Measuring low and unstable malaria transmission in Ethiopia: strategies for malaria surveillance and epidemic detection

RUTH ANNE ASHTON

Thesis submitted in accordance with requirements for the degree of Doctor of Philosophy of the University of London

September 2014

Department of Disease Control
Faculty of Infectious and Tropical Diseases
London School of Hygiene and Tropical Medicine

Funded by the John-Henry Credland Memorial Foundation. The United States Agency for International Development, Presidents' Malaria Initiative provided funding to Malaria Consortium to conduct the operational research activities in Ethiopia (Cooperative Agreement 663-A-00-09-00404-00) upon which this thesis is based.

DECLARATION BY CANDIDATE

I, Ruth Ashton, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:  _____ Date:23/09/14_____

Ruth Ashton

ABSTRACT

In Ethiopia, malaria transmission is seasonal and epidemic-prone, with both *Plasmodium falciparum* and *Plasmodium vivax* being endemic. Such spatial and temporal clustering of malaria only serves to underscore the importance of regularly collecting up-to-date malaria surveillance data to inform decision-making in malaria control and improve responsiveness to potential epidemics.

This thesis compares indicators and strategies used for the monitoring and surveillance of malaria in Ethiopia. Cross-sectional school-based surveys were conducted throughout Oromia Regional State, generating data on malaria prevalence by microscopy, risk factors for infection and intervention use. Filter paper blood samples collected during these school surveys were subsequently tested to determine exposure to malaria based on presence of anti-*Plasmodium* antibodies, and Bayesian geostatistical modelling was employed to predict *P. falciparum* and *P. vivax* seroprevalence across Oromia. In southern Ethiopia, a school-based syndromic surveillance system was piloted, exploring the utility of school absenteeism as a complementary indicator of malaria epidemics at community level. Finally, findings from the school surveys, measured and modelled seroprevalence, as well as data from the national Malaria Indicator Survey in 2011 were compared with spatially congruent estimates of malaria incidence collected from health facilities and to modelled parasite rate from the Malaria Atlas Project.

Findings from this thesis demonstrate the limitations of microscopy as a primary indicator of malaria infection in cross-sectional surveys in areas of very low transmission. The work highlights the potential of serological indicators of *Plasmodium* exposure for inclusion in periodic large-scale malaria monitoring activities and develops a first ever geostatistical risk map based on serological indictors. This was supported by comparative analysis of a range of survey and modelling indicators against estimates of incidence from passive surveillance,

indicating the inadequacy of cross-sectional surveys estimating population parasitaemia to reflect the spatial extent and temporal variability of transmission. The piloted syndromic surveillance system indicates that monitoring school absenteeism has potential as a complementary epidemic alert system, operating alongside the existing system at health posts, but is limited by low school enrolment in the piloted setting.

The findings of this thesis indicate that existing periodic monitoring strategies and tools are insufficient to fully describe the extent of malaria in settings where *Plasmodium* transmission is spatially and temporally variable. Modifications to monitoring strategies are recommended, including incorporation of serological indicators and spatial modelling.

.

ACKNOWLEDGEMENTS

I would firstly like to thank my supervisor Simon Brooker for his guidance over the years prior to and during my PhD. Thank also to Richard Reithinger, Chris Drakeley and Rachel Pullan for their valuable contributions as part of my advisory committee.

Jan Kolaczinksi, Sylvia Meek, James Tibenderana and Prudence Hamade have provided support, encouragement and fresh perspectives as I've grown within Malaria Consortium, and I'm most grateful to have had the opportunity to pursue a PhD alongside my work.

I'd like to express my gratitude to the entire Malaria Consortium team in Ethiopia, including Agonafer Tekalegne, Gezahegn Tesfaye, Zelalem Kebede, Esey Batisso and Tessema Awano, but not forgetting all of the support staff who keep things running behind the scenes. Special mention must go to Takele Kefyalew for his unwavering dedication to this research, and his calm and adept handling of the setbacks we faced along the way.

I'm indebted to collaborators from Oromia Regional Health Bureau and the Southern Nations, Nationalities and People's Regional State Health Bureau, for their oversight and facilitation of study activities as well as kindly sharing data for analysis.

Thanks to Ali Rand for support in the laboratory both in London and Addis Ababa, and the malaria team at the Ethiopian Health and Nutrition Research Institute for serological analysis and microscopy quality assurance. The Ethiopian Health and Nutrition Research Institute also kindly granted permission for inclusion of data from the 2011 Malaria Indicator Survey data in the comparative analysis presented in Chapter 5.

I am very grateful of the support provided by the John-Henry Memorial Fund for my PhD studies, and to the United States Agency for International Development, Presidents' Malaria Initiative, who provided funding to Malaria Consortium to conduct the operational research

activities in Ethiopia upon which this thesis is based. Thanks to Gunawardena Dissanayake, Sheleme Chibsa, and Hiwot Teka of USAID/PMI and Joe Malone of CDC for their generous contributions of time, trust and enthusiasm for these studies.

From LSHTM, I'd like to thank Kate Halliday and Jenny Smith for their camaraderie in times both good and bad. Many more people have supported me at various points over the years with a friendly ear or place to stay when travelling. And finally, thanks to David for his patience, rationality, and willingness to come along for the next adventure.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

1.1 BACKGROUND AND CONTEXT

Investment in malaria prevention and control in Africa has increased dramatically since the Abuja declaration of 2000, in recognition of the ongoing detrimental effect of *Plasmodium* infection on public health [1]. Scale up in coverage of insecticide-treated mosquito nets, renewed interest in indoor residual spraying using alternative insecticides, as well as the availability of effective treatment for *Plasmodium falciparum* malaria using artemisinin combination treatments are all contributing to reductions in malaria morbidity and mortality in Africa [2-5].

Both the increased investments and documented reduction in malaria morbidity and mortality have prompted recognition of the need to invest in surveillance, monitoring and evaluation to track changes in malaria burden and progress towards key targets [6]. Monitoring and surveillance needs are particularly high in the 34 countries targeting malaria elimination [7] since the aim is to break transmission, therefore, there is a need to identify *Plasmodium* infections rather than just clinical cases, and indicators should have fine spatial and temporal resolution [6]. As the 64 countries [7] currently in the 'controlling malaria' phase successfully transition from high to low malaria endemicity, the need for evidence-based and appropriate monitoring and surveillance tools for low and unstable transmission settings will only increase.

Malaria epidemiology in areas of both recently reduced and historically low transmission is diverse and complex, presenting unique challenges in targeting interventions to locations or populations at greatest risk, diagnosing *Plasmodium* infections, monitoring malaria control programme impact and developing effective surveillance systems for timely epidemic detection.

Periodic surveys such as the Demographic and Health Surveys (DHS), as well as Multiple Cluster Indicator Surveys (MICS) and Malaria Indicator Surveys (MIS) are valuable in monitoring access to and use of key malaria interventions, as well as the reported burden of malaria and number of malaria-attributable deaths. These tools were developed at a time when many malaria-endemic countries did not routinely generate nationally representative indicators, and there was a desire to develop standardised but adaptable tools for monitoring and evaluation across countries to track performance of malaria control programmes [8]. Suggested approaches to data collection in the Roll Back Malaria framework included the routine health information system, demographic surveillance systems, community surveys, health facility surveys and review of existing documents. While large-scale household surveys are useful to gather data from the community on knowledge of malaria, access to diagnosis and treatment as well as ownership and use of mosquito nets, the justification for inclusion of malariometric indicators, particularly focussing on children less than five years of age and pregnant women, is equivocal for low transmission settings [9].

Epidemiological surveillance, particularly in areas of low transmission, requires survey methodologies sufficiently powered to measure the extent of malaria transmission and parasitaemia within the population. Sampling strategies used in DHS, MICS and MIS are able to generate nationally-representative estimates of parasite prevalence, but are severely limited in low transmission settings by the use of light microscopy or rapid diagnostic test (RDT), due to poor sensitivity in detection of low density *Plasmodium* infections [10] and difficulties in capturing the temporally dynamic nature of malaria using cross-sectional survey methods. Use of molecular diagnostics may offer benefits to malariometric surveys due to their higher sensitivity in detection of *Plasmodium* parasites [10], as may the use of serological tools to describe population exposure to *Plasmodium* as opposed to current infection [11].

An effective health information system, which reports timely and accurate data to a central level and generates feedback to those collecting the data, is also crucial to malaria monitoring and surveillance. Spatial and temporal heterogeneity, presence of multiple *Plasmodium* species, presence of low density infections, as well as the potential for devastating epidemics to occur are all additional considerations to be addressed in development of effective malaria surveillance systems in low transmission areas.

School-based surveys were used during historical malaria reconnaissance activities [12-15], and their use as an alternative platform for large-scale periodic monitoring surveys has been demonstrated successfully more recently in Kenya, a moderate to low transmission setting [16]. School surveys offer practical advantages over household surveys, yet there are limited data to describe whether school surveys are a reliable alternative to standard household surveys in low and unstable transmission settings, and whether this approach may offer improved value for money. Schools have also been explored as a complementary infectious disease early warning system in high income countries, particularly for pandemic influenza [17-19], however this approach of school-based surveillance and epidemic detection has not been adapted for resource-constrained settings where routine health facility-based systems commonly underperform.

A major development in infectious disease monitoring has been the application of geostatistical methods to model associations between available malaria data and environmental predictors in order to predict these malaria indicators together with estimates of precision in areas lacking data. To date, geostatistical modelling and prediction have not used estimates of exposure to malaria to explore malaria endemicity in settings with temporally unstable transmission.

Considering the continued popularity of household surveys and availability of new tools such as model-predicted endemicity maps, and alternative platforms for monitoring and

surveillance, there is a need to compare the data generated by each tool to allow countries to make informed and evidence-based decisions of the strategies that they should prioritise in an environment of limited resources.

This chapter provides further information of the biology of *Plasmodium* and the pathology, transmission dynamics and epidemiology of infection with this parasite. Information is presented on various strategies for malaria surveillance, including routine passive surveillance at health facilities, periodic cross-sectional surveys, epidemic detection methods and the potential of syndromic surveillance. Next, an overview is given of the indicators that may be of use for malaria surveillance in low and unstable transmission settings. Finally, I provide a summary of the epidemiology and control of malaria in Ethiopia, where data used in this thesis are collected.

1.2 BIOLOGY, EPIDEMIOLOGY AND CONTROL OF MALARIA

1.2.1 Parasite lifecycle

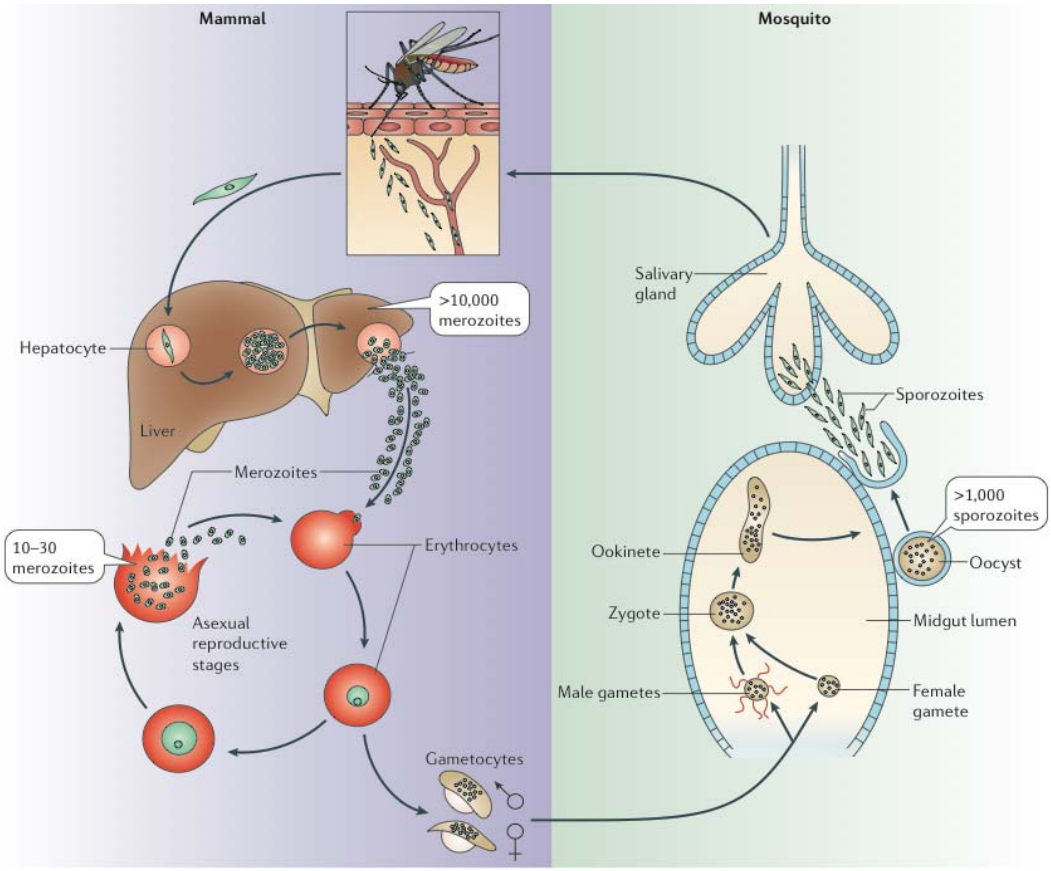
Malaria is caused by infection with protozoan parasites of the *Plasmodium* genus. Human malaria is caused by five species of *Plasmodium*: *P. falciparum*, *P. vivax*, *P. malariae, P. ovale* and *P. knowlesi*. *Plasmodium falciparum* has been generally regarded to have the greatest public health impact of all the *Plasmodium* species, particularly in sub-Saharan Africa. While *P. falciparum, P. vivax, P. malariae and P. ovale* are all found in sub-Saharan Africa [20-22], this thesis will focus on *P. falciparum* and *P. vivax*.

Malaria is a vector-borne disease, transmitted by female *Anopheles* mosquitoes. Different combinations of *Anopheles* species are responsible for *Plasmodium* transmission across the diverse environments of the globe where malaria is present [23]. The lifecycle of *Plasmodium*

has developmental stages in both the vector and human hosts, with sexual reproduction taking place in the vector and asexual reproduction in humans.

Male and female *Plasmodium* gametocytes ingested by female *Anopheles* during a blood meal fuse in the mosquito stomach to form a zygote, which develops into a motile ookinete. The ookinete moves through the mosquito stomach wall, forming an oocyst on the outer surface of the stomach wall. Up to one thousand sporozoites develop inside a single oocyst over a period of eight to 35 days. Sporozoite development time is dependent on both parasite species and external temperature. When mature, the sporozoites penetrate the wall of the oocyst and migrate to the mosquito salivary glands, where they are injected to a new host when the mosquito takes a blood meal (Figure 1.1).

Figure 1.1 - Lifecycle of *Plasmodium* spp. Adapted from Ménard *et al.* [24]



In the human host, sporozoites enter hepatocytes and develop into exo-erythrocytic schizonts. When mature, over 10,000 merozoites will be released into the bloodstream upon rupture of the infected hepatocyte. Merozoites invade erythrocytes and undergo cyclical asexual replication. Inside the erythrocyte over a period of 48-72 hours, the parasite develops into a trophozoite and then mature schizont, at which point the erythrocyte ruptures and releases between eight and 30 erythrocytic merozoites to invade further erythrocytes.

A fraction of merozoites released from erythrocytes will develop into gametocytes, the transmissible parasite form ingested by female *Anopheles* mosquitoes when taking a blood meal. The time for appearance of gametocytes varies between species: they can usually be found approximately three days after first identification of asexual *P. vivax* parasites, and after approximately ten days for *P. falciparum*. Immature gametocytes are sequestered in the bone marrow or spleen, and released to the circulation once mature. Gametocytes are found in both low- and high-density infections, symptomatic and asymptomatic. Gametocytes typically circulate at very low densities, but submicroscopic gametocytaemia is known to be infectious to mosquitoes [25,26].

Plasmodium vivax and *P. ovale* differ from *P. falciparum* in the ability to form hypnozoites, a parasite stage which develops from sporozoites but persists in hepatocytes for months or years. During this time, hypnozoites may develop further; generating merozoites and commencing asexual replication cycles in erythrocytes, leading to relapse in malaria.

1.2.2 Malaria pathology

Classic uncomplicated malaria is described as a cycle of chills, then fever, headache and vomiting, and finally sweating. However, the majority of *Plasmodium* infections have more variable symptoms, including some combination of fever, muscle and joint pain, headache, sweating, chills and anorexia. Historical controlled human malaria infection studies indicated that symptoms cycle, with attacks occurring every two days for *P. falciparum, P. vivax* and *P.*

ovale, while attacks due to *P. malariae* occur every three days [27]. In practice, however, most naturally acquired infections are not so clearly cyclical in presentation of symptoms. Symptoms of uncomplicated malaria are associated with rupture of erythrocytic schizonts and the immune response to release of toxic contents of the lysed cells [28]. After several cycles of *P. falciparum* asexual reproduction, anaemia and splenomegaly may develop [29].

The case fatality rate for non-immune adults and young children infected by *P. falciparum* can reach 10-40% [30]. Patients who progress to severe malaria are more likely to die than those with uncomplicated infection. Severe malaria is the result of organ failure or abnormalities in the patient's blood or metabolism as a result of *P. falciparum* infection and parasite sequestration. Symptoms of severe malaria include acute encephalopathy, respiratory distress, renal failure, hypoglycaemia, lactic acidosis, severe anaemia, coagulation defects and jaundice [31].

While *P. vivax* had long been considered a relatively benign infection compared to *P. falciparum*, an increasing body of evidence describes instances of severe malaria caused by *P. vivax,* indicating that the public health burden of *P. vivax* malaria may have been underestimated [32-34]. Mechanisms contributing to severe disease in *P. vivax* infection include destruction of uninfected erythrocytes leading to severe anaemia and cytokine-related changes in alveolar permeability causing respiratory distress [35].

1.2.3 Transmission dynamics of malaria

Historical categorisation of malaria endemicity into holoendemic, hyperendemic, mesoendemic and hypoendemic was defined according to the proportion of a population with a palpably enlarged spleen. However this classification was challenged by some who believed that defining malaria transmission as either stable or unstable was more appropriate in consideration of transmission dynamics of this vector-borne parasite [9].

The Ross-Macdonald model of malaria transmission defines the reproductive number (R_0) of malaria as the number of new infections arising from a single infected person in the absence of immunity and malaria control, after one generation of the parasite:

$$R_0 = \frac{ma^2bcp^n}{r(-\ln p)}$$

The components of the R_0 calculation are the ratio of anopheline mosquitoes to humans (*m*), human biting rate of anophelines (*a*), transmission efficiency of anophelines to humans (*b*) and humans to anophelines (*c*), number of days for recovery by humans from infection (*r*), proportion of mosquitoes surviving one day (*p*) and the number of days required for sporogeny (*n*) [36].

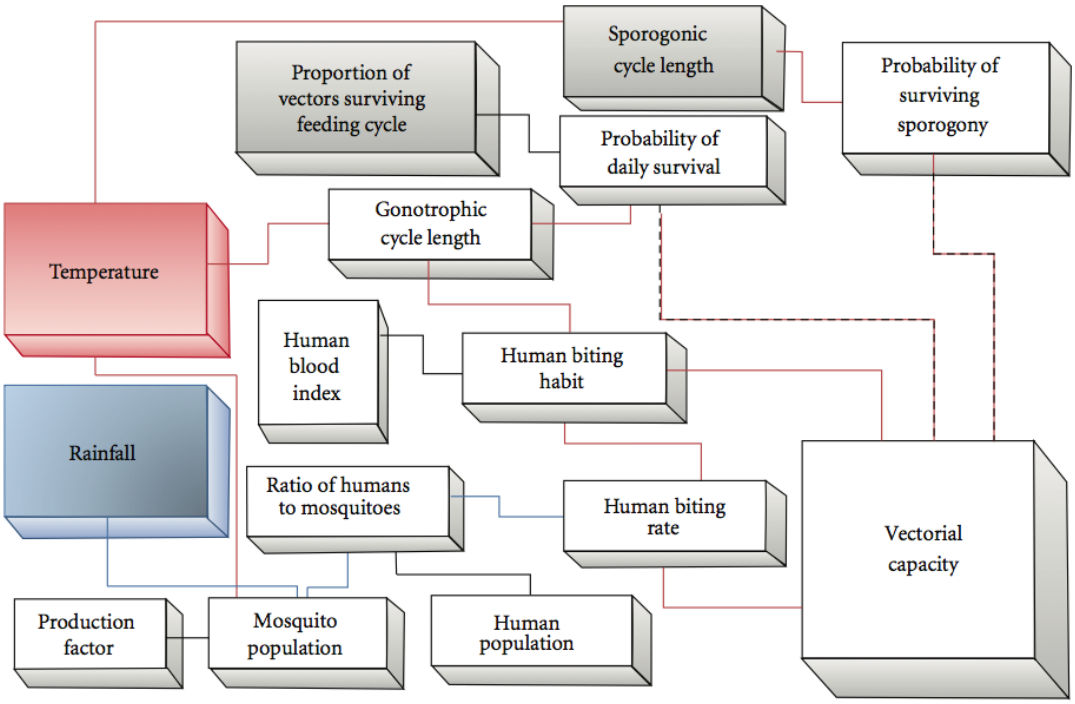
A major assumption of the Ross-Macdonald model is that of homogeneous transmission in a well-mixed population. In practice, the Ross-Macdonald model is violated by presence of immunity in a population, biasing infectivity of humans to mosquitoes and *vice versa* [36]. In addition, vector biting is often heterogeneous with 80% of infectious bites received by 20% of people, introducing a sampling bias between selection of humans by vectors and selection for inclusion in a study [37]. Contemporary malaria transmission models have attempted to incorporate heterogeneity at different scales, from small scale where human and mosquito behaviour result in heterogeneous biting, to larger scale where vector composition and dynamics are influenced by ecological factors [38]. A further key innovation was incorporation of host-parasite interactions and immuno-epidemiology into transmission models [38].

The basic reproductive number is generally interpreted to be $R_0>1$ in situations of increasing transmission, and $R_0<1$ to mean declining transmission since each infection leads to less than one subsequent infection on average. Estimating R_0 can be valuable in malaria control

planning, assist in setting achievable targets and identifying priorities, however it is rarely estimated using data from the field. Indicators such as parasite rate and entomological inoculation rate are more commonly generated metrics, but can be combined with other parameter estimates to approximate R_0 for local settings [36].

While R_0 describes interactions between the human, vector and parasite populations, the vectorial capacity describes the number of subsequent infectious bites arising from a single person-day of exposure. Vectorial capacity also describes those components of transmission which are temperature dependent (Figure 1.2).

Figure 1.2 - Diagrammatic representation of vectorial capacity model used as an early warning system for malaria epidemics. The model demonstrates how temperature and rainfall (red and blue connector lines, respectively) can trigger epidemics by increasing vectorial capacity [39]



The various species of *Anopheles* involved in malaria transmission across the world have preferences in their breeding sites, resting locations, as well as in favoured biting species,

time and location [23,40]. These preferences show some plasticity [41], often in response to changes in human behaviour or activity. Examples include changes in peak biting time as a result of use of insecticide-treated mosquito nets while sleeping [42], as well as shifting from indoor to outdoor biting and resting locations following indoor-residual spraying [43].

Minimum temperature, rainfall levels and humidity are among factors that determine the suitability of an environment for mosquito breeding and survival [39,44]. The suitability of a habitat for mosquitoes determines the probability that a mosquito will survive sufficiently long for ingested *Plasmodium* gametocytes to develop to the infective sporozoite stages and be transmitted to another human host. Altitude is commonly used as a proxy for suitability for transmission, since the fall in minimum temperature with increasing altitude is often the limiting factor in vector survival and malaria transmission in highland areas. Numerous studies in the highlands of East Africa have demonstrated this inverse relationship between altitude and indices of malaria transmission or burden [45-48]. Sporogeny for both *P. falciparum* and *P. vivax* takes eight to ten days at 28°C, but increases to 16 days when the temperature falls to 20°C. The minimum temperature at which *P. falciparum* sporogeny will take place is 16°C, but *P. vivax* can generate sporozoites at a minimum temperature of 14.5°C [49]. The ability of *P. vivax* to generate sporozoites at a lower temperature than *P. falciparum* results in potential for *P. vivax* transmission at higher altitudes, and therefore likely different spatial extents of transmission for the two species [50].

While much of sub-Saharan Africa has environmental and climatic conditions that support perennial malaria transmission, arid areas and highlands typically experience seasonal transmission since low temperature or rainfall limits mosquito survival and transmission potential [51,52]. In these settings, malaria transmission tends to peak following seasonal rainfall, then declines during the dry season [53-56].

In unstable transmission settings, epidemics may occur when environmental conditions become favourable for increased transmission, or when infection is introduced into susceptible populations [52,57-60]. Malaria epidemics are broadly defined as unusual increases in the burden of malaria illness, that are "clearly in excess of normal expectancy" [61]. Therefore, in areas which are largely malaria-free, a single locally-acquired case may be considered as a potential epidemic. While in low-endemic or seasonal settings, an epidemic may have a more subjective or programmatic definition, such as being more cases than can be managed by routine health service capacity [62]. Malaria epidemics may be due to *P. falciparum* or *P. vivax*, but generally occur in populations without protective immunity against *Plasmodium* [63].

1.2.4 Epidemiology and burden of malaria

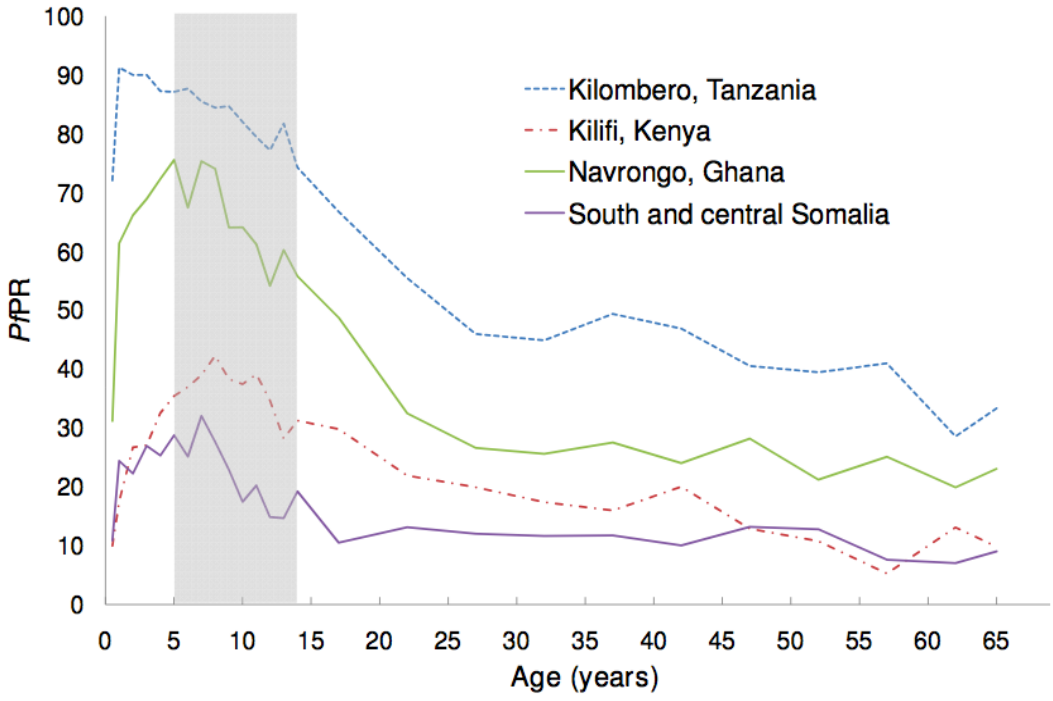
The World Health Organization (WHO) has estimated that 207 million cases of malaria occurred worldwide in 2012 and 627,000 malaria deaths, incorporating both recorded cases and those which area estimated to occur but are not captured by health information systems [64]. Geostatistical modelling estimated that 2.57 billion people were at risk of *P. falciparum* worldwide in 2010 [20], and 3.5 billion at risk of *P. vivax*. Africa was estimated to contribute 31% of the global population at risk of *P. falciparum*, but only 3.5% for *P. vivax* due to the widespread Duffy negative phenotype in sub-Saharan Africa [21]. Of these total populations at risk for each species, 44% occupy areas of unstable *P. falciparum* and 61% of population at risk of *P. vivax* reside in areas of unstable transmission.

Malaria transmission intensity influences age-specific risks of infection, clinical disease, and mortality. In areas of intense malaria transmission, individuals acquire protective immunity as a result of exposure, but with reducing transmission intensity and therefore exposure, functional protective immunity develops at older ages, until low transmission settings where the population generally do not have protective immunity against *Plasmodium.*

The focus of many malaria programme evaluation indicators has been to explore morbidity and mortality in children under five years of age, however modelling using data from a range of malaria transmission intensities indicates that age patterns of clinical malaria, malaria-diagnosed deaths and hospital admissions with malaria are less biased toward younger ages in areas of seasonality and low transmission [65].

Figure 1.3 demonstrates the relationship between age and parasite rate across high to low transmission intensities. Where transmission is most intense, parasite rate increases rapidly up to age two, remaining high until age ten and then declining in adulthood, attributable to protective acquired immunity. However even in areas of low transmission intensity, moderately higher parasite rates can be seen in children compared to adults [66]. The association between age-specific parasite rate and transmission intensity has been demonstrated in settings where transmission has reduced due to control interventions, where a right shift occurred in *Plasmodium* prevalence by age [67].

Figure 1.3 - The relationship between age and *P. falciparum* parasite rate (*Pf*PR) across various transmission intensities, from very low in Somalia to high in Tanzania. The grey box indicates the usual age of primary school children in Africa, 5-14 years. Adapted from Brooker *et al.* [68]



Identification of individuals with *Plasmodium* infection is also complicated by presence of asymptomatic and low density infections. Increasing evidence demonstrates that asymptomatic and low density infections are common even in low transmission settings [10,69,70], contradicting previous assumptions that a non-immune population would experience symptomatic and high-density infections due to their lack of acquired immune response. Identification of low density *Plasmodium* infections is crucial when countries are moving towards pre-elimination and transmission control, where it is estimated that submicroscopic carriers are the source of 20-50% of all transmission from humans to mosquitoes [71].

1.2.5 Malaria control strategies

The current recommended first-line treatment for uncomplicated *P. falciparum* malaria is artemisinin combination therapy (ACT), following development of resistance to previously used drugs including quinine, chloroquine and sulphadoxine-pyrimethamine (SP). While ACT is effective in the majority of settings in clearing asexual parasite forms and alleviating symptoms, it does not kill all gametocytes. *Plasmodium falciparum* resistance to artemisinin has been identified in several foci in the greater Mekong sub-region, and containment of artemisinin resistance has been designated a global priority [72]. ACTs are being increasingly used in areas where both *P. falciparum* and *P. vivax* are endemic [73], but the majority of countries where *P. vivax* transmission takes place continue to use chloroquine to treat *P. vivax* mono-infection.

Use of primaquine is being considered in some countries due to its gametocytocidal action, although there is no conclusive evidence that addition of primaquine is effective in reducing of *P. falciparum* transmission [74]. Primaquine is also of interest as a radical cure for *P. vivax* due to its action in clearing hypnozoites [75], but has not been widely adopted due to

haemolytic effects in individuals with glucose-6-phosphate dehydrogenase (G6PD) deficiency [76-80].

The development of long lasting insecticide-treated mosquito nets (LLINs), which do not require re-treatment with insecticide, led to a global drive for increased access to vector control. Early studies in the Gambia indicated that the use of insecticide-treated mosquito nets was associated with a 60% reduction in mortality among children aged one to four years [81]. LLINs are impregnated with pyrethroid insecticides, and expected to remain effective for up to five years. However the useful life of LLINs may be reduced due to physical damage, with households preferring not to use old nets that have become severely torn [82,83]. Consistent use of LLINs within households has also proven challenging, with a wide range of contextual factors contributing to the likelihood of net use every night, as well as the decision of which household members use the available nets [84,85]. In addition to providing personal protection for the individuals sleeping under the mosquito net, insecticide-treated nets also have a community effect as a result of reduction in the mosquito population [86-88].

Indoor residual spraying (IRS) has been demonstrated to be effective in reducing malaria prevalence within a community [89,90], by killing or reducing the lifespan of mosquitoes resting on indoor walls of the household prior to or after taking a blood meal. IRS may also elicit a repellent effect on mosquitoes seeking a blood meal, reducing the number of mosquitoes entering households to feed.

The evidence for additionality of both IRS and high coverage of ITNs in the same area is conflicting. A meta-analysis indicates that there is no additional effect of implementing both IRS and ITNs [91]. However, a subsequent cluster randomised trial in Tanzania found that implementing both interventions resulted in a decrease in mean PfPR among children [92]. This result may be partly attributable to only moderate (36 to 50%) use of ITNs in this

population, although ITN users were found to be additionally protected if their houses received IRS.

1.3 PLATFORMS FOR MONITORING, EVALUATION AND SURVEILLANCE

Collection of data to describe the implementation, outputs and impact of malaria control programmes is essential to ensure that the programme remains relevant, effective and responsive to needs of the population and context. Both monitoring and surveillance are included in this section and generally referred to as surveillance, although is it acknowledged that the formal definitions may differ: surveillance generally understood to be continuous and ongoing, while monitoring is interpreted to be intermittent or episodic collection of data [61].

1.3.1 Routine surveillance

Routine recording, reporting and analysis of clinical data from health facilities is a core component of an effective health system. Routine data are reported in a number of different formats, the primary and original system being the Health Management Information System (HMIS). HMIS is usually a paper-based system whereby reports on mortality, morbidity, health resource and preventative indicators are generated quarterly from public hospitals and health centres. HMIS also includes reporting of the level of completeness of available data. Due to the widespread adoption of HMIS, these data are often available to describe longer time periods than other surveillance data.

The Integrated Disease Surveillance and Response (IDSR) system was launched to improve timeliness of reporting on major endemic diseases of public health importance, diseases targeted for elimination or epidemic-prone diseases [93]. The IDSR aims to strengthen the capacity of countries to conduct effective surveillance activities, integrate multiple

surveillance systems to improve efficiency of surveillance resources, and to improve the flow of surveillance information between and within levels of the health system. The principle of community participation in detection and response to public health problems, along with increasing involvement of clinicians are also key components of IDSR.

There are limitations to the utility of routine data from health facilities to estimate the impact of disease control programme efforts on population health [94]. Population access to health services is one potential bias in routine data from health facilities; inequities may exist as a result of distance to health facilities, socio-economic status and ability to pay for transport to facilities, or cultural norms which limit the ability of sub-populations to access health services. Validity and representativeness of data may also be restricted, and should be acknowledged when interpreting results from analysis of routine health facility data. Use of sentinel sites, can provide an interim solution to enable timely epidemic detection and response as well as programmatic evaluation in settings where routine surveillance system are inadequate [95].

Quality of routinely collected data from health facilities may be limited by lack of feedback on submitted data, as health workers lose the motivation to invest sufficient time in completing data accurately. Duplication and redundancy in data reporting as a result of multiple recording systems can also impact on quality of routine data. For example, health centres collect data in outpatient registers, integrated management of childhood illness (IMCI) registers for children under five years of age, and laboratory registers of diagnostic tests conducted. Resolving differences between these data sources can be challenging and time consuming for staff compiling data.

One systematic approach to evaluate performance of a surveillance system is to determine the completeness of data submitted, or spot-checks may be conducted during supervisory visits to compare facility records with information submitted to the central surveillance system. To improve performance of health workers, some countries have adopted strategies

whereby money or goods are transferred to the health worker, conditional on achieving defined targets or taking measurable action. While this strategy of performance-related pay has been shown to improve maternal and child health services in Rwanda [96], evidence of the impact of this strategy in other countries is inconclusive [97].

Statistical analysis methods which take into account the spatial and temporal heterogeneity in malaria cases have been presented which estimate the likely values of missing routine data from health facilities, allowing generation of more reliable estimates of malaria burden from routine data [98]. While this is likely too complex a strategy to be widely used at national level to overcome limitations in the routine data, this method may have value in periodic retrospective analysis of health facility data with specialist statistical support. In Zimbabwe, clinical data from health facilities were used to model malaria risk, generating smoothed maps of seasonal trends in malaria burden [53]. Alternative strategies have been proposed whereby routine data are combined with cross-sectional prevalence data to estimate the force of infections in a low transmission setting using a reversible catalytic model [99].

In elimination settings, reactive case detection strategies have been trialled whereby routine surveillance data from health facilities is used as a trigger to conduct reactive screening and treatment around the index case in the community [100]. In elimination settings using a reactive case detection strategy, investments must be made in quality of malaria diagnostic services to ensure that symptomatic *Plasmodium* infections are captured by the surveillance system, and reactive screening quickly implemented. The spatial and temporal clustering of malaria cases in very low transmission settings indicates that foci of transmission could be identified by tracing the residence of passively identified cases. Reactive screening and treatment is intended to prevent onward transmission, but due to the high proportion of infections in these elimination or pre-elimination settings which will be low density and asymptomatic, it is advantageous to use diagnostic tools with high sensitivity to detect low density infections, as well as provide treatment with gametocytocidal drugs to block onward

transmission. This strategy has been piloted in Swaziland and Senegal, where individuals living in the same or neighbouring household to an index case diagnosed at the health facility were screened by RDT, but both pilots found the strategy to be operationally demanding, resource intensive, and identified few additional infections [101,102].

1.3.2 Periodic monitoring surveys

Large-scale cross-sectional surveys are used in most malaria transmission settings to gather information on population health, including access to services and preventative measures. Malaria is no exception, with a range of cross-sectional survey strategies for use across all transmission settings, designed to measure parasitaemia, reported malaria morbidity and mortality, as well as access to and use of key malaria interventions. Key considerations of cross-sectional surveys in low transmission settings are the indicator to which the survey is powered to measure, whether temporal and spatial heterogeneity are captured, the target population, frequency of data collection and cost.

Demographic and Health Surveys are nationally-representative household surveys which collect monitoring and impact evaluation indicators for a range of population, health and nutrition factors [103]. DHS are usually conducted every five years, and countries choose appropriate modules to include in the DHS, such as anaemia, child health, education, family planning, malaria, maternal health, nutrition and wealth. The sample size for DHS is usually between 5,000 and 30,000 households, in order to generate nationally representative indicators, with the whole survey process requiring on average 18-20 months to complete. DHS usually includes a household questionnaire, as well as a separate questionnaire for women of reproductive age. Biomarkers such as blood samples for haemoglobin measurement and identification of *Plasmodium* infection by microscopy may be included in the DHS, but are not a core component. Key data collected from DHS relevant to malaria

control programme monitoring are household ownership and individual reported use of mosquito nets.

Malaria Indicator Surveys are a tool developed by the Roll Back Malaria Partnership (RBM) to allow national malaria control programmes to generate standard indicators for monitoring coverage of malaria control interventions [104]. MIS are nationally-representative household surveys, designed to gather information on core household indicators defined by Roll Back Malaria [105]. The key themes that MIS are designed to collect data on include coverage of LLINs and IRS, use of mosquito nets by pregnant women and children under five years of age, intermittent preventative treatment during pregnancy, diagnosis and treatment of malaria among children under five, all cause under five mortality, and morbidity indicators from children under five years (anaemia and parasitaemia). The RBM guidance on design and implementation of MIS suggests that parasitological testing of children aged six to 59 months should take place in areas of stable malaria transmission [104].

An alternative to DHS and MIS is UNICEF's multiple indicator cluster survey (MICS) [106]. Similar to DHS and MIS, MICS are nationally representative household surveys, but use an alternative sampling strategy to select households for inclusion in the survey. MICS generates indicators related to health, education, child protection and HIV/AIDS, harmonising indicators with DHS and MIS where possible. Countries can choose the modules that are most relevant for inclusion in their questionnaire, but the aim is to monitor progress toward national and global commitments on the situation of children and women, such as the Millennium Development Goals.

While the DHS, MICS and MIS do provide a nationally-representative estimation of key indicators at the time of the survey, a major limitation to the use of these monitoring surveys is their periodic implementation; as a result of the significant investments that must be made in implementing these activities, they are usually conducted at intervals of three to five years.

An alternative strategy termed "rolling MIS" has been proposed, but not widely adopted. The rolling MIS generates data at a much finer temporal scale, by adapting a cross-sectional evaluation tool into a continuous monitoring tool, to more closely monitor changes in malaria burden as a result of rapid scale up in coverage of malaria interventions [107].

A further strategy for periodic monitoring surveys is sampling individuals within schools, rather than conducting household surveys [16]. School malaria surveys have been used historically for malaria reconnaissance in sub-Saharan Africa [12,13] and the Americas [14,15,108,109]. School surveys are logistically attractive since all eligible individuals for random selection are gathered in a single location. Furthermore, in areas of high school enrolment and attendance the school-attending population should be representative of the wider community. This allows school surveys to be completed more quickly and at lower cost than standard household surveys: school surveys usually require only one day to sample 100 children at each site, while household surveys require census, randomisation, and sampling stages and often require two or three days to complete one site. There are however, potential biases in the use of school-attending children for health surveys, since children attending school may differ from non-attending children by wealth, health status, or area of residence within a community.

School surveys have been demonstrated to generate reliable estimates of community coverage of insecticide-treated mosquito nets in Uganda [110]. In Kenya, a study found that estimates of parasitaemia from testing using RDTs correlate between school and community surveys conducted at the same locations. However, although the estimates correlate, they were statistically discordant, and school survey parasite rate was consistently higher than parasite rate in community surveys [111]. A review of the potential uses of schools for malaria surveillance and programme evaluation highlighted use of school surveys for estimating coverage of interventions and parasite prevalence, as well as epidemic alert systems and active case finding [68].

1.3.3 Epidemic detection

Malaria epidemics are usually defined by assessment of the number of cases of malaria identified at a health facility compared to the expected burden at that time. Definitions of "normal" burden of malaria vary, but the upper and lower limits of normality are often defined as being two standard deviations around the mean number of cases for a facility in a defined time period, after excluding previous epidemic periods [62]. In order to successfully identify and respond to malaria epidemics, the temporal resolution of indicators becomes critically important, as it is the ongoing collection, analysis and feedback of data that enables responses to be mounted sufficiently early to prevent large-scale morbidity.

Epidemic detection systems collect similar indicators to HMIS, a passive surveillance system, but the number of indicators is reduced and frequency of collation and reporting is increased. The definition of an epidemic requires a threshold for the expected or normal number of cases of malaria to be defined. Where malaria case data are available from previous years, it is recommended that these data are used to define the threshold. A common technique for epidemic definition is the quartile method [62], where the threshold is the third highest weekly total confirmed malaria cases for the current calendar week, taken from the previous five years' data for that health facility. Another method using historical data from the facility to define the epidemic threshold is the cumulative sum, or c-sum approach [112]. The c-sum method generates a "base year" describing the expected number of cases using the mean value for that month from the previous five years' data, but also incorporates the mean of the preceding and following month. In practice, health facilities may not have five years' complete data to be able to define the epidemic threshold using this method. Alternative strategies to define epidemic thresholds include doubling the previous year's number of cases for the same week at the facility, or the Cullen method, whereby the threshold is the annual mean number of cases plus the standard deviation multiplied by two [113,114].

Incidence thresholds, usually the weekly total number of cases by district which have been identified through the routine health facility surveillance system, expressed per 1000 people resident in the district, are used for other epidemic-prone infectious diseases such as meningococcal meningitis [115], but it is important that such incidence thresholds are locally-defined and appropriate, to improve responsiveness to potential epidemic years at a district level [116].

While literature exists comparing the sensitivity of various surveillance algorithms for rare and notifiable diseases in high-income countries [117-119], few examples exist for resource-poor settings comparing different strategies to define epidemic thresholds. A study in Kenya compared the use of the Cullen, c-sum and quartile methods to define epidemic thresholds for malaria data from health facilities [120]. While the lack of gold standard definition for an epidemic limits the ability to formally compare the methods, use of the Cullen threshold correctly identified the highest burden years at more facilities than the other thresholds. A similar comparative study in Ethiopia found a simple percentile cut-off value to be as useful in defining epidemics as more complicated algorithms [121].

In settings where the limiting factor in epidemic identification is timeliness of data submission, analysis and response, temporal resolution can be improved by use of mobile phones to report weekly or even daily number of cases of key infectious diseases including malaria. In Madagascar, a pilot network of sentinel general practitioners submitted at least daily text messages by mobile telephone to a central management team, reporting fever cases, RDT-confirmed malaria, influenza, arboviral syndromes and diarrhoeal diseases. The system identified ten clusters of febrile illness which were not identified by the traditional surveillance system [122]. In Zambia, health centres piloted a weekly short-message service (SMS) reporting system for malaria, submitting the number of individuals tested and total confirmed malaria cases, with the aim to identify foci of infection or even index cases in areas of low malaria transmission [123].

A complementary strategy to identification of epidemics by analysis of passive case detection data from health facilities, is through use of a climate-based early warning system [124,125]. Climate-based early warning systems provide earlier alerts than are generally possible using malaria case data, enabling targeting of resources to at-risk areas and control interventions to be implemented earlier than may be otherwise possible [126]. An internet-based Malaria Early Warning System is available, which identifies rainfall anomalies across malaria epidemic-prone areas, with a 10-day resolution [127,128]. Lack of internet access by district health staff in resource-constrained countries limits the utility of this alert system, however it may still be possible for the system to be accessed at national level and alerts disseminated to local staff through a cascade system when necessary. Various studies in areas of unstable malaria transmission have developed statistical models which demonstrate associations between remotely-sensed climatic data and temporal changes in malaria burden reported at health facilities, including epidemics [129-131].

While use of mobile telephones to report confirmed malaria cases from health facilities has the potential to improve surveillance system timeliness, there remain challenges in identification of *Plasmodium* infections in the community, either as a result of limitations in diagnostic tool performance, lack of availability of diagnostics or poor access to health services by the population. The use of pre-diagnostic indicators or even surrogate data may offer an alternative surveillance indicator for malaria.

1.3.4 Syndromic surveillance

Classical epidemic detection systems including those described in the previous section can be limited in effectiveness due to delays in reporting, incomplete data recording or use of inaccurate data, all of which can contribute to delays in identifying and responding to malaria epidemics [132-134]. A complementary system for surveillance is the use of pre-diagnostic indicators of clinical disease, whereby daily or weekly data are reported from health facilities

without the need to wait for confirmatory tests. This surveillance strategy can also be expanded to incorporate the use of surrogate non-clinical data indicating early illness, particularly exploiting data sources and indicators that are already routinely recorded or easily accessible. These approaches are often described as syndromic surveillance [135].

While there are differing interpretations of syndromic surveillance, the Centers for Disease Control's definition appears to be the most widely accepted: “an investigational approach where health department staff, assisted by automated data acquisition and generation of statistical alerts, monitor disease indicators in real-time or near real-time to detect outbreaks of disease earlier than would otherwise be possible by traditional public health methods” [136]. The key aspect of syndromic surveillance is, therefore, to improve temporal resolution and responsiveness, allowing faster responses to potential epidemics. These systems are generally intended to run in parallel to the more sensitive and specific surveillance systems reporting indicators of confirmed disease, since syndromic surveillance systems have low specificity and therefore may generate false positive alerts.

A syndromic surveillance system was developed for use in the Pacific islands and territories, acknowledging the challenges that geographically isolated communities with limited diagnostic capacity have when attempting to implement traditional data-intensive surveillance systems requiring confirmatory results for notifiable diseases. The syndromic surveillance system implemented generated a set of syndromic case definitions relating to the epidemic-prone diseases of interest (Table 1.1), thereby negating the need for laboratory confirmation before reporting and limiting the number of indicators on which to report [137].

Table 1.1 – Example case definitions in a syndromic surveillance system in the Pacific islands and territories [137]. Note that fever is defined as 38°C of higher, or fever or chills reported by caregiver or patient if no thermometer is available.

Syndrome	Case definition	Potential causative disease
Acute fever and rash	Sudden onset of fever, with acute non-bleeding rash	Measles, dengue, rubella, meningitis, leptospirosis
Diarrhoea	Three or more loose or watery stools in 24 hours	Viral and bacterial gastroenteritis including cholera, food poisoning, ciguatera fish poisoning
Influenza-like illness	Sudden onset of fever, with cough and/or sore throat	Influenza, other viral or bacterial respiratory infections
Prolonged fever	Any fever lasting three or more days	Malaria, typhoid fever, dengue, leptospirosis, other communicable diseases

Reporting based on a defined set of syndromes reflecting key reportable (e.g. acute flaccid paralysis, haemorrhagic fever) or epidemic-prone infectious diseases (acute watery diarrhoea, bloody diarrhoea, prolonged fever, acute fever and rash) has also been used in India, Papua New Guinea and South Africa [138-140]. In Madagascar a syndromic reporting system was piloted with a focus on diarrhoeal disease and febrile illness, but also included the reporting of RDT-confirmed *Plasmodium* infections [122]. An alternative syndromic reporting system was used in French Guiana, developing an index of febrile patients who are confirmed to not have malaria as a proxy for possible dengue outbreaks [141].

Expanding access to and use of technology, particularly expanding coverage of mobile telephone networks, is facilitating an increased interest in syndromic surveillance in resource-poor settings [142]. In Papua New Guinea, mobile reporting of syndromic case definitions was found to be more sensitive than monthly hospital-based surveillance in detecting a measles epidemic, but had potentially reduced sensitivity for malaria compared to the standard paper-based national surveillance system [139]. In Zambia, mobile reporting of data was combined with global positioning system coordinates to incorporate a spatial component in the data generated by the surveillance system [123].

In high-income countries that have robust health information systems, syndromic surveillance is applied to improve timeliness of epidemic alerts through the use of pre-clinical data to identify changes in population health prior to any increase in access to health services for diagnosis and treatment. This method has proven particularly popular for the identification of influenza epidemics [143]. Mining of existing data is the basis for this type of syndromic surveillance systems favoured by high-income countries, either by tracking pharmacy sales of non-prescription drugs, internet search engine terms, social media posts or school attendance [17-19,144-151]. However, this type of surveillance using surrogate indicators may also be appropriate in resource poor-settings, should appropriate surrogate indicators be available which are routinely recorded yet reflect health events occurring within the community.

1.4 INDICATORS FOR MALARIA SURVEILLANCE

The choice of indicator used for a surveillance system is influenced by the transmission intensity, quality of routine health services and reporting, as well as the temporal scale at which data are required. A key concept relevant to all indicators that are reported according to a defined schedule is that of zero reporting, where reports are generated and submitted even if no cases are identified during the reporting period [152].

1.4.1 Pre-diagnostic and surrogate indicators

The potential of syndromic surveillance platforms in resource-poor settings has been discussed in section 1.3.4. Surrogate and pre-diagnostic indicators are presented in the context of syndromic surveillance systems. Case definitions for pre-diagnostic syndromes are used primarily in resource-poor settings, and focus on epidemic-prone infectious diseases, or those pathogens which are targeted for elimination and therefore notifiable, and may require

confirmatory diagnosis at a specialist laboratory. Diseases considered within syndromic case definitions in these settings include cholera, meningococcal disease, poliomyelitis, measles, dengue, plague, viral haemorrhagic fevers and malaria, among others [122,137,138,140,141].

In moderate- and high-income countries, syndromic surveillance has been used to exploit non-clinical surrogate data that indicate changing health status of the population. A data source used by syndromic surveillance systems in countries including France, Sweden and the United States of America is records of non-prescription drug sales from pharmacies, and prescription drug reimbursement requests from health insurance records [150,153,154]. Particularly for diseases such as influenza in higher income countries, not all individuals affected will seek diagnosis from health facilities and may instead choose to self-treat with non-prescription medication. Increase in cold and influenza medications can therefore indicate a potential increase in influenza cases within the population. Records of access to and information sought from telephone triage systems used in some countries can also indicate changing burden of diseases such as influenza within the population [143,150].

Some individuals who are experiencing symptoms of illness in high income countries may choose to use the internet to search for information on possible diagnoses and treatment based upon their symptoms [155]. Google has developed tools which provide updated estimates of dengue and influenza activity worldwide based upon influenza and dengue-related internet search queries [156,157]. One study from Thailand has reported models developed to identify likely internet searches related to malaria [158]. Social media posts by individuals are a further indicator of interest for monitoring diseases such as influenza and cholera [159-161].

School absenteeism is an indicator that has been used in syndromic surveillance, since schools routinely record daily attendance of students, and absenteeism is hypothesised to increase when outbreaks of infectious disease occur within a community. The application of

school absenteeism as an indicator for syndromic surveillance has primarily been for detection of influenza epidemics in high-income countries [17-19,144]. These countries have high levels of school enrolment and attendance, therefore the school-attending population are expected to be representative of the wider population. School enrolment is, however, showing significant increases across lower income countries [162], and as a result, school absenteeism is a potential indicator for syndromic surveillance in resource-poor settings for a broad range of infectious diseases. Examples of school absenteeism as an indicator of infectious disease outbreaks in resource-poor settings are few, but include pilots in rural China and Cambodia [146,148]. While school absenteeism has not been formally explored as an indicator of malaria epidemics, absenteeism is acknowledged to increase during malaria epidemics [163].

