UNIVERSITY OF Southampton

University of Southampton Research Repository
ePrints Soton

Copyright © and Moral Rights for this thesis are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given e.g.

AUTHOR (year of submission) "Full thesis title", University of Southampton, name of the University School or Department, PhD Thesis, pagination

http://eprints.soton.ac.uk

UNIVERSITY OF SOUTHAMPTON

FACULTY OF SOCIAL AND HUMAN SCIENCES

GEOGRAPHY AND ENVIRONMENT

Developing empirical space-time models of health services for the treatment of malaria to estimate disease incidence

By

Victor Adagi Alegana

Thesis for the degree of Doctor of Philosophy

February 2015

ABSTRACT

Background: There has been a substantial decline in malaria burden in the last decade owing to an increase in funding for malaria control. At every stage of the malaria elimination and control pathways, maps of malaria risk are required for planning and for resource allocation. These have traditionally been modelled from parasite prevalence data. However, in low malaria transmission, parasite prevalence surveys from household surveys are insufficient because they are of low sample sizes. Undertaking parasite rate surveys of adequate sample sizes is expensive and remains unaffordable by most national malaria control programmes. Moreover, these point prevalence surveys are not suitable for tracking changes in burden because malaria becomes highly seasonal driven by climatic conditions. In this thesis, alternative approaches of estimating risk are explored. These approaches model malaria incidence using Health Management Information System (HMIS) data.

Methods: Three low transmission countries were selected as case studies namely: Namibia, Eritrea and Afghanistan. HMIS data was assembled from the respective national malaria control programmes as well as nationally representative household surveys providing information on febrile cases and treatment seeking behaviour. For each case study, analysis of healthcare utilisation pattern for fever treatment from national representative household surveys was undertaken to derive denominators (population in health facility catchments). Data on malaria cases from health facility were combined with the catchment population and environmental drivers of malaria transmission to model incidence using a hierarchical Bayesian spatio-temporal conditional autoregressive model (CAR). Facility level data was adjusted based on reporting rates, the rate of utilisation and a slide positivity rate applied to suspected cases.

Results: The proportion of febrile children seeking treatment decreased with increasing distance to the nearest public health facility and this rate was different in the three case studies. In terms of catchment population, the majority of population was within three hours of travel to nearest health facility. This translated to coverage rates of 67% in Namibia, 79% in Eritrea and 85% in Afghanistan. The mean *Plasmodium falciparum* malaria incidence in Namibia was 12.5 (95% Crl 10.4-15.5) per 1000 population. *P. vivax* was the major malaria parasite in Afghanistan with an incidence of 5.4 (95% Crl 3.2-9.2) per 1000 population compared to *P. falciparum* incidence of 1.2 (95% Crl 0.4-2.9) per 1000 population. In Eritrea, the incidence of *P. falciparum* and *P. vivax* was 3.4 (95% Crl 2.2-5.2) per 1000 population and 2.5 (95%Crl 1.5-3.9) per 1000 population, respectively. Malaria Incidence in the three countries tended to be higher in the border areas. For Namibia, there was elevated incidence at the border with Angola. In Eritrea, incidence was higher in regions that bordered Ethiopia while for Afghanistan these were in districts bordering Pakistan. Relating the modelled incidence to current maps of parasite prevalence showed areas with higher incidence also exhibited high prevalence.

Conclusion: This thesis provides a novel approach to using household and health facility case data to model malaria incidence more precisely in countries with very low malaria transmission intensity. The modelling approach is vital for disease mapping in countries aiming for elimination in reduced malaria transmission, increasing level of parasitological diagnosis and improved level of reporting through HMIS. By using incomplete HMIS, the thesis demonstrates its usefulness in producing reliable estimates of malaria incidence as well as identifying high burdened regions to direct malaria interventions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DECLARATION OF AUTHORSHIP

I, **Victor Adagi Alegana**,
declare that this thesis and the work presented in it are my own, and has been generated by me as the result of my own original research.

Developing empirical space-time models of health services for the treatment of malaria to estimate disease incidence

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Either none of this work has been published before submission, or parts of this work have been published as:

Alegana, V. A., Atkinson, P. M., Wright, J. A., Kamwi, R., Uusiku, P., Katokele, S., Snow, R. W. and Noor, A. M. (2013) Estimation of malaria incidence in northern Namibia in 2009 using Bayesian conditional-autoregressive spatial-temporal models. *Spatial and Spatio-temporal Epidemiology,* 7**,** 25-36.

Alegana, V. A., Wright, J. A., Nahzat, S. M., Butt, W., Sediqi, A. W., Habib, N., Snow, R. W., Atkinson, P. M. and Noor, A. M. (2014) Modelling the Incidence of *Plasmodium vivax* and *Plasmodium falciparum* Malaria in Afghanistan 2006 - 2009. *PLoS One,* 9**,** e102304.

Signed: …….

Date: ……...……

ACKNOWLEDGEMENTS

Many individuals and institutions have contributed to the success of this PhD. My sincere gratitude goes to the Commonwealth Scholarship Commission, the Wellcome Trust, and the University of Southampton for providing financial and institutional support for this PhD. The PhD was tenured at Southampton University, UK, through commonwealth grant #KECS-2012-601. Additional support was obtained through the Wellcome Trust strategic award to KEMRI-WTRP and fellowship awards to Professor Abdisalan Noor (#WT095127) and Professor Robert Snow (#WT079080).

The years I have spent at the Kenya Medical Research Institute-Wellcome Trust Research Programme (KEMRI-WTRP) under the mentorship of Professor Noor laid the foundation for this PhD. I would like thank Professor Noor for his continued role as a supervisor on this thesis. I am also thankful to my other supervisors, Professor Peter Atkinson and Dr Jim Wright at the University of Southampton. I am extremely grateful to have had Professor Atkinson's humble introduction to Bayesian modelling. The advice, insight, and mentorship offered by my supervisors throughout the writing of this thesis were invaluable. I would also like to thank Professor Snow for his advice on different aspects of the thesis.

My gratitude also goes to the National Malaria Control Programmes of Namibia, Afghanistan and Eritrea who shared their data. I also thank various people who worked behind the scenes to provide various information and to the staff at the University of Southampton University and KEMRI-Wellcome Trust Research.

Finally, I most grateful to my family! My parents instilled in me the value of education, and the freedom to pursue my dream. I hope I can instil the same in my children and inspire them to pursue their dreams, one just ever so new to this world. My special thanks go to my spouse Harriet for being so understanding and supportive unconditionally. Thank you.

CHAPTER 1: Introduction and literature review

1.1 Background

Malaria is a major contributor to mortality and morbidity across sub-Saharan Africa (SSA) (Black *et al.*, 2003, Hay *et al.*, 2008, Snow and Marsh, 2010). While considerable progress has been made globally on reducing malaria related deaths according to the 2013 Millennium Development Goals (MDG) report, progress in SSA remains slow and disproportionate (Noor *et al.*, 2014). In SSA countries, the burden of malaria remains a significant public health problem and a realistic assessment is that the 2015 MDG health targets are unlikely to be achievable due to slow economic development and poverty (Snow *et al.*, 2010b, Snow *et al.*, 2012). The sustainable development goals (SDGs) integrate the MDGs into a post-2015 agenda following the Rio de Janeiro declaration (UN, 2012a). To sustain progress on the main health indicators, it is important to define the population at risk and improve surveillance, planning and cost-effective allocation of resources guided by country level maps of malaria transmission (Noor *et al.*, 2014).

Malaria endemic countries have had sustained investment in malaria control since early 2000s (Snow *et al.*, 2010b, World Health Organization, 2013b) and coverage of key interventions has increased with reports showing a significant decline in malaria infections and disease burden across several sites suggestive of an epidemiological transition (Ceesay *et al.*, 2008, O'Meara *et al.*, 2010). Many countries have responded to the call by the Roll Back Malaria (RBM) initiative for universal coverage of malaria prevention strategies and some now aim for malaria elimination (Global Health Group, 2007, Malaria Elimination 8 Ministerial Meeting, 2009, The Malaria Elimination Group, 2009). Each stage on the control-elimination continuum requires accurate epidemiological assessment of infection risk to adapt operational strategies and revise impact predictions (Hay *et al.*, 2008, Smith *et al.*, 2009). Maps of malaria risk are important tools

for countries to define the changing risks, but, there is little research evidence on the development, integration and application of these maps in the context of sustained low risk or malaria elimination. Community parasite prevalence of 1% is currently considered the benchmark for countries to decide between sustaining conditions of low endemic control or moving towards an agenda that includes elimination (World Health Organization, 2007a, Snow and Marsh, 2010). Current mapping approaches to define these benchmarks in areas of low transmission, however, are faced with both data and methodological challenges that require research into developing more reliable, efficient and less costly mechanisms of assessing the shrinking malaria map (The Malaria Elimination Group, 2009, Feachem *et al.*, 2010).

In this thesis, determinants of health facility access, utilisation and disease modelling approaches were reviewed. Household survey data for treatment of fever were assembled for three low transmission countries (Namibia, Afghanistan and Eritrea) and used to model spatial catchments to derive catchment populations. A model based on self-reported fever at household level was linked to population and facility characteristics to derive, spatially, the treatment-seeking patterns within the facility catchment zone. For Afghanistan, the second case study, this analysis was stratified by health facility type. The purpose of stratification was to compare a mean utilisation pattern based on all the health facilities (i.e. the distance decay model) to the patterns derived based on facility type. The Afghanistan analysis suggested that a mean distance decay pattern based on facility type stratification was similar to that derived from all health facilities without stratification. Subsequently, the analysis in Eritrea was not stratified by health facility type.

Population within the health facility catchments were calculated and adjusted using the probability to seek treatment for fever before being used to derive malaria case incidence, via a Bayesian spatio-temporal condition autoregressive (CAR) model, using cases observed at health facilities. Data on population were obtained from Worldpop (Worldpop, 2010) while disease cases were obtained from various national malaria control programmes (NMCPs). The Bayesian spatio-temporal CAR model complexity increased from one case study to the next. Thus, there was a variation in model set-up based on experiences of previous case study. For example, the Namibia CAR model did not incorporate random effects at regional levels compared to the Afghanistan and Eritrean model, while, nonlinear priors were used in Eritrea compared to a fixed effect assumption in both Namibia and Afghanistan. These approaches used in this thesis together with their limitations are subsequently described.

This chapter reviews health goals and the MDGs in SSA outlining the burden of malaria and fever (section 1.2). Section 1.3 provides an overview of healthcare delivery in SSA with section 1.3.2 providing a description of Health Management Information Systems (HMIS) followed by roles of national household and health facility surveys for measuring and monitoring population health (section 1.3.4 and section 1.3.5). Methods for measuring access and utilisation of healthcare services are reviewed in section 1.4. Geographic access (section 1.4.2), distance as a metric of access (section 1.4.3) and catchments (section 1.4.5) form the body of this section. Lastly, disease-modelling approaches are reviewed (section 1.5) starting with historical developments in spatial and spatio-temporal disease mapping methods (sections 1.5.1 and 1.5.2) then the spatial-only methods used for point data (section 1.5.4) and areal data (section 1.5.5).

The latter part focuses on spatio-temporal approaches and concludes by outlining the scope of the thesis, providing a justification for selection of the study countries.

1.2 Health in developing countries

1.2.1 The Millennium Development Goals (MDGs) and Sustainable Development Goals (SDGs)

Over the past two decades, the human global development agenda has been shaped by two components. The first is the Millennium Development Goals (MDGs), initially mooted in 1990 and adopted by the 189 UN member states in 2000 (UN General Assembly, 2000). The second, the Sustainable Development Goals (SDGs) came to the fore at the Earth summit in Rio de Janeiro in 1992 and recently took centre stage at Rio + 20 summit in 2012 to generate concepts similar to the MDGs but looking beyond 2015. The current aim is to merge these two parallel but very similar concepts when it comes to human sustainable development agenda post-2015 (UN, 2012b). This section briefly outlines both the MDGs and SDGs particularly focusing on health goals and the agenda beyond 2015.

In general, there are eight broad MDGs (poverty, food security, education, health and family planning, infrastructure (energy, housing, water and sanitation), environment, security, and governance). The MDGs were agreed upon by UN member states after it was realized that many low-income countries could not achieve the necessary economic growth and eliminate widespread poverty. Thus, the main strength of the MDGs is that they constitute direct and measurable goals and this ignited interest from developed countries to put resources forward to foster growth and development. Consequently, funding initiatives and interventions were set up

with the aim of improving economic development, health, sanitation, access to education, better housing and basic infrastructure in developing countries (UNDP, 2003).

The health goals focused on reducing the child mortality rate by two thirds (MDG 4), improving maternal health (MDG 5) and combating HIV/AIDS, malaria and other diseases (MDG 6). MDG 4 aims at reducing, by two thirds, the rate of child and infant mortality. MDG 6 targets include (a) halting and reversing the spread of HIV/AIDS, (b) achieving universal access to antiretroviral therapy for HIV/AIDS patients, and (c) halting and reversing the incidence of malaria and other major infectious diseases. Indicators for malaria include the prevalence and deaths associated with malaria and the proportion of population at risk of malaria using appropriate preventive and treatment methods. Some of the recommended strategies at achieving the malaria targets by RBM partnership include the universal coverage and use of LLIN, indoor residual spraying, the use of diagnostic testing under the T3 (Test-Treat-Track) initiative and preventive therapies during pregnancy, in infants and young children (World Health Organization, 2014d).

According to the WHO malaria report 2014, there has been, globally, a reduction in malaria death by 47% and by 54% in Africa between 2000 and 2013. This trend goes hand in hand with an overall reduced parasite prevalence in Africa (Noor *et al.*, 2014). However, national and sub-national variation exists, especially in SSA (O'Meara *et al.*, 2010, Snow *et al.*, 2010b). The 2014 MDG report suggest that, in SSA, close to half of the population (48%) live on less than $1.25 per day, child mortality rate remains high at 98 deaths per 1000 live births with annual rates of decline well below 8% (the required rate). Thus, poverty and economic development remain a major challenge since most national governments in these low-income countries are unable to

meet the financial requirements to attain the MDG targets (Fortney *et al.*, 2001, Neuman *et al.*, 2011). Outside SSA, Bangladesh provides an example of a country on track to achieve MDG 4 (Chowdhury *et al.,* 2012) (World Health Organization and UNICEF, 2012) and China, despite significant strides in reducing poverty, is still ranked second behind India with 13% of the global extreme poor population (UN, 2014).

The Lozano multi-country study (Lozano *et al.*, 2011) pooled data from household surveys and censuses including complete vital registration data, surveillance, Complete Birth Histories (CBH), Summary of Birth Histories (SBH) and maternal mortality rates extrapolated beyond 2015 using rates of change between 1990 and 2011. Results from this study suggested few countries were likely to achieve MDG 4 or MDG 5 in SSA by 2015 based on observed rates of change. Eight countries were likely to meet these targets within 10 years after 2015 (Lozano *et al.*, 2011). The example of Bioko Island, Equatorial Guinea, showed remarkable gains on MDG 4 following scale up of interventions (Kleinschmidt *et al.*, 2009). Similar observations have been highlighted in other country-specific studies such as in Ghana (Nakamura *et al.*, 2011, Zere *et al.*, 2012), Niger (Amouzou *et al.*, 2012) and Papua New Guinea (PNG) where substantial differences in child mortality rates were observed based on the 2000 national census and 2006 DHS (Bauze *et al.*, 2012).

The main criticism of the MDG concept is that it is viewed as narrow in scope that culminates with specific targets to be evaluated by 2015. Thus, many view the MDGs as short and medium term goals with a focus on development and poverty at a micro-level and in some case with metrics that are difficult to assess. For instance, MDGs fail to address in a comprehensive way

issues of sustainability, youth unemployment, violence and conflict, good governance and human rights. Many of these issues emerged later after the 2000 declaration. There are other issues inherent in the MDGs related to metrics and how these are evaluated in terms of comparison with baseline measures and to other developed countries. The United Nations General Assembly, in 2010, put in place a high level panel to coordinate activities and consultations beyond 2015. This was followed by the launch of a task team by the UN Secretary-General in 2012 to coordinate activities of a high-level panel on policy beyond 2015. The post-2015 agenda is shaped by the political document (UN, 2012a) produced as one outcome of the Rio +20 conference on sustainable development (UN, 2012a) and reaffirms political commitment to achieving various MDG goals post-2015. The post-2015 agenda addresses some of the weaknesses in the MDGs and integrates the strength into the SDGs (UN, 2012b).

The SDGs attempt to address these issues by incorporating the micro-level metrics (MDGs essentially) at a global (macro) level and in a sustainable way (long-term). Some of the challenges to be addressed in general stem from the changing population age-structure dynamics, environmental sustainability, human resource and labour, global markets and governance. The Rio +20 declaration raised more than 20 targets and in terms of health, the indicators on child mortality, malaria, HIV/AIDS, Tuberculosis remain unchanged post-2015 as specified in the MDGs (UN, 2012a, b). The SDGs however emphasise the need to strengthen health systems, and promote preventive and effective treatment to non-communicable diseases (NCD) since NCDs pose a major challenges to sustainable development in both developed and developing countries.

1.2.2 Review of progress in malaria control in SSA post Global Malaria Eradication Plan (GMEP) era

The failure of the GMEP in the 1950s and 1960s, particularly in stable transmission areas in Africa (Nájera *et al.*, 2011) coupled with resistance to both parasite and vector, led to resurgence of the disease in the 1980s through to the 1990s (Nájera *et al.*, 2011, Snow *et al.*, 2012). Malaria control during and post the GMEP era involved largely the use of chloroquine for treatment and spraying using insecticides such as dichloro-diphenyl-trichloroethane (DDT) (Snow *et al.*, 2012). By 1978, several countries, particularly in Europe and the Americas (27 in total), had been declared malaria-free (Mendis *et al.*, 2009). In Asia, in countries such as India and Sri Lanka, the burden decreased significantly during the GMEP era. For example, Sri Lanka reduced the number of reported cases within a 20 year period from several millions in the 1940s to just 18 cases by the late 1960s (Organizastion, 1969). In October 1992, a ministerial conference constituted by the World Health Organisation (WHO) with participants from 102 malaria endemic countries adopted a declaration on malaria control and expressed commitment to implementing a global strategy aimed at reducing morbidity, mortality and addressing emerging resistance. The Roll Back Malaria (RBM) programme was later launched in 1998 to mobilize support (Roll Back Malaria partnership, 2011). Malaria was also included in the global MDGs in the year 2000 (United Nations, 2000).

With the exception of 36 countries at the margins of stable transmission, the majority of the malaria endemic countries (99) focus on control (Feachem *et al.*, 2010, Tatem *et al.*, 2010). 79 countries have eliminated malaria, most during the GMEP (1955 to 1969) (Global Health Group, 2011). The current gains are as a result of funding commitments of international, multilateral and

bilateral organisations to support malaria control activities. This include funding from the Department for Internal Development (DFID), the World Bank, the launch of the Global Fund in 2002 and the President's Malaria initiative (PMI) in 2006 (Snow *et al.*, 2010b). Although funding in general has increased in the past decade (from < US$ 100 million in 2000 to the projected US$ 1.9 billion in 2013 (World Health Organization, 2013b)), it is likely to fall short of the US$ 5 billion required to achieve global targets (Snow *et al.*, 2010b, World Health Organization, 2011). In addition, it is unlikely that domestic financing will bridge the resource needs gap in the near future. The Global Fund currently accounts for nearly half of funding commitment to endemic countries. Other support is from the Presidential Malaria Initiative (PMI) and the United Kingdom Department for International Development (DFID) and other agencies. In addition to a call for universal coverage of ITNs in 2008 by the UN-Secretary general, there is a renewed focus on elimination (defined as the state where endemic local transmission has been interrupted to zero incidence with risk of re-establishment due to imported cases minimized (The Malaria Elimination Group, 2009, Feachem *et al.*, 2010) and eradication (i.e. complete or permanent removal of incidence or malaria parasites) (Roberts and Enserink, 2007, The Malaria Elimination Group, 2009). In 2007, the Malaria Elimination Group (MEG) was launched to support and provide to countries that intend to eliminate malaria (Global Health Group, 2007) while the Malaria Eradication and Research Agenda (maIERA) (The malERA and Diagnostics, 2011) mandate is to lead a research agenda. Presently, 36 countries aim at eliminating malaria (Global Health Group, 2011). Nine countries in Asia Pacific and in Europe-Middle East-Central Asia; and 11 countries in Asia Pacific, are pursing elimination. In Africa, seven countries are pursuing elimination: four in southern Africa (Namibia, Botswana, South

Africa and Swaziland (Southern African Development Community, 2012)); Algeria in North Africa; Cape Verde; and São Tomé and Príncipe (Figure 1.1)).

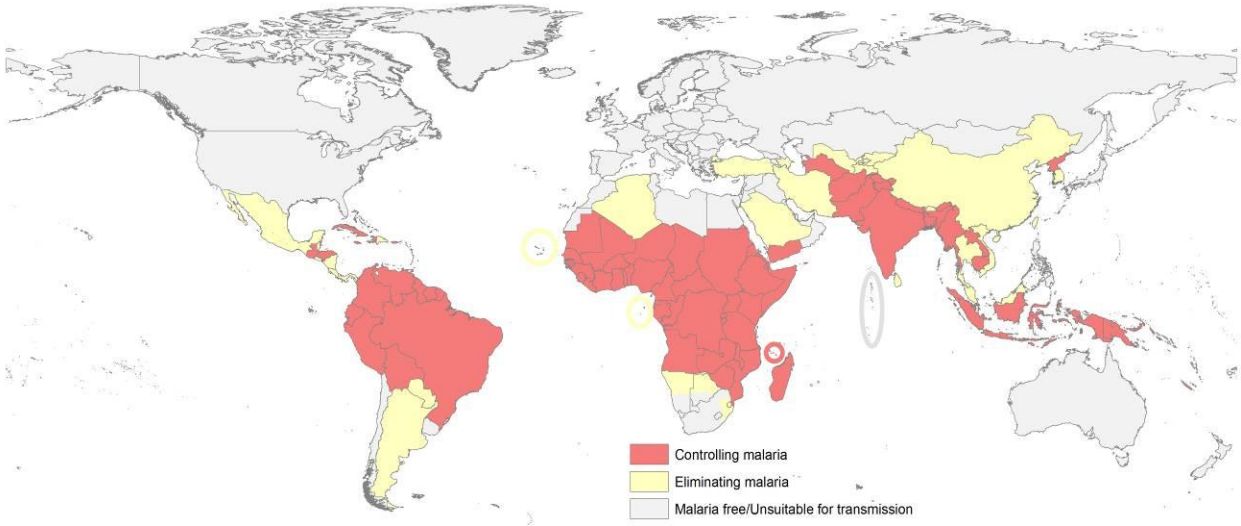

Figure 1.1: Countries eliminating or controlling malaria
Countries that are free of malaria (grey colour), controlling malaria (red) and are eliminating malaria or have declared elimination (light yellow). Data from the Global Health Group (Global Health Group, 2007).

Improved funding for malaria control has, for instance, increased the coverage of ITNs in Kenya to 44% by the end of 2007 (Noor *et al.*, 2009a) compared to only 3% in 2003, based on the Multiple Indicator Cluster Survey (MICS). Another example in SSA is Bioko Island which achieved MDG goal 4 in a single year due to a scale up of interventions (ITN and Indoor Residual Spraying (IRS)) (Kleinschmidt *et al.*, 2009) and in Rwanda which observed a decrease in malaria incidence following scale-up of ITNs (Ministry of Health, 2010). Countries in the WHO Eastern Mediterranean and European Regions stepped up efforts of elimination in the 1990s and by 2008 only Algeria of the Northern Africa countries reported an autochthonous case (World Health Organizastion, 2009). Further, United Arab Emirates and Morocco were certified by the WHO as malaria free in 2007 (World Health Organisation, 2007, Regional Office for Eastern and Mediteranean Region, 2013) and Egypt, Armenia, Turkmenistan, Syria Arab

republic all reported zero local acquired cases by the mid 2000s (Mendis *et al.*, 2009). Indeed within the WHO/EMRO region three countries (Saudi Arabia, Iraq and Islamic republic of Iran, Afghanistan) have significantly reduced burden in the last 10 years. Djibouti has also reduced its burden significantly in recent years (Noor *et al.*, 2011). The remaining countries (Pakistan, Somalia, Sudan, South Sudan and Yemen) still witness a varying burden of Malaria (World Health Organisation, 2007). The burden has also reduced in Europe and recent evidence indicates the region is on track of achieving elimination by 2015. Similarly there is reduced burden in South-East Asia (10 countries with ongoing transmission) and the western Pacific regions (Mendis *et al.*, 2009). In SSA several countries such as Eritrea, Djibouti and Namibia have reduced burden significantly (by >75%) since 2000. Ethiopia and Zambia are also projected to reduce burden by at least 50% by 2015.

ACTs are now recognized and used as first line treatment of *Plasmodium falciparum* malaria in every endemic country and the majority (for example in Ethiopia, Kenya and the Gambia) are in the process of achieving high coverage (World Health Organization, 2010a). The WHO recommends diagnosis of all febrile cases before treatment with ACTs to avoid treatment of non-malaria cases. As a result, malaria endemic countries have or are in the process of scaling up the use of rapid-diagnostic tests (RDTs) in health facilities with some deployment through community health workers (CHW) (DOMC, 2010, Ministry of Health, 2010, Ministry of Health and Social Services, 2010c).

This increase in the coverage of interventions has led subsequently to a decrease in the burden of malaria at national and sub-national levels, for example, in Kenya and the Gambia (Ceesay *et al.*,

2008, Okiro *et al.*, 2009) and in Disease Surveillance Sites (DSS) (O'Meara *et al.*, 2008). Despite the reported decline, variation in burden exists between and within countries.

A recent publication by Noor *et al.* (2014) highlighted reduction in *falciparum* malaria across SSA. The study utilised parasite rate surveys across the 49 malaria endemic countries in Africa and predicted age-standardised (2-10 years) parasite prevalence ($PfPR_{2-10}$) at fine spatial (1 x 1 km) and temporal resolution (2000 and 2010). One of the major findings of the study suggested a reduction in population at risk in 2010 when compared to 2000, although, this varied within and across the countries. In some areas, malaria transmission had reduced to a level requiring a re-orientation of national malaria control programme focus from sustained control to elimination or pre-elimination. These included regions where $PfPR_{2-10}$ remained at <3% in 2010 when compared to 2000 (Noor *et al.*, 2014). Some of the countries where national mean $PfPR_{2-10}$ is < 3% included Namibia, Eritrea, Swaziland, Djibouti, Rwanda, Mayotte and Cape Verde with the first two countries being case studies in this thesis.

In such low transmission areas, the distribution of risk is highly focal, or clustered in time and space, and identifying foci requires considerably greater sampling effort and cost using the method for mapping malaria risk used in Noor *et al*. (2014). The surveys are single time-point, cross-sectional community-based parasite prevalence surveys. Other measures that can be used to define risk include the use of health facility reported cases and become more valuable to define risk in space and time when asymptomatic infection prevalence becomes rare. However, there are few examples of using imperfect health facility reported data to map malaria risk in space and time that is a subject of this thesis (Gething *et al.*, 2008). Further, there are no

examples of where these have been used in combination with other household surveys to adjust for utilisation rate of febrile cases. This thesis applies a Bayesian hierarchical model-based to HMIS data to estimate malaria incidence at national level combined with a novel approach of estimation the denominator (catchment population) based on malaria treatment seeking behaviour. This modelling strategy should support the ambitions of national malaria control programmes in low malaria transmission and support elimination strategies.

1.2.3 Rationale for using fever for studying treatment seeking behaviour

Fever is one of the clinical symptoms associated with malaria and is a common presentation in health facilities, and has been the basis of treatment of uncomplicated malaria in endemic regions (Einterz and Bates, 1997, Snow *et al.*, 2003). This makes fever an entry point in studying malaria treatment practices. The identification of fever as a morbid event and its treatment varies within and between communities (Einterz and Bates, 1997, Beiersmann *et al.*, 2007, Chibwana *et al.*, 2009, Oyakhirome *et al.*, 2010). For example, the word '*homa*' is common to Kenya and Tanzania (Winch *et al.*, 1996), '*asra*' to Ghana (Agyepong, 1992) and '*oluludi*' to Namibia (Davies, 1994). This variation of terminology also means that the symptoms associated with its description may also vary between communities. For example *'asra*' (mild fever) not only refers to a rise in body temperature but may also be associated with bitterness in the mouth, yellow eyes and deep coloured urine, all symptoms of severe malaria. In the same local context, in Ghana, significantly high fever is identified as *'asraku*' (Agyepong and Manderson, 1994).

The nature and occurrence of malaria and related fever varies according to the intensity of transmission (Snow and Marsh, 2002, Guerra *et al.*, 2008). Thus, the number of fever cases attributed to malaria will vary depending on the transmission intensity. Figure 1.2 shows a fever

treatment protocol common in most settings. Fever in most patients can be a mild event and subsequently resolve itself without any treatment. In many other settings, fever is usually first managed at home or informally by purchasing medicines from drug shops, vendors or pharmaceuticals (McCombie, 1996, 2002, Amin *et al.*, 2003, Goodman *et al.*, 2007). Past studies have shown that mothers or care givers' education level, socio-economic status and availability and quality of care determine the use of formal health facilities (Kazembe *et al.*, 2007, Chuma *et al.*, 2009a, Rutebemberwa *et al.*, 2009). A review by Goodman and others in countries across SSA estimated that over 50% of fever cases are usually treated informally (Goodman *et al.*, 2007). Breman (2001) referred to this phenomenon as "*ears of hippopotamus*" and it is also commonly known as *'the iceberg effect'* in some public health literature (Donaldson and Gabriel, 2009), where only a few cases are formally treated in the public sector and a significantly larger burden is untreated or managed at community level (Breman, 2001). Many findings from studies carried out in Kenya (Amin *et al.*, 2003, Guyatt and Snow, 2004), Rwanda (Saksena *et al.*, 2011) and Malawi (Kazembe *et al.*, 2007) support this observation. In the Sudan, a country with low infection rates, a household survey showed that only about 40% of the surveyed population sought treatment from the public sector where the self-reported two week fever prevalence rate was approximately 20% (Elmardi *et al.*, 2011). Currently, many national malaria control programmes discourage presumptive treatment of fever following the WHO recommendation to the use of parasitological diagnosis using either microscopy or RDTs (World Health Organization, 2010b, Elmardi *et al.*, 2011, Gitonga *et al.*, 2012).

Severe or complicated malaria can be grouped largely into severe malaria anaemia and cerebral malaria (Taylor and Molyneux, 2002). The former is characterized by presence of both *P.*

falciparum as well as a haemoglobin level of less than or equal to 5 g $100ml^{-1}$ or up to 7 g $100ml^{-1}$ in low transmission settings (Taylor and Molyneux, 2002, World Health Organization, 2010a). For cerebral malaria *P. falciparum* is present in addition to clinical features of convulsions and loss of consciousness, usually coma (Marsh and Snow, 1999, Taylor and Molyneux, 2002). Marsh & Snow (1999) caution that these clinical descriptions although useful, are rather coarse because in high endemicities, *P. falciparum* is usually present while at the same time there are possible multiple causes of anaemia. Other forms of the disease have previously been shown to overlap with severe respiratory infections such as pneumonia (Akpede *et al.*, 1992, Rooth and Bjorkman, 1992, Akpede *et al.*, 1993, English *et al.*, 1996, Crawley *et al.*, 1998, Kallander *et al.*, 2004, Thurmond *et al.*, 2005).

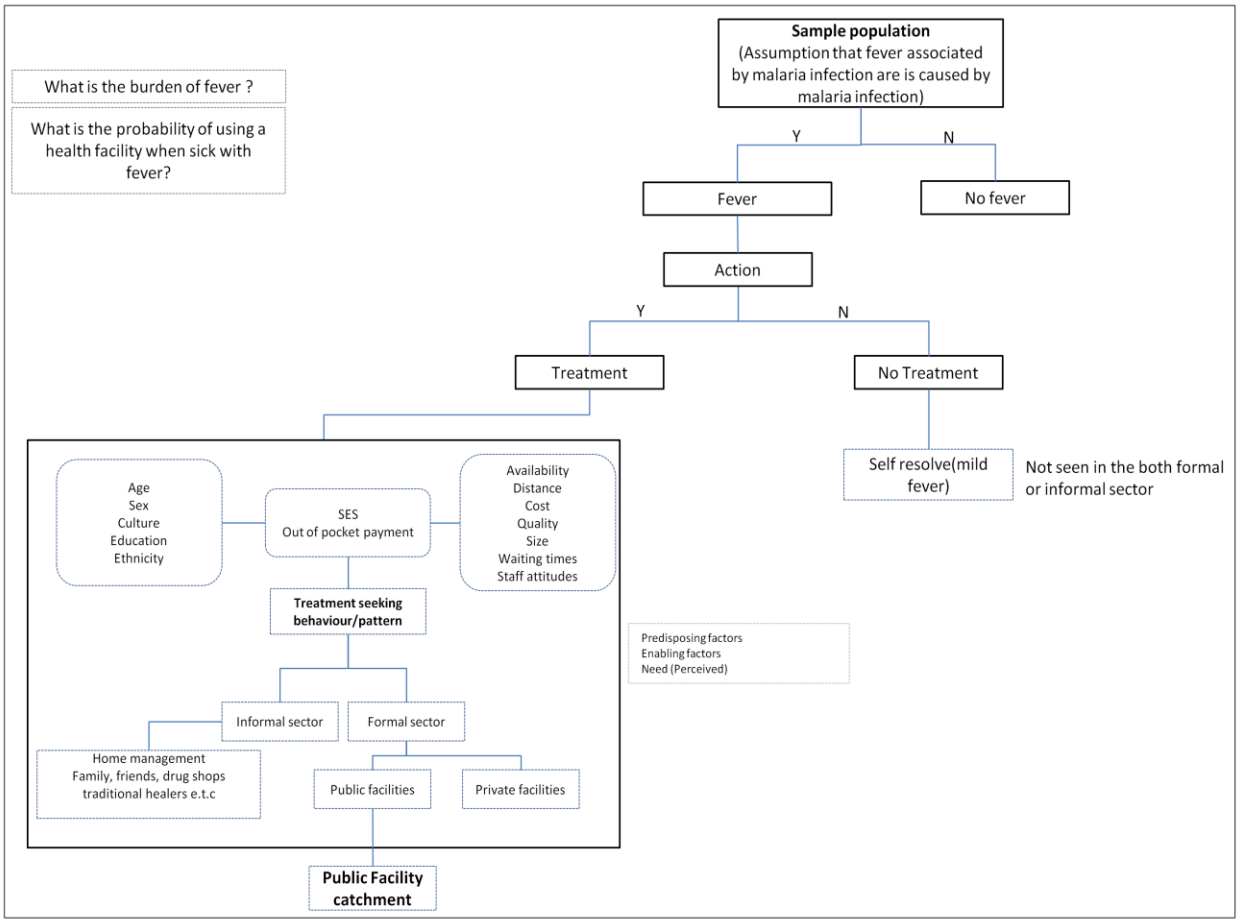

Figure 1.2 Fever burden treatment protocol
Framework for studying fever treatment behaviour showing the different pathways to treatment and some of the fever cases self-resolve with no action taken. Treatment may be sought based on several factors broadly classified as predisposing (age, sex), enabling (socio-economic status (SES) and need (availability, cost) (Aday and Andersen, 1974a). The choice (formal sector or informal sector) vary between different settings. Those that seek treatment in formal public sector are classified within the public health facility catchments.

1.2.4 Measuring malaria morbidity: Epidemiological models

Malaria risk refers to exposure to the malaria parasite (infection) through the bite of the female *Anopheles* mosquitoes that have sporozoites in their salivary glands (Ross, 1910). The morbidity of a disease is often measured using prevalence or incidence indices (Donaldson and Gabriel, 2009). The former is effectively defined as a ratio of cases observed in a population (point prevalence), while the latter deals with new cases arising over a time frame (Donaldson and Gabriel, 2009). Prevalence is usually stated as a rate (Gething *et al.*, 2011b), while incidence

indices, such as the Annual Parasite Index (API), is useful in identifying the determinants of the disease in the general population and expressed as number of cases per 1000 population (Pull, 1972, Ray and Beljaev, 1984). This section reviews the common approaches of measuring transmission.

1.2.6.1 Spleen rate (SR) and parasite rate (PR)

The oldest technique for measuring malaria prevalence in a population (*endemicity*) is the occurrence of enlarged spleens, splenomegaly, introduced in 1847 by Demster in India (Gilles, 2002). One of the main functions of the spleen in the human body is to remove infected cells (Abdalla and Pasvol, 2004). Thus, the prevalence of enlarged spleen or *'spleen rate'*, usually 2-3 times the normal size, is a useful indicator of intensity of malaria transmission (Walter Carr, 1892, Shukla *et al.*, 2011). Historically, the highest incidence of hyper-reactive malaria splenomegaly (HMS), was reported in West Africa (Adedoyin and Fagbule, 1992, Allam *et al.*, 2008), in Zambia (Lowenthal *et al.*, 1980) and for Kassalla province in the Sudan (Allam *et al.*, 2008). Some previous studies have also shown that spleen enlargement can vary based on genetic factors (Greenwood *et al.*, 1987). Spleen rates have lost importance in modern practice with advances in malaria parasitological molecular and serological measures that can isolate the malaria parasites (Shukla *et al.*, 2011).

A popular measure of population prevalence is the percentage of blood smears with malaria parasites measured using microscopy or RDTs (Hay and Snow, 2006, Hay *et al.*, 2008). Microscopy is regarded as the gold standard in identifying, detecting and quantifying malaria parasites (World Health Organizastion, 2011). The advantage of parasite prevalence surveys is that they can potentially cover large geographical areas (Guerra *et al.*, 2007) even though they

may fail to provide precise estimates of infection rates as transmission declines. At a national level, surveys such as the Malaria Indicator Survey (MIS) and the Demographic Health Survey (DHS) incorporate parasitaemia testing, even though they are primarily powered to provide information on the coverage of malaria interventions (Roll Back Malaria Monitoring and Evaluation Reference Group *et al.*, 2005, MEASURE DHS, 2011). Microscopy or RDTs are used to measure parasitaemia in these surveys. One drawback of surveys in low transmission settings is the need for larger sample sizes as well as a need for a greater temporal frequency to adequately capture infection rates (Yekutiel, 1960, Beier *et al.*, 1999).

1.2.6.2 The Force of infection (FOI)

The Force of Infection (FOI), defined as the rate of new infection in the population (Bekessy *et al.*, 1976, Charlwood *et al.*, 1998), is used as an alternative measure of transmission intensity although in practice it requires a follow up of a specific population group over a certain period of time (Yukich *et al.*, 2012). FOI can be measured using different approaches. One approach is the infant conversion rate, the rate at which prevalence increases in young children. This has been demonstrated in studies carried out since the 1970s (MacDonald, 1950) and in the example of the Gambia (Snow *et al.*, 1997). Alternatively, molecular measures have been used (Mueller *et al.*, 2012). This is usually by genotyping parasites of infected individuals such that an occurrence of a super-infection, can uniquely be isolated and monitored (Falk *et al.*, 2006, Yukich *et al.*, 2012). In some low transmission settings, serology has been used to derive FOI using prevalence of antibodies, since infected individuals can remain sero-positive for a long time after infection (Drakeley *et al.*, 2005, Corran *et al.*, 2007, Cook *et al.*, 2011). Yukich *et al.* (2012) show that a reversible catalytic model (Corran *et al.*, 2007) can be used to transform the age profiles of previously infected populations to history of infection, thus, deriving the FOI index.

1.2.6.3 Entomological indices

The reproductive number (R_0) is defined as the number of new individuals infected as a result of introducing a single case to a susceptible population covering a given period, usually annually (Dietz, 1993, Smith *et al.*, 2007). The index is therefore better at measuring spread of disease, and has a historical basis in population and demography (Harrison, 1978) (Macdonald, 1956). The R_0 has important implications for malaria control, especially when it comes to estimating the effort required to eliminate the disease. Several authors have shown that if R_0 is kept low based on control ($R_0 <1$) then the disease can be eliminated (Smith *et al.*, 2007, Gething *et al.*, 2011b). Smith *et al.* (2007) discuss the factors that affect the calculation of R_0 such as population density, population movement, larvae habitat and vector host seeking behaviour including factors that drive individual biting rate and provide a mathematical model for estimating R_0 based on Entomological Inoculation Rate (EIR) and parasite prevalence.

The EIR is defined as the number of infective mosquito bites per individual for a defined period of time, for example per annum or day (Macdonald, 1956, 1957). The EIR is therefore measured as the product of the number of infectious mosquitoes (the sporozoite rate) and the biting rate (Beier *et al.*, 1999, Hay *et al.*, 2000b) but in practice it is challenging to use this index because of the challenges posed by non-standardised entomological techniques (Hay *et al.*, 2000b). For instance, the most direct method of measuring the biting rate is by using human bait and catching the number of mosquitoes that attempt to feed on the person (Hay *et al.*, 2000b). This, however, raises ethical concerns around exposing individuals to infections while alternative techniques, such as light traps, are less attractive because they do not involve direct contact between the mosquito and humans (Le Goff *et al.*, 1997, Hay *et al.*, 2000b, Smith *et al.*, 2005). In addition, there are other uncertainties related to the sporozoite rate. For instance, not all sporozoite-

infected mosquito bites lead to an infection because the amount of sporozoites injected may be insufficient (Smith *et al.*, 2005). Moreover, there are difficulties in ascertaining whether the additional bites in an already parasitemic population (in high endemicity) (Snow and Marsh, 2002, Smith *et al.*, 2005) result in super-infections (Charlwood *et al.*, 1998). In low transmission settings, the reduced number of mosquito catches reduces the effectiveness of using EIR (Githeko *et al.*, 1996, Hay *et al.*, 2000b).

Of the two indices, EIR is more commonly used compared to $R_{0,}$ despite it being labour intensive to estimate and challenging to compare across different sites (Gething *et al.*, 2011b). One study in two communities in Senegal compared the infant mortality rate (IMR) and the EIR (estimated through field entomological surveys) and showed that the two were directly proportional, suggesting a reduction in EIR was related to reduction in IMR (Smith *et al.*, 2004).

1.2.5 The role of surveillance as a method for assessing morbidity and control

The evolution of the word surveillance started in the 1950s as part of eradication programmes and was used as a means of preventing re-emergence of malaria (World health Organizastion, 1957). According to the WHO, surveillance included the identification of infections, investigation, elimination of transmission and prevention as well as cure (World Health Organization, 2007a, 2012a). This could be done on a routine basis in high endemicity areas, thus, constituting a tool for malaria control (Mueller *et al.*, 2011). The WHO recommended surveillance to focus interventions in order to eliminate foci transmission in very low endemicities or where prevalence has been reduced to very low levels to the point that the disease is not considered a major public health problem (WHO/Regional Office for Africa, 2001,

WHO/Regional Office for South-East Asia, 2003). Examples include countries in the pre-elimination or in early consolidation phase (Pull, 1972). There are two broad areas with regard to surveillance. The first is concerned with determining the incidence of disease including the identification of cases and foci infections while the second deals with elimination of the identified cases (Pull, 1972, Ray and Beljaev, 1984).

Identification of incidence comprises both active and passive case detection as well as parasitological screening of both the febrile and non-febrile cases in foci areas (Ray and Beljaev, 1984). Screening of all individuals ensures that even asymptomatic infections are identified (Yekutiel, 1960). In high endemicity areas febrile cases are a common phenomenon and have for a long time been treated as symptomatic infections (Owusu-Agyei *et al.*, 2001). Currently, 41 out of 44 countries in Africa have adopted the use of parasitological diagnosis of suspected fever cases prior to treatment under the 2012 WHO T3 (Test, Treat, Track) initiative (World Health Organization, 2013b, Bastiaens *et al.*, 2014). Approximately 26 countries in SSA have deployed the use of RDTs at community level. Active case detection involves screening febrile cases at household or community level at regularized time intervals (weekly or monthly) during transmission months. Passive case detection identifies cases through health facilities at tertiary level (hospitals), secondary and at primary level facilities (dispensaries and clinics) (Jie *et al.*, 1998, Perry *et al.*, 2007). In some health systems community health workers attached to primary level facilities and private practitioners are additional sources of case identification (Jones *et al.*, 2008). The current WHO guideline recommends treatment of confirmed uncomplicated cases of *P. falciparum* using ACTs. Different guideline exists for special groups such as pregnant women in the first trimester while a combination of quinine plus tetracycline or doxycycline is

recommended as the second line (World Health Organization, 2010a). For severe *P. falciparum* malaria, intravenous artesunate or quinine is acceptable for both children and adults (World Health Organization, 2010a). The recommended treatment for *P. vivax* is a combination of chloroquine (CQ) and primaquine.

1.3 Review on delivery of healthcare in low income countries

1.3.1 Delivery of healthcare in Sub-Saharan African countries: Public or private providers?

Healthcare in both high and low income countries is delivered mainly through a mixture of public and private sectors. Both these sectors refer to organisations as well as institutions that are responsible for both provision and financing of health services (Bennett, 1992). Thus, public-sector institutions are usually fully within state control compared to the private sector, where the state has no exclusive control. Debate has recently developed over the balance of public and private sector provision with calls to recognize the role of the private sector in developing countries (Basu *et al.*, 2012). Basu and others report two major lines of arguments with those in favour of strengthening the public sector pointing to inequalities resulting from private sector provision since private provision is tailored on the ability to pay (Basu *et al.*, 2012). In contrast, those championing the private sector system suggest that government-based institutions are unable to provide sufficient coverage of health services in addition to factors such as poor quality of services within government facilities (Prata *et al.*, 2005). In terms of utilisation for fever treatment, approximately 40% of the population globally use the private sector based on estimates from the household surveys, although the actual estimates vary by country (World Health Organization, 2013b). Although reliable data on private sector is not usually available,

estimates from malaria indicator surveys in some countries suggest that the uses of parasitological diagnosis of suspected fever cases is lower compared to the public sector. Lack of data on cases seen in the private sector affects the estimates of malaria burden which results in uncertainty in estimating appropriate treatment needs and surveillance. This rest of this section outlines the mechanisms and interaction of these two systems in countries in SSA, focusing mainly on provision rather than financing.

Inefficiency, declining quality of services, long waiting times, lack of hospital beds and irresponsiveness of providers are some of the challenges facing the public sector in SSA countries (Bennett, 1992, Quaye, 2010, Flessa *et al.*, 2011, Laudicella *et al.*, 2012). The increase in demand for services (e.g. from the middle-class population (Berman *et al.*, 1995)) coupled with reduced expenditure on health has undermined the ability of many public facilities to provide efficient medical care (Barnighausen and Bloom, 2009, Zikusooka *et al.*, 2009). Figure 1.3 (Page 24) shows general expenditure on health in Sub-Saharan African countries and suggests public sector expenditure is the smallest component of the total health allocations in SSA. Policies such as cost sharing and user fees have had a negative impact on utilisation of publically funded facilities (Gilson, 1997, Nabyonga-Orem *et al.*, 2008, Chuma *et al.*, 2009b, Hadley, 2011). For instance, public health facilities are limited in coverage in some countries and operate at regional or district level (Noor *et al.*, 2009b, Noor *et al.*, 2009d). A survey of providers in three districts in Uganda showed 4.3% constituted publically funded facilities (Konde-Lule *et al.*, 2010). Noor *et al.* (2009) showed, in Kenya, the distribution of providers closely followed the population distribution, but the northern sparsely populated regions were underserved (Noor *et al.*, 2009b). Another study in three districts in Somalia showed that the

public sector comprised approximately 20% of facilities surveyed, but the majority were poorly equipped, lacked essential medicine and the majority of staff lacked essential training (Noor *et al.*, 2009d). The latter highlighted some of the challenges faced by the public sector in many other settings. In Kenya, for example, current policy allows practitioners to render specialist consultative services to the public within public facilities, after attempts to remove an earlier directive that allowed physicians to operate private clinics failed (Muthaka *et al.*, 2004). Lessons from Kenya and other African countries showed that outsourcing of services may often result in competition in tendering and amongst providers for various services which also may lead to unfair practices (Bennett, 1992).

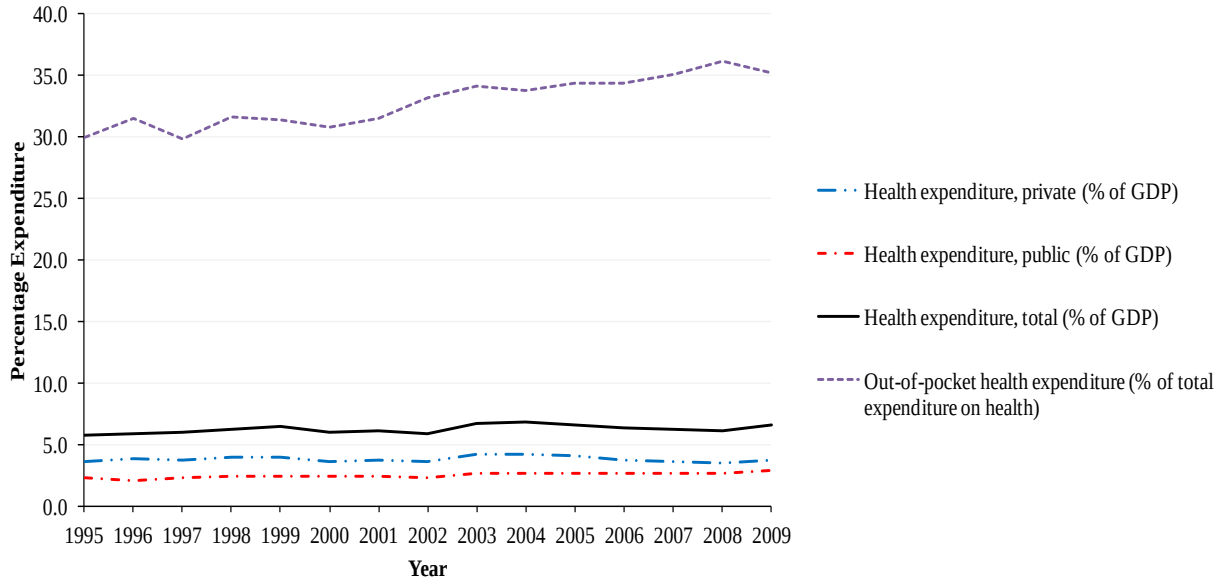

Figure 1.3: Percentage of public, private and out-of-pocket expenditure on health in SSA
Sub-Saharan African data showing the percentage of public, private and out-of-pocket expenditure on health as a percentage of total GDP at all income levels. Data source: World Bank (World Bank, 2014).

There are two main categories of private providers namely: those for profit, and not-for-profit. The private for profit sector comprises of independent providers that aim to provide services based on the ability to pay (Prata *et al.*, 2005). In SSA countries the sector accounts for 30% to

40% of total health sector expenditure (Figure 1.3 above) and includes private clinics, privately owned hospitals and nursing homes, pharmaceuticals, drug shop, vendors and informal traditional practioners, healers or birth attendants (Stekelenburg *et al.*, 2005, Konde-Lule *et al.*, 2010, Basu *et al.*, 2012). These service providers tend to be located in urban centres with their location determined by supply-and-demand factors. Carapinha *et al.* (2011) reviewed the private healthcare sector in five different countries (Kenya, Uganda, Tanzania, Ghana and Nigeria) along with financing. In Nigeria, for example, the majority of providers were located in urban zones and operated independently from government in terms of provision as well as control (such as staff employment) (Carapinha *et al.*, 2011). Government involvement in private owned facilities in countries such as Zimbabwe and Tanzania has been through social franchising (such as training, accreditation, certification and voucher schemes), where clinics and other entities operated within defined guidelines and subsidies (Kumaranayake *et al.*, 2000, Fiedler and Wight, 2003, Prata *et al.*, 2005). However, regulation by the state has in other settings impacted negatively on the private sector. For instance, Kumaranayake *et al.* (2000) indicated that hospitals reduces staff numbers or removed services in response to various government regulations aimed at regulating costs. Examples outside Africa on successful application of franchising were demonstrated in Vietnam (Ngo *et al.*, 2010) and Myanmar (O'Connell *et al.*, 2011) on reproductive health. In the African context, therefore, the suggestion is to strengthen regulations and legislation governing healthcare practices (Muthaka *et al.*, 2004).

The private not-for-profit sector consists of charitable, religious and non-governmental organisation-based healthcare facilities. The private not-for-profit sector may be privately financed by a charitable or religious organisation but provide services to the general public in a

similar way to government hospitals (Bennett, 1992). Therefore, there may be some level of collaboration with state-owned facilities, since the objectives of the charitable or religious organisations tend to be similar to those within state control. The sector has largely grown in many countries for various reasons. For instance, the majority are viewed to be more effective in service delivery compared to government facilities and there is elevated confidence by external donors, who provide financial support, due to increased accountability (Bennett, 1992, Berman *et al.*, 1995). In addition, these providers tend to be located in remote rural areas that tend to reach out to the marginalized populations. Studies that have examined performance of the religious or charitable facilities have identified duplication as a major concern due to poor coordination within the sector (Brugha and Zwi, 1998, Kumaranayake *et al.*, 2000, Muthaka *et al.*, 2004, Flessa *et al.*, 2011). Another group of facilities that may be identified within this category includes mobile units, special treatment facilities as well as health driven programmes (Muthaka *et al.*, 2004). These facilities are usually listed as part of health management information systems.

1.3.2 The role of Health Management Information Systems (HMIS) in Africa

Typically, a Health Management Information System (HMIS) coordinates the routine acquisition of data from health facilities (public and private) and compilation of these data through district, regional and national levels (Abouzahr and Boerma, 2005, Gething *et al.*, 2006, Boerma and Stansfield, 2007) (Figure 1.4). An ideal HMIS, therefore, requires all health facilities to submit reports in all months throughout the year. These include routine disease morbidity data through facilities (passive case detection) and any cases detected at household level (active case detection), mortality rates, determinants such as access, coverage, quality of care, costs and

expenditure. Such data form an integral part of healthcare delivery and are useful for planning, resource allocation as well as assessment of interventions and disease monitoring (WHO/AFRO, 2001, Abouzahr and Boerma, 2005). Within different healthcare systems, tools such as registration forms, patient admission records and clinic data collection forms, district and household surveys are used routinely to gather such data (Teich, 1998, Husk and Waxman, 2004). In reality, however, HMIS are often incomplete in many African countries and many health facilities never report (Gething *et al.*, 2006). In addition, only a proportion of cases present in the formal sector a phenomenon known as the '*iceberg effect'* (Donaldson and Gabriel, 2009). Some of the factors contributing to low facility utilisation include availability of health services, financial factors, geographic access and waiting times at facilities (Breman, 2001). Studies carried out in Kenya suggested cost, distance and opening times as some of the main factors influencing choice and decision to seek treatment in either a public or private sector (Noor *et al.*, 2006, Chuma *et al.*, 2010). Another study in four sites of varying endemicity in Ethiopia by Mustafa *et al.* (2009) found that only about 27% of suspected febrile cases sought treatment from the informal sector and a similar percentage chose self treatment while in a rural district in Zambia, Kalabo district, Stekelenburg study suggested high preference to seek treatment from a traditional healer (62% for women), although not as a result of fever (Stekelenburg *et al.*, 2005). Indeed, in many settings in SSA, multiple treatment from multiple providers is a common occurrence (Kizito *et al.*, 2012) and these in addition to under-reporting make the use of HMIS data difficult for applications such as estimating the disease burden or commodity needs.

Recently, efforts have been placed on identifying the type, extent and causes of failings of HMIS in developing countries and on developing strategies for improvement. Examples of these include the Health Metrics Network (HMN) (World Health Organization, 2008), PARIS21(OECD, 2012), the INDEPTH network (INDEPTH, 2012) and the Integrated Disease Surveillance and Response (IDSR) initiative (WHO/AFRO, 2001). These initiatives are aimed at strengthening health systems in developing countries to be able to generate timely health information useful for decision making at facility, district or national levels. An example of such a collaboration that was launched in 2006 was between the Zambia Ministry of Health (MoH) and the HMN. The HMN supported the Zambia and Eritrea in improving the status of its health information system (HIS) and formulating plans to streamline the HIS to improve performance (Ministry of Health and HMN, 2007). However, until the HMIS improves, various ministries of health in sub-Saharan Africa have little choice but to make critical public health decisions based on grossly inadequate data, often using crude estimates of national and regional burdens (Cibulskis *et al.*, 2011, Mueller *et al.*, 2011). To overcome some of the failings of HMIS or vital registration systems nationally representative demographic and health household surveys have been conducted in almost all countries in SSA since the 1980s.

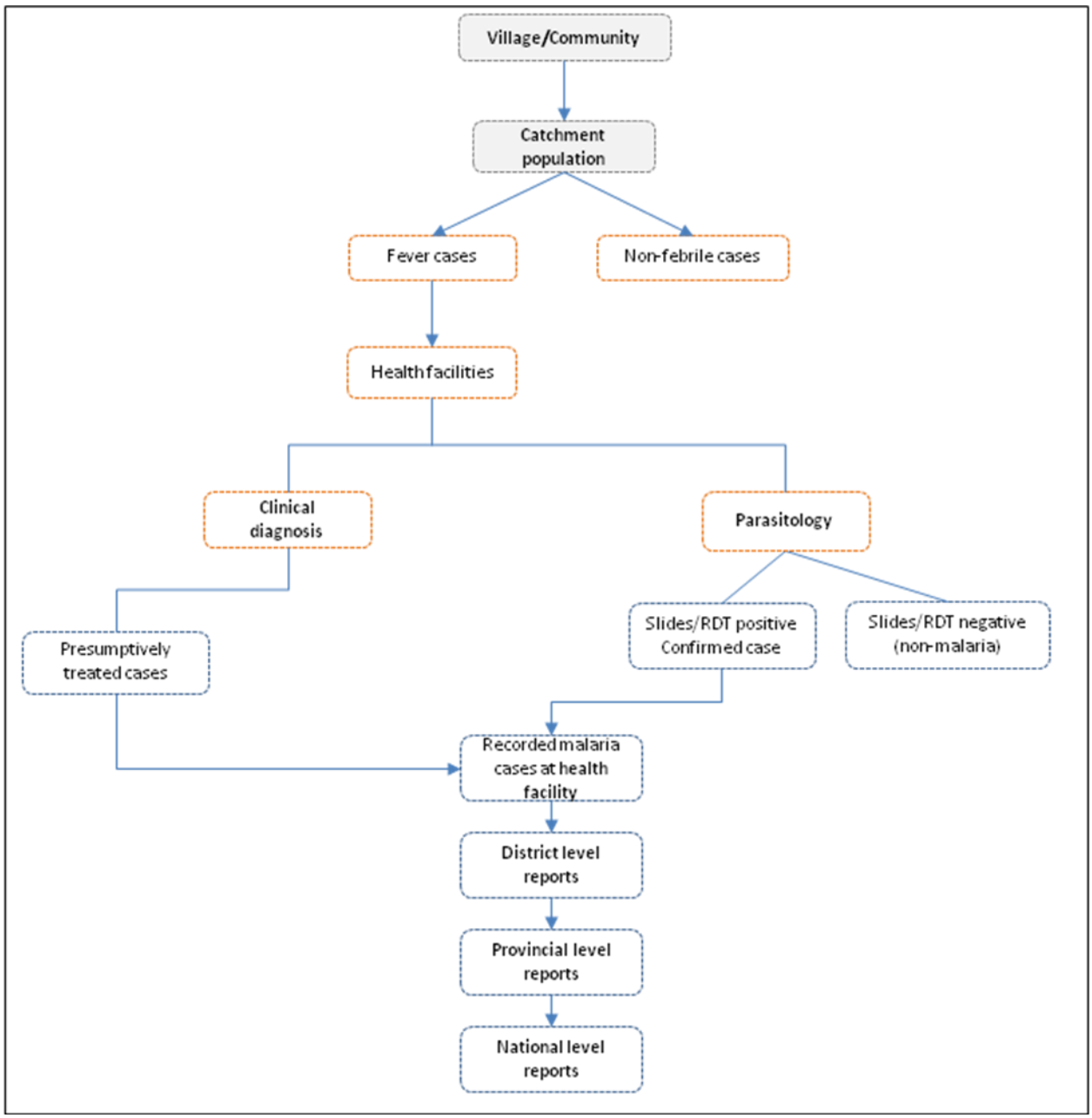

Figure 1.4 HMIS data flow diagram

Data flow within the HMIS showing data assembly flow. The number of cases assemble within the HMIS is affected by rate of utilisation, and nature of reporting to the national system

1.3.3 The role of health facilities as a component of the health system and in disease burden estimation

Health facilities form an important component of a health system as they serve as the main centres for provision of services and also are used to depict system organization spatially (World Health Organization, 2000a). They provide information on the supply side. Information on healthcare need is predominantly provided through household surveys such as the DHS (Lindelow and Wagstaff, 2003). Increased attention is now being targeted at health facilities, in a bid to understand inefficiency, quality of service provision, inequalities of distribution and financing (Lu *et al.*, 2014). Policies such as direct financing are being piloted in an effort to improve utilisation (Chuma and Okungu, 2011).

Although the use of facilities is influenced by different factors, the extent to which individuals interact with formal healthcare facilities reflects their availability and accessibility. In most rural settings in low income countries, family and friends provide first care before treatment is sought either from the private sector or the public sector (Clausen *et al.*, 2000, Dzator and Asafu-Adjaye, 2004, Deressa, 2007, Abu-Mourad *et al.*, 2008). Understanding of issues of planning and inequality in service provision can be improved with knowledge on service providers.

Health facilities provide an alternative data source for measuring disease burden and complement the existing community-based parasite prevalence. They also form the first line surveillance mechanisms for detecting disease epidemics within the population. In low transmission malaria settings, passive cases recorded at health facility can complement the deficiencies of the parasite rate approach. One advantage in using health facility data for disease burden estimation relates to

data being abundant in space since the spatial distribution of health facilities is likely to be congruent with the population distribution, for example, in Kenya (Noor *et al.*, 2009b). The other advantage is data are often collected in an ongoing manner (Mueller *et al.*, 2011). The implication is that data are likely to cover a wider geographic area, relate to the population and be useful in identifying seasonal dynamics of disease.

However, there are challenges in using health facility data. First, available health facility data from most malaria-endemic countries are based on clinical diagnosis of malaria. Approximately 26 countries out of 44 malaria-endemic countries in SSA have only rolled out the use of RDTs. Secondly, even where HMIS is better, only a sub-set of health facilities regularly report data and of these even fewer report every month of year and mainly from the public health sector (Gething *et al.*, 2006, Boerma and Stansfield, 2007, Gething *et al.*, 2007) resulting in incomplete data both spatially and temporally. Additionally, only a subset of fever cases is seen in the formal sector. The variation in utilisation (public and private) affects malaria burden estimation. Utilisation is driven by both human and health system factors.

1.3.3.1 Organisation and spatial location of health facilities

There is a difference between the location of a health facility and location of health services. Health facility location deals with the physical location while the latter requires an additional input of services offered at the facility (Cromley and McLafferty, 2002). Services could, therefore, operate within the confines of other services. The World Health Organisation 2006 report suggested that poor organisation contributes to poor delivery of healthcare (World Health Organizastion, 2006). One of the reasons for poor service delivery is that facility location may be based on pragmatic decisions that fail to consider user behaviour, distance and other

determinants of utilisation (Kiwanuka *et al.*, 2008). In Kenya, the northern region had the least coverage of health facilities even after a considerable growth in the number of facilities over a five year period (Noor *et al.*, 2009b). In Bangladesh a survey in 1997 showed that only about 40% of the population were covered in 1997 through a special government survey (Government of Bangladesh, 1998, Rahman and Smith, 1999) and the results were subsequently used to improve primary healthcare (Omer *et al.*, 2011). The study in Bangladesh demonstrated that provider location information can be useful in understanding healthcare need and can be used in the allocation of health workers to specific population groups. Location can be optimized based on provider-to-population ratios or using algorithms that minimize the distance between population and health facility (Cromley and McLafferty, 2002).

1.3.4 The role of national household surveys as a source of data for health in low income countries

Household surveys are increasingly being used to provide information on demographics and disease burden (Donaldson and Gabriel, 2009). These surveys, for example, the DHS and the MICS, can be used to compare various demographic and health indicators between different countries. Other household surveys are carried out by the national statistical bureaus such as the integrated household budget surveys, welfare monitoring surveys and economic surveys that provide specific information useful to national planning departments. This section will examine the DHS in depth along with the MIS.

The DHS were initiated in mid-1984 by the United States Agency for International Development (USAID) as an extension to the World Fertility Surveys (WFS) and the Contraceptive Prevalence

Surveys (CPS) to monitor key population and health indicators in developing nations (MEASURE DHS, 2011, Short Fabic *et al.*, 2012). The WFS and CPS mainly provided indicators on reproductive health in the 1970s and early 1980s (Short Fabic *et al.*, 2012). The DHS are nationally representative with large sample sizes, usually more than 30,000 individuals, targeting both gender groups using standardised questionnaires. Currently, the DHS is conducted in 90 countries worldwide, 44 in Sub-Saharan Africa (Figure 1.5 Page 35). Other regions include central, south and south east Asia, Latin America and the Caribbean as well as some countries in the Euro-Asia region such as Turkey (MEASURE DHS, 2011). To date, over 250 surveys have been completed, providing information on reproductive health, fertility, population demographics and general health status, nutrition, household characteristics, socio-economic status and infant and child mortality rates (MEASURE DHS, 2011). The HIV module was introduced in 2001 as an additional component while a malaria parasitaemia module was added in 2006 (Short Fabic *et al.*, 2012). The MEASURE DHS programme also supports other surveys such as the AIDs Indicator Survey (AIS), the MIS, Key Indicator Survey (KIS) and the Service Provision Assessment survey (SPA) (MEASURE DHS, 2011).

The AIS provide data for monitoring HIV/AIDS indicators and may include blood testing but usually has a smaller sample size compared to DHS. Some AIS surveys also incorporate malaria testing, for example, the Tanzania 2007-2008 AIS (Tanzania Commission for AIDS (TACAIDS) *et al.*, 2008). The tools used in the MIS module were developed by the Monitoring and Evaluation Reference Group (MERG) or Roll Back Malaria (RBM) (Roll Back Malaria Monitoring and Evaluation Reference Group *et al.*, 2005). In most countries MIS are carried out during malaria transmission months and they also provide key information on malaria

interventions such as the use of mosquito nets amongst high risk groups (children under the age of five years as well as pregnant women), access and use of effective anti-malarial drugs and coverage of IRS (Roll Back Malaria Monitoring and Evaluation Reference Group *et al.*, 2005). Over 20 MIS surveys have to date been completed since 2006 and data are provided through the malaria surveys data portal as well as through the Measure DHS. KIS is designed to provide information about health programs at regional or district levels on planning, child and maternal health and infectious diseases but not relevant for providing data on treatment seeking behaviour for fever (MEASURE DHS, 2011).

Modern DHS survey design is based on two-stage cluster sampling in which clusters, usually Enumeration Areas (EAs), are selected in a first stage and a sample of households selected at the second stage from a household list within the selected cluster (Macro International Inc., 1996). Sampling is usually based on proportion-to-population size in the cluster although in some countries clusters are oversampled in sparsely populated regions to obtain sufficient estimates of indicators, for example, in the Malawi 2010 DHS (National Statistical Office (NSO) and ICF Macro, 2011). These may also apply to urban clusters for areas where urban centres are small (Macro International Inc., 1996). A cluster usually consists of approximately 15 to 30 households geo-referenced using a global positioning systems (GPS) receivers (Macro International Inc., 1996) with an induced positional error of up to 5 km for rural clusters and 2 km in the urban clusters (MEASURE DHS, 2011). It is worth noting that some MIS surveys provide geographic coordinates at the household level. As data from these surveys become available in the public domain, they have increasingly been used to study different research questions (Short Fabic *et al.*, 2012).

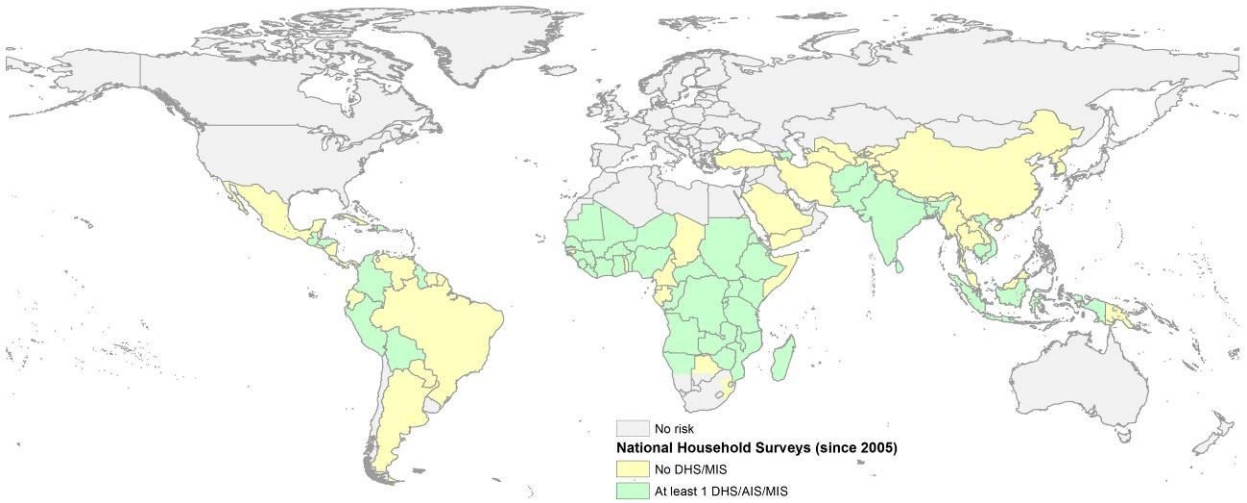

Figure 1.5: Global coverage of nationally representative household surveys
Map showing countries that have conducted a DHS, MIS or AIS survey since 2005 and those with no malaria risk masked out in grey.

1.3.5 The role of Service Provider Assessment surveys (SPAs) in developing countries

Unlike the MIS, DHS or AIS that are carried out at the household level, the SPA surveys are carried out in health facilities. SPAs are part of the MEASURE DHS and provide information on health facility characteristics, child and maternal health, family planning, diseases such as malaria, TB and HIV/AIDS and sexually transmitted infections (STIs) (MEASURE DHS, 2011). The aim of SPAs is to assess the ability of health facilities to provide various services to the population.

Usually a sample of approximately 400-600 facilities is selected from a national census list of health facilities for countries with a greater density of health facilities. In countries with a small number of providers, such as Namibia, all health facilities are usually surveyed (Ministry of Health and Social Services, 2010a) (MEASURE DHS, 2011). Survey tools include: (a) facility audit questionnaires meant to provides information on drug availability, equipment, cost and infrastructure; (b) an observational questionnaire that provides information on patient-physician

interactions and family planning; (c) a health worker questionnaire aimed at providing information on training, qualifications and supervision and (d) the patient exit interview questionnaire on client perception and satisfaction with health services (MEASURE DHS, 2011). In some surveys, additional country-specific questions are included, for example, in Zambia where HIV/AIDs modules were incorporated and in Kenya on obstetrics services (MEASURE DHS, 2011). SPA datasets are disseminated in a similar manner to the DHS and MIS surveys. In this thesis, Namibia analysis utilise data from SPA survey.

1.4 Access and utilisation

1.4.1 Definitions and concepts of healthcare access

In many sub-Saharan countries, poor access to healthcare services contributes to poor health outcomes (World Health Organization, 2000b, O'Meara *et al.*, 2009, Moisi *et al.*, 2011). An important factor in measuring the performance of a healthcare system is the understanding of provision, access and use of healthcare facilities. Aspects of informal healthcare delivery (treatment at home) were discussed in section 1.3. This section outlines in general the concepts of access to formal healthcare service through public or private sectors.

Access is a concept in geographic health that refers to the ability and willingness to use a healthcare facility when there is a need, thus gaining entry into the healthcare system (Aday and Andersen, 1974a, Cromley and McLafferty, 2002). This definition is adopted in this thesis and encompasses the five concepts defined by Penchansky and Thomas (1981). This include: availability, geographic (physical) accessibility, accommodation, affordability and acceptability (Penchansky and Thomas, 1981). *Availability* relates to the ability of various health authorities

or governments to supply health services to the population when in demand (Aday and Andersen, 1974a, Joseph and Phillips, 1984, Cromley and McLafferty, 2002). It includes not only the physical location of a health facility but also the services offered at that facility (Higgs, 2004). *Accessibility* refers to physical or geographic barriers such as distance or travel time as well as direct costs such as those incurred from commuting (Gething *et al.*, 2004, Noor *et al.*, 2006, Schuurman *et al.*, 2006, Tanser *et al.*, 2006, Moisi *et al.*, 2011). *Accommodation* refers to the organisation of these services in order to meet demand, often measured using waiting times (Aday and Andersen, 1974a, Cromley and McLafferty, 2002). Waiting time not only refers to the delay in seeking care but could also refer to the time taken at a health facility before seeing a physician (Dzator and Asafu-Adjaye, 2004, Deressa, 2007). *Affordability* is the ability to meet financial obligations related to medical services (Aday and Andersen, 1974a, Penchansky and Thomas, 1981). Modes of financing healthcare vary based on national health policies and control from the state (Joseph and Phillips, 1984, Bullen *et al.*, 1996). Policies such as out-of-pocket payment may impact negatively on healthcare provision (Dzator and Asafu-Adjaye, 2004, Chuma and Okungu, 2011). Lastly, *acceptability* generally refers to choice based on gender, culture, ethnicity and many other social factors (Haas *et al.*, 2004). These factors influence the behaviour and ability to access health services in different contexts. Of these five dimensions, accessibility and availability are *spatial* measures. Availability relates to provider physical location while accessibility can be derived spatially using metrics such as distance. The other three dimensions of access are considered to be *a-spatial* because they are dependent on non-spatial metrics such as cost, socio-demographics and health system organisation (Guagliardo, 2004).

Utilisation is a far less well understood concept. There are several methods of analysing utilisation, one of them being the calculation of potential use as an indicator of probable entry into a healthcare system (Shannon *et al.*, 1975, Khan, 1992, Khan and Bhardwaj, 1994, Guagliardo, 2004, Apparicio *et al.*, 2008). Previous research on utilization has been carried out mainly at the village level or in specific lower level administrative boundaries such as districts or counties (Leonard *et al.*, 2002, Noor *et al.*, 2006, Tanser *et al.*, 2006). In addition, the focus is usually on a single or combination of factors categorized as 'enabling', 'predisposing' or 'need' as proposed by Andersen (1983) (Andersen *et al.*, 1983). Integrating these factors into a single model is challenging. For example, it is difficult to quantify factors based on individual perceived level of illness or disease severity, while quality of care, for instance, has been defined based on patient satisfaction or health system infrastructure (Rosenberg and Hanlon, 1996, Kizito *et al.*, 2012). A study in Nigeria defined technical quality of care as effectiveness in achieving desired health gains, showing that when quality of care is based on patient satisfaction, this does not necessarily imply quality (Onwujekwe *et al.*, 2011). Even though the actual utilisation is defined by a complex interaction of several factors, most users seek treatment from multiple sources. Secondly, treatment seeking often starts with the use of the informal sector such as drug shops and home management (Littrell *et al.*, 2011).

As a result, measuring access and utilisation are usually dependent on the level at which data are collected and they may not incorporate all behavioural aspects of use. Existing national survey data usually lack information on the name of facility used by household members although this may be modelled by triangulating data on use with information on location reported type of health facility to derive catchments.

1.4.2 Review of geographic access concept

People make the decision to seek medical care when there is a perceived need based on the nature of an illness (Rosenberg and Hanlon, 1996). In most cases, it depends on the nature of the illness and knowledge which may determine whether diagnosis and subsequent treatment will be sought (Girt, 1973, Young, 2004). The choice of a particular provider depends on several factors such as health provider characteristics, financing (user fees) or travel cost, availability of health insurance, cultural and other socio-demographic factors (Nemet and Bailey, 2000, Cromley and McLafferty, 2002). Patient flow from households to providers creates a spatial pattern in geographic space. The pattern can be characterised based on different approaches such as distance or travel times (Cromley and McLafferty, 2002). Proximity is one such measure of spatial access (Stouffer, 1940) and the distance decay curve, a plot of distance (*x*-axis) against probability of using a healthcare provider (*y*-axis) has previously been demonstrated in numerous studies and reviews (Jehlik and McNamara, 1952, Shannon *et al.*, 1969, Girt, 1973, Connor *et al.*, 1994, Apparicio *et al.*, 2008). The hypothesis of using this curve is that utilisation declines with increasing distance between provider and client (patient) (Cromley and McLafferty, 2002). Thus, it is highly unlikely that patients will use a facility located farther from place of residence due to the increasing cost of travel.

Several metrics that use distance have been devised to measure access. The first is distance to the nearest provider using a straight line (Euclidean) measure between patient location (household) and health provider. A review by Shannon *et al.* (1969) suggests that this is the oldest technique in using distance that has remained popular since the 1950s (Wilson and Metzler, 1938, Jehlik and McNamara, 1952, Ciocco *et al.*, 1954). Wilson and Meltzer (1938) discuss utilization of

healthcare facilities in the Arkansas area and demonstrate distance as the main limiting factor using a *decay function*. Jehlik and McNamara (1952) describe the utilization of facilities amongst rural and semi-rural populations, demonstrating the association of distance with incidence of morbidity in Missouri. Similar studies were also carried out by Ciocco and Altman (1954) in Pennsylvania albeit with use of a sophisticated technique of a hyperbolic function while investigating distances travelled by patients in accessing medical services and physicians. The notable contribution by Ciocco and Altman was that utilisation varied inversely with distance (Shannon *et al.*, 1969). Girt (1973) investigated the use of distance using a similar technique in Newfoundland, Canada, employing three different distance decay curves while characterizing patients' consultation patterns (Girt, 1973). Ingram and colleagues (1978) investigated aspects of time as well as distance decay while examining the relationship between access to emergency services and distance. A distance decay curve was subsequently used in the study to derive a catchment area as a concentric circle, around Humber memorial hospital (Ingram *et al.*, 1978).

In Africa, Stock (1983) used distance models in Kano state, Nigeria, to characterise different access patterns based on facility ownership, gender and seasonality. Stock's study depicted different effects of distance on use of health facilities (Stock, 1983). Okafor (1984) similarly used straight line distances to measure accessibility to general hospitals and subsequently delineated rural local government health regions in Bendel state, Nigeria. Muller *et al.* (1998) also used the concept of distance decay in Papua New Guinea, observing that attendance dropped by 50% after 3.5 km to the facility and subsequently introduced the use of non-linear curves in measuring utilization (Müller *et al.*, 1998).

Studies carried out in the 1990s continued to demonstrate the importance of distance in access to health service even after controlling for other socio-demographic factors. Haynes and others (1999) controlled for age, nature of illness (acute, psychiatric, emergency), socio-economic factors and showed that inpatient utilisation dropped with increasing distance (Haynes *et al.*, 1999). An investigation amongst the elderly population in Orleans County in Vermont State, U.S, showed that the majority of the elderly population that sought care travelled shorter distances (Nemet and Bailey, 2000). Buor (2003) also showed that distance is an important factor explaining utilization after controlling for service and transport cost, education level of patient, income and waiting time in Ghana (Buor, 2003). Studies in Kenya by Gething *et al.* (2004) used straight line distance (Thiessen polygons) while delineating catchments for various facilities based on fuzzy logic while Noor *et al.* (2006) compared straight line distances to those calculated using a sophisticated transect algorithm. The latter study by Noor and colleagues highlighted deficiencies in the use of Euclidean distances as they tended to overestimate coverage (Gething *et al.*, 2004, Noor *et al.*, 2006). In recent times, the use of drive times, distance along roads, and cost surface has been used in access based models (Schuurman *et al.*, 2006, Tanser *et al.*, 2006, Apparicio *et al.*, 2008) (Schuurman *et al.*, 2006, Owen *et al.*, 2011, Alegana *et al.*, 2012). These approaches highlight proximity as a factor in utilisation but the approach of using the Euclidean model, as illustrated by Noor *et al.*, (2006) amongst other studies, assumes the phenomenon of interaction with other service providers and may overestimate coverage or remoteness within the defined geographic space.

A different approach of using provider-to-population ratios does not measure distance directly but requires description of the population served by a provider. The number of providers in a

predefined geographic space is usually combined with population estimates to create a ratio of number of people to one provider or physician. Such ratios were used by De Vise (1966) in a study of the distribution of personnel and health facilities in Chicago, U.S., (DeVise and Chicago, 1966) and by Schonfeld *et al.* (1972) in assessing the number of physicians required. The latter study was used to propose measures such as decreasing physicians' patient noncontact time and increasing the number of medical students in an attempt to increase coverage (Schonfeld *et al.*, 1972). In America, provider-population ratios were prescribed as one of several criteria for allocating human resources (stated as 3,500 people to 1 physician) by the Health Education and Welfare (HEW) department in 1978 (Dutt *et al.*, 1986). This approach was also used in the NHS while identifying the local health authority areas as reviewed by Bullen *et al.* (1996). The disadvantage of this method, however, relates to assigning precise population estimates for a spatial region (Bullen *et al.*, 1996). Census data are usually assigned to a polygon while the actual population may vary spatially (Briggs *et al.*, 2007), thus, a single ratio for each polygon is a crude measure of actual case loads per physician or provider. In addition, the actual flow of patient from place of residence to provider is not taken into account when using this approach. The latter will also apply where a mean distance to *n* number of providers is used (Apparicio and Seguin, 2006, Apparicio *et al.*, 2008), and can also be used as a density measure (Cromley and McLafferty, 2010). Density approaches also vary depending on the size and shape of polygons (Parenteau and Sawada, 2011). Raster approaches using the kernel density method (Guagliardo, 2004) are not appropriate for rural areas where providers are sparsely distributed (Cromley and McLafferty, 2010).

The flow of patients from place of residence to provider is spatially measured using spatial interaction models (Shannon *et al.*, 1969, Bailey and Gatrell, 1995, Apparicio *et al.*, 2008). Bailey and Gatrell (1995) compare such spatial models to Newton's gravitational model of the form:

$$Y_{ij} = \sum (s_j / d_{ij}^{\beta})$$

Where, a household member could travel from the origin i to destination j; the attractiveness s is analogous to the masses in a gravity model and the distance component d depends on a coefficient β which also dictates the shape of the decay curve. Thus, the greater the distance or travel time, the smaller the interaction term (Y_{ij}). The summation in the equation is analogous to area potential. An early attempt at such a model was by Ciocco and Altman (1954), in Pittsburgh, while investigating flow between counties using hyperbolic functions of the form $Y = a/x^b$ with Y in the equation representing the frequency of visits to general practioners, which varied inversely to the distance (x) and some exponential parameter b (negative if the frequency decreased with increasing distance) (Anderson, 1956). If no provider characteristic is used, the numerator (a) reduces to unity. Girt (1973) discusses the exponential quadratic forms deriving conditional probabilities based on consultation patterns in sample data from Newfoundland. Linear forms were discussed by Anderson (1996) by introducing an exponential constant. Thill and Kim (2005) illustrated the versatility of the gravity modelling approach using different forms (exponential, log-exponential, power and Gaussian) in characterizing travel by individuals. Statistical methods have in the past been used to depict the optimal spatial pattern (Thill and Kim, 2005) some approaches favouring the log-logistic decay form (De Vries *et al.*, 2009) while

others like in Kwan (1998) proposing a Gaussian instead of power or exponential models (Kwan, 1998, Wang, 2007, Cromley and McLafferty, 2010).

1.4.3 The measurement of distance in GIS

The distance component is the most important in any selected geographic access model. It represents the separation between the origin and destination. The simplest form reviewed in the preceding section is the straight line distance, calculated using spatial locations (x_a, x_b) and (y_a, y_b) as:

$$D = \sqrt{\left[(x_a - x_b)^2 + (y_a - y_b)^2\right]}$$

The above equation applies for two points close enough in space such that the Earth's curvature can be ignored. Otherwise, spherical distances are usually calculated using polar coordinates with an additional correction for the Earth's curvature (Cromley and McLafferty, 2010). The weakness in this form of distance calculation is that physical barriers (topographic effects) are not taken into account. Network distances based on network grids are an alternative (Schuurman *et al.*, 2006, Apparicio *et al.*, 2008) or Manhattan distances measured along a coordinate grid (Cromley and McLafferty, 2010). The Manhattan distance is calculated as:

$$D = |(x_a - x_b)| + |(y_a - y_b)|$$

Internet based tools such as Google Maps ® and Bing Maps ® can estimate the network-based distance readily in addition to mainstream software applications such as ArcGIS (ESRI, Redlands, CA). The disadvantage of the internet tools is that the light-weight software only handle single query calculations, for example, estimating journey distance from one origin to a destination. The ArcGIS (ESRI, Redlands, CA) network analyst extension tool can calculate service areas for facilities with a road network grid as input data, although it is difficult to

calculate actual distance over complex networks with overpasses and intersections or when information on terminal changes, bus stops or waiting times is unavailable (Martin *et al.*, 1998, Martin *et al.*, 2008).

Other studies have preferred the use of measures such as drive times (Jordan *et al.*, 2004). Clark *et al.* (1969) used an economic index model, analogous to the gravity model (Clark *et al.*, 1969), while calculating the economic potential of regions in Western Europe. The study derived maps showing line of equal economic potential (Clark *et al.*, 1969). The study by Jordan and others (2004) in south west England showed a high correlation between drive time to health services and straight line distance. Modern use of trip data in developed countries aims at using *modal splits* where trips made using a train, for example, are assigned to train routes, buses to bus routes or walking to a limited road network (de Dios Ortzar and Willumsen, Black, 1981). The method, thus relies heavily on the nature of the transport network within a country and availability of information on different types of transport modes used by a patient. In other approaches, friction surfaces have been preferred (Leonard *et al.*, 2002, