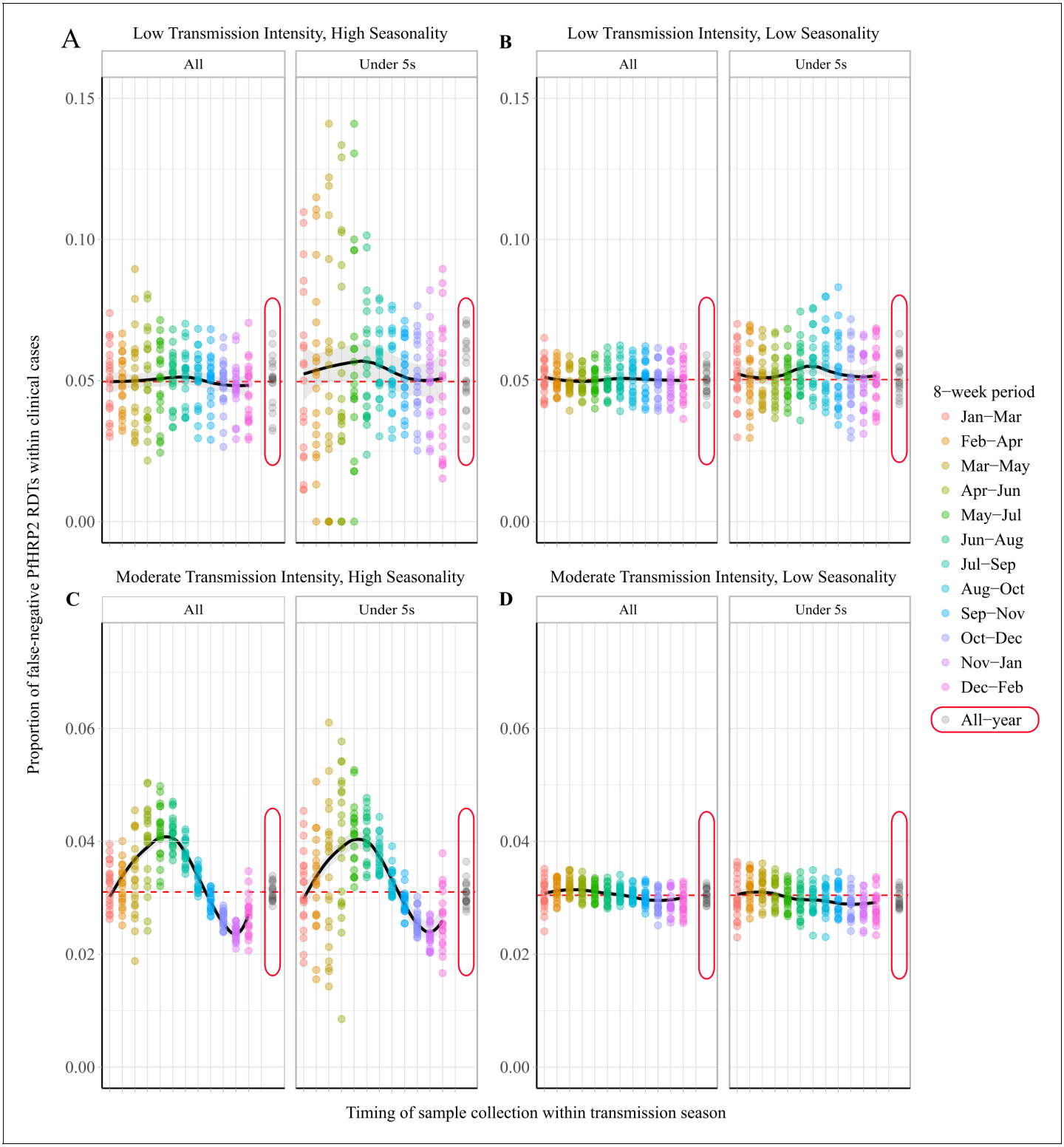


Figure 2. Observed proportion of false-negative PfHRP2 RDTs within clinical cases during 8-week intervals. Graphs show the proportion of clinical cases yielding false-negative PfHRP2 RDTs at 8-week intervals within a transmission season for both moderate (**C, D**) and low (**A, B**) transmission settings and high (**A, C**) and low (**B, D**) seasonality. In each panel, the observed proportion *pfhrp2*-negative clinical cases is shown for the whole population and within children aged under 5 years old. Ten stochastic realisations are represented by the points in each plot, with the mean relationship throughout the transmission shown in black with a locally weighted scatterplot smoothing regression (loess). The annual average proportion of false-negative RDTs

Figure 2 continued on next page

Figure 2 continued

due to *pfhrp2* gene deletions is shown with the horizontal dashed red line, and a sampling scheme that occurs throughout the year, with samples collected proportionally to clinical incidence, is shown with grey points circled in red.

DOI: https://doi.org/10.7554/eLife.40339.005

The following figure supplement is available for figure 2:

Figure supplement 1. Impact of a selective advantage for *pfhrp2*-deleted parasites on the observed proportion of false-negative PfHRP2 RDTs within clinical cases during 8-week intervals.

DOI: https://doi.org/10.7554/eLife.40339.006

observations from Dioro in the Ségou region of Mali, where in 2012 more than 80% of smear-positive individuals had false-negative RDTs when collected at the end of the dry season (***Koita et al., 2013***). The proportion of false-negative RDTs then rapidly decreased to 20% within 3–4 weeks after the start of the rainy season. It is, however, likely that a proportion of these false-negative RDTs were due to the increased observation of lower parasitaemia at lower transmission intensities such as at the end of the dry season (***Okell et al., 2012***). In addition, findings from Eritrea also support our model-predicted outcomes. Eritrea is a region with lower malaria prevalence compared to the Ségou region of Mali. The resultant decrease in transmission intensity is likely to result in an increased proportion of monoclonal infections throughout the transmission season. Consequently, we would predict less variability in the number of false-negative RDTs due to *pfhrp2* gene deletions at any given period within a transmission season. We also expect the observed prevalence of *pfhrp2* deletions to be more stochastic due to the lower effective population size of the parasite. Indeed, infections due to *pfhrp2*-deleted parasites identified in Eritrea between November 2013 and November 2014 were not more likely to have occurred after periods of lower transmission intensity (p=0.56, n=144, pfhrp2 deletions at 9.7%) (***Menegon et al., 2017***).

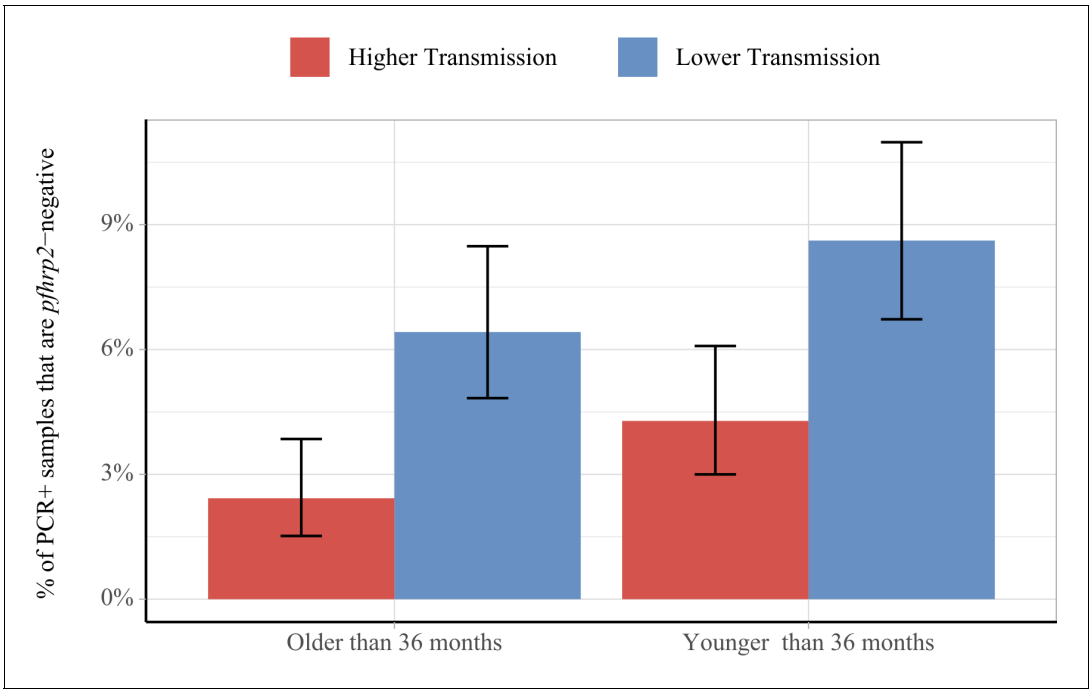


Figure 3. Impact of age and transmission intensity upon *pfhrp2* deletion in the Democratic Republic of the Congo (DRC), 2013–2014. Graphs show the percentage of PCR-positive *P. falciparum* samples taken from children under the age of 5 years from the 2013–2014 Demographic and Health Survey in DRC that are *pfhrp2*-negative. Children who are younger than the median age in the 2752 samples are grouped within the younger category. In addition, samples are classified as lower transmission if the incidence of malaria in the month prior to sample collection is lower than the median clinical incidence. The 95% binomial confidence intervals are indicated with the vertical error bars.

DOI: https://doi.org/10.7554/eLife.40339.007

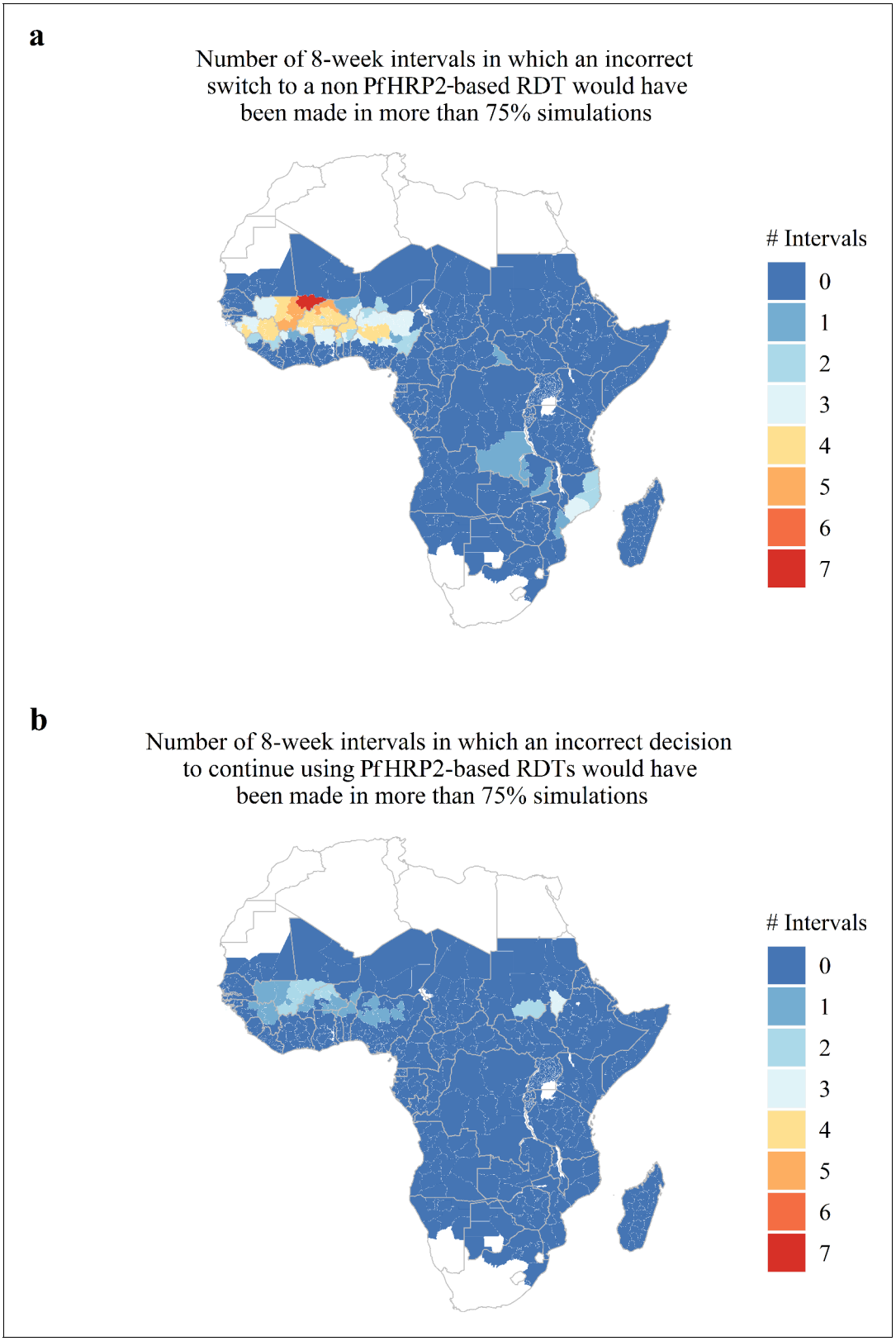


Figure 4. Predicted areas with the potential for collected estimates of the proportion of false-negative PfHRP2 RDTs due to *pfhrp2* deletions to be unrepresentative of the annual average. The maps show (**A**) the number of 8-week intervals at which an administrative region would prematurely swap to a non PfHRP2-based RDT due to overestimating the proportion of false-negative PfHRP2 RDTs due to *pfhrp2* gene deletions in more than 75% of simulations. In (**A**) the opposing trend is shown, with the number of 8-week intervals at which an administrative region would prematurely continue to

Figure 4 continued on next page

Figure 4 continued

use PfHRP2-based RDTs due to underestimating the proportion of false-negative PfHRP2 RDTs due to *pfhrp2* gene deletions in more than 75% of simulations.

DOI: https://doi.org/10.7554/eLife.40339.008

Similar to the original publication (***Watson et al., 2017***), there are a number of modelling assumptions in this study. Firstly, there are modelling uncertainties when predicting the dynamics of false-negative RDTs due to *pfhrp2*-deleted parasites. To account for this uncertainty in this analysis, we have controlled for the drivers characterised in our earlier study by assuming there was no selective advantage associated with *pfhrp2*-deleted parasites and recording the number of individuals who would have been *pfhrp2*-negative and subsequently misdiagnosed. The absence of a selective advantage in this way enabled the frequency of *pfhrp2* deletions to remain constant, which ensured that any observed dynamics in the estimates of false-negative RDTs due to *pfhrp2* deletions were due to the seasonality of transmission and not due to an increase in the population frequency of *pfhrp2* deletions. However, we are aware that there is likely a selective advantage for *pfhrp2* deleted parasites and subsequently we repeated the analyses with the selective advantage included. In these simulations, we predicted a substantial increase in the frequency of *pfhrp2* gene deletions (***Figure 1—figure supplement 2Q-T***), however clear seasonal dynamics, with an increased proportion of false-negative RDTs due to *pfhrp2* deletions at the beginning of the transmission season, were still observed (***Figure 2—figure supplement 1C***). However, the observed dynamics were less clear in settings with the greatest increase in the frequency of *pfhrp2* deletions (***Figure 2—figure supplement 1B***).

Secondly, we assessed the potential for a region to yield unrepresentative estimates of the proportion of false-negative RDTs due to *pfhrp2* deletions through comparisons to the annual average proportion. This decision reflected firstly the monitoring period defined in the WHO technical guidance, with follow-up studies recommended after two years if the 95% CI for the proportion of *P. falciparum* cases with false-negative HRP2 RDT results due to *pfhrp2/3* deletions is less than 5%, or one year if it does include 5%. It also reflected our modelling assumption that the population frequency of *pfhrp2* deletions is not increasing over time. However, in simulations in which a selective advantage to *pfhrp2* deleted parasites was included, a comparison to the annual average proportion is less suitable. For example, in ***Figure 2—figure supplement 1B***, because we started our simulations in January the optimum sampling interval is simply the interval in the middle of the year, reflecting the constant increase in *pfhrp2* deleted parasites. In these scenarios, it could be argued that the correct comparison would be to the average proportion of false-negative RDTs due to *pfhrp2/3* gene deletions in the year after sampling, which reflects how many cases could be misdiagnosed between sampling rounds. Unfortunately, this comparison is difficult without knowing how the proportion of false-negative RDTs due to *pfhrp2/3* gene deletions will change over time. However, we believe that it is more important to focus on the assumption that the strength of selection is negligible (see ***Figure 5***). Our rationale for this is that it is only in areas with a low selective pressure, for which the frequency of *pfhrp2/3* deletions is constant over time, that one could repeatedly make an incorrect decision with regards to whether to switch RDT (***Figure 5A***). In areas with a selective pressure, it is still possible to incorrectly estimate the annual average for the following year; however, the presence of the selective pressure is likely to cause any decision made to be simply premature as the frequency of *pfhrp2/3* deletions and subsequently false-negative PfHRP2 RDTs will increase over time (***Figure 5B***).

Lastly, it is important to note again that the true strength of selection is unknown. The precise strength of selection is dependent on a number of factors such as the magnitude of any fitness costs associated with *pfhrp2* deletion, the degree to which microscopy-based diagnosis is used, the level of non-adherence to RDT results, the treatment coverage and the prevalence of malaria in the region considered. Consequently, our results should not be interpreted as precise predictions of how unrepresentative future samples may be. They should instead be used to support surveillance efforts and to reinforce the need for longitudinal measures conducted at the same point within a transmission season. In addition, we recommend that if possible, sample collection in highly seasonal regions should not occur at the beginning of the transmission season, as this is predicted to lead to

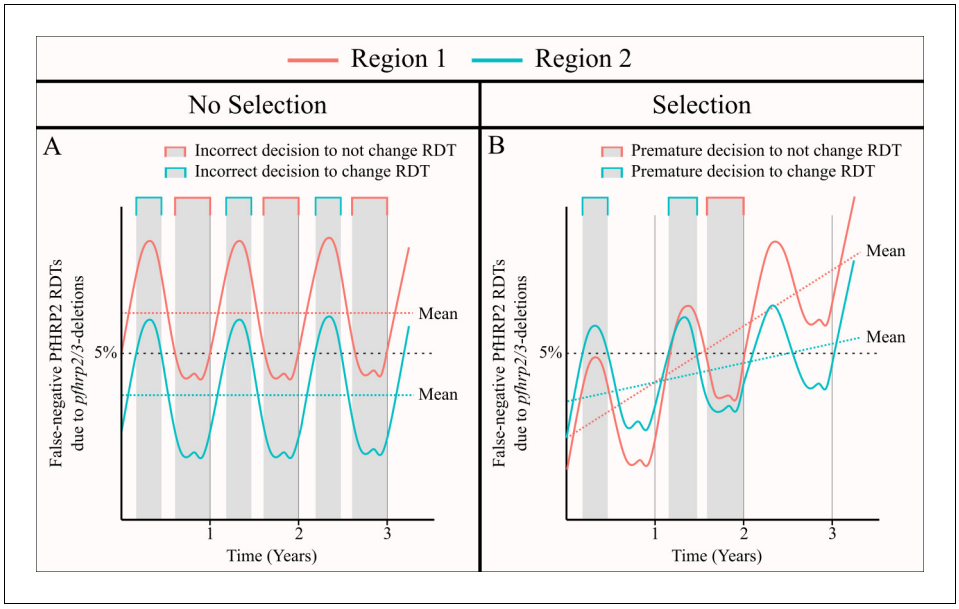


Figure 5. The impact of an assumed selective pressure for *pfhrp2/3*-deleted parasites on the decision to switch RDT. The graphs show two hypothetical scenarios with two different regions shown in red and blue for each region. In (**A**) there are strong seasonal dynamics but no selective pressure. The absence of a selective pressure causes that the mean proportion of false-negative RDTs due to *pfhrp2/3* deletions over a 1 year period to be constant and is shown with a horizontal dashed line. Consequently, there are time periods in which an incorrect decision to switch RDT could be made for the region in blue, and an incorrect decision to not switch RDT could be made for the region in red. In (**B**), there are both seasonal dynamics and a selective pressure, which results in an increasing annual mean proportion of false-negative RDTs due to *pfhrp2/3* deletions over time. As in (**A**), there are periods in which the observed proportion of false-negative RDTs due to *pfhrp2/3* deletions is both higher and lower than the rolling mean shown. However, decisions made in these periods are premature rather than definitively incorrect as the selection pressure would eventually cause the proportion to be greater than 5%.
DOI: https://doi.org/10.7554/eLife.40339.009

premature decisions to switch RDT irrespective of the strength of selection. It will, however, be possible after the samples have been collected to estimate the likely frequency of *pfhrp2* gene deletions by incorporating estimates of the multiplicity of infection within the sampled population. This frequency could then be used to estimate how the proportion of false-negative RDT results due to *pfhrp2* deletions could increase in response to decreases in the prevalence of malaria.

In summary, our extended model predicts that highly seasonal dynamics in malaria transmission intensity will cause comparable dynamics in the observed proportion of false-negative RDT results due to *pfhrp2* gene deletions. The observed proportion of false-negative RDTs due to *pfhrp2* deletions is higher when monoclonal infections are more prevalent, with the highest prevalence observed when sampling at the start of the rainy season as individuals are less likely to already be infected. Similarly, the observed proportion of false-negative RDTs due to *pfhrp2* deletions is higher in younger individuals who have lower clinical immunity, as they are more likely to present with clinical symptoms after their first infection event. As the rainy season progresses, individuals are more likely to be superinfected and acquire wild-type parasites, resulting in positive PfHRP2-based RDT results and a decrease in the observed proportion of false-negative RDTs due to *pfhrp2* deletions. In response to these dynamics, it may be sensible for national malaria control programmes conducting surveillance for *pfhrp2/3* deletions to choose a sampling interval towards the end of the transmission season, which is predicted to be most representative of the annual average proportion of false-negative RDTs due to *pfhrp2* deletions. To support surveillance efforts, we have published an online database detailing the optimum sampling interval as well as the fluctuations throughout the transmission season for each administrative region.

Materials and methods

Extensions to the *P. falciparum* transmission model

In our previous publication (**Watson et al., 2017**), we presented an extended version of an individual-based model of malaria transmission to characterise the key drivers of *pfhrp2* deletion selection; however, it did not capture seasonality. To address this, we incorporated seasonal variation in malaria transmission intensity through the inclusion of seasonal curves fitted to daily rainfall data available from the US Climate Prediction Center (**National Weather Service Climate Prediction Center, 2010**). Rainfall data was available at a 10 × 10 km spatial resolution from 2002 to 2009, with data missing for only two days. The data was subsequently aggregated to a series of 64 points per year, before Fourier analysis was conducted to capture the seasonal dynamics within this time period (**Cairns et al., 2012**). The first three frequencies of the resultant Fourier transformed data were used to generate a normalised seasonal curve. This inclusion alters the rate at which new adult mosquitoes are born, with the differential equation governing the susceptible adult stage of the mosquito population now given by:

$$\frac{dS_M}{dt} = \theta(t)\mu_M M_v - \mu_M S_M - \Lambda_M S_M$$

where μ_M is the daily death rate of adult mosquitoes, M_v is the total mosquito population, that is $S_M + E_M + I_M$, Λ_M is the force of infection on the mosquito population and $\theta(t)$ is the normalised seasonal curve, with a period equal to 365 days. The rest of the model equations remain the same as in our original study (**Watson et al., 2017**).

All extensions to the previous model code have been made using the R language (RRID:SCR_001905) (**R Development Core Team, 2016**) and are available through an open source MIT licence at https://github.com/OJWatson/hrp2malaRia (**Watson, 2019**; copy archived at https://github.com/eLifeProduction/hrp2malaRia_2019). In addition, these extensions have been included in the pseudo-code description of the model (**Supplementary file 1**).

Characterising the impact of seasonal transmission intensities upon the proportion of false-negative RDTs due to *pfhrp2* gene deletions

The impact of seasonality was examined by recording the proportion of clinical incidence that would have been misdiagnosed due to *pfhrp2* gene deletions across the year. This proportion was summarised at 12 8-week intervals, that is January – March, February – April, December – February. This proportion was recorded in both high and low seasonality settings, characterised by a Markham Seasonality Index = 80% and 10%, respectively (**Cairns et al., 2015**). These settings were examined at both low and moderate transmission intensity (EIR = 1 and 10 respectively), with the starting proportion of *pfhrp2*-deleted parasites in the whole population set equal to 6% in agreement with previous observations of *pfhrp2* gene deletions in the DRC (**Watson et al., 2017**) The proportion of symptomatic cases seeking treatment was assumed to be 40% ($f_T$pfhrp2 = 0.4). In all simulations, 10 stochastic realisations of 100,000 individuals were simulated for 60 years to reach equilibrium first, before setting the frequency of *pfhrp2* deletions. Initially, we assumed there was no assumed fitness cost or selective advantage associated with *pfhrp2* gene deletion. This was modelled by assuming that individuals who are only infected with parasites with *pfhrp2* gene deletions will still be treated. This decision allowed us to control for selection within our investigation by ensuring that the changes observed in the observation of PfHRP2-negative clinical cases are only due to seasonal variation in transmission intensity, and not due to an increase in the frequency of *pfhrp2* gene deletions due to the selective advantage by evading diagnosis. As a result, when reporting the proportion of clinical cases that were misdiagnosed resulting from a false-negative PfHPR2-negative RDT we are reporting the proportion of cases that are infected with only *pfhrp2*-deleted parasites, that is individuals who would have been *pfhrp2*-negative and subsequently misdiagnosed. We also assume that 25% of individuals who are only infected with *pfhrp2*-deleted parasites will still be *pfhrp2*-positive due to the cross reactivity of PfHRP3 epitopes causing a positive PfHRP2-based RDT result (**Baker et al., 2005**).

Model predictions were subsequently compared to data collected from the Democratic Republic of Congo as part of their 2013–2014 Demographic and Health Survey (DHS). In overview, 7137 blood samples were collected from children under the age of 5 years old, which yielded 2752

children diagnosed with *P. falciparum* infection by real-time PCR targeting the lactate dehydrogenase (*pfldh*) gene. The RDT barcodes for the 2752 samples were identified and matched to the DHS survey to identify both the age of the children and the date of sample collection. The collection date was used to predict the mean clinical incidence from the previous 30 days for each sample. This was estimated using the deterministic implementation of our model fitted to the observed PCR prevalence of malaria from the DRC DHS 2013–2014 survey (**Meshnick et al., 2013**), incorporating the seasonality and treatment coverage for each province. Children who were younger than the median age in the 2752 samples were grouped within a younger category. In addition, samples were classified as lower transmission if the clinical incidence of malaria in the month prior to sample collection was lower than the median clinical incidence. The counts of *pfhrp2*-negative samples within each group were subsequently compared using the Pearson chi-squared test with Rao-Scott corrections to account for the hierarchal survey design implemented within DHS surveys (**Jnk and Scott, 1984**). Pearson chi-squared tests were used in a similar analysis that was conducted using samples collected from the Gash-Barka and Debug regions in Eritrea between 2013 and 2014, for which the dates of sample collection were made available to us (**Menegon et al., 2017**).

Finally, the seasonal profiles for 598 first-level administrative regions across sub-Saharan Africa were used to characterise the potential for estimates of the proportion of false-negative PfHRP2 RDTs due to *pfhrp2* gene deletions to be unrepresentative of the annual average. For each region, 100 simulation repetitions were conducted for 60 years to reach equilibrium first before fitting the frequency of *pfhrp2* gene deletions in each simulation such that the annual average proportion of false-negative RDT results due to *pfhrp2* deletions is equal to 5%. Each repetition was subsequently simulated for two further years, with 7300 individuals seeking treatment sampled from each 8-week interval. This number approximates the recommended sample size within the WHO protocol for *pfhrp2* deletion prevalence at 5 ± 0.5%. For each sample, the proportion of false-negative PfHRP2-based RDTs due to *pfhrp2* gene deletions was recorded. For each sample, a binomial confidence interval was calculated and the resultant percentage of intervals that did not include the annual prevalence of 5% was calculated. For each region, the number of 8-week intervals for which a premature decision to either swap from a PfHRP2-based RDT or continue using a PfHRP2-based RDT was made in more than 75% of simulations was recorded and mapped. The raw results of this analysis were subsequently used to create a database that details the optimum sampling intervals for estimating the annual proportion of false-negative RDT results due to *pfhrp2* deletions.

Acknowledgements

We thank Dr Michela Menegon for sharing the sample dates for the 2013–2014 study in Eritrea and the administrators and participants of the Demographic and Health Surveys.

Additional information

Funding

Funder	Grant reference number	Author
Wellcome Trust	109312/Z/15/Z	Oliver John Watson
Medical Research Council	MR/N01507X/1	Robert Verity
Department for International Development		Azra C Ghani
Medical Research Council		Tini Garske
National Institute of Allergy and Infectious Diseases	R01AI132547	Steven R Meshnick
American Society for Tropical Medicine and Hygiene	Burroughs Wellcome Fund-ASTMH Postdoctoral Fellowship in Tropical Infectious Diseases	Jonathan B Parr

Burroughs Wellcome Fund	Burroughs Wellcome Fund-ASTMH Postdoctoral Fellowship in Tropical Infectious Diseases	Jonathan B Parr
Imperial College London		Hannah C Slater

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Oliver John Watson, Conceptualization, Data curation, Software, Formal analysis, Investigation, Methodology, Writing—original draft, Project administration, Writing—review and editing; Robert Verity, Formal analysis, Supervision, Visualization, Methodology, Writing—review and editing; Azra C Ghani, Jane Cunningham, Supervision, Methodology, Writing—review and editing; Tini Garske, Data curation, Software, Methodology, Writing—review and editing; Antoinette Tshefu, Melchior K Mwandagalirwa, Steven R Meshnick, Resources, Formal analysis, Writing—review and editing; Jonathan B Parr, Conceptualization, Data curation, Formal analysis, Supervision, Methodology, Writing—review and editing; Hannah C Slater, Conceptualization, Formal analysis, Supervision, Investigation, Visualization, Methodology, Writing—review and editing

Author ORCIDs

Oliver John Watson https://orcid.org/0000-0003-2374-0741

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.40339.016
Author response https://doi.org/10.7554/eLife.40339.017

Additional files

Supplementary files

- Supplementary file 1. Simulation model pseudocode. Mathematical style pseudocode description of the simulation model.

DOI: https://doi.org/10.7554/eLife.40339.010

- Transparent reporting form

DOI: https://doi.org/10.7554/eLife.40339.011

Data availability

All data generated are provided within the online database, hosted through a shiny application at https://ojwatson.shinyapps.io/seasonal_hrp2/. The raw data for the application is available within the GitHub repository at https://github.com/OJWatson/hrp2malaRia (copy archived at https://github.com/eLifeProduction/hrp2malaRia_2019).

The following previously published dataset was used:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Bhatt S, Weiss DJ, Cameron E, Bisanzio D	2015	PfPR2-10 in Africa 2000-2015	https://map.ox.ac.uk/explorer/#/explorer	Malaria Atlas Project, explorer

References

Baker J, McCarthy J, Gatton M, Kyle DE, Belizario V, Luchavez J, Bell D, Cheng Q. 2005. Genetic Diversity of *Plasmodium falciparum* Histidine-Rich Protein 2 (PfHRP2) and Its Effect on the Performance of PfHRP2-Based Rapid Diagnostic Tests . *The Journal of Infectious Diseases* **192**:870–877. DOI: https://doi.org/10.1086/432010

Berhane A, Anderson K, Mihreteab S, Gresty K, Rogier E, Mohamed S, Hagos F, Embaye G, Chinorumba A, Zehaie A, Dowd S, Waters NC, Gatton ML, Udhayakumar V, Cheng Q, Cunningham J. 2018. Major threat to malaria control programs by *Plasmodium falciparum* Lacking Histidine-Rich Protein 2, Eritrea. *Emerging Infectious Diseases* **24**:462–470. DOI: https://doi.org/10.3201/eid2403.171723, PMID: 29460730

Cairns M, Roca-Feltrer A, Garske T, Wilson AL, Diallo D, Milligan PJ, Ghani AC, Greenwood BM. 2012. Estimating the potential public health impact of seasonal malaria chemoprevention in african children. *Nature Communications* **3**:1–9. DOI: https://doi.org/10.1038/ncomms1879

Cairns ME, Walker PG, Okell LC, Griffin JT, Garske T, Asante KP, Owusu-Agyei S, Diallo D, Dicko A, Cisse B, Greenwood BM, Chandramohan D, Ghani AC, Milligan PJ. 2015. Seasonality in malaria transmission: implications for case-management with long-acting artemisinin combination therapy in sub-Saharan africa. *Malaria Journal* **14**:1–13. DOI: https://doi.org/10.1186/s12936-015-0839-4, PMID: 26283418

Cheng Q, Gatton ML, Barnwell J, Chiodini P, McCarthy J, Bell D, Cunningham J. 2014. Plasmodium falciparum parasites lacking histidine-rich protein 2 and 3: a review and recommendations for accurate reporting. *Malaria Journal* **13**:283. DOI: https://doi.org/10.1186/1475-2875-13-283, PMID: 25052298

Gatton ML, Dunn J, Chaudhry A, Ciketic S, Cunningham J, Cheng Q. 2017. Implications of parasites lacking plasmodium falciparum Histidine-Rich protein 2 on malaria morbidity and control when rapid diagnostic tests are used for diagnosis. *The Journal of Infectious Diseases* **215**:1156–1166. DOI: https://doi.org/10.1093/infdis/jix094, PMID: 28329034

Jnk R, Scott AJ. 1984. On Chi-Squared tests for multiway contingency tables with cell proportions estimated from survey data. *Annals of Statistics* **12**:46–60. DOI: https://doi.org/10.1214/aos/1176346391

Koita OA, Ndiaye J-L, Nwakanma D, Sangare L, Ndiaye D, Joof F. 2013. Seasonal changes in the frequency of false negative rapid diagnostic tests based on histidine rich protein 2 (HRP2). *The American Journal of Tropical Medicine and Hygiene* **89**:1.

Menegon M, L'Episcopia M, Nurahmed AM, Talha AA, Nour BYM, Severini C. 2017. Identification of plasmodium falciparum isolates lacking histidine-rich protein 2 and 3 in Eritrea. *Infection, Genetics and Evolution* **55**:131–134. DOI: https://doi.org/10.1016/j.meegid.2017.09.004

Meshnick S, Janko M, Thwai K, Levitz L, Emch M. 2013. *Demographic And Health Survey (DRC-DHSII) Supplemental Malaria Report 2015*: Adrian Pingstone. https://dhsprogram.com/pubs/pdf/FR300/FR300.Mal.pdf.

National Weather Service Climate Prediction Center. 2010. Climate Prediction Center - International Desks . https://www.cpc.ncep.noaa.gov/products/international/ [Accessed July 1, 2010].

Okell LC, Bousema T, Griffin JT, Ouédraogo AL, Ghani AC, Drakeley CJ. 2012. Factors determining the occurrence of submicroscopic malaria infections and their relevance for control. *Nature Communications* **3**:1237. DOI: https://doi.org/10.1038/ncomms2241, PMID: 23212366

Parr JB, Verity R, Doctor SM, Janko M, Carey-Ewend K, Turman BJ, Keeler C, Slater HC, Whitesell AN, Mwandagalirwa K, Ghani AC, Likwela JL, Tshefu AK, Emch M, Juliano JJ, Meshnick SR. 2017. Pfhrp2-Deleted plasmodium falciparum parasites in the democratic republic of the Congo: a national Cross-sectional survey. *The Journal of Infectious Diseases* **216**. DOI: https://doi.org/10.1093/infdis/jiw538, PMID: 28177502

R Development Core Team. 2016. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.r-project.org/

Watson OJ, Slater HC, Verity R, Parr JB, Mwandagalirwa MK, Tshefu A, Meshnick SR, Ghani AC. 2017. Modelling the drivers of the spread of *plasmodium falciparum hrp2* gene deletions in sub-Saharan Africa. *eLife* **6**:e25008. DOI: https://doi.org/10.7554/eLife.25008, PMID: 28837020

Watson OJ. 2019. *hrp2malaRia. GitHub*. 4d992a0. https://github.com/OJWatson/hrp2malaRia

World Health Organization. 2018a. *World Malaria Report*: World Health Organization. https://www.who.int/malaria/publications/world-malaria-report-2018/en/.

World Health Organization. 2018b. Malaria threats map. http://www.who.int/malaria/maps/threats-about/en/ [Accessed May 1, 2018].

World Health Organization. 2018c. Protocol for estimating the prevalence of pfhrp2/pfhrp3 gene deletions among symptomatic falciparum patients with false-negative RDT results. http://www.who.int/malaria/publications/atoz/hrp2-deletion-protocol/en/ [Accessed March 16, 2018].