All pre-diagnostic and surrogate indicators explored here are low specificity and prone to bias, but the use of these indicators offers improved temporal sensitivity of surveillance systems for epidemic-prone infectious diseases. These syndromic surveillance systems are designed as an adjunct to traditional surveillance using confirmed diagnosis, whereby an alert from the syndromic surveillance system should prompt examination of confirmatory diagnosis data from the same geographical location or population, or if these data are not available, a supervisory visit by district health staff to the location of the syndromic alert to conduct a situation assessment.

1.4.2 Clinical indicators

Indicators of clinical disease for malaria have long been included in passive surveillance systems at health facilities, but are becoming less common as access to parasitological diagnosis increases. Health facilities continue to routinely record the number of suspected malaria cases; those who have clinical indicators of malaria. However the exact constellation of symptoms resulting from *Plasmodium* infection is highly variable between individuals, and

clinicians may have different interpretations of when a patient exhibits signs and symptoms of malaria. A study in Tanzania investigated the causes of fever among children under ten years reporting to two outpatient clinics in an area of low malaria transmission [164]. Acute respiratory infection was the most common case of fever, and viral infection was more common among children with fever than bacterial or parasitic infection.

Reported and measured fever are also frequently included in cross-sectional surveys of malaria. Axilliary temperature measurement is a simple and non-invasive method to gather non-subjective indicators of current fever. Measurement of tympanic temperature using digital devices is an alternative which may be prove more reliable then axilliary temperature, if field staff taking axilliary temperature struggle to correctly position the thermometer. Asymptomatic infections are common across the range of *Plasmodium* transmission settings, and further reduce the sensitivity of clinical indicators of malaria [10]. Clinical indicators may, however, be valuable in identifying potential epidemics if unusual and sudden increases in febrile illness in the population are identified.

1.4.3 Parasitological indicators

Detection of parasites in blood is the most widely used malaria surveillance indicator. Passive surveillance systems at health facilities involve reporting of the total number of microscopy- or slide-confirmed malaria cases over a defined time period, often classified by age; either under five years or five years and older. Cross-sectional surveys, however, report the parasite rate (PR) calculated as the proportion of all surveyed individuals found to have *Plasmodium* parasites by a defined examination method. In settings of moderate and high transmission, PR varies by age as a result of protective acquired immunity against *Plasmodium* infection. As a result, age-standardised PR is often reported by studies, with the age range two to 10 years being the most commonly used for *P. falciparum* [9,66]. The relationship between *Pf*$PR_{2\text{-}10}$ and other indicators of malaria transmission, such as entomological inoculation rate (EIR) and

R_0, has been shown by modelling and field studies [37,66,165,166]. Parasite rate has been used in development of global predictive maps of malaria [20,21], as well as in redefining classifications of malaria endemicity and progress to elimination according to age-standardised parasite rate [66,166].

Parasite rate can be determined by microscopy, rapid diagnostic test or by nucleic-acid detection. Microscopy is considered the "gold standard" for identifying parasitaemia, but has practical and technical limitations. In a healthcare setting, microscopy can be limited by the ability of the microscopist, quality of slide staining, need for electricity, as well as high patient burden. From a technical perspective, the sensitivity of blood film microscopy is limited by the number fields routinely scanned by a microscopist, as well as the volume of blood which is used to prepare a blood film. If 100 high-power fields are screened during microscopy slide-reading, 0.1-0.25µl of blood will be examined [71]. As a result, microscopy has poor sensitivity to detect low density *Plasmodium* infections. A systematic review estimated the sensitivity of microscopy (against a gold standard of polymerase chain reaction, PCR) to be 53%, with 95% confidence interval 40% to 66% [10]. The limit of detection for routine microscopy is generally assumed to be 50 parasites per microlitre of blood (p/µl), while expert microscopists may be able to identify infections at 20 p/µl density.

Rapid diagnostic tests RDTs have great value as a simple tool that can identify *Plasmodium* infection without the need for specialist equipment or extensive training, however, they also have reduced sensitivity for low density infections [167-169]. RDTs are generally able to reliably detect infections at a minimum parasite density of 100 p/µl, but estimates of sensitivity vary between RDT type and endemicity setting [167]. A range of antigen combinations are available in RDTs, allowing *Plasmodium* species identification, but quantification of parasitaemia is not possible using this tool. RDTs are also prone to give false positive results for individuals who are still harbouring parasite antigen soon after treatment [170,171]. RDTs detecting *P. falciparum* histidine-rich protein 2 (HRP2) are also known to give

false negative results in some patients with very high parasite densities, due to the prozone effect [172].

A number of molecular methods to detect nucleic acid from parasites are available, which offer improved limits of detection compared to microscopy and RDT. Nested polymerase chain reaction (PCR) amplifies parasite DNA from a blood sample, therefore is able to detect very low parasitaemia [173]. Species-specific PCR methods are available through use of different amplification targets and primers [174]. Pooling of blood samples can improve the efficiency of sample screening in very low transmission settings, by combining samples for a first round of PCR, then conducting a second round of PCR on the positive pools only [175]. Where quantification of *Plasmodium* infection is important, real-time quantitative PCR methods have been developed which are highly sensitive, genus-specific and have very low limits of detection [176]. The loop-mediated isothermal amplification method (LAMP) does not require thermocycler equipment [177], and has allowed highly sensitive nucleic acid amplification methods to be conducted outside a national or international reference laboratory setting [178,179].

1.4.4 Indicators of prior exposure to *Plasmodium*

In settings of seasonal malaria transmission, it is challenging to coordinate cross-sectional surveys to ensure that the peak of transmission is captured. Indicators of *Plasmodium* infection taken from cross-sectional surveys in such settings of temporal and spatial heterogeneity may, therefore, not be fully representative of the true extent of transmission over a longer period of time. In these contexts, determining population exposure to malaria through detection of anti-*Plasmodium* antigens offers an alternative surveillance indicator for cross-sectional surveys.

Serological surveys are not a new concept, and detection of anti-*Plasmodium* antibodies by various methods including complement fixation test, indirect haemagglutination assay and

immunofluorescent antibody tests have been used since the eradication era of the 1960s [11]. Epidemiological applications of antibody detection from this period included demonstrating the assumed altitudinal limits of malaria transmission in Ethiopia [180], and evidence for local elimination of transmission in Mauritius by sampling children under five years to demonstrate lack of recent malaria transmission following an intensive campaign of indoor residual spraying with DDT (1,1,1-dichloro-2,2-bis(p-chlorophenyl)ethylene) [181].

The duration of antibody persistence continuous to be contentious. Antibodies have been shown to persist for years in an individual who is not subject to re-exposure [182], and reappear rapidly in individuals who are re-exposed as a result of epidemics [183]. One approach to overcome uncertainties of the duration of seroconversion and reversion, and individual differences in immunological responses, is to use serological indicators to describe exposure to *Plasmodium* among a representative sampled population, rather than individual exposure. The persistence of antibodies means that by measuring seroprevalence of population, it is possible to infer the exposure that a population has been exposed to over time, offering improved sensitivity compared to parasite detection in areas of low transmission.

Development of an indirect enzyme-linked immunosorbent assay (ELISA), and the successful elution of antibodies from finger-prick blood samples stored on filter paper has led to a resurgence of interest in seroepidemiology [184]. Furthermore, the assay can be adapted to different transmission settings by use of different antigens, including for different *Plasmodium* species [185].

Seroconversion rates measured in Tanzania were demonstrated to be representative of entomological inoculation rates, and are therefore considered to be representative of malaria transmission intensity [186]. Producing age-seroprevalence curves can also provide information about changes in transmission intensity as a result of intervention scale-up [187],

or differences in force of infection between adults and children, as a result of behavioural changes in exposure [188].

A further application of seroepidemiology is in spatial mapping of serological responses. Studies comparing seroprevalence between different locations have been conducted at a range of spatial scales from approximately 1 to 100 km, to provide information about spatial heterogeneity in exposure to malaria and therefore transmission intensity [189-194]. Serology may also have value in settings approaching elimination, whereby selectively testing of young children (but excluding children under 12 months who may have maternal anti-*Plasmodium* antibodies) can provide information about recent malaria transmission. These data could be collected through cross-sectional surveys, or using a passive approach, whereby sample collection takes place at health facilities when children provide blood for other diagnostic investigations, provided full parental informed consent is granted for the additional blood testing. Follow-up investigations could be conducted for any seropositive child to explore the source of their exposure to *Plasmodium* infection.

Crucial for selection of the most appropriate indicators for use in malaria monitoring and surveillance is an understanding of the epidemiology of *Plasmodium* in the location of interest.

1.5 MALARIA IN ETHIOPIA

Malaria epidemiology in Ethiopia is diverse as a result of the varied topography and ecology across the country. However, a large proportion of the Ethiopian population lives in areas which are generally understood to be at the fringes of the malaria map; characterised by strongly seasonal and epidemic-prone malaria, and at risk of both *P. falciparum* and *P. vivax*. The Ethiopian setting presents complex challenges to malaria surveillance as well as to

periodic monitoring and evaluation of the malaria control programme, and is further constrained by the limited resources available in this low-income country. This section presents an overview of malaria epidemiology in Ethiopia and highlights some malaria surveillance approaches specific to Ethiopia.

1.5.1 Transmission patterns across Ethiopia

Plasmodium falciparum has generally been presumed to be the most common *Plasmodium* species causing malaria in Ethiopia, with the ratio of *P. falciparum* to *P. vivax* ratio usually reported as 60:40 and cross-sectional surveys identifying more *P. falciparum* than *P. vivax* infections [195,196]. However Ethiopia reports more *P. vivax* infections than any other country globally [64]. Instances of both *P. malariae* and *P. ovale* have been indicated in Ethiopia, but they are generally understood to result in clinically mild symptoms and can be easily treated by chloroquine [197-199]. A shift in the relative importance of *P. falciparum* and *P. vivax* has been hypothesised, with some evidence indicating that a higher proportion of outpatient malaria cases reported since 2005 have been due to *P. vivax*. This transition from *P. falciparum* to *P. vivax* dominance has been demonstrated in other settings, and is thought to be a result of effective malaria control interventions reducing *P. falciparum* transmission [200,201].

Malaria transmission is seasonal in most areas of Ethiopia, with the exception of the far western lowlands bordering South Sudan. Transmission peaks following seasonal rains from February to March and from June to August, although with some variation in the peak and duration of malaria transmission across the country. Variations in altitude, mean annual temperature and rainfall have led to the classification of Ethiopia into estimated epidemiological strata according to these three variables (Figure 1.4). The seven strata include areas that are expected to be malaria-free, areas affected by occasional epidemics, settings with seasonal transmission and settings with intense transmission.

Clinical records of symptomatic *Plasmodium* infections indicate that *P. falciparum* shows stronger seasonality than *P. vivax* transmission [202]. *Plasmodium vivax* infections tend to be maintained at lower densities than *P. falciparum* due to the *P. vivax* parasite's host-cell preference for reticulocytes [203]. Lower density *P. vivax* infections may result in less overt morbidity and lower probability of the affected individual seeking diagnosis and treatment than for a high-density *P. falciparum* infection. Other potential contributing factors to the less seasonal nature of *P. vivax* include the wider spatial range, as sporogeny is able to take place at lower temperatures than for *P. falciparum*, generation of hypnozoites by *P. vivax* and consequent potential for relapsing infections [204], and the ability of *P. vivax* to generate gametocytes soon after initial infection and at relatively high proportions in low density infections [203].

In Ethiopia, the major vector of *Plasmodium* is *Anopheles arabiensis*, with secondary vectors including *An. funestus* and *An. nili*. Their primary environment, preferred source of blood meals, biting and resting locations, and biting times are reported in Table 1.2. In addition to these three *Anopheles* species, *An. pharoensis* has also been reported as a secondary vector in some settings in Ethiopia [205,206], but limited data are available regarding its habitat and feeding preferences.

Malaria is generally assumed not to be present at altitudes above 2500m [207], and the Federal Ministry of Health in Ethiopia considers highlands over 2500m to be malaria-free [208]. In addition to temperature and altitude influencing probability of mosquito survival and development of sporozoites, the local environment also influences malaria transmission. Small-scale variation in transmission intensity has been described within communities, where increased malaria risk is associated with residence close to local water bodies, including irrigation systems and dams [209-211].

Figure 1.4 - Epidemiological strata defined by the Federal Ministry of Health and World Health Organization, according to altitude, rainfall and temperature.

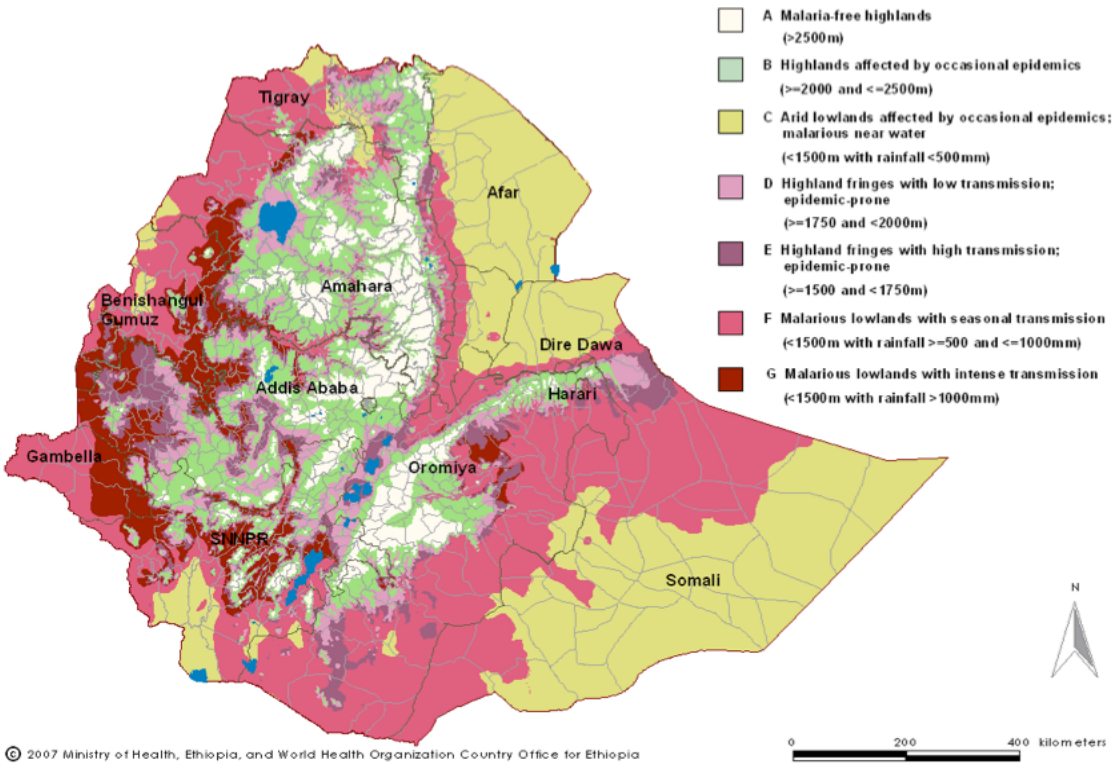


Table 1.2 - Key characteristics of major malaria vectors in Ethiopia, including their primary habitat, preferred blood meal source and biting locations, biting time, and resting location [40,212]

Species	Primary environment	Preferred blood meal	Preferred biting location	Preferred resting location	Biting time
An. arabiensis	Dry savannah, sparse woodland	Both zoophilic & anthropophilic	Exophagic, but may be endophagic	Both endophilic & exophilic	Evening, night and dawn
An. funestus	Swamps, lake edges with emergent vegetation	Anthrophophilic, but may be zoophilic	Endophagic, but may be exophagic	Both endophilic & exophilic	Evening, night and dawn
An. nili	Edges of fast-flowing streams and rivers, in degraded forest and savannah	Anthrophophilic, but may be zoophilic	Both endophagic and exophagic	Mainly exophilic in Ethiopia	Evening and night

1.5.2 Ethiopia's burden of malaria

It has been widely quoted that 68% of the Ethiopian population are at risk of *Plasmodium* infection, this being the proportion of the population who live in areas less than 2000 metres elevation [213]. Before the scale up in malaria interventions, Ethiopia is reported to have experienced up to 10 million cases of clinically suspected malaria a year [213]. A previous retrospective analysis of malaria cases presenting to hospitals and health centres from 2001 to 2006 has reported 900,000 clinical and 560,000 confirmed malaria cases during the five-year period, 60% of which were identified as *P. falciparum* [214]. While Ethiopia is a low transmission setting, and numbers of cases and malaria-attributable deaths are lower than in other countries of sub-Saharan Africa, the burden of disease during epidemics in Ethiopia can be devastating.

Ethiopia was affected by a major epidemic in 1958, which had an attack rate of 30% of those at risk nationally, and case fatality rate estimated at 5-10%, increasing to 20% in areas affected by food insecurity [215]. The 1958 epidemic was attributed to an unusually extended rainy season, with rainfall levels exceeding previous records, along with uncommonly high temperatures. Subsequent major malaria epidemics in East Africa have also been attributed to unusual climatic conditions, and their severity exacerbated by nutritional crises and reduced efficacy of first-line malaria treatment [133,134,216-219], with the most recent major epidemic in Ethiopia occurring in 2003 [220,221]. In Ethiopia, epidemics have historically occurred on a cyclical basis every five to eight years, potentially a result of global climatic fluctuations such as El Niño events [222], parasite resistance to first-line drug treatment or population movements [133,134,216].

1.5.3 Malaria control in Ethiopia

Parasitological diagnosis of all suspected malaria cases is targeted by the Ministry of Health [213]. Expansion of microscopy services at mid-level health facilities and availability of RDTs

at community-level health facilities are enabling this transition from clinical to parasitological diagnosis. From 2009, RDTs able to detect only HRP2 began to be phased out, replaced with combination HRP2 and pan-*Plasmodium* lactate dehydrogenase (panLDH) tests, allowing identification of both *P. falciparum* and non-falciparum *Plasmodium* infections by RDT.

The current first-line treatment for *P. falciparum* is artemether-lumefantrine, an ACT, which replaced SP in 2004. In 2012, the treatment guidelines were once again updated to improve case management of severe malaria by recommending intravenous or rectal artesunate at health posts as pre-referral treatment, as well as promoting the use of parasitological diagnosis rather than presumptive treatment of fever with antimalarials [213]. The first line treatment for *P. vivax* remains chloroquine, but mixed infections (including RDT HRP2 and panLDH positive cases) should receive ACT. There are some indications of developing resistance of *P. vivax* to chloroquine in Ethiopia [223-226]. The national treatment guidelines recommend radical cure with primaquine for patients with *P. vivax* infection residing in non-endemic areas who are being treated at health centres or hospitals, but primaquine is not recommended for use at the health post level due to the risk of haemolysis and lack data on prevalence of G6PD deficiency.

Ethiopia has no historical culture of use of mosquito nets while sleeping. Vector control for malaria in Ethiopia has long been focussed on indoor residual spraying, targeted to areas defined by expert opinion and local health authorities as those at highest risk. Other vector control strategies such as larval source control have not been widely used in Ethiopia.

Ethiopia first introduced insecticide-treated mosquito nets in 1997, and drastically increased coverage with mosquito nets from 2004 with support from the Global Fund to Fight AIDS, Tuberculosis and Malaria. From 2005 to 2011 more than 47 million LLINs were imported to Ethiopia and distributed free. Malaria Indicator Surveys reported an increase in the proportion of households with access to an LLIN from 5% in 2004 to 55% in 2011.

Indoor residual spraying has been a component of malaria control in Ethiopia since the Malaria Eradication Service was established in 1959. For many years DDT was the primary insecticide used for IRS. The Malaria Eradication Service successfully reduced the burden of malaria in many locations in Ethiopia, although epidemics still occurred periodically. Due to emergence of resistance to DDT among the vector population, malathion was introduced as an alternative to DDT for IRS in the 1990s. Due to documented vector resistance to both DDT and malathion, alternative insecticides have been used for IRS since 2009; initially deltamethrin, but expanded to include propoxur, bendiocarb and fenithrion. The newer insecticides used for IRS have a shorter protective duration (two to three months) than DDT, and are more expensive, limiting coverage. An integrated vector control strategy has since been developed, to coordinate insecticides used in LLINs and for IRS to minimise development of resistance to additional insecticides among the vector population. Coverage of IRS has increased from protecting less than 4.2 million people annually in 2005 to more than 20 million people protected in 2011.

The application of malaria intervention tools to the epidemiological strata (defined by rainfall and altitude) is described in Table 1.3, together with the estimated population in 2012 resident within each stratum.

Table 1.3 - Characteristics of the Federal Ministry of Health defined malaria epidemiological strata, the interventions targeted to each strata, and parasite rate (all ages) measured in each strata during Malaria Indicator Survey 2011.

Epidemiological strata	Population 2012 (% national)	Altitude range (m)	Mean annual rainfall (SD)	Mean temperature (SD)	Interventions applied		
					CM	IRS	LLIN
Highly endemic: stable lowland region	2,272,763 (3)	372-1000	935 (298)	25 (1)	Y	Y	Y
Moderate risk: unstable epidemic prone semi-arid regions	3,450,561 (4)	<1000	296 (114)	27 (2)	Y	Y	Y
Moderate risk: unstable epidemic prone midland regions	23,131,634 (27)	1001-1750	1173 (375)	21 (1)	Y	Z	Y
Highly epidemic prone: low risk unstable highland fringe regions	18,071,725 (21)	1751-2000	1225 (350)	19 (1)	Y	Y	S
Epidemic prone, periodic malaria transmission	13,222,382 (16)	2001-2500	1222 (356)	17 (0.7)	Y	Z	N
Malaria-free	24,217,644 (29)	2501-4522	1197 (258)	15 (2)	Y	N	N
Total malaria risk	*60,149,065 (71)*	*-214-2250*	*1168 (382)*	*20 (2.5)*	-	-	-
Grand total	*84,366,709*	*-214-4522*	*1176 (353)*	*18 (3)*	-	-	-

CM = case management; IRS = indoor residual spraying; LLIN – long-lasting insecticidal mosquito net

Y = intervention is provided

N = intervention is not provided routinely

S = LLINs applied in high malaria risk urban areas and where there is vector resistance to IRS insecticides

Z = IRS only in emergencies or epidemics

1.5.4 Malaria surveillance in Ethiopia

Malaria surveillance activities in Ethiopia include both periodic monitoring surveys and surveillance through generation and analysis of data from health facilities.

A DHS was conducted in 2011 in Ethiopia, including both household and women's questionnaires, haemoglobin measurement, and collecting blood samples for HIV testing at central level [227]. No specific malariometric components were included in the DHS, but questions were included to capture information on coverage of key malaria interventions and treatment seeking behaviours for children less than five years with fever. It is also possible to calculate all-cause child mortality from information generated during DHS.

The 2011 Malaria Indicator Survey sampling frame was designed to generate an estimate of levels of indoor residual spraying that was representative nationally and within the administrative Regions with highest population (Oromia, Tigray, Amhara and Southern Nations, Nationalities and People's Regional State). Blood sampling for haemoglobin measurement, RDT and microscopy examination was included in the survey [228]. Households were also interviewed to determine mosquito net ownership and use by household members, and information collected on treatment seeking behaviour by caregivers of febrile children under five years.

The HMIS system in Ethiopia is a paper-based system incorporating a wide range of health indicators collected from individuals presenting at government hospitals and health centres. While HMIS generates rich data, the lag between data recording and analysis as well as feedback means that HMIS is more useful as a tool to monitor the burden of disease over time, rather than as a more responsive epidemic-detection tool. To meet those needs of improved timeliness, the IDSR system was adopted in Ethiopia. IDSR reports are generated monthly from public hospitals and health centres and are submitted on paper to the supervising local health office, where they are entered into a database for use at higher levels [229].

In 2009, further adaptations were made to IDSR, enabling incorporation of data from health posts in a new Public Health Emergency System for surveillance. The policy change was partly

made in recognition of the high levels of population coverage that the health extension system has achieved, and utility of health posts' data for epidemic detection. Two strategies for epidemic detection are recommended by the Federal Ministry of Health in Ethiopia [230].

The first epidemic detection strategy involves comparison of weekly total malaria cases at a health facility against a defined threshold, based on historical morbidity data for the same calendar week [213]. This is referred to as the epidemic monitoring chart method in Ethiopia. Where five years' data are available, the threshold is the third highest weekly total cases for the specific calendar week over the previous five years (the quartile method). In practice, health posts often do not have five years' complete data and are permitted to use their own judgement to determine when "unusual" levels of malaria are presenting at the health post (a subjective threshold). The epidemic monitoring chart method is therefore limited by the need for complete weekly data from the previous five years at each health facility, or is dependent on subjective decision of the health extension worker as to whether the current burden is abnormal. The epidemic monitoring chart is also limited by an assumption that transmission peaks occur in the same weeks each year, while in practice the annual peak of transmission will likely fluctuate, and therefore week-on-week comparisons over multiple years are less informative if the epidemic monitoring chart is strictly applied. The standard epidemic monitoring chart system is designed for settings with relatively higher transmission.

An alternative method undergoing gradual introduction in areas of very low transmission in Ethiopia is a cluster mapping strategy [213]. Cluster mapping is not anticipated to involve any active case detection, but involves health extension workers noting the household location of passively detected malaria cases on a kebele map. If more than three confirmed malaria cases presenting at the health post are found to reside within a 1km radius of each other, and all are diagnosed within 28 days, then this is assumed to indicate local transmission and termed a "micro cluster". Information on the rate of case accumulation within micro clusters, as well as changes in the spatial extent of micro clusters over time can provide information to

the health extension workers on the changing malaria burden and whether additional resources are likely to be needed to reduce onward transmission.

1.6 AIMS AND OBJECTIVES

The primary aim of this thesis is to evaluate the operational feasibility and agreement between existing and alternative strategies and indicators used for malaria monitoring and surveillance in Ethiopia, with a view to generating recommendations of the most informative and appropriate tools to meet future surveillance needs in areas of low and unstable malaria transmission. My research is conducted in Oromia Regional State and in Southern Nations, Nationalities and Peoples' Regional State as part of operational research conducted by Malaria Consortium with funding by the President's Malaria Initiative.

The specific objectives include:

- To describe the epidemiology of malaria in Oromia Regional State, Ethiopia, by use of school-based surveys.
- To evaluate the utility of serological indicators of malaria exposure measured among school children to describe and predict the spatial heterogeneity of malaria transmission across Oromia Regional State.
- To investigate the potential of syndromic and surrogate data for malaria epidemic detection in schools in Southern Nations, Nationalities and Peoples' Regional State.
- To examine correlation between malaria indicators collected through a range of commonly used surveillance strategies, including routine health facility data, cross-sectional survey estimates, and modelling predictions.

1.6.1 Thesis outline

Chapter 2 describes the design and implementation of large-scale cross-sectional school-based surveys in Oromia Regional State, Ethiopia. Findings reported include anaemia levels, parasite rate according to microscopy, risk factors for *Plasmodium* infection, and geographical distribution of *Plasmodium* parasitaemia. Estimates of infection prevalence reported in Chapter 2 were too low to capture the expected diversity in malaria transmission across Oromia, therefore Chapter 3 presents findings of serological analysis of samples collected during school surveys and generation of predictive risk maps based upon modelled seroprevalence, using Bayesian geostatistical modelling. In Chapter 4, the potential use of schools for syndromic surveillance and epidemic detection was investigated in southern Ethiopia, focussing on the use of school absenteeism as a supplementary indicator of potential malaria epidemics, alongside existing epidemic monitoring systems at community health facilities. Chapter 5 draws upon a range of data collected through multiple surveillance platforms, including routinely recorded data at government health facilities, cross-sectional data from school surveys and community surveys, as well as modelling predictions of parasite rate. These data are compared across a several dimensions including correlation between indicator values in spatially matched locations, temporal resolution, and endemicity classification. Chapter 6 discusses the primary findings and future implications of the work presented in this thesis.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

COVER SHEET FOR EACH 'RESEARCH PAPER' INCLUDED IN A RESEARCH THESIS

Please be aware that one cover sheet must be completed for each 'Research Paper' included in a thesis.

1. For a 'research paper' already published

1.1. Where was the work published? ..... Malaria Journal .....

1.2. When was the work published? ..... February 2011 .....

1.2.1. If the work was published prior to registration for your research degree, give a brief rationale for its inclusion

.....

.....

.....

1.3. Was the work subject to academic peer review? ..... Yes .....

1.4. Have you retained the copyright for the work? **Yes / No**
If yes, please attach evidence of retention.
If no, or if the work is being included in its published format, please attach evidence of permission from copyright holder (publisher or other author) to include work

2. For a 'research paper' prepared for publication but not yet published

2.1. Where is the work intended to be published? .....

2.2. Please list the paper's authors in the intended authorship order

.....

2.3. Stage of publication – Not yet submitted / Submitted / Undergoing revision from peer reviewers' comments / In press

3. For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)

Contributed to study design and conceptualisation. Led field data collection, ensured quality of data entered. Conducted frequentist analysis, prepared first draft of paper and revised to final draft following co-author inputs.

NAME IN FULL (Block Capitals) ..RUTH ASHTON.....

STUDENT ID NO: .....181337.....

CANDIDATE'S SIGNATURE .....RAA..... Date ..12/04/14.....

SUPERVISOR/SENIOR AUTHOR'S SIGNATURE (3 above) .....

CHAPTER 2. SCHOOL-BASED SURVEYS OF MALARIA IN OROMIA REGIONAL STATE, ETHIOPIA: A RAPID SURVEY METHOD FOR MALARIA IN LOW TRANSMISSION SETTINGS

2.1 OVERVIEW

There is a need to regularly collect up-to-date malaria data to inform decision-making in malaria control. Cross-sectional surveys are one monitoring tool that has been widely used, and this chapter presents findings from a large-scale school-based cross sectional malariometric survey, using the standard diagnostic indicator of microscopy. This chapter aims to describe the spatial extent of malaria in Oromia Regional State, and to identify risk factors for *Plasmodium* parasitaemia.

This chapter has been peer reviewed and published in the *Malaria Journal: Ashton RA, Kefyalew T, Tesfaye G, Pullan RL, Yadeta D, Reithinger R, Kolaczinski JH, Brooker S, 2011. School based surveys of malaria in Oromia Regional State, Ethiopia: a rapid survey method for malaria in low transmission settings 10:25.* I contributed to the study design, which was conceived by Simon Brooker, Jan Kolaczinski and Richard Reithinger. I planned and led all field implementation, and conducted all frequentist data analysis. Rachel Pullan conducted Bayesian analysis of these data.

2.2 BACKGROUND

Following the recent achievements in global malaria control [231], there is increased emphasis on monitoring these achievements and on refining the epidemiological landscape in order to determine intervention needs and guide implementation [207]. Household surveys, including Malaria Indicator Surveys (MIS) [195], Demographic Health Surveys [227] and Multiple Indicator Cluster Surveys [106] are commonly used to achieve these surveillance and monitoring goals, but they are expensive, time-consuming and technically complicated to undertake. A complementary, inexpensive framework for malaria surveillance may be provided by school malaria surveys [68], which were an important component of early, particularly colonial, malaria reconnaissance, and more recently have contributed towards a nationwide assessment of malaria in Kenya [16].

Building on the Kenyan experience, this chapter presents results from a large-scale school survey of malaria in Ethiopia. Malaria transmission in Ethiopia is temporally and spatially dynamic [9], with transmission unstable, seasonal, and linked to environmental variables such as altitude and rainfall [202]. In recent years, there has been a marked scale-up of the distribution of long-lasting insecticidal nets (LLINs) and indoor residual spraying (IRS) in Ethiopia [232]. To track this progress and to capture the inherent heterogeneities of malaria transmission in the country, various community-based malaria surveys have been carried out at national and sub-national levels [195,196,233]. The aim of the present work was to generate data for Oromia Regional State to assist in targeting malaria control interventions across this heterogeneous transmission setting. Specifically, the objectives were to investigate risk factors for *Plasmodium* infection, and to generate parasitological data for use in geostatistical modelling to generate a risk map predicting *Plasmodium* prevalence in Oromia.

2.3 METHODS

2.3.1 Study setting

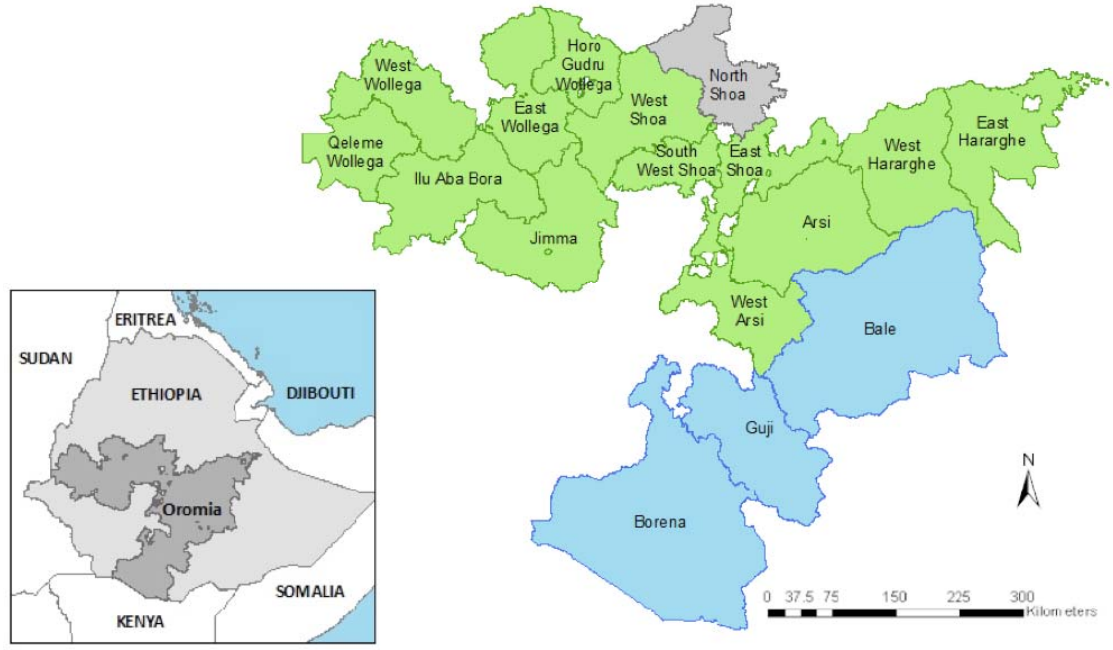
This study was undertaken throughout Oromia Regional State, the largest of Ethiopia's 11 regional states. Oromia covers approximately one third of the country's landmass (Figure 2.1) and has a population of 27 million [234], an estimated 17 million of whom are at risk of malaria [235]. It is divided into 17 administrative zones, as defined in Central Statistics Agency (CSA) 2007 census data [234], with each zone divided further into woredas (i.e. districts) followed by kebeles (i.e. municipalities).

Oromia is geographically diverse, encompassing arid lowlands, fertile and well-vegetated areas with high rainfall, and cool mountainous regions. The study was conducted in two phases in order to coincide with the historical peak of the malaria transmission season one-two months after the main rainy season: schools in the southern zones of Borena, Guji and Bale were surveyed in May 2009, while schools in all other zones of Oromia were surveyed between October and December 2009 (Figure 2.1).

2.3.2 Sample size and school selection

Oromia was divided into ecological strata defined according to elevation and rainfall ranges estimated to confer differences in suitability for *Anopheles* breeding and survival, and therefore endemicity of *Plasmodium*. These classifications are used by the Federal Ministry of Health and the World Health Organization (WHO) Ethiopia (Table 2.1) to assign malaria prevention and control interventions. Malaria transmission is assumed not to occur in arid areas (<500mm annual rainfall) and highlands (>2,500 metres) [207] and so these strata were not sampled.

Figure 2.1 - Location of Oromia Regional State within Ethiopia (inset) and administrative zones surveyed in May 2009 (blue) and October to December 2009 (green). North Shoa (grey) was not surveyed as no schools in this zone were randomly selected for inclusion



The aim of the sample size calculation was to estimate the prevalence of *Plasmodium* infection in each ecological zone, and to have sufficient power to detect changes in infection prevalence due to the interventions provided by the national control programme. A two-stage sampling design was employed, whereby schools (primary sampling unit) were selected using probability proportional to size, then within schools a fixed number of children (secondary sampling units) were randomly selected. Therefore, the number of schools sampled from each ecological zone was proportional to the number of schools in each zone, with the exception of the 'highland occasional epidemic' zone (2,000 – 2,500 metres), which was under-sampled as a result of low expected prevalence and a need to maximize the power of the survey in more stable transmission areas. These criteria excluded one administrative zone from the sampling frame (North Shoa, positioned at >2,500 metres), while the remaining 16 administrative zones in Oromia were included in the survey. The fixed number

of children sampled per school (100 plus ten reserves) was the maximum number of children feasible to sample in a single day, based on experience in similar surveys in Kenya [16].

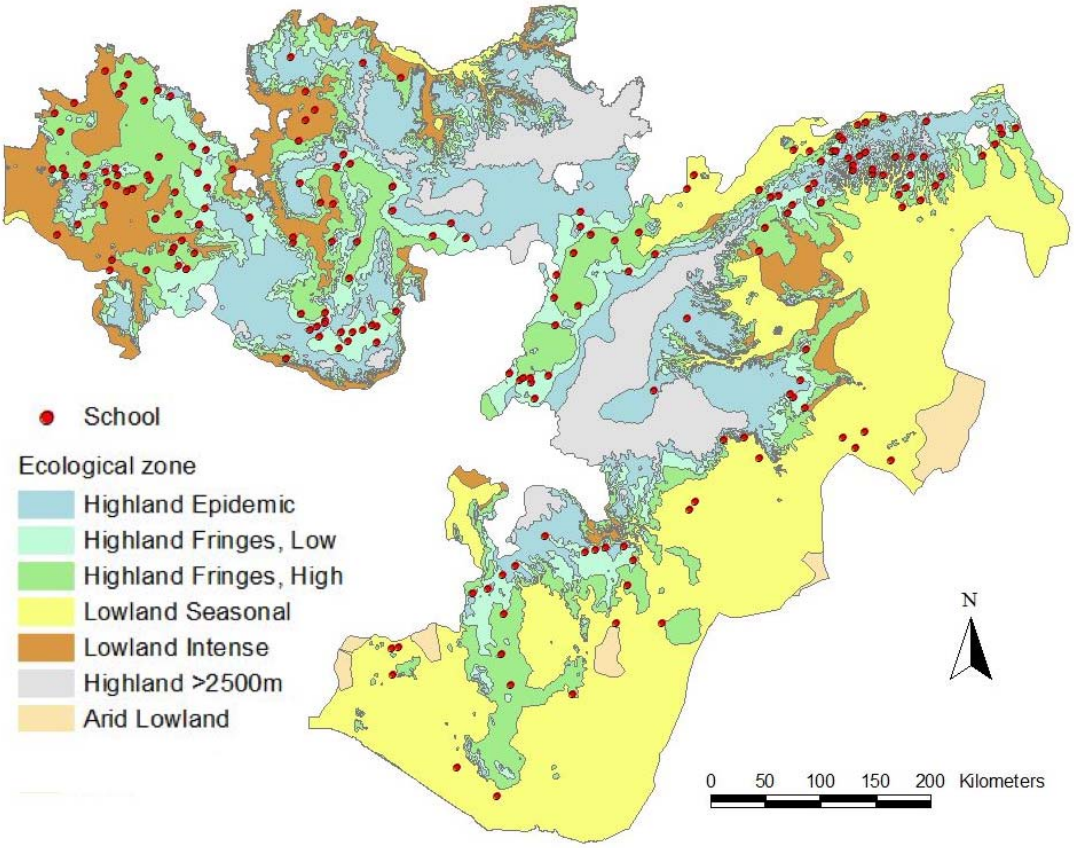
The sample size was determined using 95% confidence limits, 80% power and assuming design effect of 2, aiming to detect a prevalence of 1% with 0.5% precision. It was consequently estimated that the sample size for each ecological zone should be 3,925 children from 40 schools. . Sampling 40 schools from each of five ecological zones would give a final sample size of 20,000 children from 200 schools. It was decided to select schools from ecological zones using probability proportional to population size, due to the uneven distribution of primary schools in Oromia between ecological zones (Table 2.1 and Figure 2.2). During the survey, two schools were found to be inaccessible, and one school director refused consent for the survey. No replacement schools could readily be found within time and, thus, 197 schools were included in the final sample.

Table 2.1 - Sampling stratification used to select schools in Oromia Regional State, Ethiopia, based on ecological zones defined according to epidemiologically significant differences in elevation and rainfall, based on classifications used by the Ministry of Health and WHO Ethiopia office [208]

Stratum description	**Elevation (m asl)[1]**	**Stratum**	**Total schools in stratum**	**Proportion of sample**	**Schools sampled**
Highland, occasional epidemic	2,000-2,500	1	1651	0.1	20
Highland fringe, low unstable transmission	1,750-2,000	2	1209	0.409	73
Highland fringe, high unstable transmission	1,500-1,750	3	1163	0.393	71
Lowland with seasonal transmission (annual rainfall 500-1000 mm)	<1,500	4	389	0.132	24
Lowland with intense transmission (annual rainfall >1000 mm)	<1,500	5	196	0.066	12
Highland	>2,500	6	592	Not included in survey	
Arid lowland (annual rainfall <500 mm)	<1,500	7	4	Not included in survey	

[1] Metres above sea level

Figure 2.2 - Distribution of selected schools in relation to ecological zones in Oromia Regional State, Ethiopia.



2.3.3 Participants

Community sensitization was conducted using a cascade approach. Oromia Regional Health and Education Bureaus gave approval for the study, and official letters were sent to the Bureaus' zonal offices, then to woreda offices and to individual school directors. Information about study procedures and schedule were provided and school directors advised to hold a meeting with the school committee and parents in advance of the designated survey day. Parents who did not want their children to participate in the study were free to refuse participation. Children who were unwilling to participate were excluded from random selection, with written (or thumbprint) assent obtained from selected children before

samples were collected. In each school, 10 boys and 10 girls (plus one reserve boy and one reserve girl) aged between five and 18 years were selected from each of grades two to six using computer-generated random number tables. In practice, this was achieved by requesting all children in each grade to form one line for girls and one for boys, standing in any order. Random number tables were used to count down the line of children and identify those for inclusion. If fewer than 110 children were enrolled in the school or present on the survey day, all children aged five to 18 present in the school were included in the survey; this, in some instances, resulted in a small school sample size. Detailed logistical and ethical considerations for this style of school survey have been presented elsewhere [16].

2.3.4 School survey procedures

Finger-prick blood samples were used to prepare thick and thin blood films for microscopy, and haemoglobin concentration was estimated to an accuracy of 1 g/L using a portable haemoglobinometer (Hemocue Ltd, Angelhölm, Sweden). In addition, blood spots were collected on filter paper for serological analysis at a later date [184,186]. Children were asked a simple set of standardized, pre-tested questions on recent fever, mosquito net use, whether IRS had been conducted in their households, key household socio-economic variables, household construction and education of the child's guardian. Children reporting fever or found to be anaemic (Hb <80 g/L) were tested with a multi-species malaria rapid diagnostic test (RDT) (CareStart® Pf-HRP2 Pan-pLDH, Access Bio, USA) to allow immediate diagnosis and treatment. This test was shown to have 85.6% sensitivity and 92.4% specificity for *Plasmodium falciparum*, and 85.0% sensitivity and 97.2% specificity for *Plasmodium vivax* in Ethiopia [168]. The location of each school was measured in decimal degrees using a hand-held global positioning device (eTREX, Garmin International, Kansas, USA).

2.3.5 Microscopy quality control

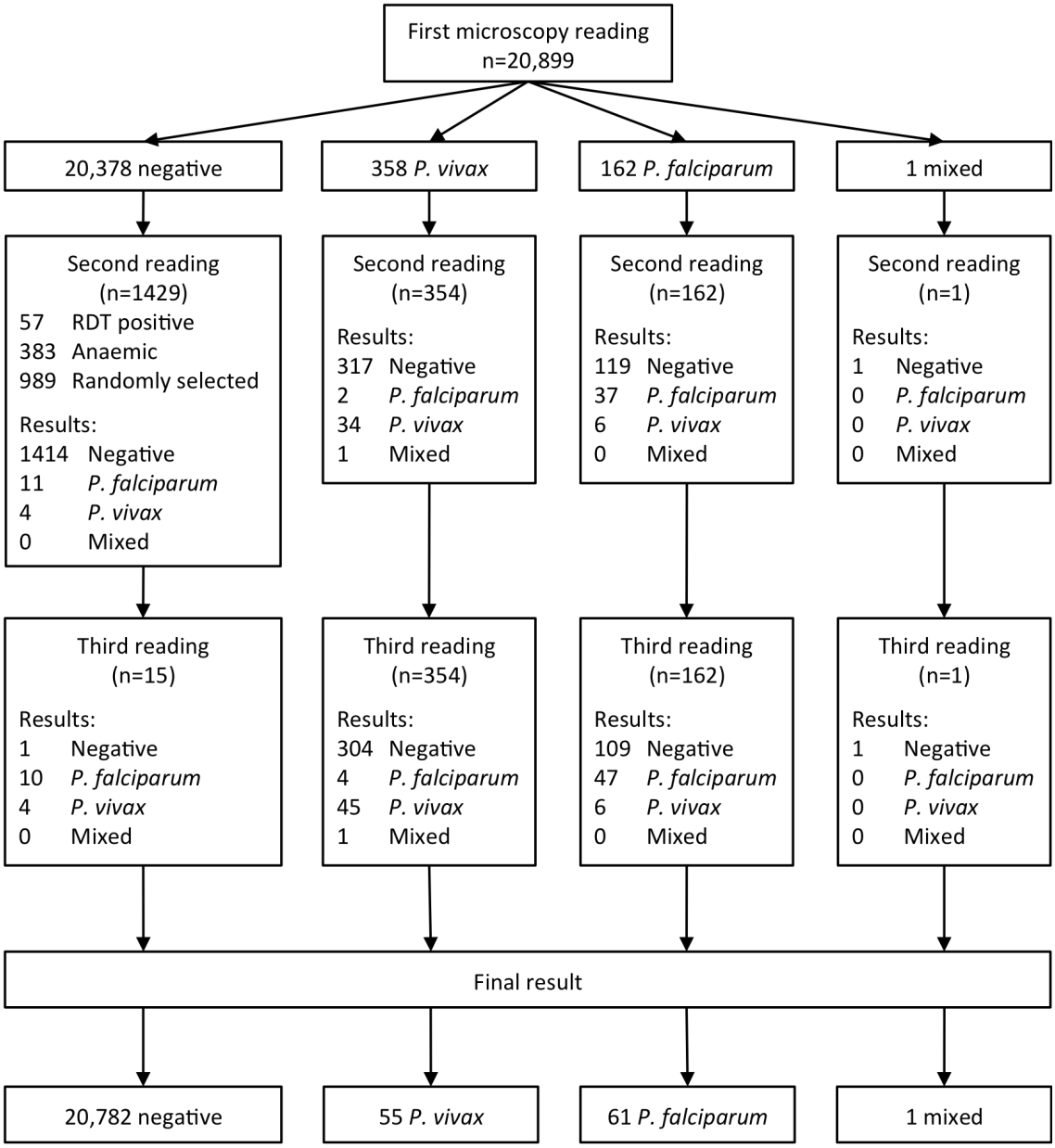
Blood films were fixed and stained at a local health centre after the survey following standard operating procedures [236], and examined after completion of field work by experienced laboratory technicians in Addis Ababa. *Plasmodium* species was recorded, but quantification of parasite density was not conducted. A second reading was carried out for a proportion of blood films, by highly experienced microscopists at the malaria reference laboratory for Oromia Regional State in Adama. Criteria for a second microscopy reading were: slides positive for *Plasmodium* spp. at first microscopy reading; individuals with discrepant microscopy and RDT results; severely anaemic individuals (<80 g/L); and a randomly selected 5% of negative slides. Slides with discrepant results between first and second readings were settled by a third, expert microscopist from the Ethiopian Health and Nutrition Research Institute, the national reference laboratory in Addis Ababa (Figure 2.3).

2.3.6 Satellite-derived environmental data

Elevation was extracted from the shuttle radar topography mission (SRTM) digital elevation model at 1km^2 resolution. Population density was extracted from gridded population of the world in 2000 (GRUMP) at 5km^2 resolution [237,238]. Land cover type was extracted from the qualitative global land cover map 2005 (defined within the UN Land Cover Classification System) using environmental satellite (ENVISAT) mission's Medium Resolution Imaging Spectrometer (MERIS) sensor at 5km^2 resolution. The distance to permanent water bodies was extracted from the World Wildlife Fund (WWF) Global 200 Ecoregions database at 5km^2 resolution [239]. Estimates of enhanced vegetation index (EVI; a proxy for vegetation coverage) and land surface temperature (LST) at 5km^2 resolution were extracted from data provided by the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument aboard the Terra (EOS AM) and Aqua (EOS PM) satellites [240], for the years 2001-2008. Data were processed by a temporal Fourier algorithm, which transforms a series of observations taken over a period of time into a set of uncorrelated harmonics that sum to the original time

series. The process filters noisy data, but preserves harmonics corresponding to biologically-relevant annual, bi-annual and tri-annual cycles of seasonal changes [241,242]. Environmental variables were linked by location to school-level parasitological data using ArcGIS 9.3 (Environmental Systems Research Institute Inc., Redlands, CA, USA). Since the resolution of most environmental data was $5km^2$, it was not deemed necessary to extract environmental data from a defined buffer zone around the school location.

Figure 2.3 – Microscopy results quality control flowchart



2.3.7 Data analysis

Microscopy results were entered into a Microsoft Excel 2007 spreadsheet (Microsoft Corporation, Seattle, USA). Questionnaire data from school surveys, including RDT results and haemoglobin measurements, were entered into a customized Microsoft Access 2007 database that had been developed to automatically conduct range and consistency checks. Any errors or inconsistencies were corrected with reference to the original paper forms. Survey data were exported from Access and Excel into a combined dataset in STATA 9.0 (Stata Corporation, College Station, TX, USA) for cleaning. Point prevalence maps were developed in ArcGIS 9.3.

Individuals aged over 18 years (n=260) or with missing parasitological data (n=270) were excluded from the school survey analysis. Anaemia was defined according to WHO classifications, adjusted by age and elevation [243]. The number, gender ratio and age distribution of children included was described, with breakdown by ecological stratum and survey period. Child age in years was classified into groups: five to nine years, ten to 14 years, and 15 to 18 years. Main outcomes, i.e. any *Plasmodium* infection, *P. falciparum* and *P. vivax* infection individually and anaemia, were presented with binomial 95% confidence intervals (CI) by sex, age group, ecological stratum and survey period. Use of malaria prevention measures, specifically LLINs and IRS, were presented by sex, age group, ecological stratum and survey period with binomial 95% CI, with associations tested using Chi squared test.

Crude univariate associations between outcomes (i.e. *P. falciparum* or *P. vivax* infection) and individual covariates were assessed by random effects logistic regression to control for clustering of infection by school. Full multivariable models to describe association between LLIN use or IRS with *Plasmodium* infection were developed, using zero-inflated Poisson (ZIP) models to account for the large proportion of schools with zero prevalence. ZIP models were favoured over standard Poisson models on the basis of the Vuong test [244].

Tests for associations between school-level prevalence and environmental covariates were performed using grouped logistic regression models taking into account clustering within schools. All significant covariates ($p<0.1$) were subsequently included in full multivariable models, and non-significant (Wald test $p<0.5$) covariates excluded sequentially in order of least significance to generate minimal adequate models. Excluded covariates were retested in the minimal model using likelihood ratio test to confirm lack of significance. Bayesian spatial multivariate models were then developed in WinBUGs version 1.4 (MRC Biostatistics Unit, Cambridge and Imperial College London, UK) to explicitly model spatial correlation between schools. The number of examined and slide-positive individuals for each species at each survey location were modelled as binomial outcomes, including covariates as described above and a geostatistical random effect that modelled spatial correlation using an isotropic, stationary exponential decay function [245].

To further investigate the distribution of *P. falciparum* and *P. vivax* in Oromia, the existence of spatial clusters of high malaria prevalence were investigated using Kulldorff's spatial scan statistic (version 7.0.2; SaTScan software [246]). A Poisson model was used, under the null hypothesis that the expected number of cases for each area was proportional to its population size. The rate ratio was defined as the observed to expected cases; significance of identified clusters was tested by likelihood ratio, based on 9,999 Monte Carlo simulations. A circular scanning window was used, but no maximum radius was set for the cluster.

2.3.8 Ethical considerations

This study received ethical approval from the national health research ethics review committee of the Ethiopian Science and Technology Ministry (RDHE/2-89/2009). Approval for the study was given by the Oromia Regional Health Bureau and the Oromia Regional Education Bureau.

Written consent for the survey was provided by each school director, but parents maintained the right to withdraw their child from the survey. Each child selected for inclusion was required to provide written assent (or thumbprint) after having the procedures explained. Schools where the director refused consent were not included in the school survey, and pupils refusing assent were excluded. Individuals with a positive malaria RDT were treated according to Ethiopian national guidelines [247]. Individuals with haemoglobin <80g/L were provided with a two-week dose of ferrous sulphate tablets and instructions on how to take this medication, and advised to attend the health centre for follow-up.

2.4 RESULTS

2.4.1 School survey participants

A total of 21,166 children, age five to 18 years (median 11, inter-quartile range 9-12 years), from 197 rural primary schools in Oromia took part in the survey, with a similar number of boys and girls included (53.2% male). A mean of 106 children were enrolled from each school (range 43-112). Blood films from 267 children were missing or unreadable. Consequently, these individuals were excluded from analysis, leaving 20,899 children (98.7%).

2.4.2 Reported use of malaria interventions

Overall, 46.0% (95% CI: 45.3-46.7%) of school children reported using a LLIN the previous night. In locations where the school elevation exceeded 2,000m, 42.0% (95% CI: 39.7-43.6) of children reported using a LLIN; however, this may be in part attributable to kebeles with large altitude ranges still falling within the National Malaria Control Programme (NMCP) objectives to target areas under 2,000m with LLIN distribution [232]. Reported LLIN use was lower amongst males than females (43.4% vs. 49.0%, p<0.001) and amongst children aged 15-18

years than other age groups (39.4% vs. 46.4% for five to nine years and 46.6% for 10-14 years, p<0.001).

2.4.3 Malaria and anaemia

The overall prevalence of *Plasmodium* infection was 0.56% (95% CI: 0.46-0.67%). Of children with *Plasmodium* infection identified by quality assured microscopy, 52.1% were infected with *P. falciparum*, 47.0% with *P. vivax* and 0.9% with mixed infections. Mixed infections were not analysed separately but included with each of the single species infections. The overall prevalence of *P. falciparum* was 0.30% (95% CI: 0.23-0.38%) and *P. vivax* 0.27% (95% CI: 0.20-0.35%). There was no evidence for a difference in *Plasmodium* prevalence between phases one and two of data collection (p=0.513), but proportion of children who were anaemic was higher in administrative zones sampled in phase one than in phase two (p=0.002).

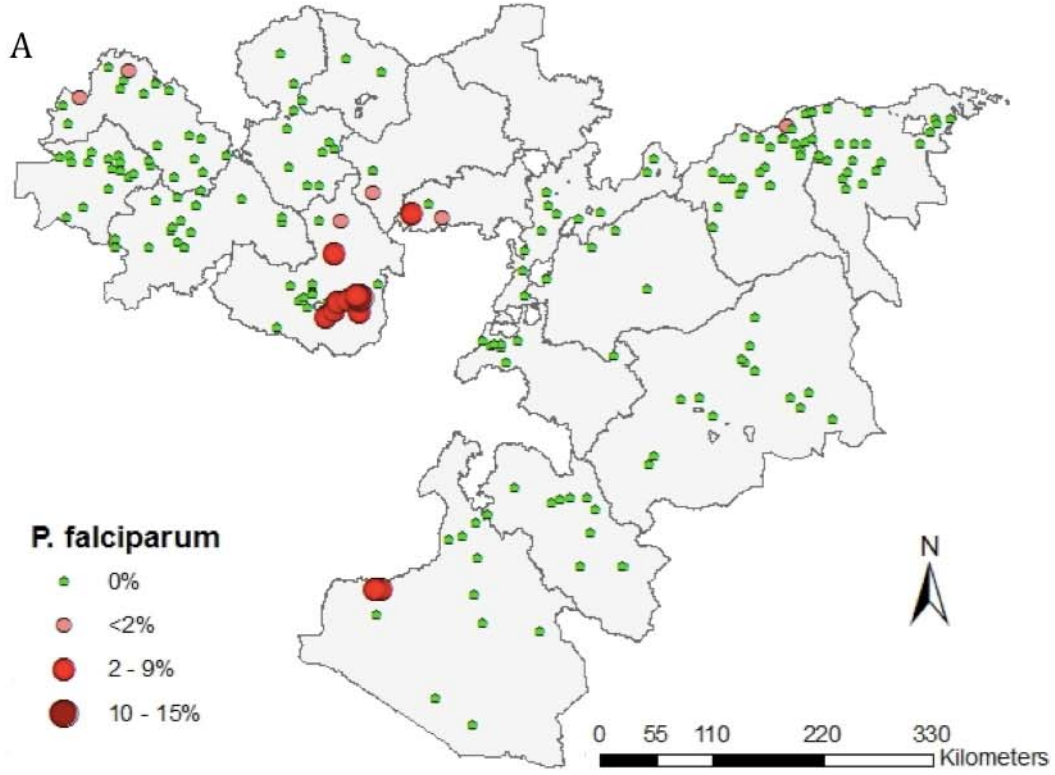
Only 18% of children with *P. falciparum* reported fever on the day of the survey; however, 72% had had fever in the past month. A greater proportion of *P. vivax* infections were asymptomatic, with only 7% of children reporting fever on the survey day, but 56% reporting to have felt fever in the past month. In total, 17.6% of children were anaemic, and the mean haemoglobin concentration was found to be 132.8 g/L (95% CI: 132.6-133.0).

Figure 2.4 shows the geographical distribution of *P. falciparum* and *P. vivax* prevalence by school. Thirty schools (15%) were found to have at least one child with *Plasmodium* spp. infection on the day of the survey (17 *P. falciparum* and 24 *P. vivax*), and prevalence by school ranged between 0 and 14.5%. However the survey sampling frame was designed to estimate the prevalence of infection by ecological zone, not within individual schools. All schools with detectable infection were located at an elevation between 1,183 and 2,187 metres above sea level. The median time taken to walk to school, an indicator of distance the child lives from school, was 30 minutes (range 0-240 minutes). The schools with highest

prevalence of infection were found in Jimma and South West Shoa administrative zones (Figure 2.4).

Plasmodium infections were found in all five ecological strata, with highest prevalence found in the stratum defined as highland fringes with low transmission (Table 2.2). The proportion of infections due to *P. falciparum* was variable between strata: from 10% in highland epidemic to 86% in lowland seasonal (p=0.001). There was strong evidence that the *P. falciparum* and *P. vivax* rates differed by ecological zone (both p<0.001). Prevalence of anaemia varied markedly between schools (Figure 2.5), ranging between 0.9% and 51.4%. Guji administrative zone was seen to have consistently high levels of anaemia at all schools sampled, while the highest prevalence of anaemia was found in the ecological stratum defined as lowland with seasonal malaria (Table 2.2).

Figure 2.4 - Prevalence of (A) *Plasmodium falciparum* and (B) *P. vivax* infection by school, Oromia Regional State, Ethiopia, 2009.



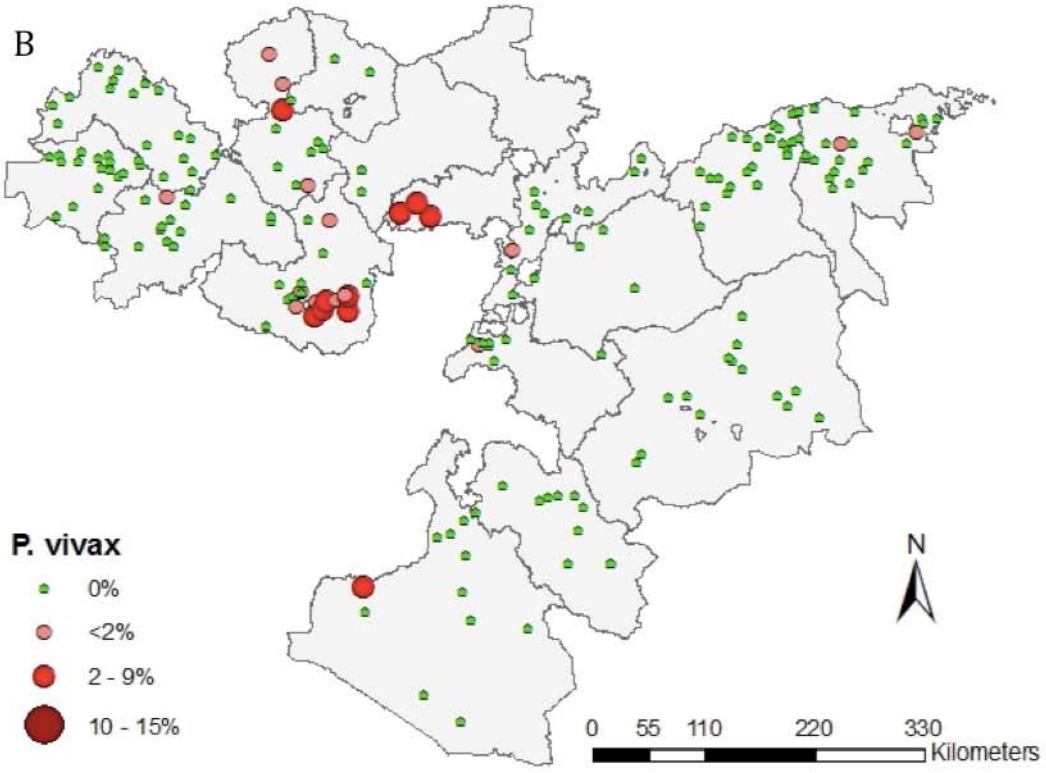


Table 2.2 - Prevalence of *P. falciparum, P. vivax* and anaemia among primary school children in 197 schools in Oromia Regional State, Ethiopia in 2009, by sex, age group, survey phase and malaria transmission zone.

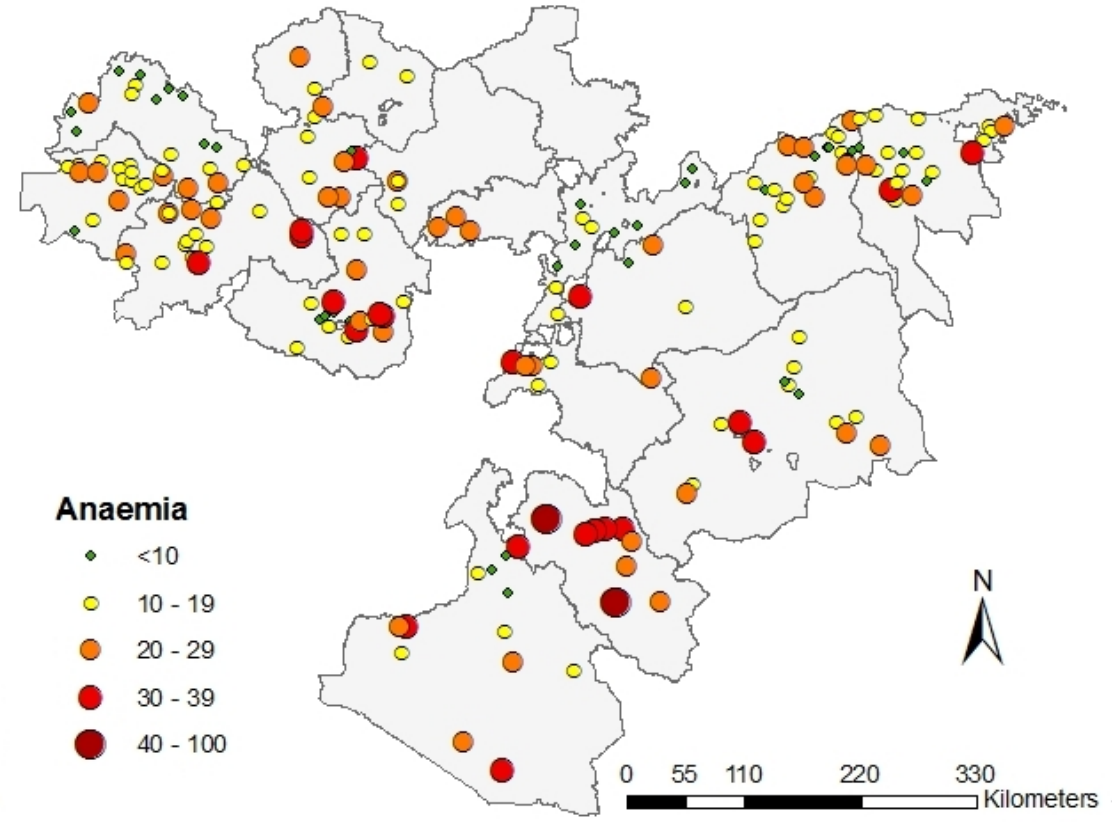
	N[1]	*Plasmodium* prevalence, % (95% CI)	% due to *P. falciparum*	Species prevalence, % (95% CI)		Anaemia, % (95% CI)
				P. falciparum	*P. vivax*	
Total	197 / 20,899	0.6 (0.5-0.7)	0.53	0.3 (0.2-0.4)	0.3 (0.2-0.3)	17.6 (17.0-18.1)
Male	11,038	0.6 (0.4-0.7)	0.57	0.3 (0.2-0.4)	0.2 (0.2-0.3)	19.2 (18.4-19.9)
Female[2]	9,731	0.6 (0.4-0.7)	0.48	0.3 (0.2-0.4)	0.3 (0.2-0.4)	15.7 (15.0-16.5)
5-9 yrs	5,471	0.4 (0.2-0.6)	0.77	0.4 (0.3-0.6)	0.1 (0.05-0.2)	18.8 (17.8-19.9)
10-14 yrs	13,890	0.6 (0.5-0.7)	0.44	0.3 (0.2-0.4)	0.3 (0.2-0.4)	16.4 (15.7-17.0)
15-18 yrs[3]	1,400	0.4 (0.2-0.9)	0.50	0.2 (0.04-0.6)	0.2 (0.04-0.6)	23.4 (22.2-26.7)
Phase 1	36 / 3,779	0.4 (0.2-0.6)	0.86	0.3 (0.2-0.6)	0.05 (0.00-0.2)	22.4 (21.1-23.8)
Phase 2	161 / 17,120	0.6 (0.5-0.7)	0.48	0.3 (0.2-0.4)	0.3 (0.2-0.4)	16.5 (15.9-17.1)
Highland epidemic	22 / 2,358	0.4 (0.2-0.8)	0.10	0.04 (0.00-0.2)	0.4 (0.2-0.7)	14.7 (13.3-16.1)
Highland fringes, low transmission	69 / 7,246	1.1 (0.8-1.3)	0.55	0.6 (0.4-0.8)	0.5 (0.3-0.7)	16.3 (15.5-17.2)
Highland fringes, high transmission	66 / 7,018	0.1 (0.06-0.2)	0.56	0.07 (0.02-0.2)	0.06 (0.01-0.1)	17.2 (16.4-18.1)
Lowland seasonal	24 / 2,540	0.5 (0.3-0.9)	0.86	0.5 (0.2-0.8)	0.08 (0.01-0.3)	24.0 (22.4-25.7)
Lowland intense	16 / 1,737	0.4 (0.2-0.8)	0.14	0.06 (0.00-0.3)	0.3 (0.1-0.8)	18.7 (16.9-20.6)

[1] Number of schools surveyed / children tested

[2] Sex data missing from 130 records

[3] Age data missing from 138 records

Figure 2.5 - Prevalence of anaemia by school, Oromia Regional State, Ethiopia, 2009.



2.4.4 Risk factors for malaria and anaemia

There were no statistical differences in *Plasmodium* infection by sex or by age (Table 2.3), whereas anaemia was more common among males than females (19.2% vs. 15.7%, p<0.001), and in children aged 15-18 years (p<0.001). From crude univariate analysis (Table 2.3), anaemia was found to be strongly associated with *P. falciparum* infection. History of fever in the previous month was associated with both *P. falciparum* and *P. vivax* infection, as well as with anaemia. IRS and LLIN use were found to be associated with increased odds of infection. Odds of *P. falciparum* and *P. vivax* infection were associated with forested areas rather than cultivated land. Some associations with vegetation (EVI) and temperature (LST) were seen, but the effect was small. In multivariate ZIP models, LLIN use was no longer found to be associated with either *P. falciparum* or *P. vivax* infection. The association between IRS in the household and increased risk of infection, however, remained in multivariate models: incidence rate ratio (IRR: transformed model coefficient estimate) of 2.91 for *P. falciparum* (95% CI: 1.36-6.21, p=0.006) and IRR of 3.92 for *P. vivax* (95% CI: 1.82-8.45, p<0.001).

Table 2.3 – Univariate analysis for associations between *P. falciparum* and *P. vivax* and potential risk factors among sampled school children, adjusting for clustering within schools

	P. falciparum			***P. vivax***		
	OR[1]	**95% CI**	**Wald p**	**OR**	**95% CI**	**Wald p**
Sex (female)	0.88	0.57-1.35	0.5	1.27	0.75-2.14	0.4
Age (linear, increasing)	0.89	0.74-1.07	0.2	1.05	0.96-1.16	0.3
Socio-economic status:						
Poorest	1	-	-	1	-	-
2nd	0.92	0.49-1.73	0.8	1.45	0.67-3.11	0.3
3rd	0.60	0.19-1.90	0.4	0.93	0.36-2.38	0.8
4th	1.14	0.49-2.66	0.8	1.10	0.58-2.06	0.8
Least poor	0.69	0.23-2.04	0.5	0.74	0.30-1.78	0.5
Anaemia	6.89	4.02-11.82	<0.001	1.36	0.73-2.54	0.3
Fever	7.89	4.26-14.62	<0.001	4.97	2.71-9.14	<0.001
LLIN use	2.01	1.11-3.64	0.02	1.21	0.67-2.18	0.5
IRS in home	2.69	1.34-5.42	0.005	2.83	1.28-6.23	0.01
Ecozone:						
Highland epidemic	1.00	-	-	1.00	-	-
Highland fringe, low transmission	12.8	1.22-134.1	0.03	1.57	0.49-5.01	0.4
Highland fringe, high transmission	1.64	0.13-20.71	0.7	0.20	0.04-0.96	0.05
Lowland seasonal	8.59	0.75-99.05	0.09	0.24	0.30-1.92	0.18
Lowland intense	1.39	0.05-38.10	0.8	0.88	0.16-4.87	0.8
Enhanced vegetation index (EVI)[2]						
Annual amplitude	1.00	0.996-1.01	0.8	1.00	0.99-1.00	0.5
Bi-annual amplitude	1.00	0.99-1.01	0.6	1.01	1.00-1.02	0.01
Tri-annual amplitude	0.99	0.98-1.01	0.5	0.99	0.99-1.01	0.9
Bi-annual phase	0.99	0.99-0.99	0.05	1.01	0.99-1.01	0.2
Quarterly phase	0.99	0.99-1.00	0.8	0.99	0.99-0.99	0.007
Land surface temperature (LST)						
Annual amplitude	1.00	0.99-1.00	0.3	1.00	0.99-1.00	0.7
Bi-annual amplitude	0.99	0.99-1.00	0.9	1.01	1.00-1.01	0.009
Tri-annual amplitude	0.99	0.98-1.01	0.23	0.99	0.98-1.00	0.1
Bi-annual phase	1.00	1.00-1.00	0.003	1.00	0.99-1.00	0.7
Quarterly phase	0.99	0.99-1.01	0.9	1.01	1.00-1.01	0.002
Distance to water (increasing)	0.67	0.41-1.10	0.1	0.58	0.39-0.86	0.007
Population density (increasing)	0.99	0.99-1.00	0.9	1.00	0.99-1.00	0.2
Land-cover type:						
Cultivated land	1	-	-	1	-	-
Forest	3.81	1.03-14.08	0.05	3.58	1.26-10.24	0.02
Shrubland	2.65	0.32-21.92	0.4	1.39	0.19-10.32	0.8
Bare / sparse	*predicts perfectly*		-	*predicts perfectly*		-

[1] OR = odds ratio; 95% CI = 95% confidence intervals

[2] Variables for EVI and LST describe the phase and amplitude of annual, bi-annual, tri-annual and quarterly cycles after processing data by a Fourier algorithm.

2.4.5 Spatial analysis

Bayesian multi-variable modelling suggested there were no significant associations between *P. falciparum* or *P. vivax* prevalence by school and population density, elevation or land cover class, once residual spatial correlation was accounted for. Parameters associated with the magnitude and timing of seasonal changes in LST and EVI did show statistically significant association with *P. falciparum* and *P. vivax* infection; although these covariates did not improve model fit (Table 2.4). The variance of the school-level random effect (σ^2_{school}; which indicates a propensity for clustering) was large for both species. Similarly, for both species the distance at which spatial correlation dropped to below 5% was very large (in excess of 200 km) when compared with other spatial models of malaria infection [238]. This suggests a slow decline of spatial correlation with distance at larger scales, and is likely to be a consequence of large areas of zero prevalence across Oromia Regional State.

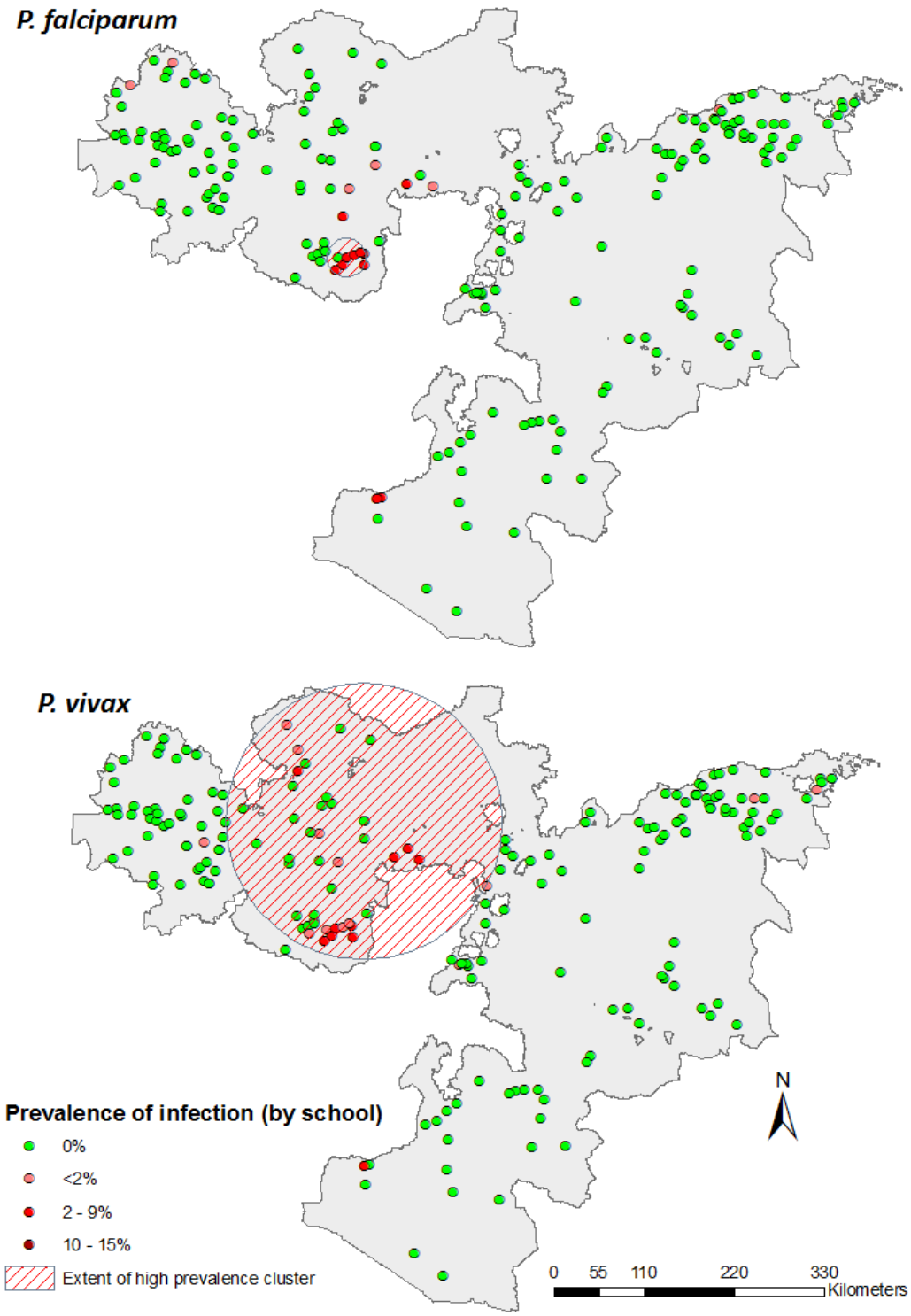
SaTScan analysis identified two clusters of high prevalence of *P. falciparum* and *P. vivax* infection (Figure 2.6). A small cluster was found for *P. falciparum* infection, with a radius of 23.3 km. The cluster contained 872 children from eight schools, with 40 *P. falciparum* infections found compared to an expected number of 2.6; the relative risk of infection in the cluster being 41.9 times higher than outside of the cluster (p=0.001). A significant cluster was seen for *P. vivax* infection, which with a radius of 169.0 km, was larger than the *P. falciparum* infection cluster. This cluster included 4,782 children from 44 schools, with 49 *P. vivax* infections found compared to an expected number of 12.5. The relative risk of infection in this cluster was 55.1 times higher than outside of the cluster (p=0.001). Parameters for seasonal changes in EVI and LST showed some association with presence inside the *P. falciparum* and *P. vivax* infection clusters, while proximity to water was associated with location inside the *P. vivax* cluster (OR=0.13, p=0.001).

Table 2.4 - Fitted parameters in Bayesian multivariate models for *P. falciparum* and *P. vivax* among school children in Oromia Regional State, Ethiopia in 2009, with and without spatial components.

	Non-spatial model		**Spatial model**	
	Parameter	**95% BCI**	**Parameter**	**95% BCI**
P. falciparum				
Without covariates				
σ^2_{school}	13.09	5.78-25.4	10.0	2.27-25.0
Range of spatial correlation (km)	-	-	271	132-794
DIC (model fit)	120.1	-	113.0	-
With covariates				
LST: bi-annual phase	OR: 2.11	1.17-4.19	OR: 2.03	0.96-4.00
σ^2_{school}	11.21	5.32-22.4	8.43	2.89-23.9
Range of spatial correlation (km)	-	-	244	119-758
DIC (model fit)	121.4	-	113.4	-
P. vivax				
Without covariates				
σ^2_{school}	6.31	2.96-13.17	7.03	2.30-19.95
Range of spatial correlation (km)	-	-	321	154-927
DIC (model fit)	162.3	-	145.5	-
With covariates				
EVI: bi-annual amplitude	OR: 1.93	0.91-4.25	OR: 2.06	0.74-4.92
LST: tri-annual phase	OR: 2.02	0.86-4.26	OR: 1.34	0.44-3.36
σ^2_{school}	6.09	2.98-11.7	8.10	2.32-29.9
Range (km)	-	-	355	162-1,589
DIC (model fit)	159.8	-	146.3	-

OR, odds ratio; 95% BCI, Bayesian credible interval; σ^2_{school} is the variance of RE; *range* is the estimated distance at which spatial correlation between sites <5%; DIC is deviance information criterion, a measure of model fit where lower values indicate better fit

Figure 2.6 - Spatial clusters of high *Plasmodium falciparum* (top) and *P. vivax* (bottom) infection, Oromia Regional State, Ethiopia, 2009.



2.5 DISCUSSION

This first application of school-based surveys to inform targeting of malaria interventions in Ethiopia has revealed a comparable prevalence of *Plasmodium* to that found in the 2007 national MIS survey [195], but lower than that reported in a 2006 survey conducted by the Carter Center (prevalence 4.1%) [196]. The study presented here highlights the marked spatial heterogeneity in infection observed in Oromia, with a small cluster of higher *P. falciparum* risk identified around Jimma administrative zone, and a larger area of elevated *P. vivax* risk across west-central Oromia. Furthermore, the results are consistent with other cross-sectional study findings [196], with malaria cases found above 2,000m, i.e. the current NMCP boundary for classification of an area as malarious and determining inclusion in control activities such as IRS and LLIN distribution [232]. While the prevalence of *P. falciparum* and *P. vivax* were found to be similar, the survey was not powered to define differences in endemicity or prevalence of the two major *Plasmodium* species in Oromia.

Reported recent fever was found to be a risk factor for both *P. falciparum* and *P. vivax* infection, but fever was not directly assessed in the surveys, and there is a risk of reporting bias in children's recall and likelihood of reporting a recent febrile episode. However, a high proportion of identified infections were asymptomatic. Cross-sectional surveys from a range of transmission settings have shown that a strong association exists between *P. falciparum* and reported fever [248]. In low transmission settings, it is often assumed that lack of acquired immunity in the population will cause all *Plasmodium* infections to elicit clinical symptoms, but the current findings dispute this. Since parasite density was not calculated in this study, it is not possible to determine if asymptomatic infections were due to very low parasite density. Asymptomatic infections will contribute to ongoing transmission in a community, but are unlikely to be detected or treated in a context where only individuals feeling unwell access diagnostic services. If Ethiopia is to achieve focal malaria elimination in

areas of current low, unstable transmission, alternative strategies may be required to identify and clear asymptomatic infections and halt transmission. In São Tomé and Príncipe, for example, mass screening by means of cross-sectional countrywide surveys, wherein all residents were screened with a RDT and RDT-positive individuals were treated with artemisinin combination therapy (ACT), has contributed to recent dramatic reductions in malaria transmission [249]. However, screening with RDT may not be sufficiently sensitive to identify all asymptomatic infections within a community.

Absence of age-dependency for infection in the current findings is consistent with lack of acquired immunity among individuals living in low malaria transmission settings [66], and findings from other surveys in Ethiopia [196]. Prevalence of anaemia was found to be higher in the present instance than in the 2005 national school health survey [250], but lower than in the 2005 Demographic Health Survey [251]. Increased odds of anaemia in males, however, was a common finding in the 2005 national school survey, and has been reported from other countries [252]. Males were also less likely to sleep under a LLIN, which may result in more frequent exposure to *Plasmodium* infection and resultant anaemia. Overall, these findings indicate that iron supplementation should be considered as a possible school health strategy, targeting boys and girls.

While documented scale-up in LLIN distribution and coverage in Ethiopia has been very successful [253], the present study shows that use of LLINs remains less than optimal among school-age children. Ownership of LLINs was not directly assessed in this study, but schoolchildren's reporting of household net ownership has been demonstrated to provide a good approximation of true household ownership in Uganda [110], therefore we do not expect that LLIN ownership in this survey was subject to reporting bias. This is consistent with other studies indicating that children of school-age are often the least likely to have access to mosquito nets owned by the household [254], as well as other data from Ethiopia indicating that net use does not directly correspond with net ownership [255]. While possession of

LLINs in a household will exhibit some indirect protective effect for individuals not sleeping under the net, Ethiopia's policy of universal coverage with LLINs in malaria risk areas [208] must be fully implemented in order to fully contribute to transmission control. There is also a need for additional behaviour-change activities linked to LLIN distribution campaigns and the routine health extension programme, to ensure consistent use of LLINs [256,257].

Somewhat surprisingly, LLIN use was associated with increased odds of malaria in crude univariate analysis. However, multivariate models did not find such association. The directionality of this crude association is likely a result of confounding from the strategy of priority LLIN distribution to highest risk areas of the Region, however this survey included areas perceived as high and low risk. Previous cross-sectional studies have found that net use is protective against malaria among school-aged children [258,259], and other surveys found that a protective effect against malaria was linked with the number of nets per household [196]. IRS, as reported by children to have been conducted in their house, was found to be associated with increased risk of both *P. falciparum* and *P. vivax* infection in multivariate models. The lack of protective effect of IRS in these findings may be a result of near-universal resistance to DDT (1,1,1-dichloro-2,2-bis(p-chlorophenyl)ethylene) in Ethiopian anopheline mosquitoes (Reithinger *et al.,* unpublished). The most likely explanation for this association between IRS and increased malaria risk is that the NMCP targets IRS to locations of known malaria endemicity; therefore living in a location where IRS is conducted is predictive of being in a malarious area.

The current surveys found a greater proportion of *Plasmodium* infections due to *P. vivax* than previously described in Ethiopia [195,196], with *P. falciparum* and *P. vivax* in equal proportion overall but *P. vivax* dominating in the highland epidemic ecological stratum (90%). The variation in species distribution may be a result of increased use of artemisinin-based combination therapy and *P. falciparum*-detecting RDTs at peripheral health facilities, impacting on transmission of *P. falciparum* and changing the epidemiology of this parasite in

Ethiopia. Alternatively, the findings may simply be due to the highly variable and unstable transmission setting, where increased *P. vivax* cases may be a result of focal epidemics in highland areas at the time of the survey, or a result of the tendency for *P. vivax* to cause long-term chronic infections and show less seasonality in transmission than *P. falciparum* [202]. The recent adoption of multi-species RDTs at health posts across Ethiopia will greatly improve the diagnosis and treatment of *P. vivax* infections. These infections are known to cause morbidity, including anaemia, malnutrition and respiratory distress [260,261], but are likely to have been under-diagnosed in the past due to use of *P. falciparum*-detecting RDTs. Challenges remaining in control of *P. vivax* include examination of drug-efficacy and potential adjustment of national policy, in light of identified foci of chloroquine resistance [223-226], as well as strategies for diagnosing and clearing asymptomatic *P. vivax* infections. Furthermore, similar to other settings, it is likely that as prevalence of *P. falciparum* in Ethiopia is reduced by effective control interventions, the burden of malaria attributable to *P. vivax* will increase [200,201].

It is the commonly held belief that in low transmission settings, a high proportion of children with malaria would be symptomatic and therefore absent from school. The present findings, however, indicate that although self-reported fever during the previous month is predictive of *Plasmodium* infection, only a minority of parasitaemic individuals identified in schools reported any fever on the day of the survey. Similar high proportions of asymptomatic *Plasmodium* infections have been found in other low transmission settings [262]. Attempts were made to ensure that all eligible children enrolled at each school were included in random sampling, but we expect a proportion of enrolled students were absent on the survey day. If there is a correlation between prevalence of infection determined from school surveys and the true prevalence of infection in the community, a correction factor could be applied to estimate prevalence of *Plasmodium* infection in the population using school surveys. Therefore the methodology presented in this study can still be applied to collect valid

epidemiological data from schools. Further investigation of the contribution of malaria to school absenteeism should be conducted to evaluate the population representativeness of parasite rates from school-based surveys.

The poor sensitivity of microscopy to detect low-density *Plasmodium* infections [263] may have affected the outcome of this study. The difficulties in correctly identifying low-density infection may have contributed to the discrepancy in microscopy results between standard and expert examination of blood films. Furthermore, these discrepancies indicate a need to implement a rigorous quality assurance system within the routine laboratory diagnostics system for malaria in Ethiopia, or alternatively, to expand the use of RDTs beyond community-level health care. However since microscopy reading of slides collected during cross-sectional surveys is usually conducted by a central team of microscopists after survey completion, and this microscopy reading is subject to quality assurance, the low sensitivity of microscopy data from this survey is expected to be spatially homogenous. The sensitivity of routine microscopy data from health facilities may not be so spatially homogenous, introducing bias to any large-scale evaluation of routine diagnostic data from health facilities Molecular techniques, such as polymerase chain reaction (PCR), have a lower detection threshold for *Plasmodium* than microscopy [10,262], and may be a more sensitive diagnostic tool in a population where low-density infections are expected. Unfortunately, PCR remains suitable only in a research context, and not as a routine diagnostic tool for malaria.

The current study was unable to create a valid model, based on environmental covariates, to predict malaria endemicity across Oromia, because strong environmental predictors for location of transmission foci were lacking. Risk mapping using similar strategies had previously been successful in Afghanistan, with a comparable prevalence of infection (0.49%) [264]. This inability to develop a risk map based on environmental correlates only indicates that there are additional factors contributing to transmission that were not captured in modelled data, and alternative covariates may have been required to test for inclusion in the

model, for example the use of climatic data relating more closely to the survey period, rather than summarising long-term trends. Modelling may also have been improved by using the average environmental covariate value for the whole school catchment area, rather than point values for the school location. The failure to develop a risk map may also be a result of the spatial and temporal variability of transmission, which was not adequately captured by a cross-sectional survey approach with microscopy diagnosis of the primary indicator. Use of a more sensitive indicator of current *Plasmodium* infection, such as polymerase chain reaction, may have improved model fit. Although it was not possible to determine the exact altitude at which infection was acquired, data indicate that most children live close to the school (median 30 minutes' walk). Therefore, it was assumed that children's homes, the site where infection is likely to have occurred, is at a similar altitude to the school. Based on this assumption it was possible to identify two clusters of infection in Oromia, using a method that has successfully described hotspots of malaria at small spatial scale in Kenya and Sudan [265,266].

Identification of all areas where malaria transmission is ongoing may be possible using an alternative diagnostic method where IgG antibodies to *Plasmodium* are detected using an enzyme-linked immunosorbent assay, reflecting exposure to infection over a longer time period. This method has been used successfully in other low and unstable transmission settings [187,191,263]. Alternatively, routinely reported malaria case data from health facilities has been used to model malaria transmission [53], but these data are subject to bias including incomplete recording and reporting, inconsistent quality of diagnostic services and variable access to health facilities across populations and localities. It may be possible to marry parasitological survey data and routine facility data to capture a reliable estimation of malaria transmission levels, and use these combined data to develop a risk map. This approach requires further investigation to ensure comparability between locations, and representativeness of the underlying population. Other information such as locations of food

insecurity can contribute to interpretation of anaemia data, and could be incorporated into mapping activities to inform programmatic decision making.

While the current cross-sectional surveys have provided data regarding the *Plasmodium* parasite rates among children attending school, there is a need to conduct a rigorous comparison to indicators determined from standard community surveys, such as MIS. This will determine if findings from school-based surveys are representative of all school-aged children in a community or, indeed, the whole community; if representative, school-based surveys could become an alternative survey method to the more costly and labour-intensive community surveys. While it is not expected that there are differences in risk of infection by age in Ethiopia, there is a need to further explore what proportion of school-absenteeism is due to malaria, as well as whether there are differences in malaria risk between enrolled and non-enrolled children. These findings will define the potential role of schools in malaria surveillance, monitoring and control in Ethiopia and other low transmission settings. Envisaged roles of schools in malaria surveillance could be to provide data on coverage of major interventions and parasite prevalence during routine school surveys, and to alert service providers of epidemics using information on school-absenteeism and from active case finding [68].

2.6 CONCLUSIONS

Results of cross-sectional school surveys in Oromia demonstrated marked spatial heterogeneity in malaria. Although several foci of infection were identified, large areas appear to be non-endemic for malaria. While these findings allow malaria control interventions to be targeted to identified endemic areas, this likely does not reflect the true extent of malaria in Oromia. Research is ongoing to further validate the use of school surveys in identifying transmission foci, as well as to investigate other potential uses of schools in

malaria surveillance including monitoring and evaluation of control programme implementation.

CHAPTER 3. GEOSTATISTICAL MODELLING OF MALARIA ENDEMICITY USING SEROLOGICAL INDICATORS OF EXPOSURE COLLECTED THROUGH SCHOOL SURVEYS

3.1 OVERVIEW

Findings from school-based cross-sectional surveys were presented in Chapter 2, with parasitaemia diagnosed by microscopy as the primary outcome. Very few *Plasmodium* infections were identified by microscopy, limiting the ability to explore spatial associations and develop predictive models. During the surveys presented in Chapter 2, blood samples were collected on filter paper and stored. These dried blood spots were analysed for presence of anti-*Plasmodium* antibodies to describe the seroprevalence at each site as a proxy for recent transmission intensity. Spatially explicit binomial models of seroprevalence were created and used to predict seroprevalence across the Region.

This chapter has been prepared for *The Journal of Infectious Diseases*, but at the date of thesis submission has not been submitted to the journal or subject to peer review. I participated in serological analysis of samples, was responsible for all processing of raw laboratory data, including decisions on samples requiring exclusion or repeat analysis. Jorge Cano extracted and processed environmental data used in this chapter. I conducted all data analysis presented.

3.2 INTRODUCTION

With the rekindling of malaria elimination goals [267,268], there is an increased need to quantify patterns of and changes in malaria risk to support evidence-based targeting of

interventions, implement surveillance strategies to monitor changes in transmission intensity, and assess feasibility of local elimination [7,269].

Low transmission settings present specific challenges to implementation of cross-sectional surveys: (i) highly seasonal transmission can result in underestimates of population parasite rates if sampling does not occur during the peak transmission period; (ii) low-density infections are frequent, and the common diagnostic tools of microscopy and rapid diagnostic tests (RDTs) demonstrate reduced sensitivity to detect these low-density infections [71,270]; and (iii) where diagnostic data are used to develop spatial prediction models, there is a risk that the true extent of transmission will be underestimated since recently cleared and low-density infections will not be included. While new strategies such as reactive case detection [101,102,271] and "rolling" cross-sectional surveys [107] have been trialled, there remains a need to develop strategies to track changes in low and unstable transmission settings.

Serum antibodies are stable when collected on filter paper, desiccated and stored below 4°C, and a methodology to detect anti-*Plasmodium* antibodies eluted from dried blood spots has been published [184,185], yielding estimates of seroprevalence and seroconversion rates that are representative of malaria transmission intensity within a community [186,272]. Since antibodies persist after infection clearance, they offer the opportunity to examine exposure to malaria over a wider time period than is typically possible through detection of parasitaemia during a cross-sectional surveys survey by means of microscopy, RDTs or polymerase chain reaction (PCR)-based methods.

Serological indicators are increasingly being used in community-based malaria epidemiological studies to assess small- and large-scale spatial heterogeneities of and changes in transmission [190,191,193,194,263,272]. Schools provide a useful alternative platform for collection and monitoring of malariometric indicators, offering logistical advantages (e.g. simplified selection of participants, high compliance and reduced survey

costs) over standard community-based cross-sectional surveys [16,68,273]. Although schools consistently yield higher estimates, school survey seroprevalence estimates have repeatedly been shown to strongly correlated with community-survey seroprevalence [111].

This study explored the use of serological indicators collected from a large-scale school-based survey to describe differences in *P. falciparum* and *P. vivax* endemicity in a low transmission setting. Spatially explicit Bayesian modelling techniques were used to explore relationships between serological indicators at population level and explanatory environmental variables, to predict estimated endemicity levels at sub-national scale.

3.3 METHODS

3.3.1 Study setting

Ethiopia has a diverse ecology and malaria transmission is known to be spatially heterogeneous, related to variables such as altitude, temperature, rainfall and presence of local water bodies or dams [130,196,202,209,210]. Malaria transmission is temporally variable due to seasonal rainfall, with a major transmission season from September to December and a minor transmission season from April to May. Cases are due to both *P. falciparum* and *P. vivax*. The Malaria Indicator Survey in 2011 demonstrated a low parasite prevalence within the population living in malaria-risk areas, estimated at 1.3% by microscopy and 4.5% by RDT in areas <2000m [228].

3.3.2 Survey data

Data presented in this paper are drawn from a large cross-sectional survey conducted in 197 government primary schools in Oromia Regional State, Ethiopia, in 2009, described in Chapter 2. Full details of school and child selection as well as sample collection are presented in

Chapter 2. Briefly, at each school 55 girls and 55 boys were randomly selected. They provided finger-prick blood samples for preparation of thick and thin blood film, haemoglobin measurement (HemoCue Ltd, Angelhölm, Sweden), and collection of blood spots on filter paper (Whatman 3MM, Whatman, Maidstone, UK). School location was measured using a global positioning satellite receiver (eTREX, Garmin International, KS, USA).

For serological analysis, samples were selected purposively from: (i) 20 schools with highest prevalence of *Plasmodium* infection detected by microscopy (range 0.9 to 14.5%); (ii) 20 schools with highest proportion of anaemic (classified according to WHO [243], including adjustment by altitude) children (range 34.2 to 51.4%); and (iii) and a random selection of remaining schools surveyed (Table 3.1). Purposive selection was conducted to capture a range of transmission settings, and since resources were not available to complete ELISA on all blood spots collected during surveys.

Table 3.1 - Number of schools and children tested by enzyme-linked immunoassay (ELISA) against each antigen, stratified by school selection criteria: high microscopy prevalence, high anaemia prevalence, randomly selected.

	Any *P. falciparum* antigen		**Any *P. vivax* antigen**	
	Schools	**Children**	**Schools**	**Children**
Total tested	62	5913	71	6609

	***Pf*MSP-1**		***Pf*GLURP**		***Pv*MSP-1**		***Pv*AMA**	
School selection criteria	**Schools**	**Children**	**Schools**	**Children**	**Schools**	**Children**	**Schools**	**Children**
High microscopy prevalence	20	2088	20	2093	20	2080	20	2074
High anaemia prevalence	20	2118	20	2092	20	2080	20	2104
Random selection	22	1614	10	1037	31	2327	10	1024
Total tested	62	5820	50	5222	71	6487	50	5202

3.3.3 Enzyme-linked immunosorbent assay (ELISA)

Blood spots from 50 schools were analysed in London against *P. falciparum* merozoite surface protein-1_{19} (*Pf*MSP-1), *P. falciparum* glutamate-rich protein-R2 (*Pf*GLURP), *P. vivax* merozoite surface protein 1_{19} (*Pv*MSP-1) and *P. vivax* apical membrane antigen-1 (*Pv*AMA). In Addis Ababa, blood spots from a further 12 schools were analysed against *Pf*MSP-1 and 21 schools against *Pv*MSP-1.

Antibodies were eluted from dried blood spots, and samples tested for IgG against *P. falciparum* and *P. vivax* antigens according to methods described elsewhere [184]. Duplicate optical density (OD) values with >20% variation were excluded. Raw ODs were corrected by blank OD and normalised between plates by fitting to the mid-point of a standard curve produced by serial dilution of hyperimmune serum (i.e. pooled hyperimmune serum from Tanzania for *P. falciparum* and NIBSC 72/096 for *P. vivax*). Normalised ODs and identification

numbers were exported into Stata 12.0 (Stata Corporation, College Station, Texas USA). Individual samples were classified as seropositive or seronegative against each antigen using a mixture model, whereby the mean of the seronegative distribution plus three standard deviations was defined as the seropositive cut-off [184,272]. Binary variables were generated to describe summary seropositivity by species: for example, *P. falciparum* seropositive samples were defined as seropositive against either *Pf*MSP-1 and/or *Pf*GLURP. In the absence of a gold standard for anti-*Plasmodium* antibody detection, it is not possible to determine the sensitivity or specificity of the ELISA, but the mixture model approach is commonly used in low transmission settings [187,191,274,275], where the population is expected to include true seronegatives and true seropositive individuals.

3.3.4 Remote sensing environmental data

Elevation data was extracted from the shuttle radar topography mission (SRTM) digital elevation model at 90m resolution [276], resampled to 250m and further processed to estimate slope in degrees. Precipitation and temperature data at 1km resolution were extracted from pre-processed data available on WorldClim [277,278]. Euclidean distance to water bodies was calculated using SRTM Water Bodies data files at 250m resolution [279], and distance to rivers and roads calculated using data from Digital Chart of the World at 250m resolution [280]. Land cover type was extracted from the qualitative global land cover map, defined within the UN land cover classification system using environmental satellite (ENVISAT) mission's Medium Resolution Imaging Spectrometer (MERIS) sensor at 300m resolution. Normalised difference vegetation index (NDVI) indicators at 1km resolution were extracted from the SPOT 5 vegetation project [281]. Population density was extracted from the AfriPop project at 100m resolution [282], and rural-urban classification at 1km from the Global Rural-Urban Mapping project (GRUMP) [283]. Input grids were either extended or clipped to match the geographic extent of a land mask template, and eventually aligned to it.

Finally, environmental data were extracted to school locations using ArcMap 12.0 (Environmental Systems Research Institute Inc., Redlands CA, USA).

3.3.5 Model development and testing

Environmental and serology data were merged and analysed using Stata 12.0. Continuous environmental variables were standardised to facilitate later model convergence. Models were developed separately to describe *P. falciparum* or *P. vivax* seroprevalence.

Univariate associations between school seroprevalence and environmental variables were explored, and colinearity (correlation coefficient >0.9) between variables tested. A school-level minimal adequate logistic regression model was developed by the backward stepwise method, whereby variables with $p>0.05$ were removed in the order of least significance; all excluded variables were subsequently re-tested in the final model. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are indicators of model fit and parsimony, but BIC has more severe penalisation for model complexity. Both AIC and BIC were used to inform model selection [284,285].

Four multivariate Bayesian binomial regression models were developed using WinBUGS 14 (Medical Research Council Biostatistics Unit and Imperial College London, London, U.K.) for *P. falciparum* and for *P. vivax.* The most complex model included the retained school-level environmental variables, school-level random effect and school-level geostatistical random effect (using an isotropic, stationary exponential decay function) [245]. Additional models excluded the environmental variables, the spatial random effect, or both.

Semi-informative priors were set for the rate of decay of spatial correlation, Φ, informed by the maximum and minimum distance between schools, and non-informative priors used for other coefficients. Models were burned-in for 10,000 iterations to achieve convergence, then nodes sampled for 10,000 iterations, thinning each ten iterations. Final model selection was

informed by examining the variance of school and spatial random effects and Deviance Information Criterion (DIC) [286].

3.3.6 Model validation

Models were internally validated by training the model on an N-5 school dataset, then predicting probability that seroprevalence thresholds (2%, 5% and 40%) are exceeded for the five excluded schools. These thresholds were estimated to describe the lowest and highest areas of seroprevalence, and hence endemicity, in order to support specific intervention targeting. The process was repeated until predictions for all schools were available. Model performance was assessed by examining the area under the curve (AUC) of the receiver operator characteristic (ROC) at each threshold [287]. AUC >0.7 indicates a reasonable discriminative capacity, and AUC >0.9 very good discriminative capacity [288,289].

3.3.7 Generating a predictive seroprevalence map

A grid of 12,048 locations at 5km spacing was generated across Oromia and environmental variables included in final models extracted to these locations. The selected Bayesian models were trained on actual school seroprevalence data, then predicted at each location by calculating the sum of the products of the covariate coefficients and the values of the covariates at each grid node, plus the interpolated geostatistical random effect, and back transforming from the logit to the prevalence scale.

3.3.8 Ethical considerations

The school surveys received ethical clearance from the Ethiopian Science and Technology Agency (RDHE/2-89/2009), with additional clearance subsequently given for serological analysis of blood spots (3.10/53/2003). Consent for participation used a passive, opt-out procedure, with school director providing written consent for the survey to proceed. Schools were requested to hold meetings in advance with parents to inform them of the survey and

allow withdrawal of children if necessary. Participating children gave written assent and were informed of their right to withdraw at any time. Children reporting fever during surveys were tested with a multi-species HRP2-panLDH RDT (CareStart, AccessBio, NJ), and any child with a positive RDT was treated according to the national guidelines [213].

3.4 RESULTS

3.4.1 Serology findings

Serology results were available for *P. falciparum* from 5,914 children from 62 schools, with a mean 95 (range 10-111) samples per school. *P. vivax* results were available from 6,609 children from 71 schools, with mean 93 (range 5-111) samples per school. Data were from children aged five to 18 years (mean 11 years).

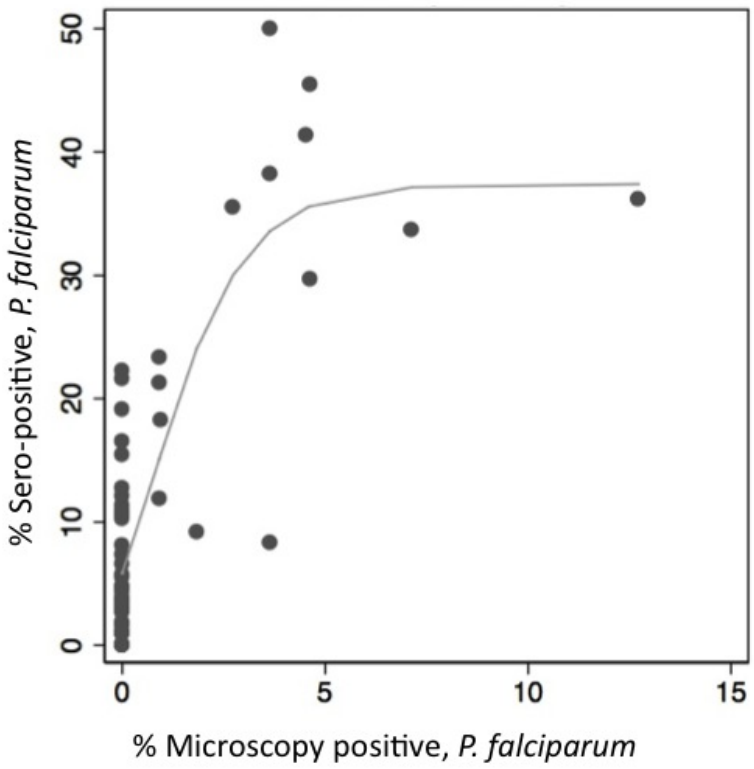
Of all children tested, 11.6% (688/5913) were *P. falciparum* seropositive and 11.1% (735/6609) *P. vivax* seropositive; 1.0% and 0.5% of the children were microscopy-positive for *P. falciparum* and *P. vivax* parasites, respectively. Cross-tabulation of microscopy and antigen-specific serology results are presented in Table 3.2. Where data were available for both species, 4.7% of 5,420 children were seropositive against both species. When restricting our analyses to schools with more than 50 children tested (56 schools for *P. falciparum*, 62 for *P. vivax*), *P. falciparum* and *P. vivax* school seroprevalence ranged from 0 to 50% and 0 to 53.7%, respectively.

Among 50 schools tested against four antigens, correlation was seen between school seroprevalence determined for *Pf*MSP-1 and *Pf*GLURP (R^2=0.84), and for *Pv*MSP-1 and *Pv*AMA (R^2=0.80). For both species, coating plates with MSP-1 resulted in higher sensitivity than *Pf*GLURP or *Pv*AMA. A strong correlation (R^2=0.84) was seen between school *P. falciparum* and *P. vivax* seroprevalence.

Table 3.2 - Description of frequency of diagnostic test (microscopy and serology) results at individual level. Combinations of microscopy and seropositivity by antigen are presented for *P. falciparum* and *P. vivax* separately. Data are only presented for individuals with results recorded for *P. falciparum* microscopy, *Pf*GLURP and *Pf*MSP-1 (N=5102), and individuals with complete results for *P. vivax* microscopy, *Pv*AMA and *Pv*MSP-1 (N=5053). Individuals tested against only one antigen for a species are excluded from the *P. falciparum* (N=783) or *P. vivax* (N=1522) sections of the table.

***P. falciparum* diagnostic tool combinations**			
		*Pf*GLURP +	*Pf*GLURP -
Microscopy Pf +	*Pf*MSP-1 +	38	5
Microscopy Pf +	*Pf*MSP-1 -	3	6
Microscopy Pf -	*Pf*MSP-1 +	217	246
Microscopy Pf -	*Pf*MSP-1 -	106	4481
***P. vivax* diagnostic tool combinations**			
		*Pv*AMA +	*Pv*AMA -
Microscopy Pv +	*Pv*MSP-1 +	4	10
Microscopy Pv +	*Pv*MSP-1 -	2	14
Microscopy Pv -	*Pv*MSP-1 +	141	381
Microscopy Pv -	*Pv*MSP-1 -	78	4423

Figure 3.1 - School-level seroprevalence and prevalence of infection by microscopy for *P. falciparum* (A) and *P. vivax* (B). Scatter plots are presented for 56 schools with *P. falciparum* data, and 62 schools with *P. vivax* data, restricted to those with serology results from ≥50 children. Non-linear regression identified a Gompertz function as best fit to *P. falciparum* (R^2=0.810), and to *P. vivax* data (R^2=0.657)



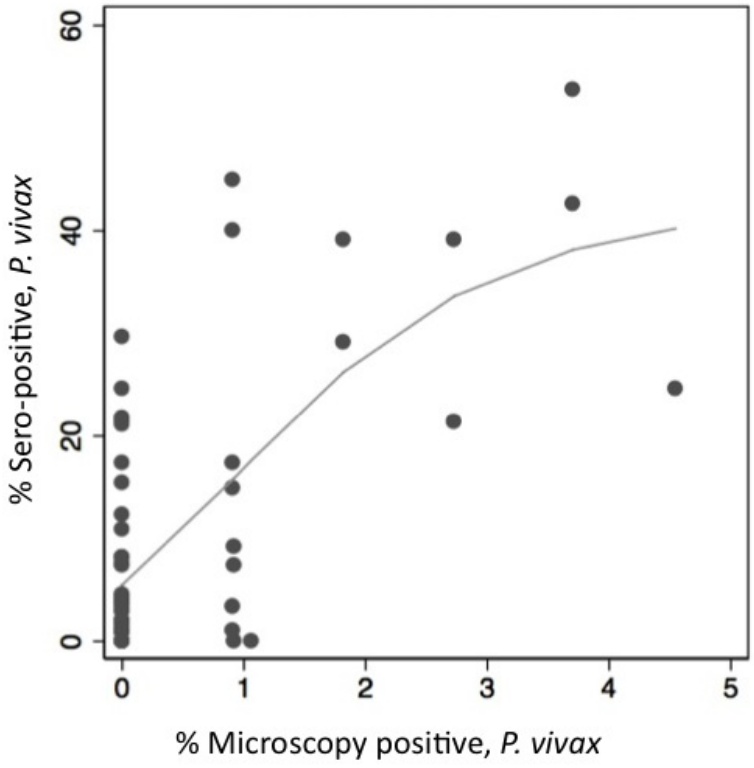


Table 3.3 - Univariate frequentist associations of key environmental variables with school seroprevalence of *P. falciparum* and *P. vivax*

	P. falciparum			***P. vivax***		
	OR	**95% CI**	**p**	**OR**	**95% CI**	**p**
Precipitation:						
Annual accumulative	0.978	0.695, 1.374	0.869	1.409	1.047, 1.898	0.024
Annual mean	0.971	0.693, 1.361	0.865	1.400	1.043, 1.880	0.025
Wettest quarter	0.991	0.723, 1.359	0.956	1.479	1.061, 2.062	0.021
Mean at peak transmission	0.781	0.588, 1.039	0.090	0.923	0.711, 1.199	0.549
Annual standard deviation	0.902	0.676, 1.203	0.483	1.336	0.928, 1.924	0.119
Land temperature:						
Annual mean	1.272	0.915, 1.767	0.152	0.866	0.657, 1.140	0.305
Mean at peak transmission	1.243	0.892, 1.734	0.199	0.890	0.659, 1.201	0.445
Altitude	0.785	0.590, 1.044	0.096	1.041	0.784, 1.381	0.782
Land gradient	0.565	0.402, 0.796	0.001	0.531	0.336, 0.838	0.007
Distance to:						
Any waterbody	0.808	0.644, 1.014	0.065	0.585	0.422, 0.811	0.001
Permanent waterbody	0.768	0.609, 0.969	0.026	0.570	0.456, 0.712	<0.001
Permanent river	0.402	0.172, 0.940	0.036	0.154	0.047, 0.506	0.002
Road	0.983	0.191, 5.057	0.984	1.059	0.206, 5.460	0.945
Land cover type:						
Shrubland	1	-	-	1	-	-
Cultivated land	5.320	2.419, 11.70	<0.001	6.782	2.907, 15.82	<0.001
Forest	9.5e-7	2.9e-7, 3.1e-6	<0.001	0.185	0.066, 0.522	0.001
Bare/sparse	4.338	1.979, 9.507	<0.001	2.236	0.848, 5.900	0.104
Normalised difference vegetation index:						
Maximum 2005-2009	0.810	0.658, 0.997	0.047	1.074	0.736, 1.569	0.711
Mean 2005-2009	0.785	0.628, 0.982	0.034	0.903	0.699, 1.166	0.434
Standard deviation 2005-2009	0.824	0.647, 1.049	0.116	1.126	0.770, 1.647	0.541
Maximum 2009	0.846	0.625, 1.146	0.280	1.269	0.831, 1.937	0.271
Mean 2009	0.739	0.571, 0.955	0.021	0.868	0.680, 1.108	0.255
Standard deviation 2009	0.928	0.695, 1.238	0.611	1.398	0.949, 2.060	0.090
Population density:						
All ages	0.994	0.720, 1.372	0.970	0.958	0.690, 1.329	0.796
Children under five years	0.981	0.696, 1.381	0.912	0.966	0.690, 1.353	0.840
Type of area:						
Rural	1	-	-	1	-	-
Peri-urban	0.468	0.182, 1.202	0.115	0.361	0.123, 1.066	0.065
Urban	0.200	0.026, 0.880	0.033	0.065	0.015, 0.283	<0.001

3.4.2 Comparing serology to microscopy

Schools with 0% positive samples by microscopy were found to have from 0 to 30% seroprevalence. While the proportion of microscopy positive and seropositive children in a school are not directly comparable, it is plausible to expect some association between the two measures. The relationship between serology and microscopy data was postulated to be a saturation growth curve, but it was not possible to fit this curve to the existing data, likely due to the small number of site with microscopy prevalence >0%. A variety of nonlinear regression functions were fitted using least squares methods to *P. falciparum* and *P. vivax* data. For both species, a Gompertz curve was found to have best fit to the data, and is presented in Figure 3.1.

3.4.3 Environmental risk factors

Colinearity was found among the precipitation, temperature and NDVI variables, and between distance to both permanent and any type of water body. Distance to water bodies and rivers, land gradient and urban areas showed univariate associations with both *P. falciparum* and *P. vivax* seroprevalence (Table 3.3).

The minimal adequate multivariate *P. falciparum* frequentist model includes elevation and angle of land slope, distance to permanent river, bare or sparse land cover, population density and urban areas. The minimal adequate *P. vivax* frequentist model includes distance to permanent river and water body, precipitation during the wettest quarter of the year (projection from 1950-2000 to allow for high spatial resolution), and mean NDVI over the preceding five years.

3.4.4 Bayesian modelling of *P. falciparum*

When comparing output from non-spatial and spatial models of *P. falciparum*, incorporating spatial structure in models was found to explain much of the variation between schools,

indicated by a reduction of σ^2_{school} when spatial random effects were included in models. A lower DIC in models including spatial structure justified retention in the final *P. falciparum* Bayesian model.

Inclusion of environmental variables in the spatial model was shown to reduce $\sigma^2_{spatial}$ and the DIC, as well as increase the rate of decay of spatial correlation (ϕ), indicating that much of the first order spatial variation can be explained adequately by the included environmental data. Therefore, the final model to describe *P. falciparum* seroprevalence in Oromia incorporates environmental covariates to explain first order deterministic spatial variation, with the spatial random effect adequately capturing second order structure (Table 3.4). Internal validation demonstrates a good discriminatory ability of the final model for 2% and 5% seroprevalence thresholds, with an AUC of 0.83 and 0.84, respectively. The model performs very well in identifying areas of over 40% seroprevalence (AUC=0.96). Actual and predicted school seroprevalence were found to be correlated (Pearson r=0.62, p<0.001). The final model was used to predict *P. falciparum* seroprevalence at 5km resolution across Oromia. The posterior mean prediction for this model is shown in Figure 3.2, with estimated probability of 2%, 5% and 40% thresholds being exceeded shown in Figure 3.3.

3.4.5 Bayesian modelling of *P. vivax*

Similar to the *P. falciparum* models, incorporating spatial structure in *P. vivax* models was found to explain much of the variation between schools, indicated by a reduction of σ^2_{school} and lower DIC. Inclusion of environmental variables in the spatial *P. vivax* model did not substantially reduce the $\sigma^2_{spatial}$ and little difference was seen in ϕ and DIC between the models with and without environmental variables. The final model for *P. vivax* is, therefore, the spatial model with no environmental covariates (Table 3.4). Internal validation of this model indicates good performance at the 2% seroprevalence threshold (AUC=0.81), and very good performance at 5% and 40% thresholds (AUC=0.91 for both). Actual and predicted

seroprevalence were correlated (Pearson *r*=0.68, p<0.001). Predictions of the final model at 5km resolution are displayed in Figure 3.4 as the posterior mean seroprevalence, and in Figure 3.5 as the probability of seroprevalence thresholds being exceeded.

Table 3.4 - Final Bayesian *P. falciparum* model developed using data from 62 schools, and *P. vivax* model developed from 71 schools' data. Both models retained school-level and spatial random effects. School and spatial variance (σ^2_{school} and $\sigma^2_{spatial}$), rate of decay of spatial correlation (ϕ), range in km at which correlation between schools falls to 5% are presented with 95% Bayesian credible intervals. The *P. falciparum* model includes parameter values and 95% BCI for standardised environmental fixed effects. No environmental fixed effects were retained in the final *P. vivax* model.

	***P. falciparum* model Parameter value (95% BCI)**	***P. vivax* model Parameter value (95% BCI)**
Altitude	-0.568 (-1.035, -0.087)	-
Slope	-0.595 (-0.996, -0.234)	-
Distance to permanent river	-0.411 (-0.774, -0.036)	-
Population density in 2010	0.418 (-0.107, 0.911)	-
Bare or sparse land (binary)	1.026 (-0.392, 2.298)	-
Urban area (binary)	-3.13 (-6.279, -0.028)	-
σ^2_{school}	0.254 (0.006, 1.250)	0.288 (0.013, 0.939)
$\sigma^2_{spatial}$	1.183 (0.177, 2.311	3.631 (1.31, 10.85)
ϕ	9.763 (2.631, 19.03)	0.866 (0.211, 2.093)
Range in km	45.57 (17.54, 127)	548.3 (160.5, 1592)
DIC	308.8	330.8

Figure 3.2 - Map of predictive *P. falciparum* seropositivity using spatial model with environmental fixed effects. Measured *P. falciparum* seroprevalence from the 62 schools used to train the model are shown by circles with size proportional to seroprevalence.

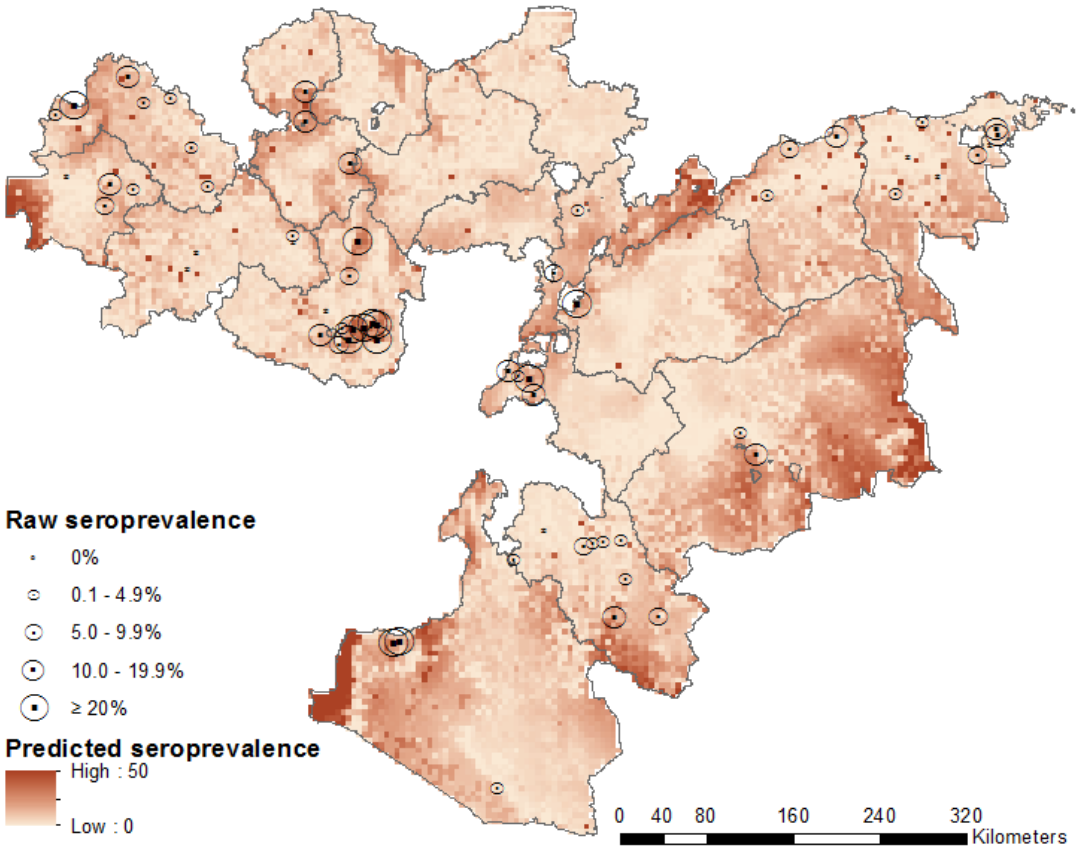


Figure 3.3 - Probability *P. falciparum* seroprevalence exceeds defined thresholds of 2% (A), 5% (B) and 40% (C) according final predictive model for *P. falciparum*. Red areas are those very likely to exceed the threshold, blue areas very unlikely to exceed the threshold, and pale yellow areas have high uncertainty

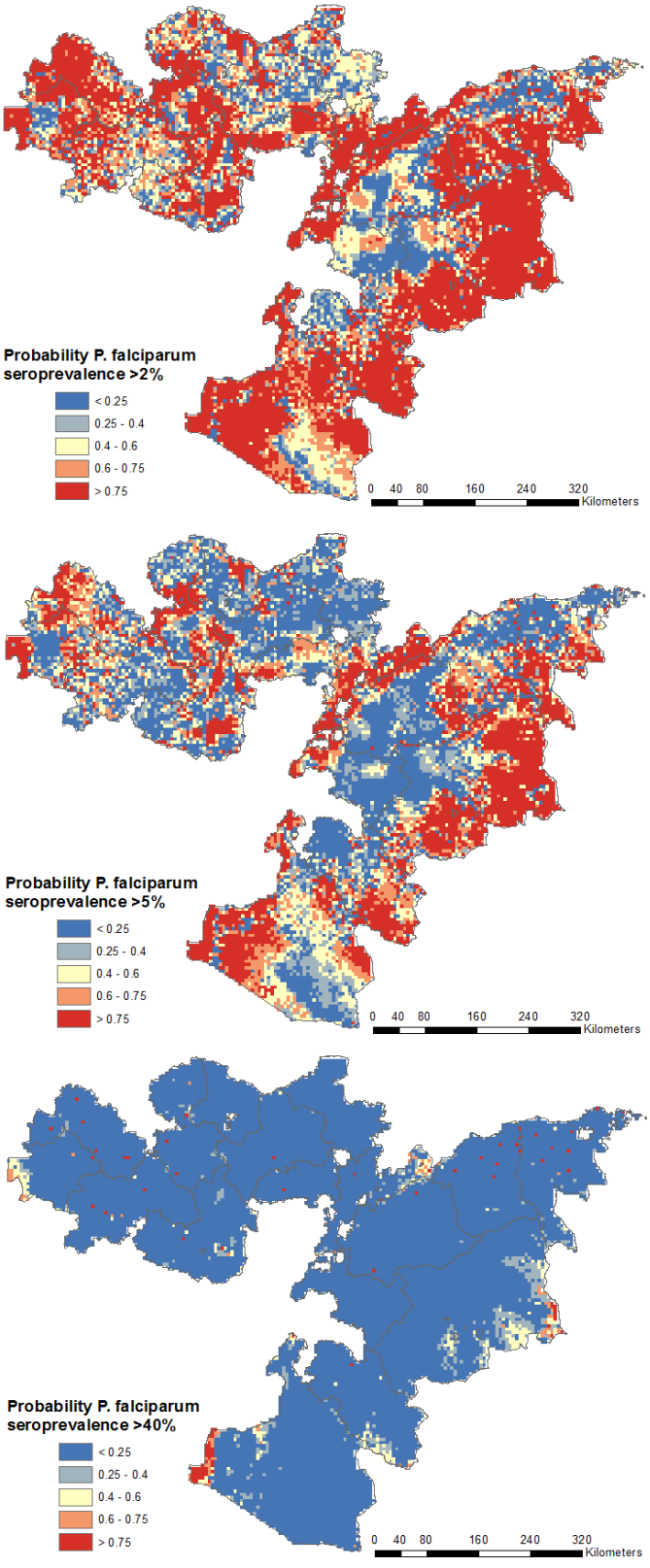


Figure 3.4 - Map of predictive *P. vivax* seropositivity, using spatial model without environmental fixed effects. Measured *P. vivax* seroprevalence from the 71 schools used to train the model are shown by circles with size proportional to seroprevalence

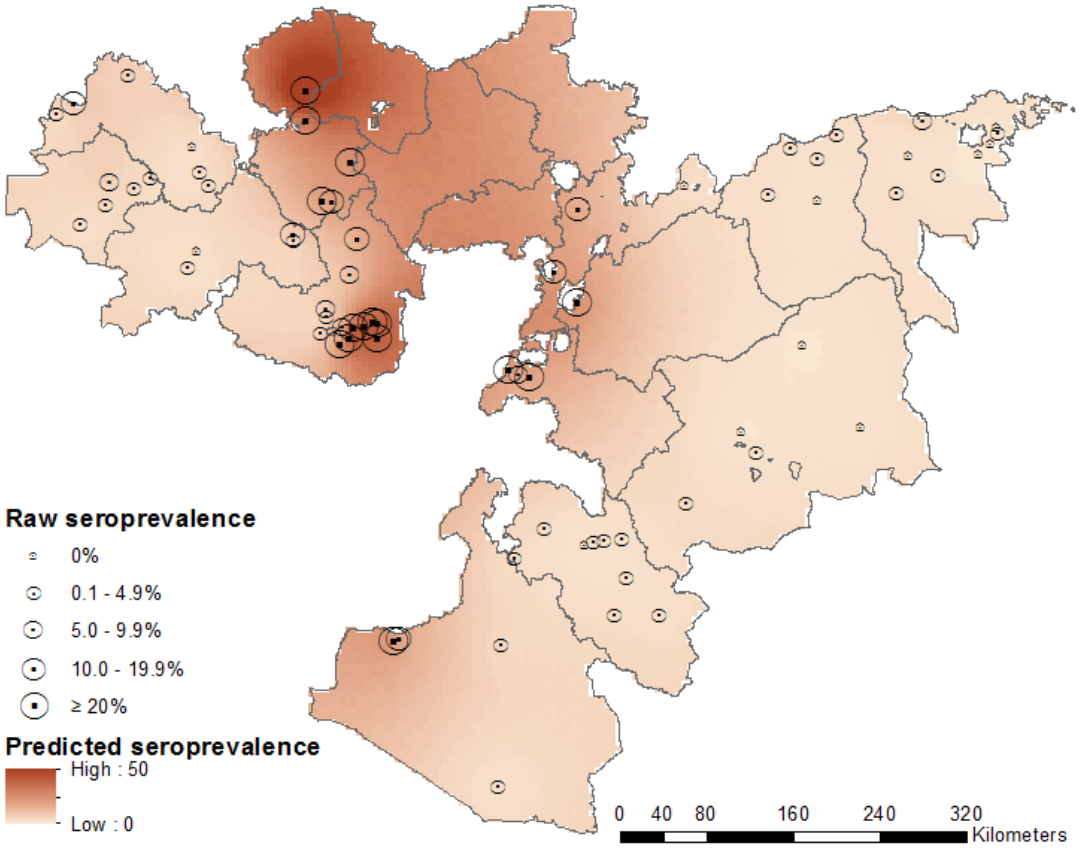
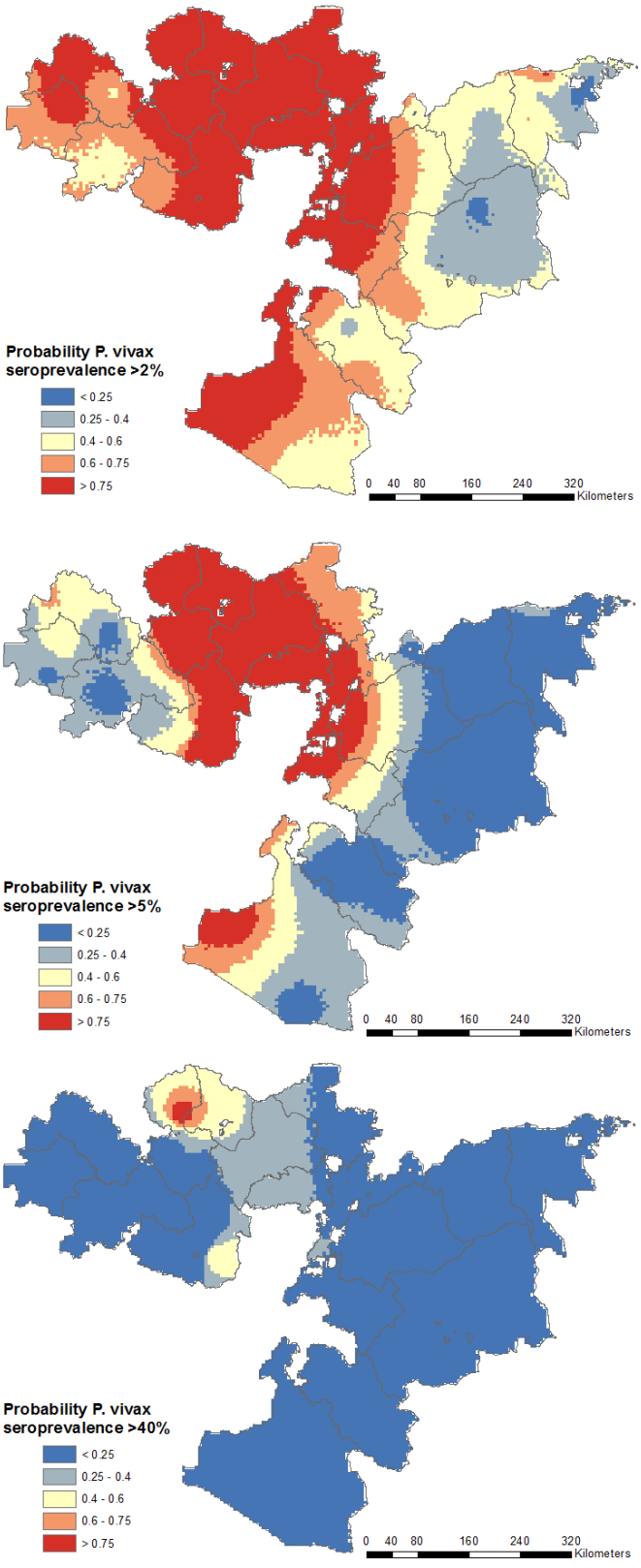


Figure 3.5 - Probability *P. vivax* seroprevalence exceeds defined thresholds of 2% (A), 5% (B) and 40% (C) according final predictive model for *P. falciparum*. Red areas are those very likely to exceed the threshold, blue areas very unlikely to exceed the threshold, and pale yellow areas have high uncertainty



3.5 DISCUSSION

This study demonstrates the capability of serological markers to detect large-scale heterogeneity in malaria transmission using samples collected during cross-sectional school surveys, in a setting with seasonal and low transmission. Seroprevalence was found to be associated with environmental variables; this relationship was used to predict seroprevalence at unsampled locations using Bayesian geostatistical modelling methods incorporating fixed and random effects.

School seroprevalence determined by different antigens showed strong correlation for each species, with MSP showing higher sensitivity for both *P. falciparum* and *P. vivax* than *P. falciparum* GLURP and *P. vivax* AMA, respectively. However, previous studies have shown AMA-1 to have higher immunogenicity than MSP-1 [185]. As transmission declines, individual antibody responses become more disparate, and therefore it is recommended that future serological analysis be conducted using multiple antigens or a whole parasite lysate.

The range of seroprevalence found across schools where no children were microscopy-positive during the cross-sectional survey demonstrates the value of serological indicators, i.e. in differentiating schools where transmission occurs but the peak transmission period was missed by surveys, from those with very low malaria risk. Differences may also be apparent if transmission has ceased in the area in recent years, before the age of the youngest school child. This is difficult to demonstrate without clinical data.

Both species' final models incorporated a spatial random effect to describe spatial autocorrelation, whereby schools located closely together are more similar than schools at greater distance. All models included a non-spatial school-level random effect. The *P. falciparum* model indicated that spatial autocorrelation was present to a distance of approximately 46km, while the *P. vivax* model showed a range of over 500km. The *P. falciparum* range is a distance at which similarities in climatic factors and ecology would be

expected, and therefore it is feasible that these areas experience similar transmission intensity. However, the very large spatial range of *P. vivax* suggests that the spatial random effect is capturing other large-scale variations not tested for inclusion in the Bayesian spatial model. A similar finding was reported in spatial modelling of malaria in Bangladesh, where environmental variables described a large proportion of spatial variation in *P. falciparum*, but little of the *P. vivax* distribution [290].

Frequentist models developed for *P. vivax* suggested biologically plausible environmental risk factors of distance to rivers and water bodies, vegetation cover and precipitation; nonetheless, these did not adequately explain the large-scale trends in *P. vivax* seropositivity after accounting for spatial dependency. The final *P. vivax* map presented here therefore simply uses spatial interpolation to predict seroprevalence at unsampled locations. The larger spatial scale of *P. vivax* may be due in part to the production of hypnozoites, since the reactivation of parasites and subsequent antibody production may occur in a different location to site of parasite acquisition, or in the absence of ongoing transmission. However, it is unlikely that recrudescent infections would have had a major confounding effect on school seroprevalence, unless large-scale population movements would have occurred. Furthermore, the wider *P. vivax* range may be due to the parasite's ability to generate sporozoites at lower temperatures and the potential to be transmitted at higher altitudes [291]. Indicators of temperature and altitude, considered to define vector survival and sporogeny, were not found to be associated with *P. vivax* seroprevalence, and not retained in the final multivariate model.

The key environmental variables identified for inclusion in the *P. falciparum* map indicate that higher risk exists in low altitude and low gradient areas close to rivers. We postulate that seasonal flooding in flatlands where floodwaters may pool and act as vector breeding sites could be the driver of this relationship.

A further extension to the current models in the future could be incorporation of intervention coverage, such as districts targeted by indoor residual spraying of households with insecticide, and estimations of long-lasting mosquito net coverage and use alongside environmental covariates.

Despite the difficulties in modelling *P. vivax* seroprevalence, our maps of both *P. falciparum* and *P. vivax* seroprevalence do show broad concordance with predictive maps developed by the Malaria Atlas Project to describe age-standardised parasite rates using model-based geostatistical prediction methods [20,21,292], with similar areas of Oromia identified as areas of highest and lowest risk for malaria. Survey locations with microscopy-positive samples in the most recent Ethiopian Malaria Indicator Survey in 2011 also correlate with our predictive map, with infections identified along the Rift Valley as well as in the far west of Oromia [228].

This study was designed to evaluate large-scale spatial heterogeneity of *P. falciparum* and *P. vivax* malaria. While logistical constraints limited the number of samples analysed, the original surveys were powered to microscopy-based parasite rate – therefore, seroprevalence rates being higher than microscopy should mean that adequate samples were examined to be able to evaluate associations with environmental variables and build the statistical model. The study was not designed to assess micro-heterogeneity in transmission within communities, which has been demonstrated in other settings with similarly low transmission levels (e.g. Somalia, The Gambia, Guinea Bissau) [191,263]. The randomisation process and use of school-attending children as a sampling frame should result in a sampled population representative of the whole school catchment area and wider community. We acknowledge that there is potential for school catchment areas in Ethiopia to have diversity in transmission intensity as a result of steep gradients and presence of local water bodies, dams and irrigation systems [209,210]. Individual differences in immune status and antibody production in response to *Plasmodium* antigen exposure are expected, and may be moderated by other parasitic infections, including helminths [293]; yet, infection risks for these are likely to be

broadly similar across all sites, and individual differences in immune response likely randomly dispersed among the population.

The Bayesian spatially explicit models developed in this study could be refined by inclusion of serology data from additional sites, both within Oromia to assist in categorising areas of high model uncertainty, as well as from other regional states to assist in developing a nationally-representative risk map. Serological analysis of filter paper blood spots included in periodic national surveys such as Malaria Indicator Surveys or Demographic and Health Surveys would be a simple strategy to collect additional seroprevalence data nationally.

Should serology become a primary indicator for malaria surveillance, it may be worthwhile to review the recommended sampling strategy for serological indicators, to ensure a cost-efficient, timely and appropriately powered survey. Further developments to this work and exploration of the utility of serological indicators as part of a package of surveillance tools in Ethiopia could be validation of measured seroprevalence and model predictions against other available data, including clinical burden recorded routinely at health facilities, and cluster-level Malaria Indicator Survey data.

These data represent the spatial integration of simple survey design with a relatively basic laboratory assay which can subsequently guide malaria control and surveillance. The approach has particular utility in low transmission settings and, therefore, has important applications for malaria elimination.

CHAPTER 4. CHARACTERISING SCHOOL ABSENTEEISM AND ITS UTILITY FOR SYNDROMIC SURVEILLANCE SYSTEMS IN LOW INCOME COUNTRIES

4.1 OVERVIEW

Temporal resolution is a critical component of surveillance systems operating in areas at risk of infectious disease epidemics. Syndromic surveillance uses pre-diagnostic and non-clinical surrogate data to increase the responsiveness of surveillance for particular epidemic-prone diseases. While a syndromic approach is not likely to replace routine surveillance strategies using confirmed disease data, due to it's improved temporal resolution, it may be a useful additional strategy in some settings. This chapter presents the first application of a school-based syndromic surveillance system for malaria in a resource poor setting.

This chapter has been prepared for submission to PLoS One, but at the date of thesis submission has not been submitted to the journal or subject to peer review. I led the study design, tool development and field data collection. I conducted all data analysis presented in this chapter.

4.2 BACKGROUND

A number of infectious diseases exhibit marked spatial and temporal trends [202,294-296], which can manifest as epidemics. Epidemics are defined as "occurrence in a community or region of cases of an illness […] clearly in excess of normal expectancy. Epidemicity is thus relative to usual frequency of the disease in the same area, among the specified population, at the same season of the year" [61]. Epidemics often result from spread of an infectious

disease to a non-immune population, and effective surveillance systems are required to identify these unexpected increases in disease burden; to minimise the onward spread of the pathogen through diagnosis and treatment of existing infections, and ensure use of preventative interventions in affected communities. In the highlands of East Africa, malaria transmission is characterised by epidemics, which occur periodically in this environment of spatially and temporally heterogeneous transmission. In Ethiopia, historical malaria epidemics have resulted in very high case fatality rates [215], While recent epidemics have been localised and associated with low mortality (PMI Ethiopia impact evaluation), reliable tools for malaria epidemic detection and response are needed. The current detection method in Ethiopia consists of plotting weekly confirmed malaria cases at health facilities against a threshold calculated using historical data [230]. However, this approach is limited by a need for five years' retrospective data, the requirement for consistent recording of weekly data against the threshold, and failure to account for year-to-year variations in the weeks of peak malaria transmission [229]. Delays in data reporting and incomplete or inaccurate data limit application of the health facility-based epidemic detection system [132-134]. While these limitations persist, alternative tools or strategies that can bridge these gaps and facilitate early identification of epidemics are needed.

Syndromic surveillance refers to the use of pre-diagnostic health indicators to allow timely detection and investigation of potential infectious disease outbreaks [135] as a supplementary approach to routine public health surveillance, by allowing early identification of clusters of illness before confirmatory data are available. In addition to use of clinical (syndrome) data, syndromic surveillance can be expanded to include surrogate non-clinical data indicating early illness, by mining available data to track possible changes in infectious diseases in the population. Surrogate data sources include prescription and over-the-counter drug sales, Internet search terms, social media [149,150,153,154,159-161] and school absenteeism [17-19,144,145]. The latter is an alternative indicator of population health that

has been applied to monitor influenza outbreaks in high-income countries, but yet to be fully explored as an approach for infectious disease surveillance in resource poor settings. Syndromic surveillance systems piloted in resource poor settings have to date used clinical signs among patients attending health facilities as their indicators [122,137,138,140,141], but two examples of school absenteeism being used as early warning of outbreaks of respiratory and gastrointestinal diseases are available from Cambodia and rural China [146-148]. The key surveillance studies using school absenteeism for outbreak detection, as well as classic applications of syndromic surveillance utilising data on clinical morbidities are presented in Table 4.1, to demonstrate the various settings, indicators, temporal resolution and complexity of these syndromic surveillance systems.

School absenteeism data may prove valuable for detecting malaria epidemics since primary school enrolment has increased dramatically in recent years in Ethiopia and other African settings [162], and school absenteeism is known to increase during malaria epidemics [163]. Febrile illness is an additional syndromic indicator that may be useful for identification of malaria epidemics. While a large proportion of infections in low transmission settings such as Ethiopia are of low parasite density and asymptomatic [69-71], it is expected that during an epidemic symptomatic illness would increase. School absenteeism during a malaria epidemic may also increase due to children taking over household chores or directly caring for a family member who has malaria.

The current study was designed to explore the usefulness of a syndromic surveillance system for detection of unusual increases in malaria at community-level in southern Ethiopia, with a particular focus on the use of school absenteeism. Absenteeism data were collected through school and community level systems. In addition, two school-based epidemic detection systems were developed and piloted at schools during the major malaria transmission season.

Table 4.1 - Selected syndromic surveillance systems reported in the literature: the setting, target diseases, indicators, system complexity and outcomes of their application. Reported studies are those which use school absenteeism as a key indicator, or systems applied in resource-limited settings for epidemic prone diseases including malaria

Setting	Target disease(s)	Indicators	Reporting frequency	Complexity of system	Surveillance system findings	Ref
Canada	H1N1 influenza	Elementary and high school absenteeism due to influenza-like illness exceeding the defined threshold of 10% of total enrolment	Daily analysis of absenteeism, reporting if threshold exceeded	Low – schools report data when indicator exceeds the threshold	Absenteeism was well correlated with hospitalisation rates for school age children and PCR positive tests for influenza. Peak absenteeism preceded peaks in hospitalisations by one week. Sensitivity high since majority of children with influenza expected to become ill, and absence due to ILI rather than all-cause absenteeism is indicator. Moderate specificity since not all ILI will be due to H1N1 infection.	[17]
United Kingdom	H1N1 influenza	School absenteeism in primary and secondary schools, comparing against telephone health hotline, general practitioner sentinel network & confirmed influenza data	Weekly mean percentage absenteeism	Low – collation of school % absenteeism data	Weekly school absenteeism peaked concomitantly with existing influenza alert systems, and would not have identified pandemic influenza earlier than other systems. Daily attendance data may have improved timeliness. Sensitivity system high since uses data from multiple sources, includes data from school children as well as general population, syndromic and confirmed indicators. Specificity moderate to high since includes confirmed influenza data rather than simply influenza-like illness.	[18]
Japan	Influenza	School influenza-related absenteeism, where child absent with confirmed diagnosis from physician	Daily school influenza-related absenteeism rate	Low – daily attendance routinely recorded and absent children require doctor's note	School influenza-related illness can be used to predict outbreaks and determine when a school should close to limit ongoing spread. Thresholds for influenza-related absenteeism proposed. High sensitivity and specificity due to requirement for physician confirmation of reason for absence, and target disease of influenza rather than H1N1 influenza	[19]
China (rural)	Respiratory infections, gastroenteritis	Symptoms reported at health clinics, over-the-counter drug sales at pharmacies and primary school absenteeism	Daily input to web-based system	High – collation and analysis of data at central level	Labour-intensive data entry to electronic system. Presentation of six months' pilot data, no validation of data from surveillance system against other sources, therefore difficult to determine sensitivity and specificity. Should have high sensitivity and specificity based upon combination of specific and non-specific indicators from school-aged and general population.	[147, 148]

Setting	Target disease(s)	Indicators	Reporting frequency	Complexity of system	Surveillance system findings	Ref
Madagascar	Malaria, influenza, dengue, diarrhoeal disease	Health facility malaria case confirmed by RDT, fever & respiratory symptoms, fever & two possible dengue symptoms, diarrhoea	Daily report by encrypted SMS. Weekly summary paper report.	Moderate – SMS reports entered to database. Temporal & spatial analysis by syndrome	Ten cases of fever clusters occurred which weren't detected by the traditional surveillance system. Five outbreaks identified: two dengue, two influenza and one malaria, indicating high sensitivity of the system. However the comment of authors that increases in self-limiting (and low priority) fever may trigger an epidemic alerts indicates that specificity of the system is low to moderate.	[122]
French Guiana	Dengue	Dengue index: percentage of patients attending the emergency department who had thrombocytopenia but were negative for *Plasmodium* infection	Weekly generation of indicators	Low – plotting of simple indicators on weekly basis, minimal analysis	Dengue index was specific – increasing during what was confirmed to be a dengue epidemic, but showing no strong increase during two respiratory infection epidemics. Total emergency department attendance with thrombocytopenia but malaria negative was also a specific indicator. System appears both sensitive and specific, correctly identifying an epidemic of targeted disease, but correctly differentiating a respiratory infection epidemic.	[141]
Pacific island countries and territories	Measles, dengue, rubella, meningitis, leptospirosis, gastroenteritis, influenza, typhoid, malaria	Hospitals report total cases for four syndromes: acute fever & rash, diarrhoea, influenza-like illness, prolonged fever	Weekly reporting of data to national level	Moderate – data reported from national to WHO regional level for analysis	The system successfully identified an outbreak of diarrhoeal disease linked to breakdown of water disinfection, and two outbreaks of influenza. The system alert was timely and allowed fast implementation of control measures. Good sensitivity according to the syndromic case definitions, but limited to patients attending hospital. Moderate specificity to identify the priority infectious diseases using each of the syndromic case definitions.	[137]
India	Cholera, dysentery, malaria, measles, meningitis, typhoid fever, and 8 others	Suspected cases (clinical diagnosis) of target diseases from public and private health facilities, except malaria, where slide-confirmation required for reporting	As clinical cases identified (daily), using prepared post cards	Low – doctors report cases on simple form to central level. Minimal analysis.	Several outbreaks were detected early and interventions applied, the most notable was cholera. Leptospirosis and acute dysentery also commonly reported. Monthly summary of reported diseases distributed to participating facilities for feedback and updates on the surveillance system. Good sensitivity to identify priority diseases using clinical case definitions.	[138]

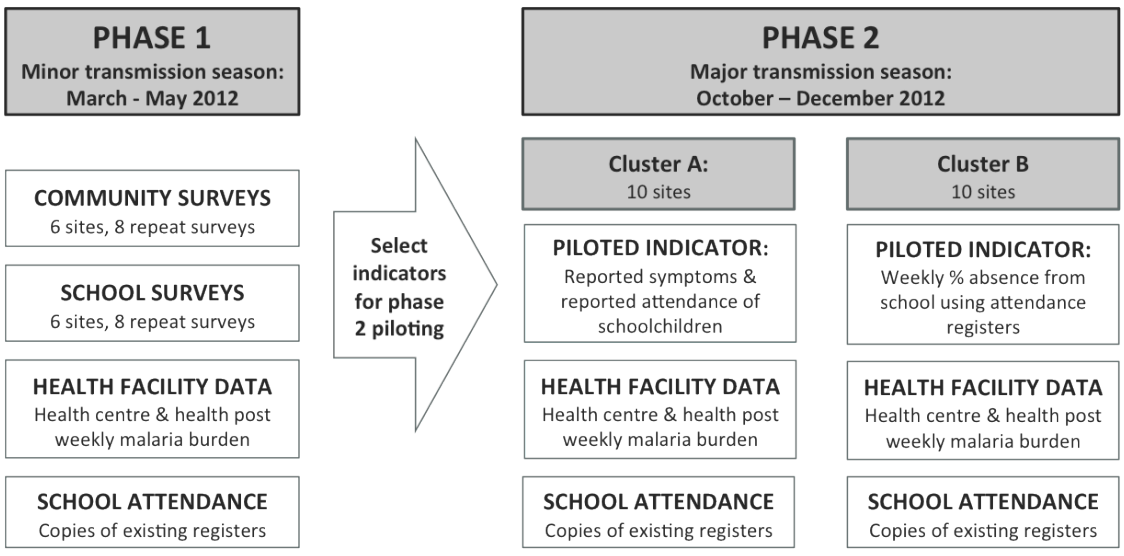
Setting	Target disease(s)	Indicators	Reporting frequency	Complexity of system	Surveillance system findings	Ref
Cambodia	Respiratory and diarrhoeal diseases	School absenteeism (aggregated daily by schools), compared against overall health facility attendance	Daily SMS report of school absenteeism due to illness, collated at weekly level for analysis	Low - daily data reported by schools to central level, compared against all cause health centre attendance	Illness-specific absenteeism identified two peaks in incidence of illness. Absenteeism data preceded peaks in health centre attendance by 0.5 weeks on average. Cross correlation analysis indicated moderate correlations between illness specific absenteeism and reference data. Sensitivity of illness-specific absenteeism was good, and had moderate specificity.	[146]
Papua New Guinea	Influenza, cholera, typhoid, malaria, poliomyelitis, meningitis, measles, dengue	Syndromes relating to target diseases identified in patients presenting to health facilities.	Weekly report by mobile phone, transcription to database	Low – health facilities submit data for analysis at provincial or national level, and automatic generation of feedback reports	System was more sensitive than the reference system for measles, but low sensitivity for malaria, due to poor case definition. Data were more timely than the reference system (mean 2.4 weeks compared to 12 weeks lag). The system had very good sensitivity for indicators of measles, but low sensitivity for indicators of malaria. Specificity of the system was moderate, due to the broad syndromic case definitions.	[139]

4.3 METHODS

4.3.1 Study design

The study was conducted in Southern Nations, Nationalities and People's Regional State (SNNPRS), Ethiopia and divided into two phases. Phase 1 comprised repeated cross-sectional school- and household-based surveys at six sites, conducted during the minor transmission season of 2012, as well as collection of routinely recorded school attendance data from all six sites and weekly summary of clinical and confirmed malaria cases identified at health facilities serving the study sites. Phase 2 involved piloting two simple school-based malaria epidemic detection systems at 20 sites in SNNPRS during the peak transmission season of 2012. A diagram of study design is presented in Figure 4.1.

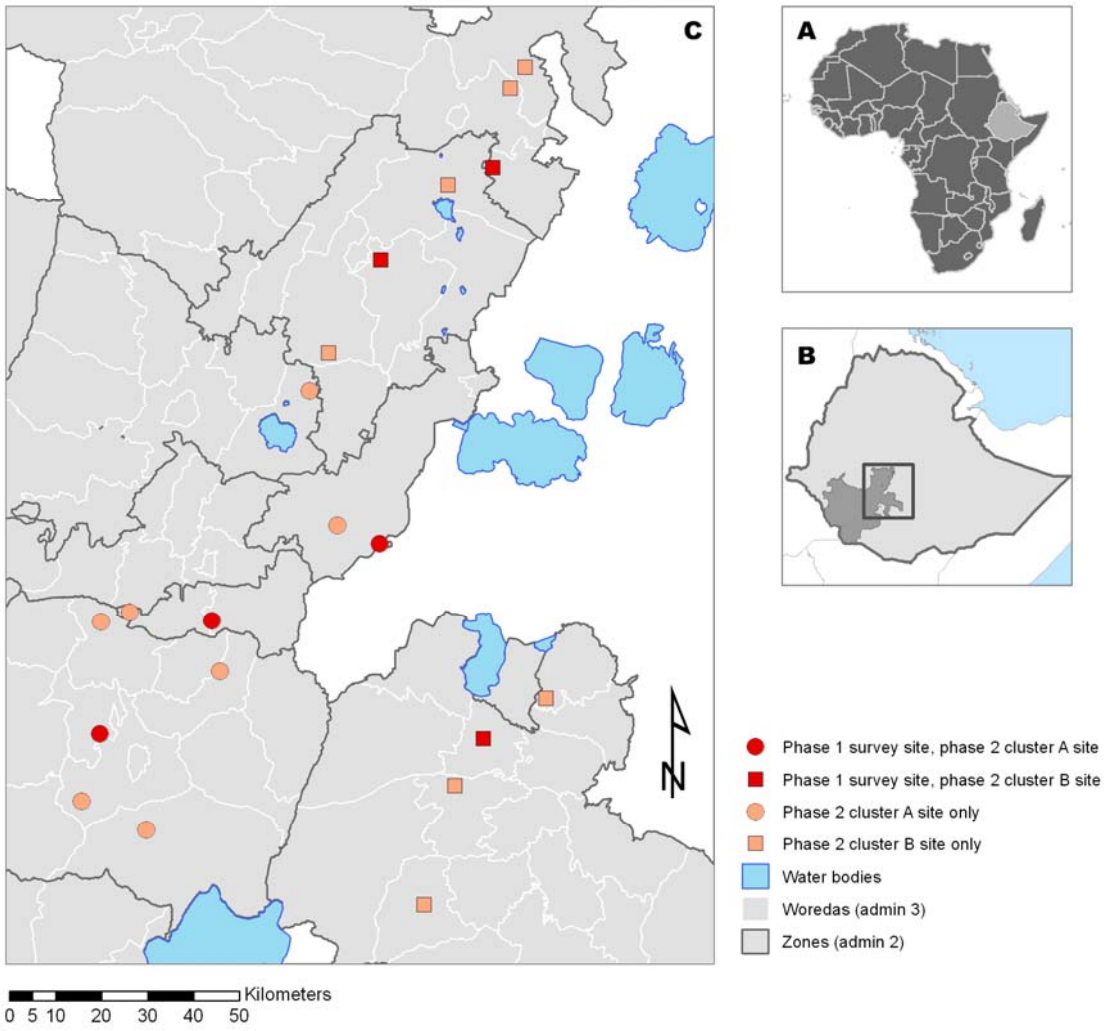
Figure 4.1 - Study design diagram indicating activities conducted during Phase 1 (school- and community-based surveys) and Phase 2 (piloting of two school-based syndromic surveillance system). Health facility and school attendance data were collected throughout Phases 1 and 2



4.3.2 Phase 1: cross-sectional surveys at school and community

Six "hotspot" woredas (administrative level 3, districts) were chosen purposively as locations for study sites for Phase 1. The "hotspot" designation is determined by the Regional Health Bureau, to indicate sites which have relatively high burden of malaria and are at risk of epidemics. The six woredas were chosen from different zones (administrative level 2) to allow coverage of different geographical and ecological settings of SNNPRS. One kebele (administrative level 4, municipalities) was chosen as the study site from each of the six woredas. Selection criteria for kebeles included presence of a government primary school with regular attendance of at least 100 children, a functional health post (community-level health facility offering a basic package of diagnosis and treatment services), an altitude range of less than 200 metres (calculated using digital elevation model in Geographic Information System), and accessibility by vehicle during the rainy season. Although the majority of kebeles in SNNPRS have primary schools and operational health posts, some woredas identified as "hotspots" for malaria are characterised by a wide range in altitude and poor accessibility. Location of selected sites is shown in Figure 4.2. Eight repeat school and community surveys were conducted at approximately ten-day intervals from March to May 2012. Due to lack of prior data describing likely range of key indicators, no sample size calculation could be completed; therefore, the number of sites and visits was maximised within the available budget and expected transmission season.

Figure 4.2 - Locator maps of Ethiopia (A) and SNNPRS (B), with map of study kebele location (C). Six sites which were included in the Phase 1 school and community surveys as well as Phase 2 pilot are indicated by red markers, while the remaining 14 sites participating in Phase 2 pilots only are indicated by orange markers. Assignment to cluster A (symptom questionnaire) during Phase 2 is indicated by circular markers, assignment to cluster B (absenteeism estimated from attendance registers) is indicated by square markers



The community survey protocol followed Malaria Indicator Survey procedures for mapping and random selection of 25 households per survey site [104]. Primary sampling units were defined as primary school catchment area rather than census enumeration area used for MIS. All six study sites and, therefore, primary sampling units, were chosen purposively. Prior to commencing surveys, all households within the school catchment area were mapped using a hand-held global positioning device (eTREX, Garmin International, Kansas USA). Parental consent for inclusion of children in school surveys was requested during household mapping. A random number table was used to select 25 households for inclusion in community survey, with random selection repeated for each of the eight survey iterations. Households were permitted to be included more than once over the course of the study. Selected households who refused participation or were absent were replaced with the next household on the mapping list.

At each household, a simple questionnaire was completed to collect socio-demographic information from all household members, fever history, school enrolment and attendance history for children of school age, as well as status of indoor residual spraying of households with insecticide (IRS), insecticide-treated net (ITN) ownership and use. All individuals aged older than one month were invited to provide a finger-prick blood sample for preparation of thick and thin blood films, malaria rapid diagnostic test (RDT, CareStart PfHRP2/panLDH, Access Bio, Somerset NJ, USA) and collection of blood spots on filter paper.

School surveys followed a standard methodology previously used in Ethiopia and Kenya [16,297]. Briefly, 100 children (plus ten reserves) were randomly selected from children in grades two to six present on the survey day, and from whom parental written consent had been received. Equal numbers of boys and girls were selected. Random selection of children was repeated at each survey iteration. Children aged over 16 years were excluded. Each child was interviewed in the local language and a standard questionnaire completed including basic socio-demographic information, fever history, frequency of school attendance, recent

absences from school and reason, use of ITN at home and status of IRS. Each child was requested to provide a single finger-prick blood sample for preparation of multi-species RDT, thick and thin blood film, and blood spots on filter paper.

In addition to cross sectional surveys, school directors were requested to give permission for photocopying of their attendance registers at the end of the Phase 1 data collection period. Registers were requested from all classes of grades two to six. For each class, weekly absenteeism was calculated as the total child-days absence recorded, divided by product of total children enrolled and number of days that attendance was recorded by the teacher.

Summary weekly malaria data are routinely collated by health facilities for reporting to the woreda level, as part of the Integrated Disease Surveillance and Response (IDSR) system on priority infectious diseases [93]. The study team collected weekly malaria data from health centres and health posts serving the six study sites at the end of the Phase 1 survey period, by manually transferring data from routine records onto duplicate forms prepared by the investigator, or photographing original forms. Health facility data included total suspected malaria cases, number tested by microscopy (health centre) or RDT (health post) and total confirmed malaria cases.

4.3.3 Blood film processing, reading and quality control

Blood films were fixed and stained at the local health centre daily during Phase 1 surveys. First slide reading was conducted at the Adama Malaria Reference Laboratory with the aim to validate RDT result. Microscopists recorded the presence or absence of asexual or sexual parasite forms, *Plasmodium* species and parasite count using standard methods. Slides from any individual with a positive RDT were read, along with slides from any individual in a household with an RDT-positive individual, and slides from all individuals in 10% randomly selected households where all individuals were RDT-negative. A random selection of 10% of RDT-negative slides from school surveys were also read. A second microscopy reading was

conducted at the Ethiopian Health and Nutrition Research Institute (EHNRI) in Addis Ababa, comprising all slides from individuals with positive RDT result, all slides found to be positive on first microscopy reading, and 10% of all remaining slides previously read. A third reading was conducted at EHNRI to settle discrepancies from first and second microscopy readings, as well as confirm conflicting RDT and microscopy results.

4.3.4 Phase 2: piloting of two school-based malaria surveillance systems

Based on findings from Phase 1 and discussions with partners, two surveillance systems were piloted during the second school semester, from October 2012 to January 2013 (Figure 4.1). Twenty woredas were purposively chosen for inclusion in Phase 2 activities, using the same inclusion criteria as during Phase 1. All sites included in Phase 1 were retained for Phase 2, and a further 14 woredas were selected from within the six zones included in Phase 1 (Figure 4.2). Woredas were assigned to participate in either the cluster A or B pilot without randomisation, in such a way that an equal number of woredas were in each cluster, and all woredas within a zone were assigned to the same cluster. This was in order to reduce confusion between the two pilot methodologies by any zone health office staff supporting implementation. Two teachers from each school (i.e. school director and one other member of staff) and a representative from the woreda health office were invited to a one-day training session for orientation in the pilot study procedures. Separate training days were held for each cluster.

Cluster A sites focussed on monitoring schoolchildren's reports of fever, absence from school for any reason and absence due to illness during the previous week. A short questionnaire including the three indicators of interest and nine masking symptoms (e.g. headache, cough, diarrhoea) was completed every Monday, immediately following completion of the routine attendance register. The teacher usually responsible for recording attendance in each class interviewed children using the questionnaire. Each child was called in turn to the teacher's

desk for the interview, to allow privacy while responding to the symptom questions. The symptom questionnaire was restricted to grades two to four, since school directors indicated that higher grades had less time to participate in piloting due to exam preparations. Interviews were rotated weekly between these three grades to minimise disruption to normal teaching.

Cluster B sites were requested to use data recorded in their usual attendance registers to complete a weekly summary across all grades of the proportion absent, with total children enrolled multiplied by number of days attendance was recorded as the denominator.

At the end of the pilot period, copies of attendance registers were collected from a convenience sample of schools from both clusters for validation purposes. Weekly health facility malaria data were also collected from 20 health centres and 20 health posts serving the Phase 2 study populations.

4.3.5 Data entry and analysis

Questionnaire data from Phase 1 surveys were entered into a customised Microsoft Access 2007 database developed to automatically conduct consistency and range checks, while microscopy results, health facility and school absenteeism extracted from attendance registers were entered into Microsoft Excel. All data were merged in Stata 12.0 (Stata Cooperation, College Station, Texas USA). Household coordinates for Phase 1 sites were imported into ArcMap 10.0 (Environmental Systems Research Institute Inc., Redlands, California USA) for display and calculation of Euclidean distance between household and school (at approximately 100m resolution). Phase 2 data were entered into Excel spreadsheets and exported to Stata 12.0 for merging and analysis.

4.3.6 Analysis of Phase 1 data

A wealth index to summarise socio-economic factors was created using principal component analysis, at household level for community survey data and at individual level for school survey data [298] (detailed in Appendix 1). Due to a lack of diversity in socio-economic factors, the wealth index was classified into three categories only for both school and community survey data.

Mixed effects logistic regression was used to develop multivariate multilevel models describing risks of non-enrolment in school, as reported by head of household during community survey, with household-level and site-level random effects. A backward step-wise method was used to exclude the least significant fixed effects one by one: a likelihood ratio test was used to re-test excluded variables for inclusion in the final model. The same modelling strategy was used to generate models of risk factors for the binary RDT result, generating separate models for data from school- and community-based surveys. Community-survey multivariate models included household and site as random effects, but school-survey models included only site as a random effect since it was not possible to link children sampled in schools to their households.

4.3.7 Analysis of Phase 2 data

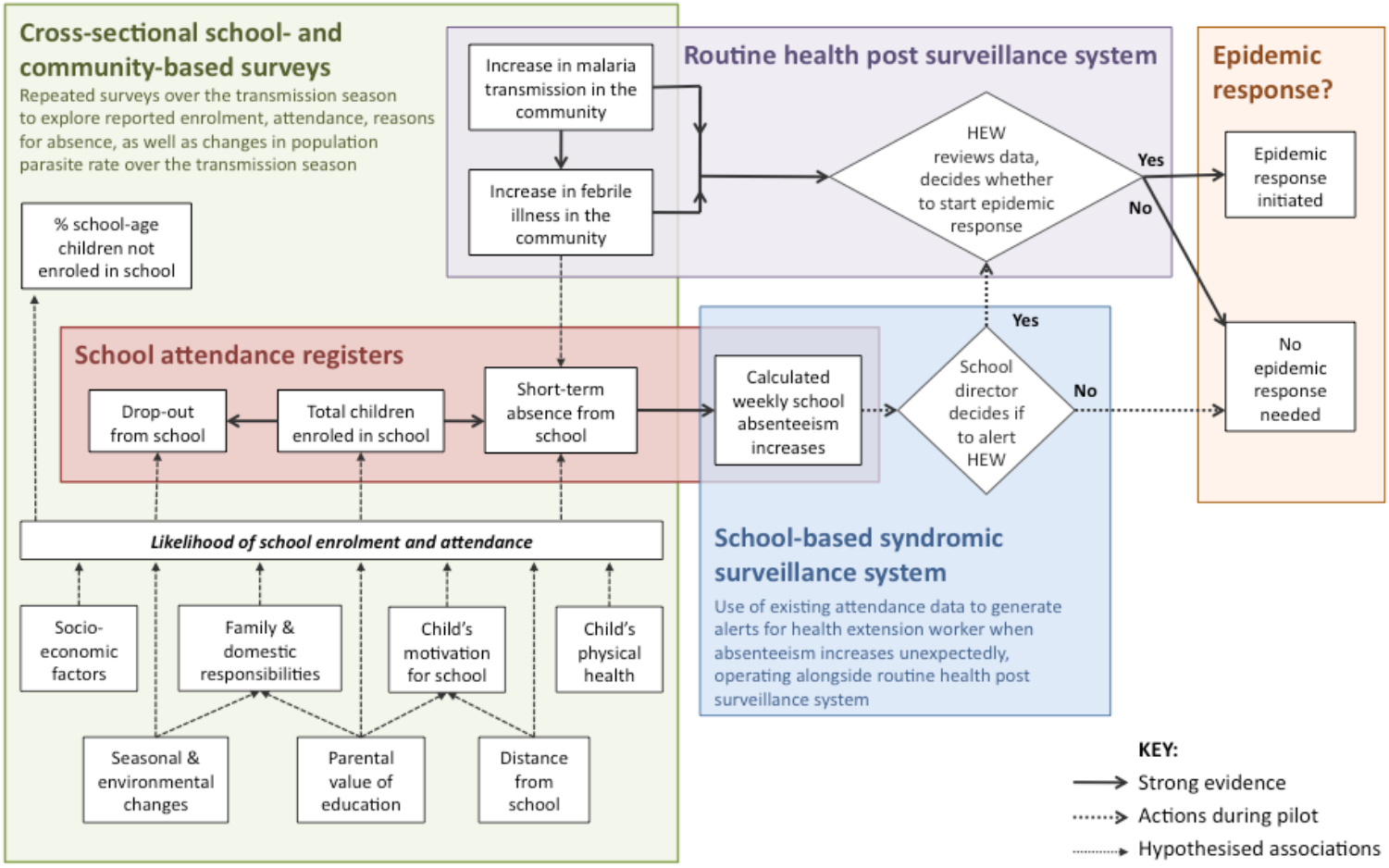
Analysis of Phase 2 data focussed upon describing the characteristics of indicators collected at schools during the pilot. Box plots and logistic regression were used to describe absenteeism by grade over the study period. To explore dropout levels in these populations, mean absenteeism was evaluated by time. Accuracy of weekly summary absenteeism calculated by cluster B sites was determined by comparing teacher-generated summaries against summaries calculated by the study team using original registers.

4.3.8 Theoretical framework for the syndromic surveillance system

An information and decision flow diagram is presented in Figure 4.3 to demonstrate the key processes involved in one of the piloted school-based syndromic surveillance systems (based on absenteeism calculated from school attendance registers). Cross-sectional surveys conducted in the community and school during Phase 1 were designed to assess the level of enrolment in school and factors associated with enrolment. Blood samples were collected from sampled populations to explore changes in malaria burden over the transmission season, and individuals were asked to report recent febrile illness. These indicators were also extracted from routine health facility records, but representing individuals attending health facilities, rather than the random sample of the population assessed in surveys. Reasons for short-term absence were collected during surveys, and attendance registers collected to assess levels of absenteeism and dropout. The syndromic surveillance system piloted in Phase 2 used the weekly absenteeism rate calculated from attendance registers, to prompt a decision by school directors if there was an unexpected increase in absenteeism, and whether to alert the health extension worker. This alert feeds into the health extension worker activities, and is compared to the routine surveillance data (clinical data from patients attending health post). The health extension worker can therefore decide to launch an epidemic response as a result of their routine data, intelligence from the school, or a combination of both.

.

Figure 4.3 - Information and decision flow chart, describing the use of a school-based syndromic surveillance system to identify malaria epidemics, operating alongside the routine health post malaria surveillance system



4.3.9 Ethical considerations

Approval for this study was granted by the London School of Hygiene and Tropical Medicine ethical committee (6003) and by the SNNPRS Health Research Ethics Review committee (P026-19/6157).

Written, informed consent for participation of children in school surveys was collected from parents during the household mapping exercise, prior to the first school survey. Parents were free to withdraw consent at any time by informing the school director. Children without written consent were not eligible for selection in school surveys. Children provided written assent for participation, and were informed of the study procedures prior to random selection, as well as their right to withdraw at any point. The head of household provided written consent for inclusion of household members in community survey at the point of data collection. Verbal assent was sought from all household members before participation. Where households were selected more than once, written consent was requested at each survey.

A health extension worker was present throughout school and community surveys to assist the survey team. Any individual with positive RDT was provided with treatment by health extension worker on the same day, according to national guidelines (i.e. artemether-lumefantrine for *P. falciparum* or mixed infections and chloroquine for *P. vivax* mono-infection).

4.4 RESULTS

4.4.1 Population participating in school- and community-based surveys

Each of the eight repeat community surveys at all six sites included 25 households, with 140 households being randomly selected for inclusion in more than one survey. Communities had mean 664 households within the school catchment area (range by site 390-831). RDT results were available from 4117 individuals participating in community surveys, aged from two months to 101 years (mean 20 years). Of the 5238 children participating in school surveys, data from 5189 children aged from seven to 16 years were retained for further analysis. "School-aged" in the current study was defined as aged seven to 16, since national policy is for children to enrol in school at seven years of age [299]. RDT results were available from 5145 of these children sampled during school surveys.

The prevalence of *Plasmodium* infection by RDT across all sites and survey iterations was 2.0% (range across site and iteration 0-12.3%) for school surveys and 2.6% (0-8.9%) for community surveys. Of the 104 RDT-positive samples from school surveys, slides were missing or unreadable from three individuals, 35 (34.7%) were positive by microscopy (15 *P. falciparum* and 20 *P. vivax*). In community surveys, slides were available from 84 of the 106 RDT-positive individuals, 32 (38.1%) of which were positive by microscopy (7 *P. falciparum* and 25 *P. vivax*). Of 501 slides read for validation purposes from RDT-negative individuals across both school- and community-based surveys, three were *Plasmodium*-positive by microscopy. The range of RDT positivity across school survey iterations and sites was 0-12.3%, and 0-8.9% for community surveys. At any single site, the maximum difference in RDT prevalence over all survey visits was 7.6%. A summary of key indicators by survey type and site, as well as from health facilities during the Phase 1 period are presented in Table 4.2. Trends in RDT prevalence in surveys and total confirmed malaria cases confirmed at health centre over Phase 1 at each study site are presented in Figure 4.4.

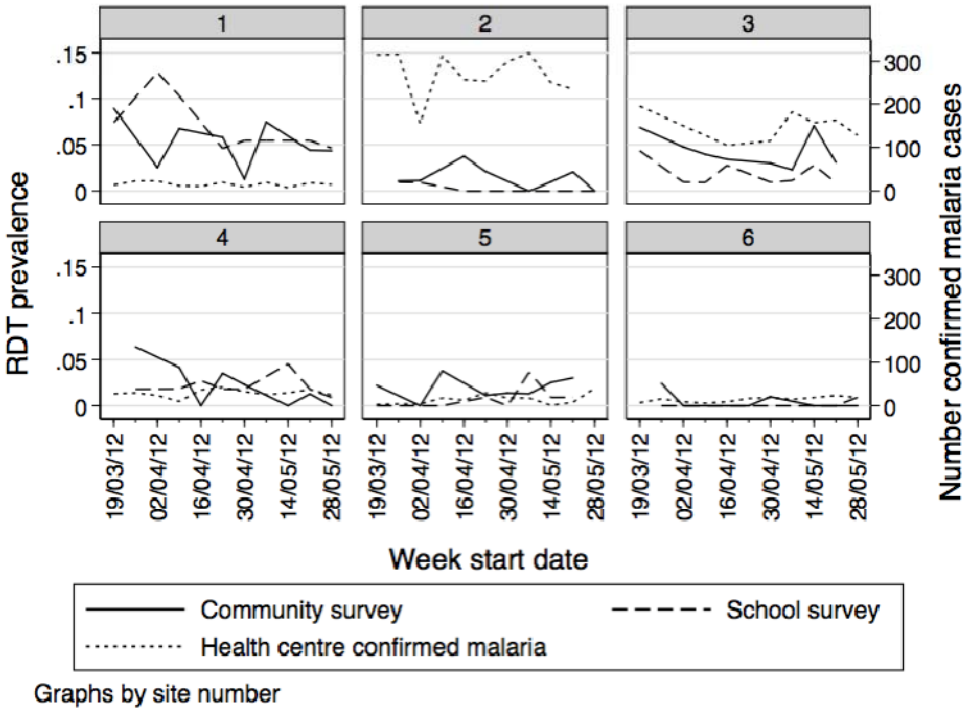
Table 4.2 - Description of key indicators collected during Phase 1 school- and community-based surveys at six sites. The number of individuals providing blood samples across all visits to each site, and range of RDT prevalence by survey iteration at each site are shown, to demonstrate changes in malaria infection in each community over the study period. In addition, the range in total clinical malaria (febrile illness), confirmed malaria and test positivity rate by week at each health centre and health post are presented for each study site

.	**Site 1**	**Site 2**	**Site 3**	**Site 4**	**Site 5**	**Site 6**
School surveys						
N children bled	855	861	820	868	872	869
% RDT positive range across survey iterations	5.5 - 12.3	0.0 - 1.0	0.9 - 4.4	0.9 - 4.6	0.0 - 3.7	0.0 - 0.9
Community surveys						
N bled (all ages)	575	718	823	675	656	673
% RDT positive range across survey iterations	1.3 - 8.9	0.0 - 3.8	2.3 - 7.6	0.0 - 6.4	0.0 - 3.8	0.0 - 2.4
Health post						
N weeks data reported	12	9	12	11	0	12
n RDT positive range	5 - 12	4 - 26	0 - 27	1 - 18	-	0 - 21
Health centre one[A]						
N weeks data reported	12	11	8	11	11	11
n febrile range	14 - 64	280 - 438	165 - 284	30 - 83	28 - 64	112 - 212
n malaria range	8 - 23	77 - 319	105 - 195	10 - 43	1 - 41	6 - 24
% test positivity range	19.6 - 100	92.9 - 98.8	54.2 - 68.7	37.9 - 75.0	1.9 - 68.2	3.3 - 15.2
Health centre two[B]						
N weeks data reported	0	12	0	12	0	12
n febrile range	-	348 647	-	30 - 64	-	54 - 174
n malaria range	-	210 - 335	-	1 - 52	-	3 - 28
% test positivity range	-	44.4 - 60.3	-	3.0 - 100	-	3.0 - 24.1
School attendance registers						
N weeks data reported	12	12	10	12	11	12
n child-days attendance reported range	247 - 3598	334 - 1634	918 - 3130	1430 - 3190	168 - 2876	1560 - 2951
% absenteeism range across school	71.5 - 83.4	43.4 - 77.9	69.5 - 92.4	65.9 - 86.1	67.7 - 82.2	60.7 - 82.5

[A] Health centre one is the health centre closest to the study site

[B] Health centre two is a larger health centre in the woreda town, also serving the study area

Figure 4.4 - Phase 1 community survey (solid line) and school survey (dashed line) prevalence of RDT positivity at each survey iteration. Dotted lines indicate total confirmed malaria cases identified at the local health centre by routine passive surveillance, plotted against secondary y-axis. Individual graphs are plotted for each Phase 1 study site.



In Figure 4.4, the lack of variation in crude total malaria cases confirmed at health centres over the Phase 1 period can be seen at sites 1, 4, 5 and 6. Sites 2 and 3 had higher burden of malaria reported from health centres, but fluctuations in burden appeared to be random noise rather than the expected seasonal transmission peak. Prevalence of infection by RDT (any *Plasmodium* species) from school- and community-based surveys similarly showed small fluctuations between survey iterations.

4.4.2 Reported primary school enrolment of school-aged children

Overall, 32.7% of the population living in sampled households were of school-age (7-16 years). Of all school-age children registered during community surveys, 54.0% (range by site 42-62%) were reported by their head of household to be enrolled at the local primary school. While it is possible that some children are enrolled at a different school, this is expected to be very unlikely since the survey was conducted in the target primary school catchment area, therefore even if another primary school exists in the kebele, the closest school is the target school for the study. Furthermore, there is no culture of private or boarding school attendance in rural Ethiopia, therefore the majority of children who attend school will do so at the closest primary school to their home. Full description of the school-aged population and reported enrolment of school-aged children in school is presented in Table 4.3. Factors associated with non-enrolment of school-aged children were investigated to assess any commonalities among the school-aged children who would not be captured by school-based malaria surveillance.

Table 4.3 - Description of school-aged population by site, and reported enrolment in primary school by health of household from community-based surveys in Phase 1. Factors which may influence likelihood of school enrolment are also described at each of the survey sites. Number in brackets are 95% confidence intervals unless otherwise stated.

	Site 1	**Site 2**	**Site 3**	**Site 4**	**Site 5**	**Site 6**
Total unique households interviewed	182	167	175	167	185	175
Total unique children 7-16 years in interviewed households	328	267	397	269	256	277
Mean household size	4.91 (4.64, 5.18)	5.29 (4.99, 5.59)	6.59 (6.28, 6.90)	5.11 (4.58, 5.20)	4.89 (4.58, 5.20)	5.19 (4.89, 5.49)
Mean number children 7-16 in household	1.80 (1.58, 2.01)	1.58 (1.39, 1.77)	2.26 (2.01, 2.50)	1.59 (1.41, 1.78)	1.38 (1.19, 1.56)	1.54 (1.37, 1.73)
% SAC reported enrolled in school	59.6 (54.1, 65.0)	61.7 (55.6, 67.6)	54.9 (49.9, 59.9)	60.2 (54.0, 66.1)	46.3 (40.0, 52.6)	41.7 (35.8, 47.8)
Median (range) wealth index	-0.05 (-1.48, 6.97)	-0.94 (-1.48, 1.95)	-0.94 (-1.48, 5.09)	0.08 (-0.94, 2.98)	-0.94 (-1.48, 2.98)	-0.94 (-1.48, 2.98)
Median (range) distance of household from school in metres	0.90 (0.00, 2.35)	1.06 (0.11, 2.45)	1.20 (0.11, 2.57)	1.12 (0.11, 2.99)	0.90 (0.00, 2.27)	1.74 (0.16, 3.24)
% household heads with no formal education	40.1 (32.9, 47.6)	28.1 (21.5, 35.6)	38.2 (30.9, 45.8)	42.2 (34.6, 50.1)	60.7 (53.2, 67.8)	79.3 (72.5, 85.1)

SAC: school-aged children

From multivariate modelling, key risk factors for non-enrolment of school-aged children in school were the distance of the household from school, and the number of children of school age in the household (Table 4.4). Odds of a child not being enrolled in school were lower in the least poor and median wealth households compared to the poorest households. Odds of enrolment also varied with education level of the head of household, with children from households where the head had attended any education having higher odds of enrolment than those from households headed by an individual with none.

Table 4.4 - Multivariate model of risk factors for non-enrolment of school-age children (as reported by head of household during community survey). Fixed effects are presented, the multilevel model included random effects at household- and study-site level. Data were available from 1794 unique children and total 908 households, sampled from six sites in SNNPRS in 2012.

	Odds ratio	**95% confidence interval**	**P**
Age (increasing)	0.91	0.88, 0.95	<0.001
Number of children 7-16 years in household	1.20	1.08, 1.33	<0.001
Distance from school in km	1.57	1.30, 1.89	<0.001
Household wealth			
Poorest	1	-	-
Median	0.73	0.49, 1.10	0.132
Least poor	0.64	0.49, 0.84	0.001
Parental education			
None	1	-	-
Primary incomplete	0.66	0.51, 0.86	0.002
Primary complete or higher	0.64	0.42, 0.96	0.030

4.4.3 Child-reported reasons for absence from school

During school surveys, children were asked their frequency of usual school attendance, if they had been absent from school in the previous two weeks and the reason for absence (Table 4.5). Across all sites, 94% of children reported usually attending school five days per week, indicating that when children are enrolled they do routinely attend. Of all reported absences by children, 28% were due to illness, while 67% of absences were in order to assist in the home or with farming activities. Variations by site were seen, with two sites reporting the majority of absences being due to illness. Where children reported absence from school due to illness, fever was the most common symptom (88%), however only 50% of those who reported fever as a reason for absence from school attended a health facility. The same questions were included in community surveys for children reportedly enrolled at school, with the majority of those enrolled attending every day (94%). Due to small numbers of children reportedly enrolled and absent from school in the previous two weeks who were interviewed in community survey, it was not possible to assess reasons for absence in-depth.

Table 4.5 - Description of frequency of usual school attendance and reasons for recent absence from school, as reported during school- and community-based surveys, by Phase 1 study site

	Site 1	**Site 2**	**Site 3**	**Site 4**	**Site 5**	**Site 6**	**Total**
Community survey							
Total unique children reported enrolled in school	195	163	218	162	119	115	972
% report usually attend 5 days/week	100.0	95.1	88.3	98.2	92.4	94.8	93.7
% report absence in previous 2 weeks	1.6	3.7	17.5	13.6	6.8	3.5	8.3
Reason for absence among those reporting absence in previous 2 weeks							
Unwell	66.7	0.0	22.2	72.7	0.0	0.0	34.7
Helping in the home or fields	33.3	33.3	36.1	9.1	100.0	100.0	37.3
Working elsewhere	0.0	33.3	2.8	4.6	0.0	0.0	4.0
Other[1]	0.0	33.3	38.9	13.6	0.0	0.0	24.0
School survey							
Total children interviewed	859	866	846	870	871	868	5180
% report usually attend 5 days/week	98.6	92.7	97.2	99.8	96.7	77.0	93.6
% report absence in previous 2 weeks	5.0	5.0	9.4	11.0	3.7	25.3	9.9
Reason for absence among those reporting absence in previous 2 weeks							
Unwell	62.8	65.1	34.2	30.2	31.3	11.0	28.3
Helping in the home or fields	27.9	32.6	58.2	57.3	68.8	88.6	67.0
Working elsewhere	0.0	2.3	0	7.3	0.0	0.0	1.6
Other[2]	9.3	0.0	7.6	5.2	0.0	0.4	3.1
% of children absent due to illness reporting fever	80.8	82.1	88.9	83.3	80.0	83.3	83.5

[1]Other reasons for absence from school (reported in community survey) included food shortage, to visit the market, and absence of teachers

[2]Other reasons for absence from school (reported in school survey) included lack of school materials, absence of teachers, travel, and menstruation.

4.4.4 Risk factors for *Plasmodium* infection during cross-sectional surveys

Considering the lower than expected school enrolment level, supplementary analysis was conducted to explore the risk factors for *Plasmodium* infection, and whether there are common risk factors for non-enrolment and malaria. Common risk factors would reduce the sensitivity of the syndromic surveillance system, if a large proportion of individuals at highest risk of malaria were not likely to be enrolled in school. Insufficient infections (29 among 1026 school-aged children tested) were found in community surveys to conduct stratified analysis to explore whether risks of malaria were similar between enrolled and non-enrolled school-aged children.

Univariate analysis indicated that reported fever during the previous two weeks and on the day of the survey, sex and household wealth were associated with RDT positivity in school surveys. The minimal multivariate model retained sex, fever in the previous two weeks and household wealth, with females having lower odds of RDT positivity, and recent fever associated with much higher odds of RDT positivity (odds ratio, OR, of 3.21, 95% confidence interval 1.92, 5.34). Unexpectedly, children from households with higher levels of wealth had higher odds of RDT positivity than children from the poorest households (Appendix 2).

Univariate analysis of community survey data indicated that age, previous fever, ITN use on previous night and household wealth were associated with RDT positivity. The minimal multivariate model for community survey data, including random effects at household and study site level retained only age and ITN use, with odds of RDT positivity declining with age (OR=0.87, 95% CI 0.74, 1.02) and declining with reported use of an ITN on the previous night (OR=0.15, 95% CI 0.03, 0.88). Full results of univariate and multivariate modelling of *Plasmodium* infection determined in community- and school-based surveys are listed in Appendix 2.

4.4.5 Phase 2 surveillance system pilot

Nine of the 10 cluster A schools submitted weekly summaries of the proportion of children reporting fever, absence from school, and absence due to illness during the previous week. Schools collected the weekly indicators for a mean 11 weeks (range 6-14), overlapping with the peak transmission season from October to December. On average 68 children were interviewed each week (range 16-170). All ten of the cluster B schools submitted weekly estimates of absenteeism, calculated by summarising absent sessions recorded in attendance registers. Schools reported a mean 12 weeks of data (range 11-13). Schools summarised attendance for an average of 508 children each week (range 118 to 914). In addition to the indicators collected from cluster A and B sites, available attendance registers were collected from a convenience sample of seven schools (five from cluster A and two from cluster B) for comparison and validation of indicators.

4.4.6 Absenteeism and drop-out recorded by school attendance registers

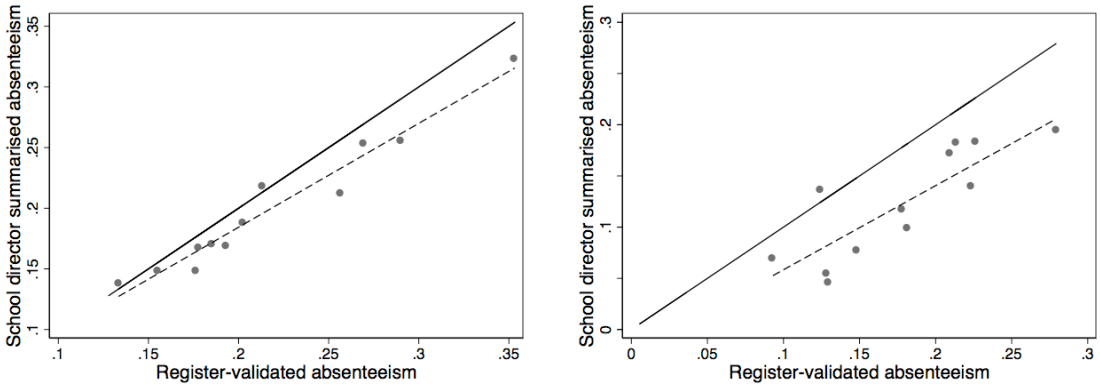
At most schools, summary rates of absenteeism calculated from attendance registers appear similar between classes. There was no evidence for a statistical difference in absenteeism across sites when comparing all grades individually (all $p>0.70$) or when grade was used as a proxy for age and assessed as a continuous linear variable ($p=0.84$).

Changes in proportion of enrolled children who are absent from school are expected to increase over the course of the semester as a result of drop-out. Strong evidence for an increase in average weekly absenteeism was found when analysing grade total absenteeism and allowing for clustering by school ($p=0.008$). Absenteeism fluctuated on a weekly basis and varied by school, but showed an overall increase of 10% during the study period.

Weekly summary attendance recorded at cluster B schools was validated at two sites by comparing teachers' calculated absenteeism against original attendance registers (Figure 4.4). One school showed good agreement between teacher and study team calculated

absenteeism. However absenteeism was consistently underestimated at the second school, likely as a result of teachers not counting absences by children who were judged to have dropped out. For any future study, it is essential that the strategy for inclusion or exclusion of children who appear to have dropped out be clarified.

Figure 4.5 - Scatterplots of weekly absenteeism recorded by school director against absenteeism calculated by study investigators using attendance registers, at two sites in cluster B, with line of unity (solid) and best fit line (dashed)



4.4.7 Do syndromic surveillance indicators from schools correlate with health facility malaria trends?

Testing of the syndromic surveillance system was hampered by the lack of strong seasonal increase in malaria cases seen at the study sites during both Phases 1 (Figure 4.6) and 2 (Figure 4.7). Health facility data from all sites demonstrated weekly fluctuations in number of cases but no clear peak in transmission and no epidemic situation.

Figure 4.6 - Phase 1 weekly proportion of children absent from school, calculated from school attendance register (solid line, primary y-axis) and total confirmed malaria infections identified at the local health centre by routine passive surveillance (dashed line, secondary y-axis). Individual graphs are plotted for each of the six Phase 1 study sites.

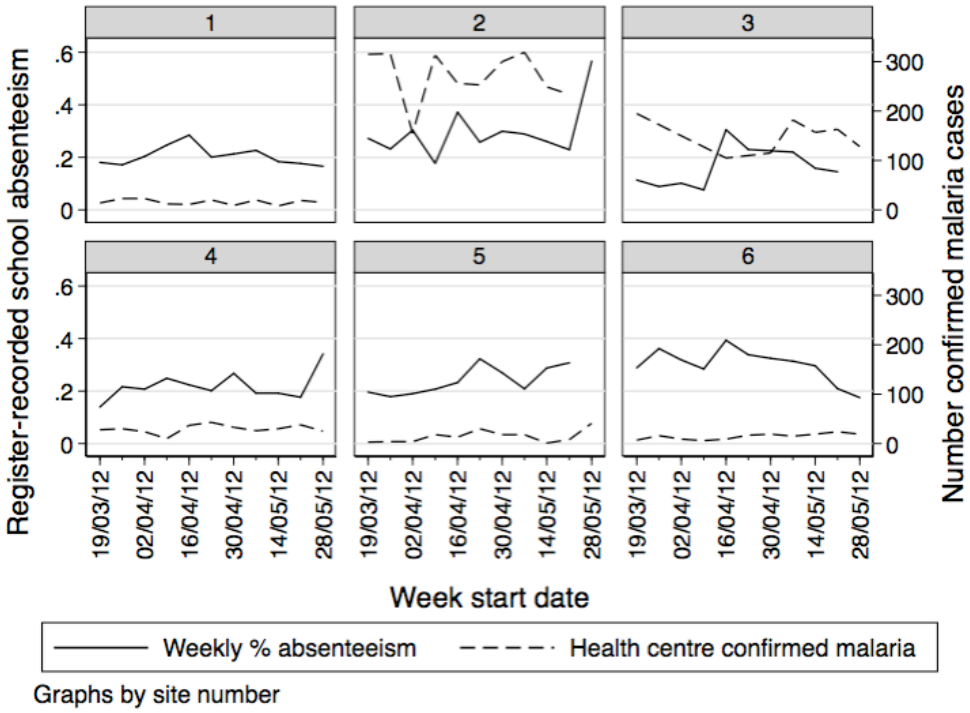
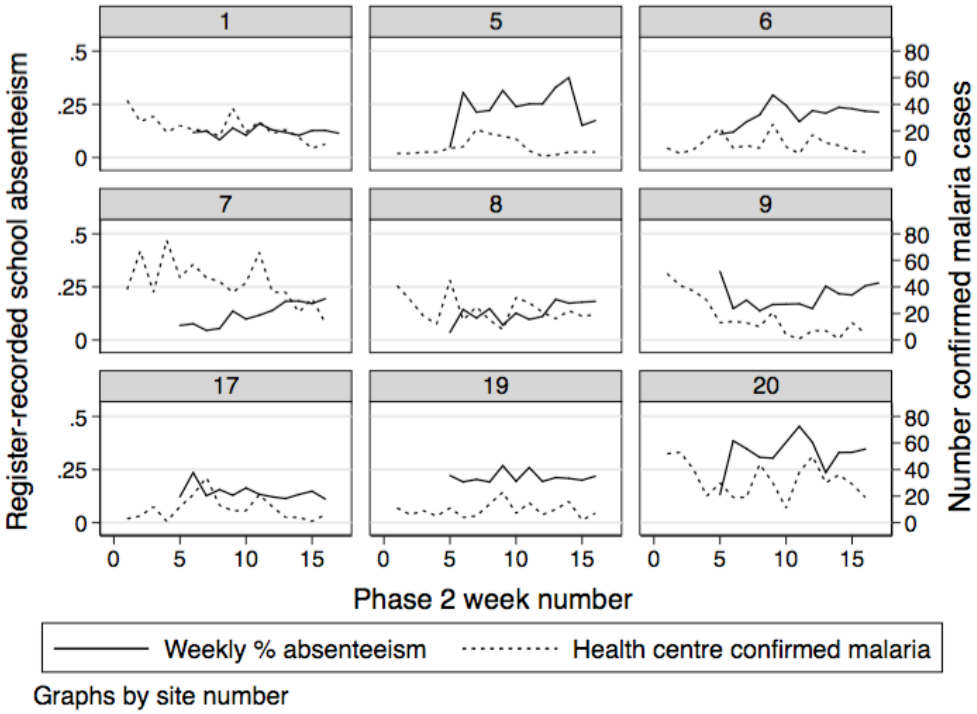


Figure 4.7 - Phase 2 weekly proportion of children absent from school, calculated by school staff from attendance registers (solid line, primary y-axis) and total confirmed malaria infections identified at the local health centre by routine passive surveillance (dotted line, secondary y-axis). Individual graphs are presented for each for the nine Phase 2 study sites in cluster B which reported data (one school failed to collect or report data).



Correlations between indicators collected in schools and health facility data were examined, but no strong correlations were seen between confirmed malaria at health facilities and the majority of the piloted indicators (child-reported fever, reported recent absence from school, absence from school due to illness, or teacher-summarised weekly absenteeism). There was evidence for an association between the proportion of child-days absent by week extracted from attendance registers collected for validation purposes and health facility total positive cases (p=0.002) or test positivity rate (p=0.028). No evidence was found for an association between confirmed malaria at health facilities and teacher-summarised absenteeism from cluster B sites (p=0.197).

This difference in association with health facility data using absenteeism data from different sources may be a result of differences in malaria transmission levels between sites, with different locations contributing data to the attendance registers collected for validation purposes from both clusters A and B, and the teacher-summarised data from cluster B sites only.

4.5 DISCUSSION

This study was the first application of a syndromic surveillance system based on school absenteeism to the context of malaria epidemic early warning in a resource-poor setting. The study was conducted in two phases: first exploring factors associated with enrolment of children in school, reasons for school absenteeism and risk factors for *Plasmodium* infection assessed by RDT, then piloting two school-based syndromic surveillance systems over the course of one semester and transmission season. Enrolment in school was lower than expected from available national statistics, and linked to a number of factors including household wealth, number of school-aged children in the household, distance of household from school and whether the head of household attended school. However, none of these factors were found to be associated with *Plasmodium* infection except wealth, which had an association in opposite directions for the two outcomes. Therefore, we hypothesise that non-enrolled children do not have different malaria risks compared to those who are enrolled. Levels of weekly absenteeism varied by site, but were consistent across grades within each site, as well as increasing over time as a result of drop out. During both the first and second phases of the study, no seasonal peaks in malaria were observed from either the health facility data or from RDT positivity in sampled individuals over the Phase 1 survey iterations. As a result, the majority of absences from school were reported to be not due to illness, and it was not possible to rigorously assess the reliability of school absenteeism as an indicator of

increasing malaria burden within communities. Findings of a qualitative evaluation of the school-based syndromic surveillance system, resulting from in-depth interviews with school staff, health extension workers and woreda health office staff will be presented in a forthcoming publication (Okello G, Kefyalew T, Batisso E, Mesele T, Ashton R, Brooker S: Teacher and health worker perceptions of a school-based syndromic surveillance system).

4.5.1 School enrolment in SNNPR

The normal age for first entry to primary education in SNNPRS is seven years, which was corroborated by our findings on reported school enrolment during community survey. National level data indicate that the net enrolment (i.e. ratio of primary school aged children who are enrolled in primary school to the total population of the official primary school age) in Ethiopia was 79% in 2012 [162], but enrolment was found to vary from 42% to 62% across sites, substantially lower than expected. Previous studies in Ethiopia have shown primary school enrolment to be more likely in female-headed households, and positively associated with educational level of the head of household, as well as household wealth [300]. Boys have been shown to be more likely to be enrolled than girls, and presence of younger children in the household has a negative impact on likely enrolment in school of older children [300]. Our data indicate that enrolment is associated with higher household wealth, the head of household having attended formal education (including incomplete primary education) and age of child. We found that the odds of enrolment were reduced with increasing number of school-aged children in the household, and increasing distance between household and school.

It was not possible to conduct a stratified analysis from the available community survey data to determine whether those children who are not enrolled in school have any increased risk of malaria, due to small number of *Plasmodium* infections identified during the course of the Phase 1 surveys. Multivariate modelling indicated that odds of RDT positivity in the

community survey decreased with increasing age, and ITN use during the previous night was strongly protective. In the school surveys, RDT positivity was associated with fever during the previous two weeks and household wealth (lower odds of infection in the poorest households). Girls were found to have lower odds of infection than boys. Therefore, there is potential for an overestimate of community malaria burden when using school-based platforms, considering the increased likelihood of enrolment for children from the richest households. There were no associations between distance of household from the school and odds of malaria, although it is acknowledged that microheterogeneity in malaria transmission may exist within a community [47,265,301], which is more complex than straight-line distance of household from school. Considering that universal enrolment and attendance at primary school has not yet been achieved in Ethiopia, it is possible that the sensitivity of a school-based syndromic surveillance system will be reduced, since the whole community will not be captured at a school-level platform. It will be necessary to further explore whether there are common socio-economic, age, gender, or geographical risk factors for enrolment or attendance at primary school and malaria risk.

4.5.2 School absenteeism

Absence in schools can be classified as temporary or permanent (i.e. dropout). The most common reasons reported in the current study for temporary absence from school were to assist with domestic chores in the home or on farmland, but the second most common reason for absence from school was illness. Additional, infrequently reported reasons included economic activities, attending market and lack of school materials (books, uniform etc). Food insecurity has also been shown to be a determinant of school absenteeism and attainment [302]. We expect that these reported reasons for absence from school are accurate, since *ad hoc* discussion with school staff during the survey period yielded similar feedback that children miss school to support their family with farming or domestic chores, and these reasons for absence are seasonal (higher during harvest and planting seasons. The

shift system which operates in many rural primary schools is recommended as an effective way to allow children to balance their schooling with work responsibilities of farming, herding or domestic chores [303]. Nevertheless, it remains likely that during periods of peak agricultural activity, children may be increasingly absent from school to support farming activities.

While absence due to illness may be an intuitively more sensitive indicator for malaria surveillance, all-cause absenteeism is a more reliable indicator to collect on a daily basis. When a child is absent it is not always possible to know the reason for that absence unless a sibling is present in the same class. It is not feasible on the first day of absence from school to either trace a sibling or neighbour who may know the reason for a child's absence, or to contact a parent or guardian at home. Therefore absence due to illness would likely have worse temporal resolution than all-cause absence, due to the need for classification of reason for absence when the child returns to school. It is not currently routine practice for schools in SNNPRS to record the reason for absence, although where it is known it may be noted in the attendance register. While teachers are expected to collect attendance data on a regular basis, levels of supervision of this activity by school director was variable, and consequently the level of completeness of attendance registered between classes and schools.

The current study collected school attendance data from different sources: from school attendance registers, reported by head of household during community surveys, reported by school-attending children during school surveys, and reported by school-attending children to their teacher as part of the syndromic surveillance pilot. Weekly absenteeism (as calculated from attendance registers) was found to be similar across grades, and absenteeism could therefore be monitored from any grade as part of future implementation of the surveillance system. Some differences were observed when comparing weekly absenteeism calculated by school staff and study team from attendance registers. This is likely due to individuals' interpretations of whether children who have dropped out from school should be counted as

absent or ignored. Excluding data from children who have dropped out is problematic, unless schools adhere to a common definition of how long a child must be continually absent for to be designated as permanent drop-out.

The Ministry of Education set a target drop-out rate for the academic year 2010/11 of 8%, but actual drop out rate was 13% [299]. At participating study sites, drop out was approximately 10% over the semester. Schools participating in the current study reported conducting community sensitisation before the start of the semester to increase enrolment, but also carrying out home visits when a child was absent from school for extended periods to try and understand why the child was persistently absent, to prevent them dropping out. While reasons for drop out were not specifically investigated in the current study, drop out can either be due to the long term challenges which also influence likelihood of enrolment, as well as by economic shocks (e.g. drought, crop failures, death and illness of family members), with these shocks influencing the proportion of a child's time which is dedicated to unpaid activities in the home or paid activities elsewhere [299]. Older children have been shown to be more likely to drop out of school in Ethiopia, and risk of drop out also increases with the number of children under five years in the household [304].

Syndromic surveillance systems in high-income countries generally use electronic data capture or web-based systems to collate reported and existing data for analysis [17,18,149,150]. Low- and middle-income countries are increasingly adopting such technology. A surveillance system for gastrointestinal and respiratory infectious diseases in China required daily submission of data using a web-based system [148]; while systems in Cambodia and India have used SMS reporting of data to a central level, where it is entered to a database, analysed and responses issued [138,146]. The surveillance system piloted in Ethiopia did make use of existing data, school attendance registers, but rather than develop a more complex system of data reporting to a central level for analysis and then response, the aim was to simply enable an alert to be passed from school to community health worker of

possible increases in illness in the community. This information would then act as a prompt for the health extension worker to assess their recent case data, or to conduct active surveillance in targeted areas of the kebele. While this system has low specificity, it builds upon existing links at community level between school and health extension workers, through the kebele committee's weekly meetings to discuss local issues. This low-tech, simple approach may prove more sustainable in the long-term than any surveillance system requiring data to be reported upwards for analysis and feedback. It also allows flexibility to respond to rumours and local opinion as well as the defined indicators. Health extension workers routinely spend a proportion of their time making home visits in the kebele, and it is credible that intelligence from the syndromic surveillance system may allow targeting of home visits at the sub-kebele level to areas of highest absenteeism.

4.5.3 Benefits and drawbacks of the piloted system

Of the two piloted syndromic surveillance systems, monitoring school absenteeism is a less time-intensive activity than weekly completion of symptom questionnaires by school-attending children, therefore, absenteeism is a more feasible indicator for long-term implementation. Absenteeism is routinely recorded by primary schools in Ethiopia, and generation of weekly summary absenteeism is a fast and simple addition to existing responsibilities. However it was apparent during the study implementation that schools do not have any standard procedures for recording absence: symbols for absent and present, follow-up of reasons for absence, defining period of absence for drop out, and summary data are not routinely generated. Additionally, class teachers usually keep their registers in their home, consequently, senior staff are often not able to not validate registers. For any future school absenteeism-based surveillance system, it is recommended to roll-out standard register formats and symbols for recording pupils' daily presence and absence, as well as regular checking and feedback on attendance register completion by senior school staff.

Differences in "normal" absenteeism rates between schools would likely remain due to systematic differences in populations across epidemic-prone areas. The usefulness of the system would be dependent on motivation of the school director and teachers to collect and assess absenteeism data, and report to health extension workers when increases occur. No thresholds would be assigned to schools, but it would be the responsibility of the school director to determine when absenteeism becomes unusual and to alert the health extension workers.

4.5.4 Future implications

In the current study, the lack of any malaria epidemic or strong seasonal peak in malaria transmission during the data collection period at the study sites resulted in inconclusive findings, preventing evaluation of the performance of a school-based syndromic surveillance system for malaria epidemic detection.

The key limitations of a surveillance system using school absenteeism are drop out during the academic year, and low rates of enrolment. In a resource-poor context, drop out is unlikely to be eliminated in the near future. Any system using school absenteeism should account for gradual increases in absenteeism over the semester. Implementing a system without fixed thresholds for alert generation, and relying on subjective identification of "unusual" increases by school staff is one approach to avoid bias due to drop out. Low school enrolment is a major limitation to the sensitivity of the piloted surveillance system, and further investigation is required to determine if individuals at higher risk of malaria, and those who may be the initial index cases of malaria in a community, are likely to be those families who are not able to enrol their children in school. However, it is expected that in the context of a malaria epidemic, the school-based system would have sufficient sensitivity to identify large increases in illness, and resultant school absenteeism in the community.

CHAPTER 5. COMPARATIVE EVALUATION OF STRATEGIES FOR MALARIA SURVEILLANCE AND MONITORING: HOW COMPLEMENTARY ARE CROSS-SECTIONAL SURVEYS, PREDICTIVE MODELS AND ROUTINE HEALTH FACILITY DATA?

5.1 OVERVIEW

Previous chapters have explored the use of cross-sectional school- and community-based surveys to describe the endemicity of malaria, as well as investigating alternative epidemic detection systems. While many tools for surveillance and monitoring of malaria are available, there have been few attempts to compare the indicators they generate and their usefulness for malaria control programmes. This chapter presents a comparative analysis of survey data using traditional and new malaria indicators with estimates of incidence using routinely recorded health facility data, and modelling estimates of malaria endemicity.

This chapter has been prepared for submission to the Malaria Journal, but at the date of thesis submission has not been submitted to the journal or subject to peer review. I led the conceptualisation and analysis presented in this chapter. Input to the analysis strategy and development of the scoring framework was provided by Simon Brooker.

5.2 BACKGROUND

In settings with low and unstable malaria transmission, data collected as part of monitoring, evaluation and surveillance activities have two key purposes. First is a need to monitor the implementation of malaria control activities and evaluate their impact on malaria burden

within the population, as well as their value for money. This process is critical to allow evidence-based design of intervention packages and targeting of resources. Evaluating changes in malaria transmission can be particularly challenging in settings with low and unstable malaria transmission, consequently, tools used by control programmes for this purpose must be carefully considered to ensure they are sufficiently sensitive and representative of the true extent of transmission [267]. The second key purpose to collect these data is for surveillance: the continuous collection and analysis of data in order to generate timely alerts and responses to increases in malaria that may develop into epidemics. The most important feature of surveillance data is its temporal resolution, since reliable data must be generated, analysed and acted upon within a sufficiently narrow timeframe to enable nascent epidemics to be identified.

Various tools are used for monitoring, evaluation and surveillance, yet in a context of declining malaria transmission and limited resources, there is a need to reconsider the relative merits and capacity of different tools and indicators to meet the needs of a malaria control programme. National malaria control programmes must therefore tailor available tools and balance their investments in different monitoring, evaluation and surveillance strategies.

As I highlight in Chapter 1, the specific challenges of malariometric surveys in low transmission settings were first emphasized during the Global Malaria Eradication Programme [305], primarily the reduced sensitivity of malariometric indices to measure changes in transmission at relatively low levels. Incorporating malariometric indicators into large-scale population surveys such as the Malaria Indicator Survey (MIS) continues to be popular and widely used across a range of transmission settings [104], but as shown in Chapter 3, parasitological surveys may not capture the small scale spatial heterogeneity in infection in low transmission surveys. As described in Chapter 4, inclusion of serological indicators of exposure to malaria and spatial modelling of data are way to adapt

malariometric surveys, which can overcome some limitations of parasitological surveys in low transmissions settings. Molecular diagnostic tools are also valuable in such parasitological surveys due to their improved sensitivity to detect *Plasmodium* parasites [193,194]. Definitions of malaria elimination strategies and indicators that can robustly measure malaria endemicity remain a subject of ongoing research and debate [9,306,307], and it is likely that a combination of malariometric surveys and surveillance systems are needed. Key to effective surveillance systems is timeliness of data reporting and responsiveness to these data, particularly in areas vulnerable to epidemics. Chapter 4 presented a syndromic surveillance system using school attendance as a surrogate indicator of malaria burden within a community, to explore the utility of syndromic surveillance to complement data collection and reporting at health posts, particularly to support timely identification of malaria epidemics. Other innovations piloted elsewhere to explore improved surveillance and monitoring strategies include a "rolling" MIS at district-level [107], and the use of sentinel health facilities to generate gold standard surveillance data [308].

The aim of this chapter is to review the main malaria monitoring and surveillance approaches, and to apply a set of performance parameters to evaluate the suitability of these monitoring or surveillance platforms and indicators in a setting of low and unstable malaria transmission. I then compare data collected by a variety of monitoring and surveillance methods to describe *Plasmodium* infection and exposure levels in Oromia Regional State, and describe the correlation between these estimates of malaria burden at spatially congruent locations. Finally, I provide recommendations as to the most appropriate malaria surveillance tools for future implementation in settings such as Ethiopia.

5.3 METHODS

5.3.1 Data sources

This chapter utilizes multiple data sources that collect data on a variety of malariometric indicators from cross-sectional surveys, passively reported malaria cases at health facilities and statistical modelling predictions. A summary of the data used is provided in Table 5.1.

5.3.1.1 Routine health facility data

Monthly malaria outpatient data from all public health facilities from health post to referral hospital were requested from woreda health offices in Oromia Regional State, for the period from June 2006 to January 2011. Health facilities routinely record the number of clinical (suspected) malaria patients and number tested and confirmed by parasitological diagnosis, including breakdown by species. In addition, the total outpatient attendance was requested from all health facilities as a denominator for all-cause health service access over time. Data were entered into an EpiData template (The EpiData Association, Odense, Denmark), which included range and logic checks to reduce potential for entry errors. Data were exported into Stata 12.0 (Stata Corporation, College Station, TX, USA) for cleaning and analysis.

Monthly data were converted from Ethiopian calendar to Gregorian calendar, approximating the month of *Tir* to January. Data reported for *Pagume*, the 13th month of the Ethiopian calendar, have been included in totals for the prior month (equivalent to August). New woredas formed after the 2007 national census are included under their former woreda boundaries [234]. Missing data were estimated by linear interpolation for instances where fewer than three months' data were missing (4.8% of all data). Health facilities with large quantities of missing data or significant inconsistencies and logic errors were excluded from the dataset (4.1% of all health facilities where any data were received). Health facility data were pooled to woreda-level annual data.

Woreda-level population estimates were extracted from the 2007 census [234], and reported annual population growth rate in Oromia of 2.9% used to estimate woreda populations in 2008, 2009 and 2010. Annual population incidence for all ages was estimated by woreda for 2007, 2008, 2009 and 2010.

5.3.1.2 School-based survey data

A large cross-sectional survey was conducted at 197 government primary schools in Oromia Regional State in 2009. A two-stage sampling design was used, selecting schools using probability proportional to size, and then randomly selecting a fixed number of children (100 plus ten reserves) from each school. Full details of school selection and sample size calculations are presented in Chapter 2. From each school, equal proportions of boys and girls were randomly selected from all present in grades two to six on the survey day. Children provided a single finger-prick blood sample for preparation of a thick and thin blood film, haemoglobin measurement (HemoCue Ltd, Angelholm, Sweden) and collection of blood spots on filter paper (Whatman 3MM, Whatman, Maidstone, UK). School location was recorded using handheld global positioning satellite receiver (eTREX, Garmin International, KS, USA). Data used from these school-based surveys in the current analysis was the prevalence of *Plasmodium* infection by site, including speciation, as determined by microscopy.

5.3.1.3 Serological indicators of exposure to malaria

Dried blood spots collected from school surveys in 2009 were used for subsequent serological analysis by enzyme-linked immunosorbent assay (ELISA) to investigate the population exposure to malaria. Schools for serological analysis were chosen purposively from 197 surveyed: 20 schools with highest prevalence of infection by microscopy, 20 schools with highest anaemia levels, and random selection of remaining schools. Full details of serological methods and results by antigen are listed in Chapter 3. Estimates of school seroprevalence

against *P. falciparum* were available from 62 schools, while *P. vivax* seroprevalence estimates were available from 71 schools.

In addition to measured seroprevalence at schools, multivariate Bayesian binomial models were developed and used to predict *P. falciparum* and *P. vivax* seroprevalence at 5km resolution across Oromia Regional State. The optimal model for *P. falciparum* included a school-level random effect and a geostatistical random effect (using isotropic, stationary exponential decay function) as well as selected environmental variables. The *P. vivax* predictive model included a school random effect and a geostatistical random effect, but no environmental variables. Full modelling methods are detailed in Chapter 3.

5.3.1.4 Malaria Indicator Survey 2011

A national MIS was conducted in Ethiopia from September to December 2011 [228], following Roll Back Malaria guidelines [104]. A stratified two-stage sample design was used, with census enumeration areas (EAs; administrative level 5, usually 150-200 households) as primary sampling unit, and households within EAs as secondary sampling units. The aim was to generate robust national level estimates of intervention coverage for areas <2,500 metres elevation, and rural- and urban-level information for <2000m, sample size calculation therefore was based upon indoor residual spray coverage. Selected administrative regions, including Oromia, were over-sampled to allow generation of sub-national estimates. 440 EAs were sampled nationally during MIS 2011, with 162 EAs in Oromia. From each EA, 25 households were selected for participation in the survey, following mapping of all households in the EA using personal digital assistant (PDA) with built-in geographic positioning satellite receiver (GPS) capacity. From each household, all children under five years of age provided a blood sample for preparation of thick and thin blood films, RDT (CareStart HRP2-panLDH) and haemoglobin measurement. Individuals of all ages from every fourth household were asked to provide blood samples for blood slide and RDT. A household questionnaire was used to

collect information on malaria knowledge and intervention use, and a women's questionnaire targeted to one woman aged 15-49 in each household.

MIS data were available from 156 EAs sampled in Oromia, encompassing 3793 households. Blood samples were taken from 4887 individuals. 51% of blood samples were taken from children under five years of age. Serological analysis of MIS samples is planned against one *P. falciparum* antigen and one *P. vivax* antigen, but had not been conducted at the time of analysis.

5.3.1.5 Malaria Atlas Project modelled parasite rate

Posterior mean predicted *P. falciparum* and *P. vivax* parasite rates generated by the Malaria Atlas Project (MAP) were downloaded [292] and imported to ArcMAP 10.2 (Environmental Systems Research Institute Inc., Redlands CA, USA). Age-standardised *P. falciparum* parasite rate ($Pf\text{PR}_{2\text{-}10}$) describes the estimated proportion of children aged from two to ten years who are infected with *P. falciparum*, averaged over the 12-month period of 2010 [20]. Estimates of $Pf\text{PR}_{2\text{-}10}$ were available at 1km^2 resolution. The modelled *P. vivax* parasite rate was age standardised to the 1-99 years age range ($Pv\text{PR}_{1\text{-}99}$), and similar to $Pf\text{PR}_{2\text{-}10}$, describes the estimated proportion of the population who are infected with *P. vivax* at any time, averaged over 2010 [21]. $Pv\text{PR}_{1\text{-}99}$ was available at 5km^2 resolution. In addition, surfaces describing the estimated transmission limits of both *P. falciparum* and *P. vivax* were downloaded [292]. Transmission limits were classified using a combination of medical intelligence and temperature and aridity masks into transmission-free, unstable transmission (annual parasite incidence [API] <0.1%) and stable transmission (API ≥ 0.1%) [20,21]. Data collected during parasitological surveys in 2009 were not included in MAP databases, and therefore did not contribute to the MAP modelled $\text{PfPR}_{2\text{-}10}$ or $\text{PvPR}_{1\text{-}99}$ estimates.

Table 5.1 - Key descriptive characteristics of the various surveillance data compiled for the current study

	Woreda annual incidence	**Malaria Indicator Survey RDT data**	**School survey microscopy data**	**School survey serology data**	**Modelled seroprevalence**	**Malaria Atlas Project modelled PR**
Data collection method	Patients attending health facilities	Cross-sectional household survey	Cross-sectional school survey	Cross-sectional school survey	Model predictions, using school survey serology data	Model predictions, using historical survey data at continental scale
Denominator	Population in health facility catchment area for all woreda facilities	All children <5 in sampled households, and all individuals in every 4th household	All children present at school on the survey day	All children present in school on the survey day	Total school-aged population resident in location	Total population resident at location
Age range of sampled population	All ages	All ages, but children under 5 over-sampled	5-16 years	5-16 years	5-16 years	Age-standardised: 2-10 years *P. falciparum* and 1-99 years for *P. vivax*
Indicator	Mean annual incidence confirmed malaria, per 1000 people	Proportion of sampled individuals in each cluster with positive RDT	Proportion of individuals in each school microscopy positive	Proportion of individuals in each school seropositive	Estimated proportion of 5-16 years olds seropositive	Age-standardised parasite rate
Diagnostic tool	Microscopy and rapid diagnostic test, as normally used at facility	HRP2-panLDH rapid diagnostic test	Microscopy examination of thick & thin blood films	ELISA against *P. falciparum* and *P. vivax* antigens	ELISA against *P. falciparum* and *P. vivax* antigens	N/A
Number clusters sampled	256 woredas	156 enumeration areas in Oromia	197 schools	75 schools	Model predictions with 5km resolution	*P. falciparum* has 1km pixel resolution, *P. vivax* pixel resolution 5km
Number individuals sampled per cluster	Mean 17,241 annual outpatient visits recorded in a woreda	Mean 31 individuals (all ages) and 16 children <5 sampled per EA	Mean 106 children sampled per school	Mean 94 children sampled per school	N/A	N/A
Range of recorded indicator by cluster	0 - 200 confirmed malaria cases / 1000 people / year	0 - 48.6% (any species, all ages)	0 - 14.5% *Plasmodium* prevalence	*P. falciparum*: 0-50% *P. vivax*: 0-89%	Range of woreda mean *P. falciparum*: 1-95%, *P. vivax*: 1-49%	Woreda mean *Pf*PR: 0.7-7.4%, woreda mean *Pv*PR: 0.6-2.2%
Mean cluster indicator value	12 confirmed malaria cases / 1000 people / year	1.3% (any species, all ages)	0.5% *Plasmodium* prevalence	*P. falciparum*: 11.0% *P. vivax*: 10.9%	*P. falciparum*: 8.6% *P. vivax*: 11.6%	*Pf*PR: 2.8% *Pv*PR: 1.2%

5.3.2 Matching spatially congruent data

Descriptive characteristics of all surveillance data utilised are presented in Table 5.1. Sampled school and EA location were recorded during the school survey and MIS implementation, respectively, and point data plotted in ArcMAP. Survey locations were matched to their administrative woreda, using 2007 census administrative classifications, in order to link woreda-level health facility indicators with survey data. To reduce possible noise in the raster values extracted to survey point locations, the mean of a three-by-three pixel array with the centred at the survey location was calculated for both the *Pf*PR_{2-10} and *Pv*PR_{1-99} rasters. Where a survey location fell in an area defined as beyond the limits of transmission (API=0) by the MAP predictions, a parasite rate of zero was assigned.

P. falciparum and *P. vivax* seroprevalence model estimates at 5 km^2 resolution were summarised by woreda, with summary statistics including minimum, maximum, mean and standard deviation of pixel values within the woreda calculated. The same method was used to generate summary statistics by woreda for *Pf*PR_{2-10} and *Pv*PR_{1-99} from MAP model raster data. Seroprevalence model and MAP model summary statistics by woreda were linked with woreda-level health facility data.

In addition to data cleaning processes described previously for survey and health facility data, further exclusion criteria were applied to woreda-summary health facility data to remove unreliable data. Any woreda with fewer than ten health facilities in total reported data (n=6) was excluded from further analysis, as well woredas with logic errors in summary data (total outpatient attendance < total recorded clinical malaria cases, total outpatient attendance < total confirmed malaria, total tested by microscopy or RDT < total confirmed malaria).

The number of spatially-matched data points for each pair-wise combination of indicators is presented in Table 5.2. Due to cluster randomisation sampling for the school survey and MIS, there are instances of more than one cluster data point existing within a single woreda, in which instance cluster data were matched for incidence using a many-to-one strategy.

Table 5.2 - Two-way matched data available from the five surveillance data. Data are matched using woreda location, therefore number of cross-sectional survey clusters and encompassing woredas are presented for each combination of data. It was possible for more than one cross-sectional survey cluster to be present in a single woreda.

	Woreda annual incidence	**Malaria Indicator Survey RDT**	**School survey microscopy**	**School survey serology**	**Modelled sero-prevalence**	**Malaria Atlas Project modelled PR**
Malaria Atlas Project modelled PR	218 woredas	136 EAs in 116 woredas	174 schools in 98 woredas	66 schools in 51 woredas	220 woredas	256 woredas
Modelled sero-prevalence	219 woredas	137 EAs in 117 woredas	174 schools in 98 woredas	66 schools in 51 woredas	264 woredas	
School survey serology	62 schools in 50 woredas	50 schools & 42 EAs in 38 woredas	75 schools in 58 woredas	75 schools		
School survey microscopy	165 schools in 97 woredas	127 schools & 78 EAs in 67 woredas	197 schools			
Malaria Indicator Survey RDT	132 EAs in 116 woredas	156 EAs				
Woreda annual incidence	230 woredas					

5.3.3 Scoring framework development

To qualitatively assess some core characteristics of the various surveillance data, a scoring framework was applied to seven parameters of the surveillance systems and the data they generate. For each parameter and dataset, a description was provided and score assigned by the author between one and three. An expert opinion scoring workshop is planned, following which, scores will be reviewed and published. The scoring system was defined as: 1 = poor; 2 = neutral; 3 = good. Directionality of parameter scoring was defined as higher parameter value conferring higher score, except for cost and complexity, where systems with lower cost or complexity received higher scores.

5.3.4 Data analysis

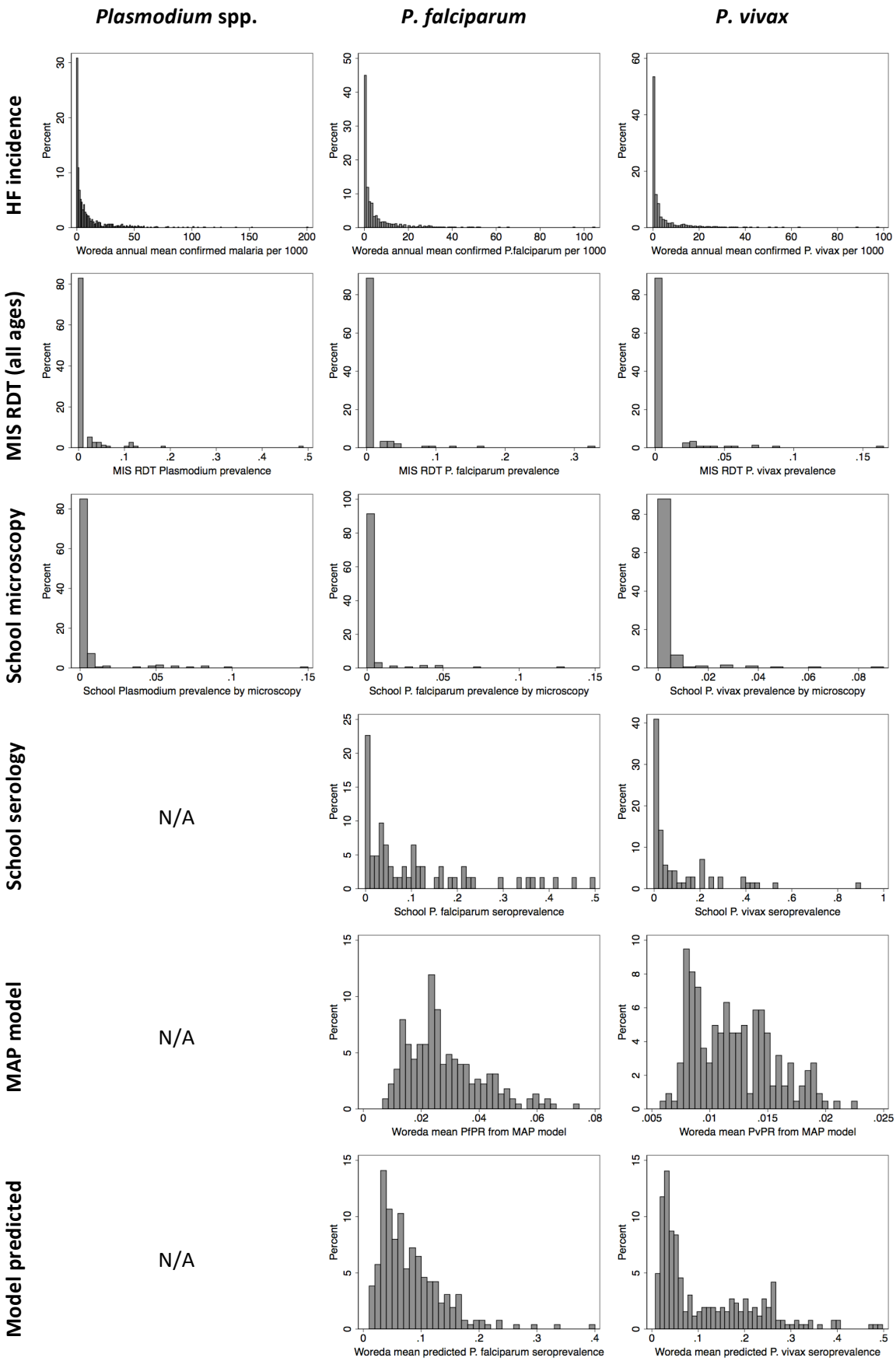
Data included in this chapter include indicators used for periodic monitoring (survey and modelling data) as well as those used for surveillance (health facility data). Considering the different objectives of surveillance and monitoring, these data are not directly comparable, and there is no universal gold standard against which to measure the accuracy of each indicator.

Associations between indicators at survey locations, as well as mean indicator values by woreda, were explored by generating scatter plots, calculating correlation and simple linear regression between the continuous measures.

Woreda annual mean incidence per 1000 people was used as the basis for endemicity classification. While there is no gold standard indicator, woreda incidence data were used to classify endemicity since these data are the most widely available. Histograms were generated from all surveillance data for each of *P. falciparum*, *P. vivax* and any *Plasmodium* infection to describe the distribution of the various indicators (Figure 5.1).

Woreda mean incidence was classified into three categories: moderate unstable transmission, low unstable transmission and very low unstable transmission. Classification was informed by the distribution of the data, as well as the range of malaria transmission settings anticipated over Oromia Regional State, to ensure that sufficient data would be retained within each of the categories. Endemicity category cut-offs were identical for each of the species and for any *Plasmodium* infection indicators. Cut-off values used for the categories were: very low unstable transmission or malaria-free being <1 confirmed cases per 1000 people; low unstable transmission being 1-10 confirmed cases per 1000 people; moderate unstable transmission being >10 confirmed cases per 1000 people. To explore the distribution of other indicators according to these endemicity categories, and therefore the level of agreement between surveillance data, box-plots were generated.

Figure 5.1 (following page) - Histograms displaying distribution of key indicators from the various surveillance data. Axes and bin width have been adjusted for each variable to optimise display of the available data. Histogram are presented for woreda annual parasite incidence per 1000 people from health facility data, MIS RDT cluster prevalence and school survey microscopy prevalence by *P. falciparum*, *P. vivax* and a summary any *Plasmodium* infection category. In addition, school *P. falciparum* and *P. vivax* seroprevalence, woreda mean Malaria Atlas Project modelled *Pf*PR and *Pv*PR, and woreda mean of modelled *P. falciparum* and *P. vivax* seroprevalence


Plasmodium spp.
P. falciparum
P. vivax
HF incidence
Percent
Woreda annual mean confirmed malaria per 1000
Woreda annual mean confirmed P.falciparum per 1000
Woreda annual mean confirmed P. vivax per 1000
MIS RDT (all ages)
MIS RDT Plasmodium prevalence
MIS RDT P. falciparum prevalence
MIS RDT P. vivax prevalence
School microscopy
School Plasmodium prevalence by microscopy
School P. falciparum prevalence by microscopy
School P. vivax prevalence by microscopy
School serology
N/A
School P. falciparum seroprevalence
School P. vivax seroprevalence
MAP model
N/A
Woreda mean PfPR from MAP model
Woreda mean PvPR from MAP model
Model predicted
N/A
Woreda mean predicted P. falciparum seroprevalence
Woreda mean predicted P. vivax seroprevalence

5.4 RESULTS

5.4.1 Parameters of surveillance system performance

Considering the different need of ongoing surveillance compared to periodic monitoring, it was decided to present scoring frameworks separately for parameters important for monitoring and for surveillance, respectively. Parameters used to score systems for their utility in programme monitoring and evaluation were sensitivity, frequency of data collection, time for feedback to peripheral levels, spatial resolution, complexity and cost (Table 5.3). The parameters crucial to identification of an optimal surveillance system were temporal resolution and spatial resolution.

5.4.1.1 Sensitivity

In this application, sensitivity refers to the ability of the surveillance system indicator to accurately represent the level of malaria within the underlying population sampled. The use of microscopy and rapid diagnostic test in cross-sectional surveys has moderate sensitivity to identify infections in the individuals randomly selected for testing, although it is likely that some low density infections are missed by both tools. These cross-sectional surveys are not powered to give representative cluster-level estimates of prevalence of malaria in very low transmission settings. In contrast, seroprevalence gives a more accurate representation of the malaria situation in a location due to the wider temporal reference period for the diagnostic indicator. Similarly, woreda annual incidence is generally representative of malaria burden within the catchment areas of the health facilities, provided that access to health services is high and diagnostic services at the facilities are quality assured. Indicators of uncertainty are available for modelled parasite rate, therefore it is possible to estimate the sensitivity of these data across the spatial extent of interest.

5.4.1.2 Frequency of data collection

All of the survey and modelling data presented are collected and re-analysed every three to five years, with the exception of woreda annual incidence. It would also be possible to generate monthly incidence using projected population estimates and summary monthly data from the woreda health facilities. Weekly data are available from health facility as part of the Public Health Emergency System, but the indicator would be crude total cases rather than population incidence.

5.4.1.3 Time for feedback to peripheral level

Effective surveillance systems include a feedback component, whereby findings from submitted data are reported back to peripheral levels to enable change, as well as providing accountability and motivation for ongoing submission of data. Cross sectional MIS and school surveys operate with a minimum six to 12 month period for analysis of data to take place, and for feedback on the findings and action required to be provided to malaria control programme personnel. Limited feedback is expected from modelled parasite rate data, since models are produced at global level, however additional data may be submitted to refine future models. Feedback is a core aspect of the routine health facility data reporting system, however in practice, limited feedback on submitted HMIS data is provided to peripheral health facilities and health workers on the quality and impact of their data.

5.4.1.4 Complexity

The complexity of a surveillance system will influence its future sustainability, as a result of the level of specialist skills and equipment required to operate the system and generate data. Modelled parasite rate has a high level of complexity to develop the model predictions, but maps are available free to download, therefore, only Internet access and an understanding of the indicator is needed for this tool to be of use. Some comprehension of defined uncertainty around the model predictions is beneficial to aid interpretation of model outputs. Cross-

sectional surveys have moderate to high complexity, primarily related to the skills required to select sites for sampling without bias, as well as data handling and analysis, particularly where weighting is required to generate indicator estimates. Routine health facility data systems have low complexity, and although understanding of key epidemiological concepts and data handling skills are required, there is often capacity within countries for these tasks, but it may not be present at the health cluster or woreda level.

5.4.1.5 Cost

The cost of a surveillance system influences its sustainability, since tools with high cost and complexity are generally only feasible when substantial additional resources are available beyond the standard levels of government funding. Routine health facility data collation is a low cost strategy, and encompassed within the usual responsibilities of health system staff. Use of modelled parasite rate data has a low cost to the end user, since maps are available for free. Of the cross-sectional survey methods presented, school surveys are estimated to be lower cost than the Malaria Indicator Survey. This is likely a result of the reduced time required at each cluster site to complete sampling; for school surveys only one day is needed while household surveys often require one day for mapping households, then up to two days to sample individuals. Serological analysis does increase the cost of school surveys, however the cost per sample analysed by ELISA is approximately one US Dollar, broadly equivalent to the cost of a rapid diagnostic test.

5.4.1.6 Spatial resolution

Malaria Indicator Survey findings have poor spatial resolution, being powered to generate reliable indicator estimates at national or regional level. Similarly, microscopy has poor spatial resolution from school surveys as a result of the moderate sensitivity of the indicator and low parasite prevalence found during the surveys. Serological indicators have improved spatial resolution, and modelling results using serology data were able to generate 5km

resolution predictions of seroprevalence with good reliability. Modelled parasite rate also has good spatial resolution, generating predictions at 5km for *P. vivax* and 1km for *P. falciparum*. Health facility data analysed here to generate woreda annual incidence is limited to a woreda-level resolution, but it may be possible to generate data at the health cluster level (one health centre and five satellite health posts), or even individual health facility level.

5.4.1.7 Temporal resolution and responsiveness to epidemics

Temporal resolution is coarse for all surveillance tools evaluated in this chapter except health facility data at woreda level, since most of the presented indicators are only collected periodically. Modelled parasite rate data has the largest temporal resolution, since the models are developed using historical data and may not reflect recent changes in malaria transmission. Serology data has a temporal resolution of months or years due to the persistence of antibodies in the population, however selective sampling and serological testing of young children can restrict the temporal reference period for serology data in population surveys. Use of RDT and microscopy as indicators in the MIS and school survey, respectively, have a fine temporal resolution and reflect presence of *Plasmodium* in sampled individuals at the time of the survey. However RDT and microscopy in cross sectional surveys would only coincidentally identify epidemics if a site with an epidemic happened to be randomly selected for inclusion in the cross sectional survey.

Table 5.3 - Description of parameters of surveillance system performance, leading to development of scoring for each parameter

	Woreda annual incidence	**Malaria Indicator Survey RDT data**	**School survey microscopy data**	**School survey serology data**	**Modelled seroprevalence**	**Malaria Atlas Project modelled parasite rate**
Sensitivity	Biased by population health-seeking behaviour and quality of routine diagnostic services at health facilities	Reasonable sensitivity at individual level, poor sensitivity to describe transmission intensity at population level	Reasonable sensitivity at individual level, poor sensitivity to describe transmission intensity at population level	Captures exposure across previous transmission seasons, representative of transmission intensity at population level	Estimates exposure across previous transmission seasons, representative of transmission intensity at population level	May not represent current PR if rapid changes in transmission intensity have occurred
Frequency of data collection	Monthly data, reported monthly or quarterly. Weekly malaria data may be accessible in the future	Periodic, repeated every 3-5 years	One-off activity, could be repeated every 3-5 years	One-off activity, could be repeated every 3-5 years	One-off activity, models could be updated as additional serology data available	Models updated approximately every 3-5 years
Temporal resolution: responsiveness to epidemics	Moderate. Weekly IDSR reporting designed to be responsive to epidemics, monthly HMIS data indicates longer term trends	Limited. Not suitable for epidemic identification unless an epidemic happens to be occurring at a location randomly selected to participate	Limited. Not suitable for epidemic identification since samples analysed after survey completion	Limited. Not suitable for epidemic identification since samples analysed after survey completion	Limited. Not suitable for epidemic identification.	Limited. Not suitable for epidemic identification
Time for feedback to peripheral level	Limited feedback in practice from central level following analysis of data	6-12 months from data collection to sample analysis and reporting of findings to local health authorities	6-12 months from data collection to sample analysis and reporting of findings to local health authorities	6-12 months from data collection to sample analysis and reporting of findings to local health authorities	12-18 months from data collection to reporting modelling findings to local health authorities	No feedback expected, other than submission of additional data for model refinements
Spatial resolution of data	Currently woreda-level, could be adapted to health cluster (one health centre & 5 satellite health posts	Surveys powered to yield national and sub-national level estimates	Poor resolution, 100 samples per site insufficient power where microscopy prevalence is very low	Moderate resolution, reliable estimates at cluster-level, models can extrapolate and predict beyond sampled locations	Predictions at $5km^2$ resolution for both *P. falciparum* and *P. vivax*	Prediction at $1km^2$ (*P. falciparum*) or $5km^2$ (*P. vivax*) scale

	Woreda annual incidence	Malaria Indicator Survey RDT data	School survey microscopy data	School survey serology data	Modelled seroprevalence	Malaria Atlas Project modelled parasite rate
Complexity	Low. Collation of monthly data is routine, simple calculation to generate incidence.	Moderate. Standard tools available from RBM, but countries usually require technical support in sampling design and data analysis	Moderate. National control programme staff may require technical support with two-stage sampling using PPS, and data analysis	Moderate. Two-stage sampling using PPS. Laboratory analysis possible at national level, but external support may be required for data processing and analysis.	High complexity for data collation and model generation, but maps shared for health authorities' use	High complexity for data collation and model generation, but national maps freely available for use
Cost	Low. No incentives paid for data submission, routine activity for salaried staff	High. Total national survey including supplies, sensitisation, implementation, and technical support ~$1.1 million. Proportion for Oromia ~$400,000	Moderate. Total survey including supplies, implementation and technical support ~$230,000	Moderate. Total costs for survey and laboratory analysis ~$280,000	Moderate. Specialist statistical skills required in addition to ~$280,000 for surveys and laboratory analysis	Maps available free. Costs of data extraction and modelling not available.

5.4.2 Scoring framework result

Considering the different needs of periodic monitoring and ongoing surveillance systems, the parameters used to score each of the evaluated systems were split into those prioritised for surveillance and those important for monitoring. Therefore each system has two summary scores, with higher scores indicating better performance against the parameter.

For periodic monitoring (Table 5.4), woreda annual incidence data had the highest assigned score, while serological indicators from school survey had the next highest score, followed by modelled seroprevalence and parasite rate. Microscopy and RDT indicators from cross-sectional surveys received the lowest scores.

Table 5.4 - Scoring framework for programme monitoring and evaluation. Scores assigned by the author, justifications and description listed in Table 5.3. Scoring system: 1=poor, 2=neutral, 3=good. Higher parameter levels are judged to be better for all variables except complexity and cost, where high scores are assigned for the lowest parameter levels.

	Woreda annual incidence	**Malaria Indicator Survey RDT**	**School survey microscopy**	**School survey serology**	**Modelled sero-prevalence**	**Malaria Atlas Project PR**
Sensitivity	3	1	1	3	3	2
Frequency of data collection	3	1	1	1	1	1
Time for feedback to peripheral level	2	2	2	2	1	1
Spatial resolution of data	2	1	1	2	3	3
Complexity	3	2	2	2	1	1
Cost	3	1	2	2	2	3
Total score (of 18)	*16*	*8*	*9*	*12*	*11*	*11*

For surveillance (Table 5.5), routine data from health facilities again received the highest score as a result of the potential for fine temporal and spatial resolution of these data. The poor temporal resolution of all other data collection platforms assessed limits their utility for ongoing surveillance, regardless of spatial resolution.

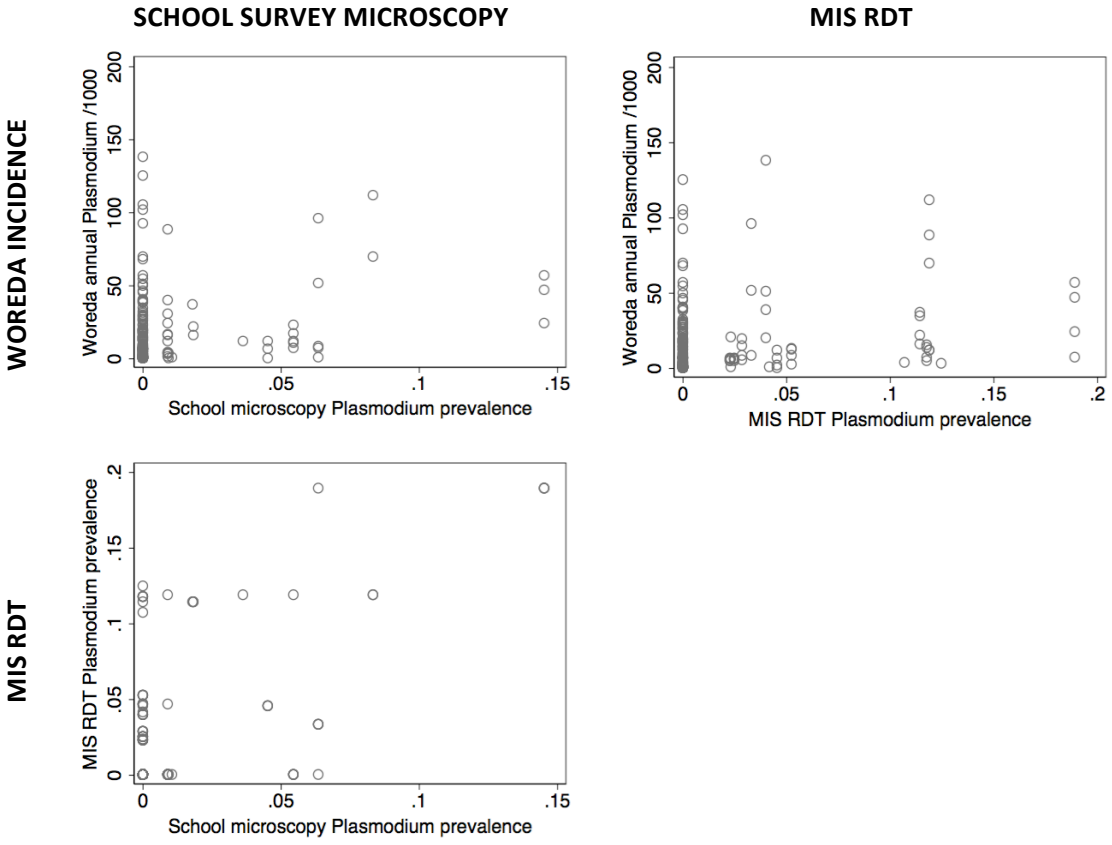
Table 5.5 - Scoring framework for surveillance. Scores assigned by the author, justifications and description listed in Table 5.3. Scoring system: 1=poor, 2=neutral, 3=good.

	Routine health facility data	**Malaria Indicator Survey RDT**	**School survey microscopy**	**School survey serology**	**Modelled sero-prevalence**	**Malaria Atlas Project PR**
Temporal resolution: responsiveness to epidemics	3	1	1	1	1	1
Spatial resolution of data	3	1	1	2	3	3
Total score (of 6)	*6*	*2*	*2*	*3*	*4*	*4*

5.4.3 Association between surveillance indicators

Scatter plots comparing all *Plasmodium* indicators, *P. falciparum* indicators and *P. vivax* indicators are presented in Figures 5.2 to 5.4, respectively.

Figure 5.2 - Scatter plots comparing indicators which describe *Plasmodium* infection, from routine health facility data summarised as annual incidence, Malaria Indicator Survey RDT positivity and school survey microscopy positivity



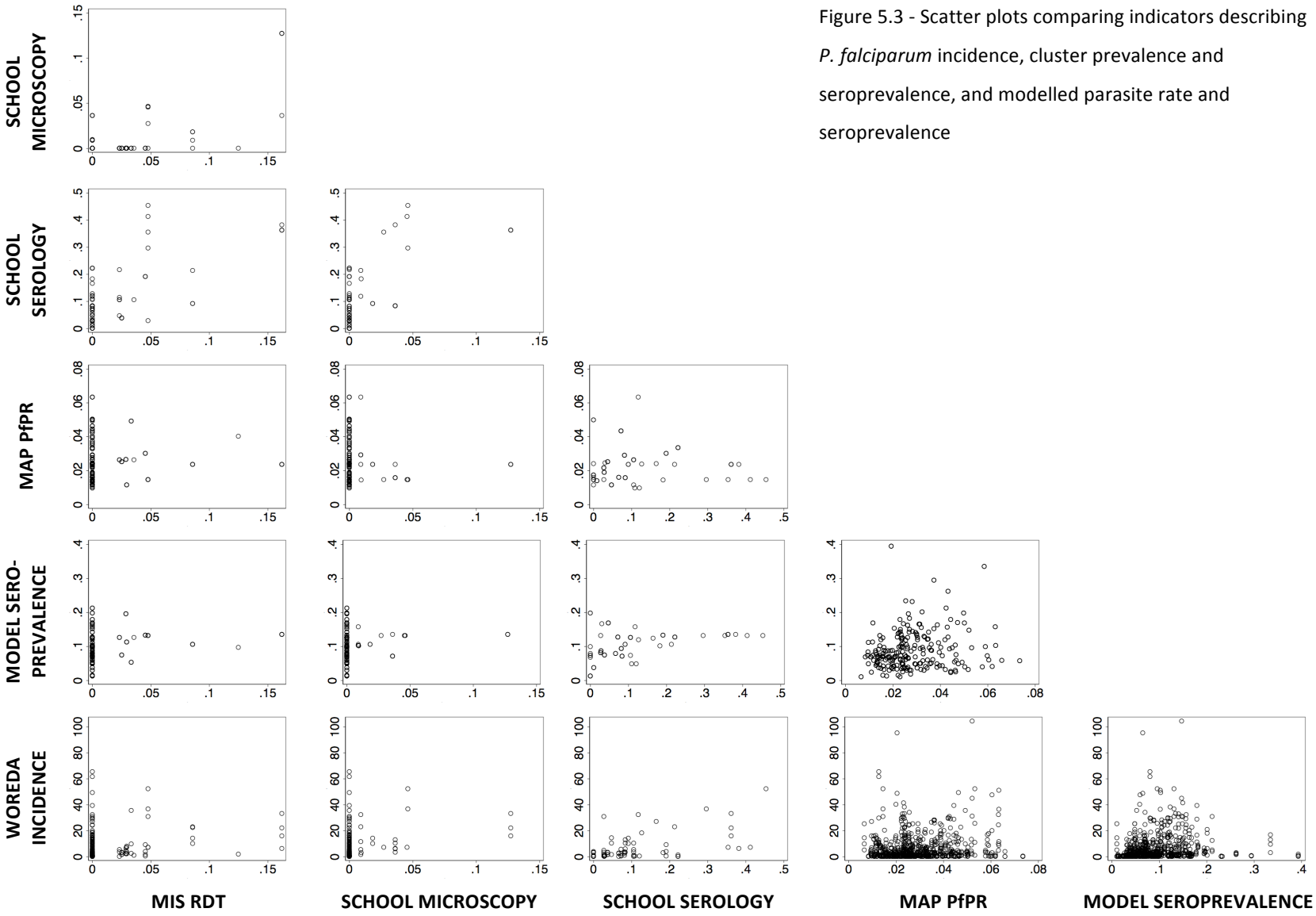


Figure 5.3 - Scatter plots comparing indicators describing *P. falciparum* incidence, cluster prevalence and seroprevalence, and modelled parasite rate and seroprevalence

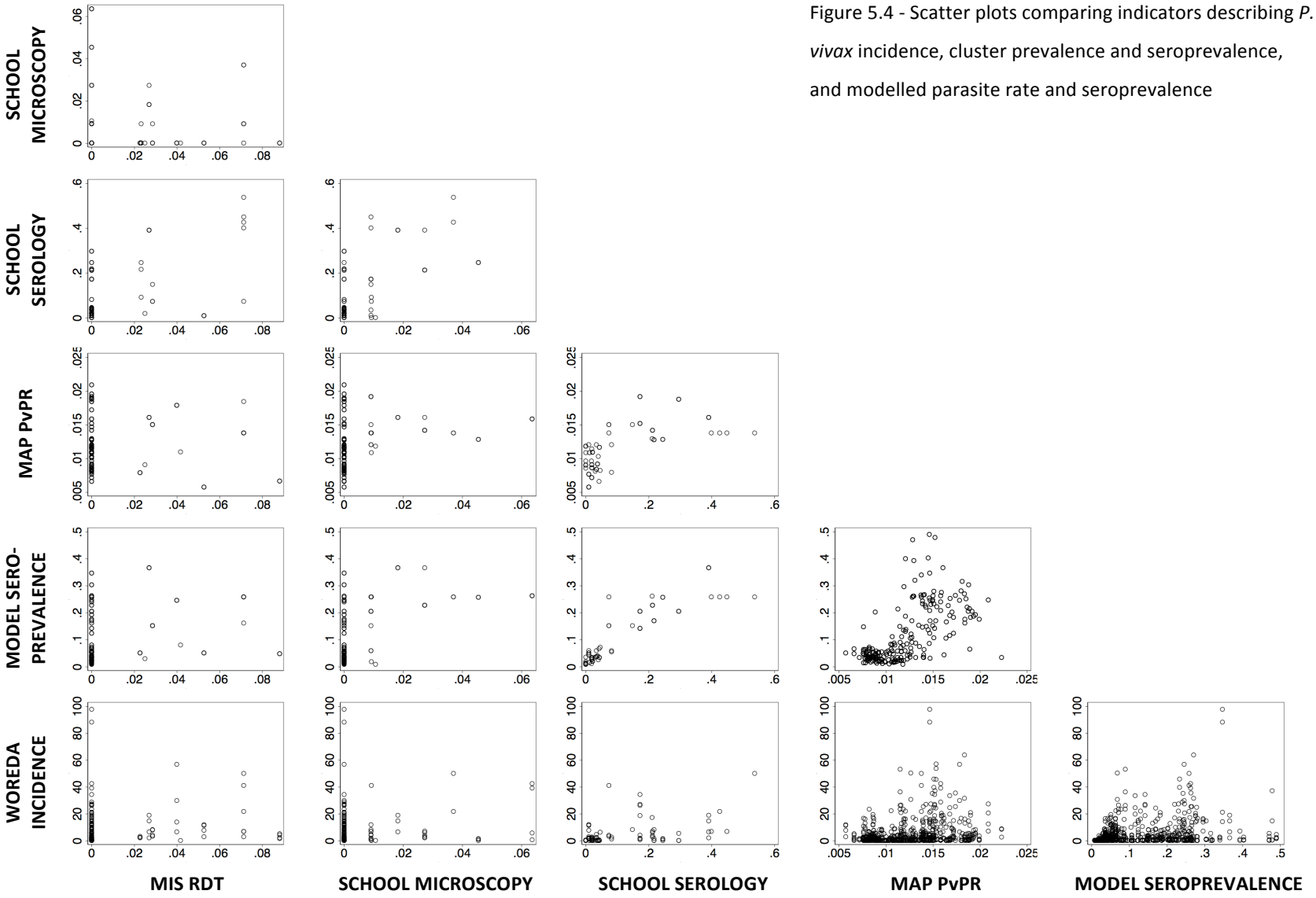


Figure 5.4 - Scatter plots comparing indicators describing *P. vivax* incidence, cluster prevalence and seroprevalence, and modelled parasite rate and seroprevalence

5.4.3.1 Cross-sectional survey malaria infection and incidence

Results of analysis of routinely recorded health facility data and annual incidence of confirmed malaria are detailed in Appendix 3. RDT positivity across sites sampled in Oromia during MIS was 1.5% among individuals of all ages, and also 1.5% among children under five. A very weak correlation (R^2=0.195, p=0.003) was seen between proportion RDT positive (all ages) during the MIS and the mean incidence of confirmed malaria at health facilities, including data from 2007-2010 (Figure 5.2), however, this should be interpreted cautiously due to the predominance of 0% RDT results in MIS data. Health facility data were not available for 2011, therefore it was not possible to directly compare facility incidence and survey prevalence for the same period. When assessing the correlation between MIS RDT positivity and incidence by year, the correlation was strongest in 2008 (R^2=0.405, p=0.002), with no correlation seen in the 2010 and 2007 data (scatter plots by year not shown).

Slide positivity during school-based survey in 2009 was very low (0.56%) and the majority of sites had no *Plasmodium* infection detected by microscopy. However a correlation was seen between proportion of children with infection detected by microscopy and mean incidence of malaria in 2009 in the woreda (R^2=0.423, p=0.001, Figure 5.2).

5.4.3.2 Cross-sectional survey malaria infection and Malaria Atlas Project modelled parasite rate

MIS RDT positivity was weakly correlated with predicted *Pv*PR_{1-99} (R^2=0.308, p<0.001), however the majority of MIS cluster prevalence results were 0% by RDT. No correlation was seen with RDT positivity and *Pf*PR_{2-10} prediction (R^2=0.025, p=0.97, Figure 5.3). No correlation was seen between school survey prevalence of *P. falciparum* by microscopy and predicted *Pf*PR_{2-10} (p=0.231), but a weak correlation was found between school survey *P. vivax* prevalence and *Pv*PR_{1-99} (R^2=0.187, p=0.003, Figure 5.4).

5.4.3.3 Serological indicators from cross-sectional surveys

While few and very weak correlations were found between cross-sectional survey estimates of prevalence of malaria infection and either annual mean incidence from routine health facility data or modelled parasite rate, use of serological indicators of malaria exposure collected from cross-sectional surveys offers an alternative indicator with a longer temporal reference period than presence of parasites on the survey day.

Upon comparing *P. falciparum* school seroprevalence and predicted *Pf*PR_{2-10}, no correlation was seen (p=0.60, Figure 5.3), however a non-linear relationship was seen between *P. vivax* seroprevalence and predicted *Pv*PR_{1-99} Figure 5.4. For locations with *Pv*Pr_{1-99} < 1.2%, school *P. vivax* seroprevalence was very low, while a far wider range of seroprevalence was seen (but with no clear linear relationship) for higher values of *Pv*PR_{1-99}.

Strong correlations were found between school seroprevalence and average annual incidence of malaria recorded by health facilities in the woreda. A correlation was seen between *P. falciparum* seroprevalence and mean annual incidence of *P. falciparum* in the woreda in 2009 (R^2=0.892, p<0.001, Figure 5.5). The association between measured school *P. falciparum* seroprevalence and incidence of *P. falciparum* at health facilities was also seen using mean incidence data from 2008 (R^2=0.510, p=0.036) and 2007 (R^2=0.831, p<0.001).

A significant correlation was also found between *P. vivax* seroprevalence and annual mean incidence of *P. vivax* in 2009 (R^2=0.691, p<0.001, Figure 5.5), however the association was not observed from incidence data in 2008 (R^2=0.406, p=0.068) or 2007 (R^2=0.342, p=0.20). Prior to 2009, *P. vivax* infection could only be confirmed by microscopy at health centres, since RDTs in use at health posts were detected HRP2 antigen only, so it is unsurprising that seroprevalence data did not have any association with *P. vivax* incidence before the availability of combination HRP2-panLDH RDTs at health posts from 2009 onwards.

Figure 5.5 - Scatter plot of (left) measured *P. falciparum* seroprevalence in school-based surveys against mean incidence of *P. falciparum* per 1000 people in 2009 in the same woreda, and (right) *P. vivax* seroprevalence against woreda mean *P. vivax* incidence per 1000 people in 2009.

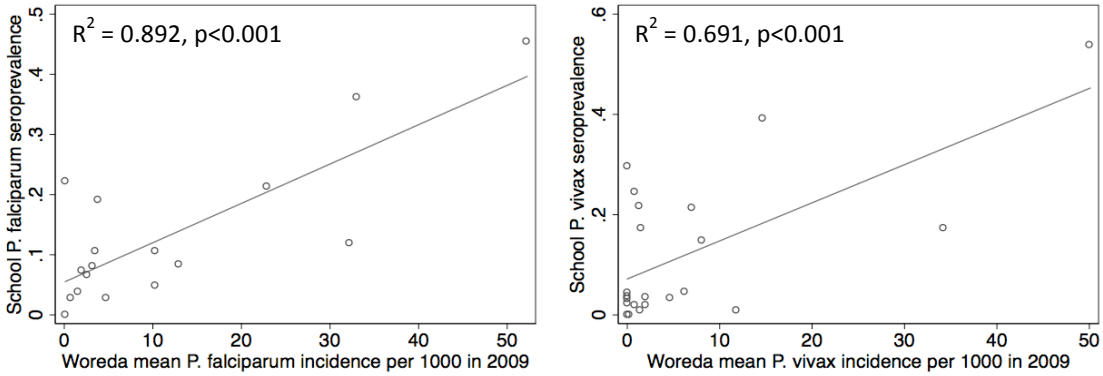
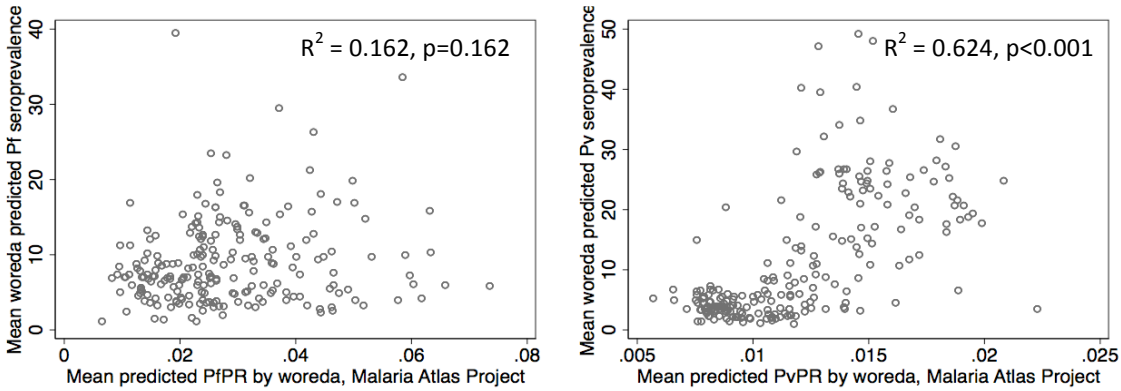


Figure 5.6 - Scatterplots of (left) woreda predicted mean *P. falciparum* seroprevalence against woreda mean predicted *Pf*PR_{2-10}, and (right) woreda predicted mean *P. vivax* seroprevalence against woreda mean predicted *Pv*PR_{1-99}. These scatter plots are also presented in the composite Figures 5.3 and 5.4



5.4.3.4 Bayesian predictive models of seroprevalence and parasite rate

The mean model prediction value by woreda for MAP age-standardised parasite rate and predicted seroprevalence were compared, to investigate if model predictions identified similar woredas as relative high and low transmission (Figure 5.6). No evidence was found for any correlation between woreda mean *Pf*PR_{2-10} and *P. falciparum* seroprevalence (R^2=0.162, p=0.162), but there was evidence of an association between *Pv*PR_{1-99} and *P. vivax* seroprevalence (R^2=0.624, p<0.001). The different types of models used to predict *P. falciparum* and *P. vivax* seroprevalence across Oromia may be contributing to the variation in association with MAP model outputs; *P. falciparum* seroprevalence was predicted by a model incorporating spatial random effects as well as environmental covariates, while the *P. vivax* model was a spatial smoothing model without environmental covariates. As a consequence of the *P. vivax* model using spatial smoothing only, the range of pixel seroprevalence estimates in a woreda was far smaller than the range of *P. falciparum* seroprevalence predicted within a woreda.

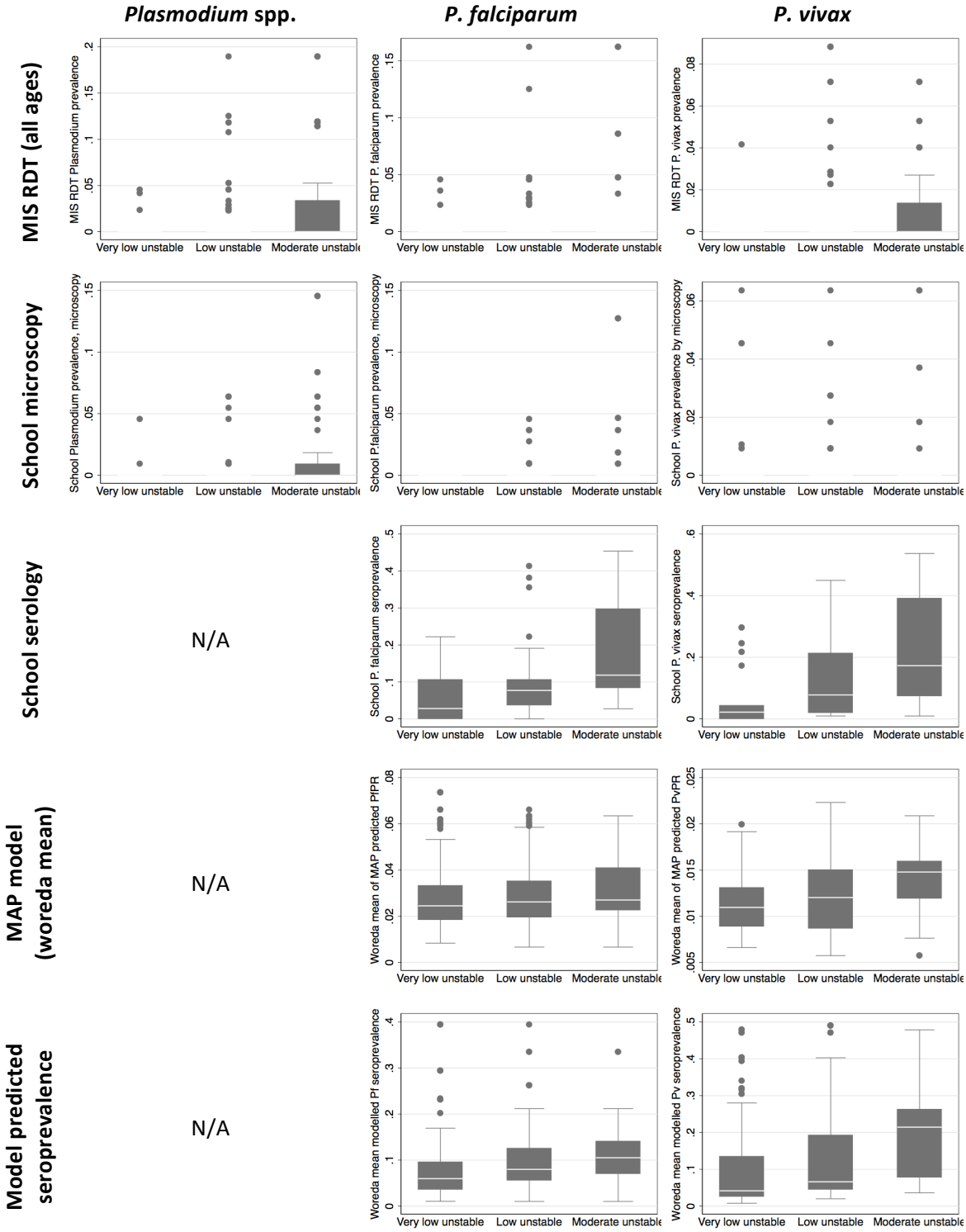
5.4.4 Classifying endemicity by woreda

Visual inspection of the distribution of woreda annual mean incidence of confirmed malaria led to the identification of endemicity class boundaries according to the incidence data. Woredas with very low unstable transmission were those with mean <1 confirmed case per 1000 people per year, woredas with low unstable transmission assigned as 1-10 confirmed cases per 1000 people per year, and those with moderate unstable transmission were all with ≥10 confirmed cases per 1000 people per year. The other surveillance data variables were not categorised into endemicity classes, but their distribution compared to incidence-defined endemicity classes using box plots (Figure 5.7).

Comparison of the distribution of school survey microscopy prevalence and MIS RDT prevalence between the endemicity categories was severely limited by the few sites which

had >0% prevalence. Median school seroprevalence increased across increasing endemicity categories for both *P. falciparum* and *P. vivax*, with the very low endemicity category effectively identifying the schools with very low *P. vivax* seroprevalence. The woreda mean seroprevalence model predictions appeared to show superior agreement with endmicity categories for *P. vivax* as median woreda seroprevalence increased across the endemicity categories, but there was little difference in predicted woreda seroprevalence by the different endemicity categories. A similar finding was observed for the MAP parasite rate predictions, being similar across the *P. falciparum* endemicity categories but an increasing median *Pv*PR into the higher endemicity categories.

Figure 5.7 - Box plots describing the distribution of each of the surveillance indicators against mean woreda annual incidence, classified as: <1 case/1000 as very low unstable transmission; 1-10 cases/1000 low unstable transmission; >10 cases/1000 moderate unstable



5.5 DISCUSSION

The relative merits and shortcomings of a range of routinely used and innovative surveillance and monitoring tools have been presented in the form of scoring frameworks, scoring tools separately against parameters which are most important for ongoing surveillance and for periodic monitoring. In addition, indicator values from different surveillance tools at spatially matched locations were compared to evaluate correlation between indicators. However, there exists no gold standard for these monitoring and surveillance data.

The scoring framework applied in this study does not evaluate the quality of data generated by each of the surveillance systems, but assumes that each system is operating correctly according to its design. Woreda summary data from health facilities was identified as the best-performing tool with the highest score, while use of RDT and microscopy data from MIS and school surveys were the lowest scoring methods. School survey seroprevalence and modelling estimates of seroprevalence and parasite rate received similar mid-ranging scores.

Routine data from health facilities scored highly on almost all parameters, but this is based upon assumptions of high levels of access to parasitological diagnosis of malaria at health facilities, that diagnosis is quality assured, data are reported accurately and in a timely fashion, as well as analysis and feedback being provided in good time. These data from health facilities are unique in having the potential to monitor the malaria situation over time, but also have sufficient temporal resolution to be valuable as a surveillance tool.

All other tools included in the scoring framework should be considered as monitoring tools only; they had reduced scores compared to health facility-generated data due to their limited temporal resolution for epidemic response. Among these monitoring tools, use of RDT and microscopy as the primary indicator in cross-sectional malariometric surveys is acceptable to generate nationally-representative estimates of parasite prevalence. However, to capture sub-national or small scale heterogeneity in malaria burden and transmission in an

environment such as Ethiopia would require large sample sizes to generate representative cluster-level parasite rate estimates. Predictive modelling using Bayesian methods and incorporating associated environmental covariates allows generation of fine spatial resolution estimates from the available survey data. Endemicity maps generated using this method have the potential to be well-used by malaria control programme staff for stratification of risk and targeting of resources, provided there is some understanding of the associated uncertainties in model estimates. While it appears that many national malaria control programmes are aware of malaria risk mapping, and incorporate various types of risk map into official documentation including strategies and programme reviews or Global Fund applications, these risk maps are rarely used to support national planning activities such as targeting of interventions and resource requirements [309]. The majority of maps presented in these documents are either expert opinion and eco-climatic stratification, or mapped data from routine or national sample surveys

The strongest correlations of all pair-wise comparisons conducted using surveillance data available in this study were between serological indicators of exposure to malaria and estimates of annual incidence at the woreda level, estimated from passively detected cases presenting to health facilities. The correlation was particularly strong for *P. vivax.* This could reflect a recent transition in malaria epidemiology in Ethiopia with reductions in *P. falciparum* but stability in *P. vivax*, as a consequence, the serological data representing transmission over a period of months to years has a stronger correlation with *P. vivax* incidence than *P. falciparum*. The correlation between woreda endemicity category defined by incidence, and the measured seroprevalence at schools and model-predicted seroprevalence was apparent from box plots, where median seroprevalence increased with increasing endemicity.

The cross-sectional survey data using standard malariometric survey tools of microscopy and RDT identified very few infections, and the vast majority of clusters were found to have zero prevalence. The predominance of zero prevalence sites limits the ability of these surveys to

differentiate between sites which are malaria free or very low risk, and those which do experience malaria but where the peak seasonal transmission may have been missed by the cross-sectional survey. In low and seasonal transmission settings, findings from cross-sectional surveys using microscopy and RDT prevalence as primary outcomes should be interpreted in conjunction with longitudinal surveillance data, such as routine data from health facilities, to confirm whether surveys were conducted at the peak transmission season.

The use of cross-sectional survey cluster-level estimates to develop predictive spatial models and endemicity maps has shown growing popularity in infectious diseases, with initiatives in place for malaria and neglected tropical diseases to collate historical data and make map outputs available for use at no cost [310-312].

In this chapter, woreda summaries of predictions from global models of *P. falciparum* and *P. vivax* parasite rate were compared to recent field-level estimates of parasite rate, and to estimates of incidence from passively detected cases presenting to health facilities. When comparing mean predicted MAP parasite rate by woreda with incidence-defined endemicity categories, little agreement was seen for *P. falciparum*, but model predictions did increase with increasing category of endemicity for *P. vivax*. The superior agreement between modelled seroprevalence and endemicity categories over the MAP predictions may be a result of the serological model having been fitted to data for Oromia Region, rather than global data, or may be due to use of more contemporary data for serology than parasite rate.

While this study was able to successfully match multiple surveillance indicators spatially, not all of the indicators could be matched temporally. In particular, health facility data were not available during the period of the MIS in 2011. This may have resulted in bias in interpretation of cross-sectional survey findings from 2011, particularly if there is substantial heterogeneity in the locations that are found to show seasonal increases in malaria infection

from year to year. The spatial resolution of health facility data would have been improved if health facility location coordinates were available, allowing generation of health cluster (one health centre and 5 satellite health post) resolution estimates of incidence. Furthermore, the health facility data collected were incomplete; some woredas were excluded due to absence of data or very incomplete or inconsistent data, therefore not all cross-sectional survey clusters could be matched to woreda incidence data.

The findings from this chapter indicate that a combination of monitoring and surveillance tools are required in low transmission settings. Routine health facility data has potential to be a very valuable source for both surveillance and monitoring needs, but practical application of these data continues to be limited by quality and timeliness. Serological indicators have the potential to be a valuable addition to the monitoring toolkit, while use of microscopy and RDT data to generate of estimates of *Plasmodium* prevalence are of declining utility. Spatial model predictions offer an additional tool to inform national malaria control programme activities, particularly stratification and intervention targeting. Generation of nationally-representative parasite prevalence estimates using microscopy and RDT data from large cross-sectional surveys are likely to continue to be a component of periodic malaria monitoring, however, in low transmission settings these data should be accompanied by additional surveillance and monitoring data offering higher temporal and spatial resolution.

CHAPTER 6. SUMMARY AND DISCUSSION OF FINDINGS

6.1 OVERVIEW

Substantial investment in interventions for malaria control and concomitant reductions in disease burden have led to a resurgence of interest in monitoring, evaluation and surveillance tools. These tools are used to demonstrate impact of the interventions applied, track progress of malaria control towards elimination, and subsequently monitor for recrudescence or reintroduction of *Plasmodium*.

Resource-limited countries continue to face challenges in availability and quality of routine data from health facilities, limiting the utility of these data for programme monitoring and for ongoing surveillance. Considering the increasing number of countries where malaria epidemiology is in transition from high or moderate transmission to low and temporally variable transmission, it is timely to evaluate the ability of standard monitoring and evaluation tools to describe programmatically and epidemiologically relevant indicators, and to pilot innovative strategies for malaria monitoring.

This thesis therefore aimed to evaluate alternative strategies and indicators used for malaria monitoring, evaluation and surveillance in Ethiopia, with a view to generating recommendations of the most informative and appropriate tools to meet future needs in similar epidemiological settings. This chapter provides a summary of the findings and their implications for future monitoring, evaluation and surveillance activities in low transmission settings, as well identifying areas for further research.

6.2 SUMMARY OF FINDINGS

The need to collect data to describe implementation, outputs and impact of malaria control programmes is essential to ensure that control programmes remain relevant, effective and

responsive to the needs of the population and context [313]. The range of monitoring and surveillance strategies for malaria programmes were reviewed in Chapter 1, in addition to description of the various indicators available for use by these systems. Low and unstable malaria transmission settings present specific challenges to monitoring and surveillance, and the majority of standardised tools for monitoring, evaluation and surveillance have been designed for settings with moderate to intense transmission. Therefore, there remains a need to define tools that are appropriate for malaria programme periodic monitoring, and ongoing surveillance for epidemic detection.

Chapter 2 presented findings from large-scale school-based cross-sectional surveys in Oromia Regional State, describing *Plasmodium* infection by microscopy, risk factors, spatial distribution and clustering of infection. The results demonstrated very low levels of *Plasmodium* infection at the time of the survey, and a large proportion of infections were found to be asymptomatic. The sampling frame of the survey was designed to generate representative estimates of parasite prevalence within each of five ecological zones, representing anticipated different malaria transmission settings. Use of insecticide-treated mosquito nets by school-aged children was found to be low, but net use was a risk factor for *Plasmodium* infection, likely as a result of successful targeting of long-lasting insecticide-treated mosquito nets (LLINs) to areas at highest malaria risk. The results of Chapter 2 adequately captured programmatic indicators such as use of LLINs, but due to both the sampling design of the survey and very low prevalence of infection by microscopy, it was not possible to generate statistical models of endemicity and spatial distribution using microscopy data.

While microscopy has long been considered as a gold standard for malaria diagnosis in a clinical context, in areas of temporally unstable malaria transmission microscopy may not be the most appropriate primary indicator for cross-sectional surveys, due to logistical challenges in ensuring that the survey is conducted at the peak transmission season in every

location. Chapter 3 presented the use of serological indicators of exposure to *Plasmodium* collected from school surveys in Oromia Regional State, as an alternative analysis tool for samples collected by cross-sectional methods in low and temporally unstable settings. By assessing previous exposure rather than current infection with *Plasmodium*, it is possible to generate more comparable estimates of malaria endemicity across sampled locations, since intra- and inter-year differences in transmission are smoothed out at each site sampled as a result of the persistence of anti-*Plasmodium* antibodies for months to years. This approach combining a simple cross-sectional survey design and relatively simple laboratory assay is particularly appropriate as a periodic monitoring tool alongside, or in place of, standard malariometric surveys generating nationally-representative estimates of parasite prevalence in pre-elimination settings.

Chapter 3 also presents the first application of Bayesian geostatistical modelling methods to malaria serological data to generate endemicity maps for both *P. falciparum* and *P. vivax*. While the optimal model for *P. falciparum* seroprevalence included a spatial random effect and various environmental indicators describing suitability of an area for vector breeding and survival, for *P. vivax* a simple spatial smoothing model was the most appropriate. Factors contributing to the large spatial scale of *P. vivax* may include relapse of previous infections, or sporogenesis at lower temperatures than for *P. falciparum* and therefore potential for *P. vivax* transmission higher altitudes.

While Chapter 3 presented an innovative tool for periodic monitoring of malaria endemicity, a novel system for ongoing surveillance and malaria epidemic detection was presented in Chapter 4, using school attendance as a proxy indicator of malaria burden within communities at sites in the Southern Nations, Nationalities and People's Regional State, Ethiopia. Syndromic surveillance using pre-diagnostic (clinical) indicators has been piloted in resource poor settings as an epidemic warning system for diseases such as dengue, meningitis and malaria, where laboratory confirmation of notifiable diseases is not widely

available. Syndromic surveillance using surrogate indicators such as school attendance has primarily been used in high-income countries to identify outbreaks of pandemic influenza. Chapter 4 explored the potential of both syndromic and surrogate indicators of malaria, piloting surveillance systems at school-level that were designed to feed into existing routine surveillance systems at health post. While school enrolment was found to be lower than expected in the pilot communities, and the lack of strong seasonal increase in malaria transmission during the pilot limited the ability to test reliability of the piloted systems, it was simple to generate a weekly all-cause absenteeism summary at pilot schools using existing attendance registers. In a setting such as Ethiopia with a strong community-level health system, but ongoing limitations to timeliness and quality of routine health information system data, the syndromic surveillance system offers a community-level solution whereby school staff have a framework to identify unusual health events which result in increased school absenteeism, and alert health extension workers to investigate further.

Chapters 2, 3 and 4 presented in-depth a selection of monitoring and surveillance indicators and strategies, however there are few studies which have compared in detail different monitoring and surveillance methods for use in low transmission settings. Chapter 5 aimed to score and compare some of the different indicators that are used for programme monitoring and malaria surveillance: routinely recorded data from health facilities; prevalence of infection by microscopy or rapid diagnostic test (RDT) from school surveys and Malaria Indicator Survey (MIS), respectively; seroprevalence from cross-sectional school surveys; model-predicted parasite rate; and model predicted seroprevalence. A scoring framework was used to evaluate the different tools against defined parameters, then spatially congruent data linked and correlation between indicators investigated. Routine health facility data received the highest scores for both monitoring and surveillance from the scoring framework, demonstrating the potential of a fully functional and high quality health information system to meet the needs of both ongoing surveillance for epidemic detection as well as monitoring

the impact of control interventions on malaria burden. However, while routine data from health facilities continues to by limited by poor timeliness in reporting and data analysis, lacks quality assurance, and where access to health services is limited in some geographic areas or population groups, alternative strategies are needed.

The challenges of malariometric cross-sectional surveys in settings where malaria is strongly seasonal and of low transmission intensity are apparent from the findings of Chapter 5, with many survey clusters finding no infections using microscopy and RDTs. For such cross-sectional surveys to have sufficient power to generate population-representative estimate of parasitaemia at sub-national of level at very low levels of transmission, sample sizes may need to be increased from those used at present. Diagnostic tools with superior sensitivity to microscopy and RDT may add additional resolution to estimates of parasite prevalence from cross-sectional surveys, since they would detect more of the very low density and asymptomatic infections present in the population, however surveys such as MIS would likely remain under-powered for generation of reliable estimates of sub-national infection prevalence.

Chapter 5 particularly highlighted the potential of serological indicators in cross sectional surveys to overcome the limitations of classical tools of parasite detection, by representing population exposure to malaria transmission over a period of months or years. This temporal smoothing effect reduces the potential bias in estimating cluster prevalence in strongly seasonal settings, where there is a risk of survey implementation not exactly coinciding with the actual transmission peak at each site, by presenting an estimate of exposure averaged over previous years. While serological analysis has limited current utility for surveillance focussed on epidemic detection, it may prove to be a significant addition to the periodic monitoring toolkit for malaria programme managers.

Upon developing the parameters for the scoring framework to assess the surveillance data available, it became clear that a single tool or indicator would be unlikely to meet the objectives of both monitoring and surveillance in Ethiopia. The needs of a monitoring tool are weighted towards sensitivity and spatial resolution, with some potential to accommodate higher complexity and cost in order to achieve this. However, surveillance tools are focussed towards temporal resolution and ease or speed of generation of feedback between those collecting data and those conducting the analysis. Therefore it is crucial for national malaria control programmes to identify their priorities in order to allocate resources appropriately, whether a system with high temporal resolution to identify epidemics, a system with high sensitivity to identify every malaria case, or a system to document large-scale changes in malaria burden over time.

6.3 RECOMMENDATIONS FOR APPROPRIATE MALARIA SURVEILLANCE AND MONITORING TOOLS IN LOW TRANSMISSION SETTINGS

Routinely recorded data at health facilities has the potential to be a highly sensitive tool for malaria monitoring and surveillance, justifying further investments to address the remaining limitations in these data. A focus on ensuring both timely and complete reporting of data from all facilities, including zero-reporting, should be an immediate priority. While the limitations in analysis and interpretation of routinely collected health facility data are well-known [94,314], it remains the most sustainable and lowest cost monitoring and surveillance tool, and exploration of new initiatives for surveillance should not reduce the investment in improved health information systems. Innovations such as reporting of data using mobile telephones may improve timeliness of reporting and simplify data entry prior to analysis, and may also have a secondary impact by providing additional motivation for submission of data and increasing completeness and compliance by health workers [315-317].

In low transmission and pre-elimination settings, malariometric surveys using standard cross-sectional survey designs and tools such as microscopy and rapid diagnostic tests designed to generate national-level estimates of parasite prevalence are costly. This is not an unexpected finding, and recommendations have been made previously advising of the imprecise nature of population parasite rate estimates using cross-sectional surveys when prevalence falls below 5% [9]. Furthermore, MIS guidance recommends that malariometric modules be included in MIS in areas of stable transmission [104]. Where *Plasmodium* prevalence is estimated in surveys, there should also be reconsideration of the sampled population groups, since most malariometric surveys currently focus on testing children under five years and pregnant women. Modelling of data from 23 African countries indicates that in low transmission settings more than 60% of infections will be in individuals over 15 years of age, supporting expansion of the age range included in blood sampling beyond children under five years [318].

An alternative to the standard MIS and other large-scale malariometric surveys may be to incorporate indicators of intervention coverage, access to diagnosis and treatment and malaria knowledge into other population-representative surveys or activities, to generate these key indicators of interest to donors and for comparison against other country programmes. Several questions relevant to malaria are already incorporated into Demographic and Health Surveys. Amending the sampling frame or reducing the scale of a questionnaire-only MIS is a further alternative to ensure that the surveys remains relevant and low cost, collecting key indicators. This may allow focus on selected epidemiologically relevant sentinel sites, to track changes in malaria transmission across the range of different transmission settings that exist in Ethiopia. Reducing the scale of such surveys may potentially allow incorporation of more sensitive and expensive diagnostic tools, or increased frequency of survey implementation to allow control programmes to be more responsive to changing needs in different settings.

A further alternative is the use of schools to monitor malaria control and intervention coverage. Across sub-Saharan Africa, net enrolment in primary school (enrolment among children of the official primary school age) in 2011 was estimated at 76% [162]. Primary schools are therefore increasingly representative of the underlying population. Schools have been used to monitor *Plasmodium* infection and mosquito net use, for intermittent screening and treatment of malaria, and to estimate local transmission intensity in Kenya [111,319,320]. While in the Gambia, school surveys have been used in the dry season to investigate hotspots of malaria which may be responsible for resurgence of *Plasmodium* transmission during and following seasonal rains [192].

Further innovation is required to refine monitoring and surveillance tools for woredas that are within the World Health Organization pre-elimination classification of less than one malaria case per 1000 population, but the focus should remain on surveillance at fine temporal scale due to the risk of epidemics. If the long-term aim is to break transmission in these areas meeting pre-elimination criteria, then identification of the hotspots of transmission and reservoirs of parasites is required. However, programme managers should balance the priorities and long-term costs of sustained control versus attempting elimination where there remains the potential for resurgence and importation of parasites from neighbouring areas.

6.4 FUTURE DIRECTIONS

The results presented in this thesis have explored the relative merits and disadvantages of several innovative strategies for malaria monitoring and surveillance in low and unstable transmission settings in Ethiopia. Chapter 2 presented the application of school-based surveys as an alternative to large-scale community surveys used to monitor population *Plasmodium* prevalence as well as access to and use of key malaria interventions such as LLINs. While school surveys have logistical advantages over community surveys, they face the

same limitations in low transmission settings as other large-scale surveys utilising standard diagnostic tools of microscopy or RDTs, in that very few infections are found and extremely large sample sizes would be required to adequately capture geographical heterogeneities in transmission intensity.

Alternative indicators may be needed if national malaria control programmes wish to continue conducting periodic large-scale cross-sectional surveys to monitor changes in *Plasmodium* transmission and burden. Highly sensitive molecular diagnostics are one potential solution, but use of this tool will also restricted by the operational challenges of targeting cross-sectional surveys to the peak transmission period at every sampled site. Periodic population monitoring using serological indicators of exposure shows much promise as a malaria monitoring tool to compare relative transmission intensity between different settings, using blood samples collected through a cross-sectional survey method. Alternatively, where cross-sectional surveys are not possible, a convenience sampling approach could be used, conducting further analysis of blood samples from individuals presenting at health facilities for a different diagnostic test requiring finger-prick blood samples. This could also be applied in a pre-elimination setting to selectively screen blood samples from children less than five years attending sentinel health facilities (or accompanying children) for antibodies to *Plasmodium*, in order to selectively investigate recent malaria transmission in areas approaching elimination. These alternative sampling strategies should only be applied if the Ministry of Health and relevant malaria control partners deem them appropriate, and provided that full informed consent is given for additional diagnostic testing of samples.

In conclusion, this thesis has presented alternative strategies for malaria monitoring, including the use of school-based cross-sectional surveys as an alternative to standard household-level population surveys, and proposed serological indicators of exposure to *Plasmodium* as an alternative indicator for such periodic large-scale monitoring activities. This

thesis specifically presented an endemicity map generated using seroprevalence data from cross-sectional surveys, but serological indicators may have further potential applications for malaria monitoring in low and unstable malaria transmission settings. In addition to new tools for malaria periodic monitoring, this thesis also presented a first pilot of a school-based syndromic surveillance system, whereby school absenteeism was proposed as a complementary indicator of malaria epidemics in the community. While the syndromic surveillance system requires further refinement and testing, it has potential as a complementary system alongside routine health post systems for malaria and other infectious disease epidemic detection. Finally, this thesis presented a comparative analysis of a range of malaria surveillance and monitoring systems, presenting recommendations of the most appropriate tools for future use in areas of low and unstable malaria transmission.

REFERENCES

1. WHO (2003) The Abuja Declaration and the plan of action. Geneva: World Health Organization. WHO/CDS/RBM/2003.46 http://whqlibdoc.who.int/hq/2003/WHO_CDS_RBM_2003.46.pdf

2. Gething PW, Battle KE, Bhatt S, Smith DL, Eisele TP, Cibulskis RE, Hay SI (2014) Declining malaria in Africa: improving the measurement of progress. *Malaria Journal* 13:39.

3. Ceesay SJ, Casals-Pascual C, Erskine J, Anya SE, Duah NO, Fulford AJ, Sesay SS, Abubakar I, Dunyo S, Sey O, Palmer A, Fofana M, Corrah T, Bojang KA, Whittle HC, Greenwood BM, Conway DJ (2008) Changes in malaria indices between 1999 and 2007 in The Gambia: a retrospective analysis. *Lancet* 372:1545-1554.

4. Ceesay SJ, Casals-Pascual C, Nwakanma DC, Walther M, Gomez-Escobar N, Fulford AJ, Takem EN, Nogaro S, Bojang KA, Corrah T, Jaye MC, Taal MA, Sonko AA, Conway DJ (2010) Continued decline of malaria in The Gambia with implications for elimination. *PLoS ONE* 5:e12242.

5. O'Meara WP, Bejon P, Mwangi TW, Okiro EA, Peshu N, Snow RW, Newton CRJC, Marsh K (2008) Effect of a fall in malaria transmission on morbidity and mortality in Kilifi, Kenya. *Lancet* 372:1555-1562.

6. The MalERA Consultative Group on Monitoring EaS (2011) A research agenda for malaria eradication: Monitoring, evaluation, and surveillance. *PLoS Medicine* 8:e1000400.

7. Cotter C, Sturrock HJ, Hsiang MS, Liu J, Phillips AA, Hwang J, Gueye CS, Fullman N, Gosling RD, Feachem RG (2013) The changing epidemiology of malaria elimination: new strategies for new challenges. *Lancet* 382:900-911.

8. Remme JH, Binka F, Nabarro D (2001) Toward a framework and indicators for monitoring Roll Back Malaria. *American Journal of Tropical Medicine and Hygiene* 64:76-84.

9. Hay SI, Smith DL, Snow RW (2008) Measuring malaria endemicity from intense to interrupted transmission. *The Lancet Infectious Diseases* 8:369-378.

10. Okell LC, Ghani AC, Lyons E, Drakeley CJ (2009) Submicroscopic infection in *Plasmodium falciparum*-endemic populations: A systematic review and meta-analysis. *The Journal of Infectious Diseases* 200:1509-1517.

11. Drakeley C, Cook J (2009) Potential contribution of sero-epidemiological analysis for monitoring malaria control and elimination: historical and current perspectives. *Advances in Parasitology* 69:299-352.

12. Schwetz J, Baumann H (1929) Study of the malaria index in young natives of school age in the settlement of Stanleyville (Congo Belge). *Transactions of the Royal Society of Tropical Medicine and Hygiene* 23:279-288.

13. Alves W (1958) Malaria parasite rates in Southern Rhodesia: May-September 1956. *Bulletin of the World Health Organization* 19:69-74.

14. Carr HP, Hill RB (1942) A malaria survey of Cuba. *American Journal of Tropical Medicine* s1-22:587-607.

15. Boyd MF, Aris FW (1929) A malaria survey on the island of Jamaica. *The American Journal of Tropical Medicine and Hygiene* s1-29:309-399.

16. Gitonga CW, Karanja PN, Kihara J, Mwanje M, Juma E, Snow RW, Noor AM, Brooker S (2010) Implementing school malaria surveys in Kenya: towards a national surveillance system. *Malaria Journal* 9:306.

17. Kom Mogto CA, De Serres G, Douville Fradet M, Lebel G, Toutant S, Gilca R, Ouakki M, Janjua NZ, Skowronski DM (2012) School absenteeism as an adjunct surveillance indicator: experience during the second wave of the 2009 H1N1 pandemic in Quebec, Canada. *PLoS ONE* 7:e34084.

18. Kara EO, Elliot AJ, Bagnall H, Foord DG, Pnaiser R, Osman H, Smith GE, Olowokure B (2012) Absenteeism in schools during the 2009 influenza A(H1N1) pandemic: a useful tool for early detection of influenza activity in the community? *Epidemiology and Infection* 140:1328-1336.

19. Sasaki A, Hoen AG, Ozonoff A, Suzuki H, Tanabe N, Seki N, Saito R, Brownstein JS (2009) Evidence-based tool for triggering school closures during influenza outbreaks, Japan. *Emerging Infectious Diseases* 15:1841-1843.

20. Gething PW, Patil AP, Smith DL, Guerra CA, Elyazar IRF, Johnston GL, Tatem AJ, Hay SI (2011) A new world malaria map: *Plasmodium falciparum* endemicity in 2010. *Malaria Journal* 10:378.

21. Gething PW, Elyazar IR, Moyes CL, Smith DL, Battle KE, Guerra CA, Patil AP, Tatem AJ, Howes RE, Myers MF, George DB, Horby P, Wertheim HF, Price RN, Mueller I, Baird JK, Hay SI (2012) A long neglected world malaria map: *Plasmodium vivax* endemicity in 2010. *PLoS Neglected Tropical Diseases* 6:e1814.

22. Roucher C, Rogier C, Sokhna C, Tall A, Trape JF (2014) A 20-year longitudinal study of *Plasmodium ovale* and *Plasmodium malariae* prevalence and morbidity in a West African population. *PLoS ONE* 9:e87169.

23. Sinka ME, Bangs MJ, Manguin S, Rubio-Palis Y, Chareonviriyaphap T, Coetzee M, Mbogo CM, Hemingway J, Patil AP, Temperley WH, Gething PW, Kabaria CW, Burkot TR, Harbach RE, Hay SI (2012) A global map of dominant malaria vectors. *Parasites & Vectors* 5:69.

24. Menard R, Tavares J, Cockburn I, Markus M, Zavala F, Amino R (2013) Looking under the skin: the first steps in malarial infection and immunity. *Nature Reviews Microbiology* 11:701-712.

25. Boyd MF, Kitchen SF (1937) On the infectiousness of patients infected with *Plasmodium vivax* and *Plasmodium falciparum*. *The American Journal of Tropical Medicine and Hygiene* 17:253-262.

26. Jeffery GM, Eyles DE (1955) Infectivity to mosquitoes of *Plasmodium falciparum* as related to gametocyte density and duration of infection. *The American Journal of Tropical Medicine and Hygiene* 4:781-789.

27. Kwiatkowski D, Nowak M (1991) Periodic and chaotic host-parasite interactions in human malaria. *Proceedings of the National Academy of Sciences of the United States of America* 88:5111-5113.

28. Kwiatkowski D, Greenwood BM (1989) Why is malaria fever periodic? A hypothesis. *Parasitology Today* 5:264-266.

29. Buffet PA, Safeukui I, Deplaine G, Brousse V, Prendki V, Thellier M, Turner GD, Mercereau-Puijalon O (2011) The pathogenesis of *Plasmodium falciparum* malaria in humans: insights from splenic physiology. *Blood* 117:381-392.

30. Heymann Dl, editor (2004) Control of Communicable Diseases Manual, 18th Edition; Washington, D.C.: American Public Health Association.

31. Mackintosh CL, Beeson JG, Marsh K (2004) Clinical features and pathogenesis of severe malaria. *Trends in Parasitology* 20:597-603.

32. Tjitra E, Anstey NM, Sugiarto P, Warikar N, Kenangalem E, Karyana M, Lampah DA, Price RN (2008) Multidrug-resistant *Plasmodium vivax* associated with severe and fatal malaria: a prospective study in Papua, Indonesia. *PLoS Medicine* 5:e128.

33. Genton B, D'Acremont V, Rare L, Baea K, Reeder JC, Alpers MP, Muller I (2008) *Plasmodium vivax* and mixed infections are associated with severe malaria in children: a prospective cohort study from Papua New Guinea. *PLoS Medicine* 5:e127.

34. Kochar DK, Das A, Kochar SK, Saxena V, Sirohi P, Garg S, Kochar A, Khatri MP, Gupta V (2009) Severe *Plasmodium vivax* malaria: a report on serial cases from Bikaner in northwestern India. *The American Journal of Tropical Medicine and Hygiene* 80:194-198.

35. Anstey NM, Russell B, Yeo TW, Price RN (2009) The pathophysiology of vivax malaria. *Trends in Parasitology* 25:220-227.

36. Smith DL, McKenzie FE, Snow RW, Hay SI (2007) Revisiting the basic reproductive number for malaria and its implications for malaria control. *PLoS Biology* 5:e42.

37. Smith DL, Drakeley CJ, Chiyaka C, Hay SI (2010) A quantitative analysis of transmission efficiency versus intensity for malaria. *Nature Communications* 1:108.

38. Smith DL, Perkins TA, Reiner RC, Jr., Barker CM, Niu T, Chaves LF, Ellis AM, George DB, Le Menach A, Pulliam JR, Bisanzio D, Buckee C, Chiyaka C, Cummings DA, Garcia AJ, Gatton ML, Gething PW, Hartley DM, Johnston G, Klein EY, Michael E, Lloyd AL, Pigott DM, Reisen WK, Ruktanonchai N, Singh BK, Stoller J, Tatem AJ, Kitron U, Godfray HC, Cohen JM, Hay SI, Scott TW (2014) Recasting the theory of mosquito-borne pathogen transmission dynamics and control. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 108:185-197.

39. Ceccato P, Vancutsem C, Klaver R, Rowland J, Connor SJ (2012) A vectorial capacity product to monitor changing malaria transmission potential in epidemic regions of Africa. *Journal of Tropical Medicine* 2012:595948.

40. Sinka ME, Bangs MJ, Manguin S, Coetzee M, Mbogo CM, Hemingway J, Patil AP, Temperley WH, Gething PW, Kabaria CW, Okara RM, Van Boeckel T, Godfray HC, Harbach RE, Hay SI (2010) The dominant *Anopheles* vectors of human malaria in Africa, Europe and the Middle East: occurrence data, distribution maps and bionomic precis. *Parasites & Vectors* 3:117.

41. Durnez L, Coosemans M (2013) Residual transmission of malaria: an old issue for new approaches. In: Manguin S, editor. Anopheles mosquitoes - New insights into malaria vectors.

42. Russell TL, Govella NJ, Azizi S, Drakeley CJ, Kachur SP, Killeen GF (2011) Increased proportions of outdoor feeding among residual malaria vector populations following increased use of insecticide-treated nets in rural Tanzania. *Malaria Journal* 10:80.

43. Padonou GG, Gbedjissi G, Yadouleton A, Azondekon R, Razack O, Oussou O, Gnanguenon V, Rock A, Sezonlin M, Akogbeto M (2012) Decreased proportions of indoor feeding and endophily in *Anopheles gambiae s.l.* populations following the indoor residual spraying and insecticide-treated net interventions in Benin (West Africa). *Parasites & Vectors* 5:262.

44. Thomson MC, Connor SJ, Milligan PJ, Flasse SP (1996) The ecology of malaria--as seen from Earth-observation satellites. *Annals of Tropical Medicine and Parasitology* 90:243-264.

45. Akhwale WS, Lum JK, Kaneko A, Eto H, Obonyo C, Bjorkman A, Kobayakawa T (2004) Anemia and malaria at different altitudes in the western highlands of Kenya. *Acta Tropica* 91:167-175.

46. Balls MJ, Bodker R, Thomas CJ, Kisinza W, Msangeni HA, Lindsay SW (2004) Effect of topography on the risk of malaria infection in the Usambara Mountains, Tanzania. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 98:400-408.

47. Brooker S, Clarke S, Njagi JK, Polack S, Mugo B, Estambale B, Muchiri E, Magnussen P, Cox J (2004) Spatial clustering of malaria and associated risk factors during an epidemic in a highland area of western Kenya. *Tropical Medicine and International Health* 9:757-766.

48. Drakeley CJ, Carneiro I, Reyburn H, Malima R, Lusingu JP, Cox J, Theander TG, Nkya WM, Lemnge MM, Riley EM (2005) Altitude-dependent and -independent variations in

Plasmodium falciparum prevalence in northeastern Tanzania. *The Journal of Infectious Diseases* 191:1589-1598.

49. Detinova TS, Bertram DS, World Health Organization (1962) Age-grouping methods in diptera of medical importance, with special reference to some vectors of malaria. World Health Organization. http://apps.who.int/iris/handle/10665/41724

50. Gething PW, Van Boeckel TP, Smith DL, Guerra CA, Patil AP, Snow RW, Hay SI (2011) Modelling the global constraints of temperature on transmission of *Plasmodium falciparum* and *P. vivax*. *Parasites & Vectors* 4:92.

51. Lindsay SW, Martens WJ (1998) Malaria in the African highlands: past, present and future. *Bulletin of the World Health Organization* 76:33-45.

52. Afrane YA, Lawson BW, Githeko AK, Yan G (2005) Effects of microclimatic changes caused by land use and land cover on duration of gonotrophic cycles of *Anopheles gambiae* (Diptera: Culicidae) in western Kenya highlands. *Journal of Medical Entomology* 42:974-980.

53. Mabaso MLH, Craig M, Vounatsou P, Smith T (2005) Towards empirical description of malaria seasonality in southern Africa: the example of Zimbabwe. *Tropical Medicine and International Health* 10:909-918.

54. Fontenille D, Lochouarn L, Diagne N, Sokhna C, Lemasson JJ, Diatta M, Konate L, Faye F, Rogier C, Trape JF (1997) High annual and seasonal variations in malaria transmission by anophelines and vector species composition in Dielmo, a holoendemic area in Senegal. *The American Journal of Tropical Medicine and Hygiene* 56:247-253.

55. Hamad AA, Nugud Ael H, Arnot DE, Giha HA, Abdel-Muhsin AM, Satti GM, Theander TG, Creasey AM, Babiker HA, Elnaiem DE (2002) A marked seasonality of malaria transmission in two rural sites in eastern Sudan. *Acta Tropica* 83:71-82.

56. Shililu J, Ghebremeskel T, Mengistu S, Fekadu H, Zerom M, Mbogo C, Githure J, Novak R, Brantly E, Beier JC (2003) High seasonal variation in entomologic inoculation rates in Eritrea, a semi-arid region of unstable malaria in Africa. *The American Journal of Tropical Medicine and Hygiene* 69:607-613.

57. Martens P, Hall L (2000) Malaria on the move: human population movement and malaria transmission. *Emerging Infectious Diseases* 6:103-109.

58. Kiszewski AE, Teklehaimanot A (2004) A review of the clinical and epidemiologic burdens of epidemic malaria. *The American Journal of Tropical Medicine and Hygiene* 71:128-135.

59. Loevinsohn ME (1994) Climatic warming and increased malaria incidence in Rwanda. *Lancet* 343:714-718.

60. Afrane YA, Zhou G, Lawson BW, Githeko AK, Yan G (2006) Effects of microclimatic changes caused by deforestation on the survivorship and reproductive fitness of *Anopheles gambiae* in western Kenya highlands. *The American Journal of Tropical Medicine and Hygiene* 74:772-778.

61. Last JM, editor (2001) A Dictionary of Epidemiology. 4th ed. New York: Oxford University Press.

62. Najera JA (1999) Prevention and control of malaria epidemics. *Parassitologia* 41:339-347.

63. World Health Organization (2006) Systems for the early detection of malaria epidemics in Africa. An analysis of current practices and future priorities. ISBN: 92415 9 488 8, 978 92 4 159488 2 http://www.who.int/malaria/publications/atoz/9789241594882/en/

64. World Health Organization (2013) World Malaria Report 2013. Geneva: World Health Organization. http://www.who.int/malaria/publications/world_malaria_report_2013/report/en/

65. Carneiro I, Roca-Feltrer A, Griffin JT, Smith L, Tanner M, Schellenberg JA, Greenwood B, Schellenberg D (2010) Age-patterns of malaria vary with severity, transmission intensity and seasonality in sub-Saharan Africa: a systematic review and pooled analysis. *PLoS ONE* 5:e8988.

66. Smith DL, Guerra CA, Snow RW, Hay SI (2007) Standardizing estimates of the *Plasmodium falciparum* parasite rate. *Malaria Journal* 6:131.

67. Schellenberg D, Menendez C, Aponte J, Guinovart C, Mshinda H, Tanner M, Alonso P (2004) The changing epidemiology of malaria in Ifakara Town, southern Tanzania. *Tropical Medicine and International Health* 9:68-76.

68. Brooker S, Kolaczinski J, Gitonga C, Noor A, Snow R (2009) The use of schools for malaria surveillance and programme evaluation in Africa. *Malaria Journal* 8:231.

69. Baliraine FN, Afrane YA, Amenya DA, Bonizzoni M, Menge DM, Zhou G, Zhong D, Vardo-Zalik AM, Githeko AK, Yan G (2009) High prevalence of asymptomatic *Plasmodium falciparum* infections in a highland area of western Kenya: a cohort study. *The Journal of Infectious Diseases* 200:66-74.

70. Ogutu B, Tiono AB, Makanga M, Premji Z, Gbadoe AD, Ubben D, Marrast AC, Gaye O (2010) Treatment of asymptomatic carriers with artemether-lumefantrine: an opportunity to reduce the burden of malaria? *Malaria Journal* 9:30.

71. Okell LC, Bousema T, Griffin JT, Ouedraogo AL, Ghani AC, Drakeley CJ (2012) Factors determining the occurrence of submicroscopic malaria infections and their relevance for control. *Nature Communications* 3:1237.

72. Fairhurst RM, Nayyar GM, Breman JG, Hallett R, Vennerstrom JL, Duong S, Ringwald P, Wellems TE, Plowe CV, Dondorp AM (2012) Artemisinin-resistant malaria: research challenges, opportunities, and public health implications. *The American Journal of Tropical Medicine and Hygiene* 87:231-241.

73. Douglas NM, Anstey NM, Angus BJ, Nosten F, Price RN (2010) Artemisinin combination therapy for vivax malaria. *The Lancet Infectious Diseases* 10:405-416.

74. Graves PM, Gelband H, Garner P (2012) Primaquine for reducing *Plasmodium falciparum* transmission. *Cochrane Database of Systematic Reviews (Online)* 9:CD008152.

75. Baird JK (2007) Neglect of *Plasmodium vivax* malaria. *Trends in Parasitology* 23:533-539.

76. Howes RE, Piel FB, Patil AP, Nyangiri OA, Gething PW, Dewi M, Hogg MM, Battle KE, Padilla CD, Baird JK, Hay SI (2012) G6PD deficiency prevalence and estimates of affected populations in malaria endemic countries: a geostatistical model-based map. *PLoS Medicine* 9:e1001339.

77. Howes RE, Battle KE, Satyagraha AW, Baird JK, Hay SI (2013) G6PD deficiency: global distribution, genetic variants and primaquine therapy. *Advances in Parasitology* 81:133-201.

78. Shekalaghe SA, ter Braak R, Daou M, Kavishe R, van den Bijllaardt W, van den Bosch S, Koenderink JB, Luty AJ, Whitty CJ, Drakeley C, Sauerwein RW, Bousema T (2010) In Tanzania, hemolysis after a single dose of primaquine coadministered with an

artemisinin is not restricted to glucose-6-phosphate dehydrogenase-deficient (G6PD A-) individuals. *Antimicrobial Agents and Chemotherapy* 54:1762-1768.

79. Alving AS, Johnson CF, Tarlov AR, Brewer GJ, Kellermeyer RW, Carson PE (1960) Mitigation of the haemolytic effect of primaquine and enhancement of its action against exoerythrocytic forms of the Chesson strain of *Plasmodium vivax* by intermittent regimens of drug administration: a preliminary report. *Bulletin of the World Health Organization* 22:621-631.

80. Baird JK, Surjadjaja C (2011) Consideration of ethics in primaquine therapy against malaria transmission. *Trends in Parasitology* 27:11-16.

81. Alonso PL, Lindsay SW, Armstrong Schellenberg JR, Keita K, Gomez P, Shenton FC, Hill AG, David PH, Fegan G, Cham K, et al. (1993) A malaria control trial using insecticide-treated bed nets and targeted chemoprophylaxis in a rural area of The Gambia, west Africa. 6. The impact of the interventions on mortality and morbidity from malaria. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 87 Suppl 2:37-44.

82. Wills AB, Smith SC, Anshebo GY, Graves PM, Endeshaw T, Shargie EB, Damte M, Gebre T, Mosher AW, Patterson AE, Tesema YB, Richards FO, Jr., Emerson PM (2013) Physical durability of PermaNet 2.0 long-lasting insecticidal nets over three to 32 months of use in Ethiopia. *Malaria Journal* 12:242.

83. Mejia P, Teklehaimanot HD, Tesfaye Y, Teklehaimanot A (2013) Physical condition of Olyset(R) nets after five years of utilization in rural western Kenya. *Malaria Journal* 12:158.

84. Pulford J, Hetzel MW, Bryant M, Siba PM, Mueller I (2011) Reported reasons for not using a mosquito net when one is available: a review of the published literature. *Malaria Journal* 10:83.

85. Singh M, Brown G, Rogerson SJ (2013) Ownership and use of insecticide-treated nets during pregnancy in sub-Saharan Africa: a review. *Malaria Journal* 12:268.

86. Hawley WA, Phillips-Howard PA, ter Kuile FO, Terlouw DJ, Vulule JM, Ombok M, Nahlen BL, Gimnig JE, Kariuki SK, Kolczak MS, Hightower AW (2003) Community-wide effects of permethrin-treated bed nets on child mortality and malaria morbidity in western Kenya. *The American Journal of Tropical Medicine and Hygiene* 68:121-127.

87. Howard SC, Omumbo J, Nevill C, Some ES, Donnelly CA, Snow RW (2000) Evidence for a mass community effect of insecticide-treated bednets on the incidence of malaria on the Kenyan coast. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 94:357-360.

88. Gimnig JE, Kolczak MS, Hightower AW, Vulule JM, Schoute E, Kamau L, Phillips-Howard PA, ter Kuile FO, Nahlen BL, Hawley WA (2003) Effect of permethrin-treated bed nets on the spatial distribution of malaria vectors in western Kenya. *The American Journal of Tropical Medicine and Hygiene* 68:115-120.

89. Kim D, Fedak K, Kramer R (2012) Reduction of malaria prevalence by indoor residual spraying: a meta-regression analysis. *The American Journal of Tropical Medicine and Hygiene* 87:117-124.

90. Pluess B, Tanser FC, Lengeler C, Sharp BL (2010) Indoor residual spraying for preventing malaria. *Cochrane Database of Systematic Reviews (Online)*:CD006657.

91. Fullman N, Burstein R, Lim SS, Medlin C, Gakidou E (2013) Nets, spray or both? The effectiveness of insecticide-treated nets and indoor residual spraying in reducing malaria morbidity and child mortality in sub-Saharan Africa. *Malaria Journal* 12:62.

92. West PA, Protopopoff N, Wright A, Kivaju Z, Tigererwa R, Mosha FW, Kisinza W, Rowland M, Kleinschmidt I (2014) Indoor residual spraying in combination with insecticide-treated nets compared to insecticide-treated nets alone for protection against malaria: a cluster randomised trial in Tanzania. *PLoS Medicine* 11:e1001630.

93. Perry HN, McDonnell SM, Alemu W, Nsubuga P, Chungong S, Otten MW, Jr., Lusamba-Dikassa PS, Thacker SB (2007) Planning an integrated disease surveillance and response system: a matrix of skills and activities. *BMC Medicine* 5:24.

94. Rowe AK, Kachur SP, Yoon SS, Lynch M, Slutsker L, Steketee RW (2009) Caution is required when using health facility-based data to evaluate the health impact of malaria control efforts in Africa. *Malaria Journal* 8:209.

95. Yukich JO, Butts J, Miles M, Berhane Y, Nahusenay H, Malone JL, Dissanayake G, Reithinger R, Keating J (2014) A description of malaria sentinel surveillance: a case study in Oromia Regional State, Ethiopia *Malaria Journal* 13:88.

96. Basinga P, Gertler PJ, Binagwaho A, Soucat AL, Sturdy J, Vermeersch CM (2011) Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. *Lancet* 377:1421-1428.

97. Witter S, Fretheim A, Kessy FL, Lindahl AK (2012) Paying for performance to improve the delivery of health interventions in low- and middle-income countries. *Cochrane Database of Systematic Reviews (Online)* 2:CD007899.

98. Gething PW, Noor AM, Gikandi PW, Ogara EA, Hay SI, Nixon MS, Snow RW, Atkinson PM (2006) Improving imperfect data from health management information systems in Africa using space-time geostatistics. *PLoS Medicine* 3:e271.

99. Yukich J, Briet O, Bretscher MT, Bennett A, Lemma S, Berhane Y, Eisele TP, Keating J, Smith T (2012) Estimating *Plasmodium falciparum* transmission rates in low-endemic settings using a combination of community prevalence and health facility data. *PLoS ONE* 7:e42861.

100. Sturrock HJW, Hsiang MS, Cohen JM, Smith DL, Greenhouse B, Bousema T, Gosling RD (2013) Targeting asymptomatic malaria infections: active surveillance in control and elimination. *PLoS Medicine* 10:e1001467.

101. Sturrock HJ, Novotny JM, Kunene S, Dlamini S, Zulu Z, Cohen JM, Hsiang MS, Greenhouse B, Gosling RD (2013) Reactive case detection for malaria elimination: real-life experience from an ongoing program in Swaziland. *PLoS ONE* 8:e63830.

102. Littrell M, Sow GD, Ngom A, Ba M, Mboup BM, Dieye Y, Mutombo B, Earle D, Steketee RW (2013) Case investigation and reactive case detection for malaria elimination in northern Senegal. *Malaria Journal* 12:331.

103. Demographic and Health Surveys. http://www.measuredhs.com/aboutsurveys/dhs/start.cfm

104. Malaria Indicator Survey. http://www.malariasurveys.org/toolkit.cfm

105. MEASURE Evaluation, MEASURE DHS, President's Malaria Initiative, Roll Back Malaria Partnership, UNICEF, World Health Organization (2013) Household survey indicators for malaria control. http://www.rollbackmalaria.org/toolbox/docs/rbmtoolbox/tool_HouseholdSurveyIndicatorsForMalariaControl.pdf

106. UNICEF (2006) Multiple Indicator Cluster Survey Manual 2005. Monitoring the Situation of Children and Women: UNICEF.

107. Roca-Feltrer A, Lalloo DG, Phiri K, Terlouw DJ (2012) Rolling Malaria Indicator Surveys (rMIS): a potential district-level malaria monitoring and evaluation (M&E) tool for program managers. *The American Journal of Tropical Medicine and Hygiene* 86:96-98.

108. Sutter VA, Zuniga H (1942) A malaria survey of El Salvador, Central America. *The American Journal of Tropical Medicine and Hygiene* 22:387-398.

109. Castellanos JB, Murrieta C, Lassman G, Ortiz C (1949) A malaria reconnaissance of the state of Veracruz, Mexico. *The American Journal of Tropical Medicine and Hygiene* s1-29:23-35.

110. Ndyomugyenyi R, Kroeger A (2007) Using schoolchildren's reports of bed net use monitored by schoolteachers as a proxy of community coverage in malaria endemic areas of Uganda. *Tropical Medicine and International Health* 12:230-237.

111. Stevenson JC, Stresman GH, Gitonga CW, Gillig J, Owaga C, Marube E, Odongo W, Okoth A, China P, Oriango R, Brooker SJ, Bousema T, Drakeley C, Cox J (2013) Reliability of school surveys in estimating geographic variation in malaria transmission in the western Kenyan highlands. *PLoS ONE* 8:e77641.

112. Center for Disease Control (1986) Comprehensive plan for epidemic surveillance. Atlanta: Centres for Disease Control.

113. Cullen JR, Chitprarop U, Doberstyn EB, Sombatwattanangkul K (1984) An epidemiological early warning system for malaria control in northern Thailand. *Bulletin of the World Health Organization* 62:107-114.

114. Roll Back Malaria Partnership (2001) Malaria Early Warning Systems: A Framework for Field Research in Africa. WHO/CDS/RBM/2001.32 http://www.rbm.who.int/cmc_upload/0/000/014/807/mews2.pdf

115. World Health Organization (2009) Control of epidemic meningococcal disease. WHO practical guidelines. 2nd edition. Geneva: World Health Organization. WHO/EMC/BAC/98/3/EN http://www.who.int/csr/resources/publications/meningitis/WHO_EMC_BAC_98_3_EN/en/

116. Agier L, Broutin H, Bertherat E, Djingarey MH, Lingani C, Perea W, Hugonnet S (2013) Timely detection of bacterial meningitis epidemics at district level: a study in three countries of the African Meningitis Belt. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 107:30-36.

117. Stroup DF, Williamson GD, Herndon JL, Karon JM (1989) Detection of aberrations in the occurrence of notifiable diseases surveillance data. *Statistics in Medicine* 8:323-329; discussion 331-322.

118. Nobre FF, Stroup DF (1994) A monitoring system to detect changes in public health surveillance data. *International Journal of Epidemiology* 23:408-418.

119. Frisen M (1992) Evaluations of methods for statistical surveillance. *Statistics in Medicine* 11:1489-1502.

120. Hay SI, Simba M, Busolo M, Noor AM, Guyatt HL, Ochola SA, Snow RW (2002) Defining and detecting malaria epidemics in the highlands of western Kenya. *Emerging Infectious Diseases* 8:555-562.

121. Teklehaimanot HD, Schwatrz J, Teklehaimanot A, Lipsitch M (2004) Alert threshold algorithms and malaria epidemic detection. *Emerging Infectious Diseases* 10:1220-1226.

122. Randrianasolo L, Raoelina Y, Ratsitorahina M, Ravolomanana L, Andriamandimby S, Heraud JM, Rakotomanana F, Ramanjato R, Randrianarivo-Solofoniaina AE, Richard V (2010) Sentinel surveillance system for early outbreak detection in Madagascar. *BMC Public Health* 10:31.

123. Kamanga A, Moono P, Stresman G, Mharakurwa S, Shiff C (2010) Rural health centres, communities and malaria case detection in Zambia using mobile telephones: a means to detect potential reservoirs of infection in unstable transmission conditions. *Malaria Journal* 9:96.

124. Connor SJ, Thomson MC, Flasse SP, Perryman AH (1998) Environmental information systems in malaria risk mapping and epidemic forecasting. *Disasters* 22:39-56.

125. Abeku TA, Hay SI, Ochola S, Langi P, Beard B, de Vlas SJ, Cox J (2004) Malaria epidemic early warning and detection in African highlands. *Trends in Parasitology* 20:400-405.

126. Teklehaimanot HD, Schwartz J, Teklehaimanot A, Lipsitch M (2004) Weather-based prediction of *Plasmodium falciparum* malaria in epidemic-prone regions of Ethiopia II. Weather-based prediction systems perform comparably to early detection systems in identifying times for interventions. *Malaria Journal* 3:44.

127. WHO (2002) Web-based tool for early warning of malaria epidemics in Africa: monitoring current rainfall anomalies in zones at epidemic risk. *Weekly Epidemiological Record* 77:276.

128. Grover-Kopec E, Kawano M, Klaver RW, Blumenthal B, Ceccato P, Connor SJ (2005) An online operational rainfall-monitoring resource for epidemic malaria early warning systems in Africa. *Malaria Journal* 4:6.

129. Wimberly MC, Midekisa A, Semuniguse P, Teka H, Henebry GM, Chuang TW, Senay GB (2012) Spatial synchrony of malaria outbreaks in a highland region of Ethiopia. *Tropical Medicine and International Health* 17:1192-1201.

130. Midekisa A, Senay G, Henebry GM, Semuniguse P, Wimberly MC (2012) Remote sensing-based time series models for malaria early warning in the highlands of Ethiopia. *Malaria Journal* 11:165.

131. Adimi F, Soebiyanto RP, Safi N, Kiang R (2010) Towards malaria risk prediction in Afghanistan using remote sensing. *Malaria Journal* 9:125.

132. Checchi F, Cox J, Balkan S, Tamrat A, Priotto G, Alberti KP, Guthmann J-P (2006) Malaria epidemics and interventions, Kenya, Burundi, southern Sudan, and Ethiopia, 1999-2004. *Emerging Infectious Diseases* 12:1477-1485.

133. Brown V, Abdir Issak M, Rossi M, Barboza P, Paugam A (1998) Epidemic of malaria in north-eastern Kenya. *Lancet* 352:1356-1357.

134. Negash K, Kebede A, Medhin A, Argaw D, Babaniyi O, Guintran JO, Delacollette C (2005) Malaria epidemics in the highlands of Ethiopia. *East African Medical Journal* 82:186-192.

135. Henning KJ (2004) What is syndromic surveillance? *MMWR Morbidity and mortality weekly report* 53 Suppl:5-11.

136. Centers for Disease Control (2004) Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC working group. RR-5 http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5305a1.htm

137. Kool JL, Paterson B, Pavlin BI, Durrheim D, Musto J, Kolbe A (2012) Pacific-wide simplified syndromic surveillance for early warning of outbreaks. *Global Public Health* 7:670-681.

138. John TJ, Rajappan K, Arjunan KK (2004) Communicable diseases monitored by disease surveillance in Kottayam district, Kerala state, India. *Indian Journal of Medical Research* 120:86-93.

139. Rosewell A, Ropa B, Randall H, Dagina R, Hurim S, Bieb S, Datta S, Ramamurthy S, Mola G, Zwi AB, Ray P, MacIntyre CR (2013) Mobile phone-based syndromic surveillance system, Papua New Guinea. *Emerging Infectious Diseases* 19:1811-1818.

140. Durrheim DN, Harris BN, Speare R, Billinghurst K (2001) The use of hospital-based nurses for the surveillance of potential disease outbreaks. *Bulletin of the World Health Organization* 79:22-27.

141. Carme B, Sobesky M, Biard MH, Cotellon P, Aznar C, Fontanella JM (2003) Non-specific alert system for dengue epidemic outbreaks in areas of endemic malaria. A hospital-based evaluation in Cayenne (French Guiana). *Epidemiology and Infection* 130:93-100.

142. Chretien JP, Burkom HS, Sedyaningsih ER, Larasati RP, Lescano AG, Mundaca CC, Blazes DL, Munayco CV, Coberly JS, Ashar RJ, Lewis SH (2008) Syndromic surveillance: adapting innovations to developing settings. *PLoS Medicine* 5:e72.

143. Cooper DL, Verlander NQ, Elliot AJ, Joseph CA, Smith GE (2009) Can syndromic thresholds provide early warning of national influenza outbreaks? *Journal of Public Health (Oxford)* 31:17-25.

144. Besculides M, Heffernan R, Mostashari F, Weiss D (2005) Evaluation of school absenteeism data for early outbreak detection, New York City. *BMC Public Health* 5:105.

145. Mook P, Joseph C, Gates P, Phin N (2007) Pilot scheme for monitoring sickness absence in schools during the 2006/07 winter in England: can these data be used as a proxy for influenza activity? *Euro Surveillance* 12:E11-12.

146. Cheng CKY, Channarith H, Cowling BJ (2013) Potential use of school absenteeism record for disease surveillance in developing countries, case study in Cambodia. *PLoS ONE* 8:e76859.

147. Yan WR, Nie SF, Xu B, Dong HJ, Palm L, Diwan VK (2012) Establishing a web-based integrated surveillance system for early detection of infectious disease epidemic in rural China: a field experimental study. *BMC Medical Informatics and Decision Making* 12:4.

148. Yan W, Palm L, Lu X, Nie S, Xu B, Zhao Q, Tao T, Cheng L, Tan L, Dong H, Diwan VK (2013) ISS--an electronic syndromic surveillance system for infectious disease in rural China. *PLoS ONE* 8:e62749.

149. Enserink R, Noel H, Friesema IH, de Jager CM, Kooistra-Smid AM, Kortbeek LM, Duizer E, van der Sande MA, Smit HA, Pelt W (2012) The KIzSS network, a sentinel surveillance system for infectious diseases in day care centers: study protocol. *BMC Infectious Diseases* 12:259.

150. Andersson T, Bjelkmar P, Hulth A, Lindh J, Stenmark S, Widerstrom M (2013) Syndromic surveillance for local outbreak detection and awareness: evaluating outbreak signals of acute gastroenteritis in telephone triage, web-based queries and over-the-counter pharmacy sales. *Epidemiology and Infection*:1-11.

151. Lawpoolsri S, Khamsiriwatchara A, Liulark W, Taweeseneepitch K, Sangvichean A, Thongprarong W, Kaewkungwal J, Singhasivanon P (2014) Real-time monitoring of school absenteeism to enhance disease surveillance: a pilot study of a mobile electronic reporting system. *JMIR mHealth and uHealth* 2:e22.

152. World Health Organization DoCDSaR (1999) WHO Recommended Surveillance Standards. 2nd edition. Geneva: World Health Organization. WHO/CDS/CSR/ISR/99.2 http://data.unaids.org/publications/irc-pub04/surveillancestandards_en.pdf

153. Bounoure F, Beaudeau P, Mouly D, Skiba M, Lahiani-Skiba M (2011) Syndromic surveillance of acute gastroenteritis based on drug consumption. *Epidemiology and Infection* 139:1388-1395.

154. Patwardhan A, Bilkovski R (2012) Comparison: Flu prescription sales data from a retail pharmacy in the US with Google Flu trends and US ILINet (CDC) data as flu activity indicator. *PLoS ONE* 7:e43611.

155. Milinovich GJ, Williams GM, Clements AC, Hu W (2013) Internet-based surveillance systems for monitoring emerging infectious diseases. *The Lancet Infectious Diseases*.

156. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L (2009) Detecting influenza epidemics using search engine query data. *Nature* 457:1012-1014.

157. Chan EH, Sahai V, Conrad C, Brownstein JS (2011) Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance. *PLoS Neglected Tropical Diseases* 5:e1206.

158. Ocampo AJ, Chunara R, Brownstein JS (2013) Using search queries for malaria surveillance, Thailand. *Malaria Journal* 12:390.

159. Signorini A, Segre AM, Polgreen PM (2011) The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PLoS ONE* 6:e19467.

160. Broniatowski DA, Paul MJ, Dredze M (2013) National and local influenza surveillance through Twitter: An analysis of the 2012-2013 influenza epidemic. *PLoS ONE* 8:e83672.

161. Chunara R, Andrews JR, Brownstein JS (2012) Social and news media enable estimation of epidemiological patterns early in the 2010 Haitian cholera outbreak. *The American Journal of Tropical Medicine and Hygiene* 86:39-45.

162. World Development Indicators: primary school enrollment. http://data.worldbank.org/indicator/SE.PRM.NENR

163. Some ES (1994) Effects and control of highland malaria epidemic in Uasin Gishu District, Kenya. *East African Medical Journal* 71:2-8.

164. D'Acremont V, Kilowoko M, Kyungu E, Philipina S, Sangu W, Kahama-Maro J, Lengeler C, Cherpillod P, Kaiser L, Genton B (2014) Beyond malaria--causes of fever in outpatient Tanzanian children. *New England Journal of Medicine* 370:809-817.

165. Smith DL, Dushoff J, Snow RW, Hay SI (2005) The entomological inoculation rate and *Plasmodium falciparum* infection in African children. *Nature* 438:492-495.

166. Beier JC, Killeen GF, Githure JI (1999) Short report: entomologic inoculation rates and *Plasmodium falciparum* malaria prevalence in Africa. *The American Journal of Tropical Medicine and Hygiene* 61:109-113.

167. Ochola LB, Vounatsou P, Smith T, Mabaso ML, Newton CR (2006) The reliability of diagnostic techniques in the diagnosis and management of malaria in the absence of a gold standard. *The Lancet Infectious Diseases* 6:582-588.

168. Ashton RA, Kefyalew T, Tesfaye G, Counihan H, Yadeta D, Cundill B, Reithinger R, Kolaczinski JH (2010) Performance of three multi-species rapid diagnostic tests for diagnosis of *Plasmodium falciparum* and *Plasmodium vivax* malaria in Oromia Regional State, Ethiopia. *Malaria Journal* 9:297.

169. Endeshaw T, Gebre T, Ngondi J, Graves PM, Shargie EB, Ejigsemahu Y, Ayele B, Yohannes G, Teferi T, Messele A, Zerihun M, Genet A, Mosher AW, Emerson PM, Richards FO (2008) Evaluation of light microscopy and rapid diagnostic test for the detection of malaria under operational field conditions: a household survey in Ethiopia. *Malaria Journal* 7:118.

170. Mayxay M, Pukrittayakamee S, Chotivanich K, Looareesuwan S, White NJ (2001) Persistence of *Plasmodium falciparum* HRP-2 in successfully treated acute falciparum malaria. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 95:179-182.

171. Swarthout TD, Counihan H, Senga RKK, van den Broek I (2007) Paracheck-Pf accuracy and recently treated *Plasmodium falciparum* infections: is there a risk of over-diagnosis? *Malaria Journal* 6:58.

172. Gillet P, Mori M, Van Esbroeck M, Van den Ende J, Jacobs J (2009) Assessment of the prozone effect in malaria rapid diagnostic tests. *Malaria Journal* 8:271.

173. Snounou G, Viriyakosol S, Zhu XP, Jarra W, Pinheiro L, do Rosario VE, Thaithong S, Brown KN (1993) High sensitivity of detection of human malaria parasites by the use of nested polymerase chain reaction. *Molecular and Biochemical Parasitology* 61:315-320.

174. Snounou G, Viriyakosol S, Jarra W, Thaithong S, Brown KN (1993) Identification of the four human malaria parasite species in field samples by the polymerase chain reaction and detection of a high prevalence of mixed infections. *Molecular and Biochemical Parasitology* 58:283-292.

175. Hsiang MS, Lin M, Dokomajilar C, Kemere J, Pilcher CD, Dorsey G, Greenhouse B (2010) PCR-based pooling of dried blood spots for detection of malaria parasites: optimization

and application to a cohort of Ugandan children. *Journal of Clinical Microbiology* 48:3539-3543.

176. Hwang SY, Kim SH, Lee GY, Hang VT, Moon CS, Shin JH, Koo WL, Kim SY, Park HJ, Park HO, Kho WG (2011) A novel real-time PCR assay for the detection of *Plasmodium falciparum* and *Plasmodium vivax* malaria in low parasitized individuals. *Acta Tropica* 120:40-45.

177. Abdul-Ghani R, Al-Mekhlafi AM, Karanis P (2012) Loop-mediated isothermal amplification (LAMP) for malarial parasites of humans: would it come to clinical reality as a point-of-care test? *Acta Tropica* 122:233-240.

178. Notomi T, Okayama H, Masubuchi H, Yonekawa T, Watanabe K, Amino N, Hase T (2000) Loop-mediated isothermal amplification of DNA. *Nucleic Acids Research* 28:E63.

179. Hopkins H, Gonzalez IJ, Polley SD, Angutoko P, Ategeka J, Asiimwe C, Agaba B, Kyabayinze DJ, Sutherland CJ, Perkins MD, Bell D (2013) Highly sensitive detection of malaria parasitemia in a malaria-endemic setting: performance of a new loop-mediated isothermal amplification kit in a remote clinic in Uganda. *The Journal of Infectious Diseases* 208:645-652.

180. Collins WE, Warren M, Skinner JC (1971) Serological malaria survey in the Ethiopian highlands. *The American Journal of Tropical Medicine and Hygiene* 20:199-205.

181. Bruce-Chwatt LJ, Draper CC, Konfortion P (1973) Seroepidemiological evidence of eradication of malaria from Mauritius. *Lancet* 302:547-551.

182. Bouchaud O, Cot M, Kony S, Durand R, Schiemann R, Ralaimazava P, Coulaud JP, Le Bras J, Deloron P (2005) Do African immigrants living in France have long-term malarial immunity? *The American Journal of Tropical Medicine and Hygiene* 72:21-25.

183. Migot F, Chougnet C, Henzel D, Dubois B, Jambou R, Fievet N, Deloron P (1995) Anti-malaria antibody-producing B cell frequencies in adults after a *Plasmodium falciparum* outbreak in Madagascar. *Clinical and Experimental Immunology* 102:529-534.

184. Corran PH, Cook J, Lynch C, Leendertse H, Manjurano A, Griffin J, Cox J, Abeku T, Bousema T, Ghani AC, Drakeley C, Riley E (2008) Dried blood spots as a source of anti-malarial antibodies for epidemiological studies. *Malaria Journal* 7:195.

185. Corran P, Coleman P, Riley E, Drakeley C (2007) Serology: a robust indicator of malaria transmission intensity? *Trends in Parasitology* 23:575-582.

186. Drakeley CJ, Corran PH, Coleman PG, Tongren JE, McDonald SLR, Carneiro I, Malima R, Lusingu J, Manjurano A, Nkya WMM, Lemnge MM, Cox J, Reyburn H, Riley EM (2005) Estimating medium- and long-term trends in malaria transmission by using serological markers of malaria exposure. *Proceedings of the National Academy of Sciences of the United States of America* 102:5108-5113.

187. Cook J, Reid H, Iavro J, Kuwahata M, Taleo G, Clements A, McCarthy J, Vallely A, Drakeley C (2010) Using serological measures to monitor changes in malaria transmission in Vanuatu. *Malaria Journal* 9:169.

188. Cook J, Speybroeck N, Sochanta T, Somony H, Sokny M, Claes F, Lemmens K, Theisen M, Soares IS, D'Alessandro U, Coosemans M, Erhart A (2012) Sero-epidemiological evaluation of changes in *Plasmodium falciparum* and *Plasmodium vivax* transmission patterns over the rainy season in Cambodia. *Malaria Journal* 11:86.

189. Wilson S, Booth M, Jones FM, Mwatha JK, Kimani G, Kariuki HC, Vennervald BJ, Ouma JH, Muchiri E, Dunne DW (2007) Age-adjusted *Plasmodium falciparum* antibody levels in school-aged children are a stable marker of microgeographical variations in exposure to *Plasmodium* infection. *BMC Infectious Diseases* 7:67.

190. Oduro AR, Conway DJ, Schellenberg D, Satoguina J, Greenwood BM, Bojang KA (2013) Seroepidemiological and parasitological evaluation of the heterogeneity of malaria infection in the Gambia. *Malaria Journal* 12:222.

191. Bousema T, Youssef RM, Cook J, Cox J, Alegana VA, Amran J, Noor AM, Snow RW, Drakeley C (2010) Serologic markers for detecting malaria in areas of low endemicity, Somalia, 2008. *Emerging Infectious Diseases* 16:392-399.

192. Takem EN, Affara M, Amambua-Ngwa A, Okebe J, Ceesay SJ, Jawara M, Oriero E, Nwakanma D, Pinder M, Clifford C, Taal M, Sowe M, Suso P, Mendy A, Mbaye A, Drakeley C, D'Alessandro U (2013) Detecting foci of malaria transmission with school surveys: A pilot study in the Gambia. *PLoS ONE* 8:e67108.

193. Noor AM, Mohamed MB, Mugyenyi CK, Osman MA, Guessod HH, Kabaria CW, Ahmed IA, Nyonda M, Cook J, Drakeley CJ, Mackinnon MJ, Snow RW (2011) Establishing the extent

of malaria transmission and challenges facing pre-elimination in the Republic of Djibouti. *BMC Infectious Diseases* 11:121.

194. Hsiang MS, Hwang J, Kunene S, Drakeley C, Kandula D, Novotny J, Parizo J, Jensen T, Tong M, Kemere J, Dlamini S, Moonen B, Angov E, Dutta S, Ockenhouse C, Dorsey G, Greenhouse B (2012) Surveillance for malaria elimination in Swaziland: a national cross-sectional study using pooled PCR and serology. *PLoS ONE* 7:e29550.

195. Jima D, Getachew A, Bilak H, Steketee RW, Emerson PM, Graves PM, Gebre T, Reithinger R, Hwang J, Ethiopia Malaria Indicator Survey Working G (2010) Malaria indicator survey 2007, Ethiopia: coverage and use of major malaria prevention and control interventions. *Malaria Journal* 9:58.

196. Graves PM, Richards FO, Ngondi J, Emerson PM, Shargie EB, Endeshaw T, Ceccato P, Ejigsemahu Y, Mosher AW, Hailemariam A, Zerihun M, Teferi T, Ayele B, Mesele A, Yohannes G, Tilahun A, Gebre T (2009) Individual, household and environmental risk factors for malaria infection in Amhara, Oromia and SNNP regions of Ethiopia. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 103:1211-1220.

197. Mueller I, Zimmerman PA, Reeder JC (2007) *Plasmodium malariae* and *Plasmodium ovale*--the "bashful" malaria parasites. *Trends in Parasitology* 23:278-283.

198. Krafsur ES, Armstrong JC (1982) Epidemiology of *Plasmodium malariae* infection in Gambela, Ethiopia. *Parassitologia* 24:105-120.

199. Alemu A, Fuehrer HP, Getnet G, Tessema B, Noedl H (2013) *Plasmodium ovale curtisi* and *Plasmodium ovale wallikeri* in North-West Ethiopia. *Malaria Journal* 12:346.

200. Lee PW, Liu CT, Rampao HS, do Rosario VE, Shaio MF (2010) Pre-elimination of malaria on the island of Príncipe. *Malaria Journal* 9:26.

201. Chaves LF, Kaneko A, Taleo G, Pascual M, Wilson ML (2008) Malaria transmission pattern resilience to climatic variability is mediated by insecticide-treated nets. *Malaria Journal* 7:100.

202. Abeku T, Van Oortmarssen GJ, Borsboom G, De Vlas SJ, Habbema JDF (2003) Spatial and temporal variations of malaria epidemic risk in Ethiopia: factors involved and implications. *Acta Tropica* 87:331-340.

203. Mueller I, Galinski MR, Baird JK, Carlton JM, Kochar DK, Alonso PL, del Portillo HA (2009) Key gaps in the knowledge of *Plasmodium vivax,* a neglected human malaria parasite. *The Lancet Infectious Diseases* 9:555-566.

204. Battle KE, Karhunen MS, Bhatt S, Gething PW, Howes RE, Golding N, Van Boeckel TP, Messina JP, Shanks GD, Smith DL, Baird JK, Hay SI (2014) Geographical variation in *Plasmodium vivax* relapse. *Malaria Journal* 13:144.

205. Kenea O, Balkew M, Gebre-Michael T (2011) Environmental factors associated with larval habitats of anopheline mosquitoes (Diptera: Culicidae) in irrigation and major drainage areas in the middle course of the Rift Valley, central Ethiopia. *Journal of Vector Borne Diseases* 48:85-92.

206. Animut A, Balkew M, Gebre-Michael T, Lindtjorn B (2013) Blood meal sources and entomological inoculation rates of anophelines along a highland altitudinal transect in south-central Ethiopia. *Malaria Journal* 12:76.

207. Guerra CA, Gikandi PW, Tatem AJ, Noor AM, Smith DL, Hay SI, Snow RW (2008) The limits and intensity of *Plasmodium falciparum* transmission: implications for malaria control and elimination worldwide. *PLoS Medicine* 5:e38.

208. Federal Democratic Republic of Ethiopia MoH (2009) National Strategic Plan for Malaria Prevention, Control and Elimination in Ethiopia, 2011-2015. Addis Ababa, Ethiopia.

209. Ghebreyesus TA, Haile M, Witten KH, Getachew A, Yohannes AM, Yohannes M, Teklehaimanot HD, Lindsay SW, Byass P (1999) Incidence of malaria among children living near dams in northern Ethiopia: community based incidence survey. *BMJ (Clinical Research Ed)* 319:663-666.

210. Lautze J, McCartney M, Kirshen P, Olana D, Jayasinghe G, Spielman A (2007) Effect of a large dam on malaria risk: the Koka reservoir in Ethiopia. *Tropical Medicine and International Health* 12:982-989.

211. Kibret S, Alemu Y, Boelee E, Tekie H, Alemu D, Petros B (2009) The impact of a small-scale irrigation scheme on malaria transmission in Ziway area, Central Ethiopia. *Tropical Medicine and International Health* 15:41-50.

212. Mosquito vectors in Ethiopia. http://www.map.ox.ac.uk/explore/countries/ETH/#!mosquito-vectors

213. Federal Democratic Republic of Ethiopia MoH (2012) National Malaria Guidelines, 3rd edition. Addis Ababa, Ethiopia.

214. Olana D, Chibsa S, Teshome D, Mekasha A, Graves PM, Reithinger R (2011) Malaria, Oromia Regional State, Ethiopia, 2001-2006. *Emerging Infectious Diseases* 17:1336-1337.

215. Fontaine RE, Najjar AE, Prince JS (1961) The 1958 malaria epidemic in Ethiopia. *The American Journal of Tropical Medicine and Hygiene* 10:795-803.

216. Warsame M, Wernsdorfer WH, Huldt G, Bjorkman A (1995) An epidemic of *Plasmodium falciparum* malaria in Balcad, Somalia, and its causation. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 89:142-145.

217. Lindblade KA, Walker ED, Onapa AW, Katungu J, Wilson ML (1999) Highland malaria in Uganda: prospective analysis of an epidemic associated with El Nino. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 93:480-487.

218. Etchegorry MG, Matthys F, Galinski M, White NJ, Nosten F (2001) Malaria epidemic in Burundi. *Lancet* 357:1046-1047.

219. Rogier C, Pradines B, Bogreau H, Koeck JL, Kamil MA, Mercereau-Puijalon O (2005) Malaria epidemic and drug resistance, Djibouti. *Emerging Infectious Diseases* 11:317-321.

220. Das P (2003) Ethiopia faces severe malaria epidemic. WHO predicts 15 million people could be infected. *Lancet* 362:2071.

221. Hay S, Renshaw M, Ochola SA, Noor AM, Snow RW (2003) Performance of forecasting, warning and detection of malaria epidemics in the highlands of western Kenya. *Trends in Parasitology* 19:394-399.

222. Kovats RS, Bouma MJ, Hajat S, Worrall E, Haines A (2003) El Nino and health. *Lancet* 362:1481-1489.

223. Ketema T, Bacha K, Birhanu T, Petros B (2009) Chloroquine-resistant *Plasmodium vivax* malaria in Serbo town, Jimma zone, south-west Ethiopia. *Malaria Journal* 8:177.

224. Yeshiwondim AK, Tekle AH, Dengela DO, Yohannes AM, Teklehaimanot A (2010) Therapeutic efficacy of chloroquine and chloroquine plus primaquine for the treatment of *Plasmodium vivax* in Ethiopia. *Acta Tropica* 113:105-113.

225. Teka H, Petros B, Yamuah L, Tesfaye G, Elhassan I, Muchohi S, Kokwaro G, Aseffa A, Engers H (2008) Chloroquine-resistant *Plasmodium vivax* malaria in Debre Zeit, Ethiopia. *Malaria Journal* 7:220.

226. Yohannes AM, Teklehaimanot A, Bergqvist Y, Ringwald P (2011) Confirmed vivax resistance to chloroquine and effectiveness of artemether-lumefantrine for the treatment of vivax malaria in Ethiopia. *The American Journal of Tropical Medicine and Hygiene* 84:137-140.

227. Ethiopian Central Statistics Agency, ICF Macro (2011) Ethiopia Demographic and Health Survey 2011, Preliminary Report. Addis Ababa, Ethiopia and Calverton, Maryland, USA: Central Statistics Agency and IRC Macro.

228. The Ethiopian Health and Nutrition Research Institute & partners (2012) Ethiopia National Malaria Indicator Survey 2011.

229. Jima D, Wondabeku M, Alemu A, Teferra A, Awel N, Deressa W, Adissie A, Tadesse Z, Gebre T, Mosher AW, Richards FO, Graves PM (2012) Analysis of malaria surveillance data in Ethiopia: what can be learned from the integrated disease surveillance and response system? *Malaria Journal* 11:330.

230. Federal Democratic Republic of Ethiopia MoH (2009) Guidelines for malaria epidemic prevention and control in Ethiopia. 3rd edition: Addis Ababa, Ethiopia.

231. O'Meara WP, Mangeni JN, Steketee R, Greenwood B (2010) Changes in the burden of malaria in sub-Saharan Africa. *The Lancet Infectious Diseases* 10:545-555.

232. Federal Democratic Republic of Ethiopia MoH (2006) National Five Year Strategic Plan for Malaria Prevention and Control in Ethiopia: Addis Ababa, Ethiopia.

233. Shargie EB, Gebre T, Ngondi J, Graves PM, Mosher AW, Emerson PM, Ejigsemahu Y, Endeshaw T, Olana D, WeldeMeskel A, Teferra A, Tadesse Z, Tilahun A, Yohannes G, Richards FO, Jr. (2008) Malaria prevalence and mosquito net coverage in Oromia and SNNPR regions of Ethiopia. *BMC Public Health* 8:321.

234. Federal Democratic Republic of Ethiopia CSA (2007) Summary and Statistical Report of the 2007 Population and Housing Census. Addis Ababa, Ethiopia.

235. USAID/PMI (2009) President's Malaria Initiative Malaria Operational Plan Ethiopia, FY 2009.

236. WHO (1991) Basic malaria microscopy: World Health Organization.

237. Center for International Earth Science Information Network, Columbia University, International Food Polcy Research Institute, The World Bank and Centro Internacional de Agricultura Tropical (2007) Global Rural Urban Mapping Project (GRUMP) alpha: Gridded Population of the World, version 2, with urban reallocation (GPW-UR).

238. Hay SI, Guerra CA, Gething PW, Patil AP, Tatem AJ, Noor AM, Kabaria CW, Manh BH, Elyazar IRF, Brooker S, Smith DL, Moyeed RA, Snow RW (2009) A world malaria map: *Plasmodium falciparum* endemicity in 2007. *PLoS Medicine* 6:e1000048.

239. WWF Global 200 Ecoregions database. http://www.worldwildlife.org/science/data/item1878.html

240. Moderate Resolution Imaging Spectroradiometer (MODIS). http://modis.gsfc.nasa.gov/

241. Hay SI, Tatem AJ, Graham AJ, Goetz SJ, Rogers DJ (2006) Global environmental data for mapping infectious disease distribution. *Advances in Parasitology* 62:37-77.

242. Scharlemann JP, Benz D, Hay SI, Purse BV, Tatem AJ, Wint GR, Rogers DJ (2008) Global data for ecology and epidemiology: a novel algorithm for temporal Fourier processing MODIS data. *PLoS ONE* 3:e1408.

243. WHO (2001) Iron deficiency anaemia. Assessment, prevention, and control. A guide for programme managers: World Health Organization.

244. Vuong QH (1989) Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57:307-333.

245. Diggle PJ, Ribeiro PJJ (2007) Model-based Geostatistics. New York: Springer.

246. SatScan software for the spatial, temporal, and space-time scan statistics. www.satscan.org/

247. Federal Democratic Republic of Ethiopia MoH (2004) Malaria diagnosis and treatment guidelines for health workers in Ethiopia: Addis Ababa, Ethiopia.

248. Okiro EA, Snow RW (2010) The relationship between reported fever and *Plasmodium falciparum* infection in African children. *Malaria Journal* 9:99.

249. Lee PW, Liu CT, do Rosario VE, de Sousa B, Rampao HS, Shaio MF (2010) Potential threat of malaria epidemics in a low transmission area, as exemplified by São Tomé and Príncipe. *Malaria Journal* 9:264.

250. Hall A, Kassa T, Demissie T, Degefie T, Lee S (2008) National survey of the health and nutrition of schoolchildren in Ethiopia. *Tropical Medicine and International Health* 13:1518-1526.

251. Central Statistical Agency Ethiopia & ORC Macro (2006) Ethiopa Demographic and Health Survey 2005. Addis Ababa, Ethiopia and Calverton, Maryland, USA: Central Statistical Agency and ORC Macro.

252. Hall A, Bobrow E, Brooker S, Jukes M, Nokes K, Lambo J, Guyatt H, Bundy D, Adjei S, Wen ST, Satoto, Subagio H, Rafiluddin MZ, Miguel T, Moulin S, de Graft Johnson J, Mukaka M, Roschnik N, Sacko M, Zacher A, Mahumane B, Kihamia C, Mwanri L, Tatala S, Lwambo N, Siza J, Khanh LN, Khoi HH, Toan ND (2001) Anaemia in schoolchildren in eight countries in Africa and Asia. *Public Health Nutrition* 4:749-756.

253. World Health Organization (2009) World Malaria Report. Geneva: World Health Organization.

254. Noor A, Kirui V, Brooker S, Snow R (2009) The use of insecticide treated nets by age: implications for universal coverage in Africa. *BMC Public Health* 9:369.

255. Shargie EB, Ngondi J, Graves PM, Getachew A, Hwang J, Gebre T, Mosher AW, Ceccato P, Endeshaw T, Jima D, Tadesse Z, Tenaw E, Reithinger R, Emerson PM, Richards FO, Ghebreyesus TA (2010) Rapid increase in ownership and use of long-lasting insecticidal nets and decrease in prevalence of malaria in three regional states of Ethiopia, 2006-2007. *Journal of Tropical Medicine* 2010:Article ID 750978.

256. Dye TD, Apondi R, Lugada ES, Kahn JG, Smith J, Othoro C (2010) "Before we used to get sick all the time": perceptions of malaria and use of long-lasting insecticide-treated bed nets (LLINs) in a rural Kenyan community. *Malaria Journal* 9:345.

257. Widmar M, Nagel CJ, Ho DY, Benziger PW, Hennig N (2009) Determining and addressing obstacles to the effective use of long-lasting insecticide-impregnated nets in rural Tanzania. *Malaria Journal* 8:315.

258. Pullan RL, Bikirwa H, Staedke SG, Snow RW, Brooker S (2010) *Plasmodium* infection and its risk factors in eastern Uganda. *Malaria Journal* 9:2.

259. Noor AM, Moloney G, Borle M, Fegan GW, Shewchuk T, Snow RW (2008) The use of mosquito nets and the prevalence of *Plasmodium falciparum* infection in rural South Central Somalia. *PLoS ONE* 3:e2081.

260. Mendis K, Sina BJ, Marchesini P, Carter R (2001) The neglected burden of *Plasmodium vivax* malaria. *The American Journal of Tropical Medicine and Hygiene* 64:97-106.

261. Williams TN, Maitland K, Phelps L, Bennett S, Peto TE, Viji J, Timothy R, Clegg JB, Weatherall DJ, Bowden DK (1997) *Plasmodium vivax*: a cause of malnutrition in young children. *QJM: monthly journal of the Association of Physicians* 90:751-757.

262. Harris I, Sharrock WW, Bain LM, Gray KA, Bobogare A, Boaz L, Lilley K, Krause D, Vallely A, Johnson ML, Gatton ML, Shanks GD, Cheng Q (2010) A large proportion of asymptomatic *Plasmodium* infections with low and sub-microscopic parasite densities in the low transmission setting of Temotu Province, Solomon Islands: challenges for malaria diagnostics in an elimination setting. *Malaria Journal* 9:254.

263. Satoguina J, Walther B, Drakeley C, Nwakanma D, Oriero EC, Correa S, Corran P, Conway DJ, Walther M (2009) Comparison of surveillance methods applied to a situation of low malaria prevalence at rural sites in The Gambia and Guinea Bissau. *Malaria Journal* 8:274.

264. Brooker S, Leslie T, Kolaczinski K, Mohsen E, Mehboob N, Saleheen S, Khudonazarov J, Freeman T, Clements A, Rowland M, Kolaczinski J (2006) Spatial epidemiology of *Plasmodium vivax*, Afghanistan. *Emerging Infectious Diseases* 12:1600-1602.

265. Bejon P, Williams TN, Liljander A, Noor AM, Wambua J, Ogada E, Olotu A, Osier FH, Hay SI, Farnert A, Marsh K (2010) Stable and unstable malaria hotspots in longitudinal cohort studies in Kenya. *PLoS Medicine* 7:e1000304.

266. Mirghani SE, Nour BY, Bushra SM, Elhassan IM, Snow RW, Noor AM (2010) The spatial-temporal clustering of *Plasmodium falciparum* infection over eleven years in Gezira State, The Sudan. *Malaria Journal* 9:172.

267. World Health Organization (2012) Disease surveillance for malaria elimination. Geneva: World Health Organization. November 4, 2013 http://whqlibdoc.who.int/publications/2012/9789241503334_eng.pdf

268. Greenwood BM (2008) Control to elimination: implications for malaria research. *Trends in Parasitology* 24:449-454.

269. Tatem AJ, Smith DL, Gething PW, Kabaria CW, Snow RW, Hay SI (2010) Ranking of elimination feasibility between malaria-endemic countries. *Lancet* 376:1579-1591.

270. Ishengoma DS, Francis F, Mmbando BP, Lusingu JP, Magistrado P, Alifrangis M, Theander TG, Bygbjerg IC, Lemnge MM (2011) Accuracy of malaria rapid diagnostic tests in community studies and their impact on treatment of malaria in an area with declining malaria burden in north-eastern Tanzania. *Malaria Journal* 10:176.

271. Searle KM, Shields T, Hamapumbu H, Kobayashi T, Mharakurwa S, Thuma PE, Smith DL, Glass G, Moss WJ (2013) Efficiency of household reactive case detection for malaria in rural southern Zambia: simulations based on cross-sectional surveys from two epidemiological settings. *PLoS ONE* 8:e70972.

272. Stewart L, Gosling R, Griffin J, Gesase S, Campo J, Hashim R, Masika P, Mosha J, Bousema T, Shekalaghe S, Cook J, Corran P, Ghani A, Riley EM, Drakeley C (2009) Rapid assessment of malaria transmission using age-specific sero-conversion rates. *PLoS ONE* 4:e6083.

273. Gitonga CW, Kihara JH, Njenga SM, Awuondo K, Noor AM, Snow RW, Brooker SJ (2012) Use of rapid diagnostic tests in malaria school surveys in Kenya: Does their under-performance matter for planning malaria control? *The American Journal of Tropical Medicine and Hygiene* 87:1004-1011.

274. Bousema T, Drakeley C, Gesase S, Hashim R, Magesa S, Mosha F, Otieno S, Carneiro I, Cox J, Msuya E, Kleinschmidt I, Maxwell C, Greenwood B, Riley E, Sauerwein R, Chandramohan D, Gosling R (2010) Identification of hot spots of malaria transmission for targeted malaria control. *The Journal of Infectious Diseases* 201:1764-1774.

275. Cook J, Kleinschmidt I, Schwabe C, Nseng G, Bousema T, Corran PH, Riley EM, Drakeley CJ (2011) Serological markers suggest heterogeneity of effectiveness of malaria control interventions on Bioko Island, Equatorial Guinea. *PLoS ONE* 6:e25137.

276. USGS-SRTM. http://srtm.csi.cgiar.org

277. WorldClim. http://www.worldclim.org

278. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25:1965-1978.

279. USGS-SRTM Water Bodies data files. http://dds.cr.usgs.gov/srtm/version2_1/SWBD

280. Digital Chart of the World. http://diva-gis.org/gdata

281. SPOT 5 VEGETATION. http://www/spot-vegetation.com

282. AfriPop Project. http://www.clas.ufl.edu/users/atatem/index_files/demographydata.htm

283. Global Rural-Urban Mapping project (GRUMPv1). http://sedac/ciesin.columbia.edu/data/sets/browse

284. Schwarz G (1978) Estimating the dimensions of a model. *The Annals of Statistics* 6:461-464.

285. Akaike H (1973) Information theory and an extension of the maximum likelihood principle. In: Petrov BN CF, editor. Budapest: Akademiai Kiado. pp. 276-281.

286. Spiegelhalter DJ, Best NG, Carlin BR, van der Linde A (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B-Statistical Methodology* 64:583-616.

287. Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* 240:1285-1293.

288. Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognition Letters* 27:861-874.

289. Clements ACA, Lwambo NJS, Blair L, Nyandindi U, Kaatano G, Kinung'hi S, Webster JP, Fenwick A, Brooker S (2006) Bayesian spatial analysis and disease mapping: tools to

enhance planning and implementation of a schistosomiasis control programme in Tanzania. *Tropical Medicine and International Health* 11:490-503.

290. Reid HL, Haque U, Roy S, Islam N, Clements AC (2012) Characterizing the spatial and temporal variation of malaria incidence in Bangladesh, 2007. *Malaria Journal* 11:170.

291. Guerra CA, Howes RE, Patil AP, Gething PW, Van Boeckel TP, Temperley WH, Kabaria CW, Tatem AJ, Manh BH, Elyazar IRF, Baird JK, Snow RW, Hay SI (2010) The international limits and population at risk of *Plasmodium vivax* transmission in 2009. *PLoS Neglected Tropical Diseases* 4:e774.

292. Malaria Atlas Project maps of malaria risk in Ethiopia. http://www.map.ox.ac.uk/explore/countries/ETH

293. Hartgers FC, Yazdanbakhsh M (2006) Co-infection of helminths and malaria: modulation of the immune responses to malaria. *Parasite Immunology* 28:497-506.

294. Mutonga D, Langat D, Mwangi D, Tonui J, Njeru M, Abade A, Irura Z, Njeru I, Dahlke M (2013) National surveillance data on the epidemiology of cholera in Kenya, 1997-2010. *The Journal of Infectious Diseases* 208 Suppl 1:S55-61.

295. Halperin SA, Bettinger JA, Greenwood B, Harrison LH, Jelfs J, Ladhani SN, McIntyre P, Ramsay ME, Safadi MA (2012) The changing and dynamic epidemiology of meningococcal disease. *Vaccine* 30 Suppl 2:B26-36.

296. Racloz V, Ramsey R, Tong S, Hu W (2012) Surveillance of dengue fever virus: a review of epidemiological models and early warning systems. *PLoS Neglected Tropical Diseases* 6:e1648.

297. Ashton RA, Kefyalew T, Tesfaye G, Pullan RL, Yadeta D, Reithinger R, Kolaczinski JH, Brooker S (2011) School-based surveys of malaria in Oromia Regional State, Ethiopia: a rapid survey method for malaria in low transmission settings. *Malaria Journal* 10:25.

298. Filmer D, Pritchett LH (2001) Estimating wealth effects without expenditure data--or tears: an application to educational enrollments in states of India. *Demography* 38:115-132.

299. Woldehanna T, Hagos A (2012) Shocks and Primary School Drop-out Rates: A Study of 20 Sentinel Sites in Ethiopia. Young Lives, Oxford Department of International Development,

Oxford University. http://www.younglives.org.uk/files/working-papers/wp88_woldehanna-and-hagos

300. Admassu KA (2011) Primary school enrollment and dropout in Ethiopia: Household and school factors. Washington D.C, USA. http://paa2011.princeton.edu/papers/111283

301. Mosha JF, Sturrock HJ, Greenhouse B, Greenwood B, Sutherland CJ, Gadalla N, Atwal S, Drakeley C, Kibiki G, Bousema T, Chandramohan D, Gosling R (2013) Epidemiology of subpatent *Plasmodium falciparum* infection: implications for detection of hotspots with imperfect diagnostics. *Malaria Journal* 12:221.

302. Belachew T, Hadley C, Lindstrom D, Gebremariam A, Lachat C, Kolsteren P (2011) Food insecurity, school absenteeism and educational attainment of adolescents in Jimma Zone Southwest Ethiopia: a longitudinal study. *Nutrition Journal* 10:29.

303. Pereznieto P, Jones N (2006) Young Lives Policy Brief 2. Educational Choices in Ethiopia: what determines whether poor children go to school? Oxford, UK. http://r4d.dfid.gov.uk/Output/173697/

304. Woldehanna T, Jones N, Tefera B (2005) Children's Educational Completion Rates and Achievement: Implications for Ethiopia's Second Poverty Reduction Strategy (2006-10). Oxford: Young Lives. http://www.younglives.org.uk/files/working-papers/wp18-children2019s-educational-completion-rates-and-achievement-implications-for-ethiopia2019s-second-poverty-reduction

305. Yekutiel P (1960) Problems of epidemiology in malaria eradication. *Bulletin of the World Health Organization* 22:669-683.

306. Cohen JM, Moonen B, Snow RW, Smith DL (2010) How absolute is zero? An evaluation of historical and current definitions of malaria elimination. *Malaria Journal* 9:213.

307. Moonen B, Cohen JM, Snow RW, Slutsker L, Drakeley C, Smith DL, Abeyasinghe RR, Rodriguez MH, Maharaj R, Tanner M, Targett G (2010) Operational strategies to achieve and maintain malaria elimination. *Lancet* 376:1592-1603.

308. Yukich JO, Butts J, Miles M, Berhane Y, Nahusenay H, Malone JL, Dissanayake G, Reithinger R, Keating J (2014) A description of malaria sentinel surveillance: a case study in Oromia Regional State, Ethiopia. *Malaria Journal* 13:88.

309. Omumbo JA, Noor AM, Fall IS, Snow RW (2013) How well are malaria maps used to design and finance malaria control in Africa? *PLoS ONE* 8:e53198.

310. Hay SI, Snow RW (2006) The Malaria Atlas Project: developing global maps of malaria risk. *PLoS Medicine* 3:e473.

311. Brooker S, Hotez PJ, Bundy DAP (2010) The global atlas of helminth infection: mapping the way forward in neglected tropical disease control. *PLoS Neglected Tropical Diseases* 4:e779.

312. Smith JL, Haddad D, Polack S, Harding-Esch EM, Hooper PJ, Mabey DC, Solomon AW, Brooker S (2011) Mapping the global distribution of trachoma: why an updated atlas is needed. *PLoS Neglected Tropical Diseases* 5:e973.

313. World Health Organization (2012) Disease surveillance for malaria control. Geneva: World Health Organization. http://whqlibdoc.who.int/publications/2012/9789241503341_eng.pdf

314. Afrane YA, Zhou G, Githeko AK, Yan G (2013) Utility of health facility-based malaria data for malaria surveillance. *PLoS ONE* 8:e54305.

315. Zurovac D, Talisuna AO, Snow RW (2012) Mobile phone text messaging: tool for malaria control in Africa. *PLoS Medicine* 9:e1001176.

316. Prue CS, Shannon KL, Khyang J, Edwards LJ, Ahmed S, Ram M, Shields T, Hossain MS, Glass GE, Nyunt MM, Sack DA, Sullivan DJ, Jr., Khan WA (2013) Mobile phones improve case detection and management of malaria in rural Bangladesh. *Malaria Journal* 12:48.

317. Githinji S, Kigen S, Memusi D, Nyandigisi A, Wamari A, Muturi A, Jagoe G, Ziegler R, Snow RW, Zurovac D (2014) Using mobile phone text messaging for malaria surveillance in rural Kenya. *Malaria Journal* 13:107.

318. Griffin JT, Hollingsworth TD, Okell LC, Churcher TS, White M, Hinsley W, Bousema T, Drakeley CJ, Ferguson NM, Basanez MG, Ghani AC (2010) Reducing *Plasmodium falciparum* malaria transmission in Africa: a model-based evaluation of intervention strategies. *PLoS Medicine* 7:e1000324.

319. Gitonga CW, Edwards T, Karanja PN, Noor AM, Snow RW, Brooker SJ (2012) *Plasmodium* infection, anaemia and mosquito net use among school children across different settings in Kenya. *Tropical Medicine and International Health* 17:858-870.

320. Okello G, Ndegwa SN, Halliday KE, Hanson K, Brooker SJ, Jones C (2012) Local perceptions of intermittent screening and treatment for malaria in school children on the south coast of Kenya. *Malaria Journal* 11:185.

APPENDICES

APPENDIX 1: GENERATION OF A WEALTH INDEX FOR SCHOOL AND COMMUNITY SURVEYS CONDUCTED IN SNNPR, CHAPTER 4

A wealth index was generated to summarise socio-economic indicators among the study population sampled for Chapter 4 using principal component analysis (PCA). Each variable has binary coding, 1 if true and 0 otherwise. For household construction, there was little diversity in households' materials for walls, roof and floor, therefore binary coding was sufficient to define the materials. Data were not standardised prior to conducting PCA, therefore correlation matrix was used rather than co-variance matrix.

For the community surveys, the percentage of covariance explained by the first principal component is 34%, the first eigenvalue is 2.39, the second eigenvalue is 1.31. For the school surveys, the percentage of covariance explained by the first principal component is 34%, the first eigenvalue is 2.40, the second eigenvalue is 1.37.

Scoring factor (eigenvector) for each factor contributing to the wealth index calculated from community survey and school survey data. The mean and standard deviation of each variable within the two datasets are also presented.

Supplementary Table 1 - Principal components (scoring factors) for the variables included in wealth index, together with population mean and standard deviation for each variable

	Community survey			School survey		
Variable	**Scoring factor**	**Mean**	**SD**	**Scoring factor**	**Mean**	**SD**
Own radio	0.404	0.220	0.414	0.346	0.532	0.499
Own mobile telephone	0.387	0.246	0.431	0.354	0.570	0.495
Own bicycle	0.235	0.056	0.230	0.241	0.229	0.420
Latrine	0.119	0.948	0.222	0.066	0.989	0.104
Brick walls in household	0.463	0.065	0.246	0.497	0.120	0.325
Iron roof in household	0.444	0.248	0.432	0.442	0.398	0.489
Cement floor in household	0.454	0.049	0.215	0.501	0.099	0.299

The principal component analysis creates uncorrelated indices, where each component is a linear weighted combination of the initial variables:

$$PC_m = a_{m1}X_1 + a_{m2}X_2 + \ldots + a_{mn}X_n$$

Where a_{mn} represents the weight for the m principal component and n^{th} variable.

The households in community data and individuals in school survey were classified into three groups according to the wealth index. Three categories were used due to the small dataset size (particularly for community survey) and limited variation at the six sites contributing this phase one data. The mean variable value for each of the three wealth index group are presented in the table below, along with mean socio-economic index in each group.

Supplementary Table 2 - Mean variable values for each wealth category, according to classification of wealth index calculated by principal component analysis. The mean wealth index value by group is also presented for both school and community surveys.

	Community survey			**School survey**		
Variable	**Poorest**	**Middle**	**Richest**	**Poorest**	**Middle**	**Richest**
Own radio	0.000	0.456	0.494	0.225	0.627	0.845
Own mobile telephone	0.000	0.537	0.546	0.207	0.734	0.879
Own bicycle	0.000	0.000	0.170	0.027	0.303	0.421
Latrine	0.920	0.966	0.987	0.979	0.993	0.999
Brick walls in household	0.000	0.000	0.195	0.000	0.001	0.421
Iron roof in household	0.000	0.007	0.751	0.000	0.349	0.997
Cement floor in household	0.000	0.000	0.147	0.000	0.000	0.348
Mean socio-economic index	*-0.987*	*-0.028*	*1.646*	*-1.303*	*-0.164*	*1.972*

APPENDIX 2: RISK FACTORS FOR *PLASMODIUM* INFECTION IN SCHOOL AND COMMUNITY SURVEYS, CHAPTER 4

Univariate associations between RDT-positive and microscopy-corrected RDT result from school and community surveys (phase one) and potential individual and household-level risk factors

Supplementary Table 3 - Univariate analysis of school survey *Plasmodium* infection

	RDT result			Slide-corrected RDT result		
Variable	**OR**	**95% CI**	**p**	**OR**	**95% CI**	**p**
Sex (female vs. male)	0.56	0.37, 0.84	0.006	0.60	0.31, 1.18	0.137
Age	0.92	0.83, 1.02	0.109	0.96	0.81, 1.13	0.601
Fever in previous 2 weeks	1.30	1.04, 1.62	0.019	1.22	0.85, 1.76	0.281
Fever on survey day	2.70	1.25, 5.80	0.011	2.38	0.71, 7.94	0.160
Net use last night?	0.79	0.52, 1.20	0.271	0.98	0.49, 1.97	0.955
Any nets in household?	1.29	0.75, 2.20	0.360	1.41	0.60, 3.34	0.433
IRS in household?	1.36	0.47, 3.95	0.571	3.38	0.51, 22.29	0.206
Household wealth						
Poorest	1	-	-	1	-	-
Median	1.86	1.01, 3.43	0.047	1.90	0.71, 5.07	0.200
Least poor	1.81	1.03, 3.18	0.038	1.78	0.70, 4.51	0.226
Education of household head						
None	1	-	-	1	-	-
Primary incomplete	0.91	0.50, 1.65	0.753	0.66	0.27, 1.63	0.370
Primary complete or higher	0.62	0.32, 1.22	0.168	0.50	0.18, 1.35	0.172
Distance to walk to school[1]						
<1 km	1	-	-	1	-	-
1-2 km	1.20	0.77, 1.88	0.429	1.02	0.47, 2.19	0.964
>2 km	0.85	0.46, 1.55	0.593	1.22	0.50, 2.96	0.660

[1]Estimated using reported time to walk to school, and assuming a child walks at 3.4 miles per hour on average

Supplementary Table 4 - Minimal multivariate model of RDT result (binary) and risk factors for infection in school-based survey, with study site random effects. Multivariate models could not be produced for slide-corrected RDT result due to small numbers.

	Odds ratio	**95% confidence interval**	**P**
Sex (female vs. male)	0.56	0.37, 0.85	0.006
Fever in previous two weeks	3.21	1.92, 5.35	<0.001
Household wealth			
Poorest	1	-	-
Median	1.98	1.06, 3.69	0.032
Least poor	1.92	1.08, 3.42	0.027

Supplementary Table 5 - Univariate analysis of community survey *Plasmodium* infection

	RDT result			Slide-corrected RDT result		
Variable	**OR**	**95% CI**	**p**	**OR**	**95% CI**	**p**
Sex (female vs. male)	0.92	0.62, 1.36	0.667	0.97	0.49, 1.93	0.940
Age (continuous)	0.97	0.96, 0.98	<0.001	0.98	0.96, 1.00	0.111
Age group						
<7 years	1	-	-	1	-	-
(school age) 7-16 years	0.86	0.54, 1.37	0.537	1.34	0.58, 3.12	0.491
>16 years	0.33	0.20, 0.55	<0.001	0.52	0.21, 1.27	0.151
Fever in previous 2 weeks	1.39	1.15, 1.69	0.001	1.33	0.99, 1.80	0.061
Fever on survey day	1.11	0.85, 1.46	0.428	1.13	0.76, 1.68	0.547
Net use last night?	0.41	0.21, 0.79	0.008	0.59	0.22, 1.61	0.304
Any nets in household?	0.77	0.49, 1.20	0.254	0.56	0.25, 1.27	0.166
Number of nets in household	0.85	0.65, 1.12	0.244	0.71	0.42, 1.20	0.202
IRS in household?	1.09	0.54, 2.20	0.803	4.05	1.25, 13.18	0.020
Household wealth						
Poorest	1	-	-	1	-	-
Median	1.63	0.89, 2.99	0.112	3.27	1.12, 9.60	0.031
Least poor	1.58	1.01, 2.47	0.046	3.32	1.39, 7.90	0.007
Education of household head						
None	1	-	-	1	-	-
Primary incomplete	1.31	0.83, 2.06	0.241	2.40	1.04, 5.56	0.040
Primary complete or higher	1.75	0.99, 3.08	0.053	1.25	0.36, 4.31	0.727
Distance from household to school						
<1 km	1	-	-	1	-	-
1-2 km	0.84	0.56, 1.27	0.406	0.54	0.25, 1.15	0.111
>2 km	0.66	0.27, 1.61	0.358	0.55	0.12, 2.55	0.445

Supplementary Table 6 - Minimal multivariate multilevel model of RDT result (binary) and risk factors for infection in community-based survey, with random effects at household- and study site-level. Multivariate models could not be produced for slide-corrected RDT result due to small numbers.

	Odds ratio	**95% confidence interval**	**P**
Age (continuous)	0.87	0.74, 1.02	0.085
Slept under net previous night	0.15	0.03, 0.88	0.036

APPENDIX 3: SUMMARY RESULTS FROM ROUTINELY COLLECTED HEALTH FACILITY DATA IN OROMIA REGIONAL STATE, 2006-2011

Of a total 6553 government health facilities (hospitals, health centres, clinics and health posts) estimated by Oromia Regional Health Bureau to be operational at the end of 2010, 5,553 (85%) facilities were included in documentation submitted as part of the retrospective data collection activity, and data from 5304 (81%) facilities in 246 woredas were suitable for inclusion in the clean dataset.

While the primary indicator of interest from health facility data was annual incidence of confirmed malaria, these data should be interpreted in the context of some important changes in malaria diagnosis and treatment in Oromia, which occurred concurrent to the data presented. Over the study period, there was a large expansion of the health extension programme, with many health posts becoming operational and beginning to submit data as part of HMIS. From January 2007 to December 2010, the number of health posts increased from 992 to 3939 (four-fold increase). The number of health centres operational and reporting data also increased (305 to 523) from June 2006 to January 2011. Mean woreda annual any-cause outpatient attendance increased from 12,732 in 2007 to 22,160 in 2010, indicating an overall 74% increase in utilisation of health facilities from 2007 to 2010.

The mean number of health facilities reporting data from a single woreda was 20, including two health centres and 18 health posts. The population of a single woreda (from 2007 census) ranges from 10,752 to 337,913. A large increase in proportion of health facilities with access to parasitological diagnosis (either microscopy or RDTs) was seen, from 26% at the start of 2007 to 57% by the end of 2010, with further increases documented since 2010. The increase in access to diagnostics is more striking when considering that the proportion of all suspected malaria cases which were tested using RDT or microscopy increased from 45% in 2007 to 81% in 2010.

It was possible to calculate annual mean incidence of confirmed malaria for 219 woredas. A map displaying annual incidence of confirmed malaria from each year from 2007 to 2010, inclusive, is presented in Supplementary Figure 1.

Supplementary Figure 1 - Annual woreda incidence of confirmed malaria per 1000 people, for all health facilities reporting data. Note that from 2007 to 2010, health facilities had increasing access to confirmatory diagnostics, both increased use of microscopy at health centres and availability of rapid diagnostic tests (changing from HRP2-only to HRP2-panLDH combination kits from 2009 onwards) at health posts. In addition, expansion of the health extension system resulted in a large increase in the number of health facilities reporting data to the Health Management Information System from 2007 to 2010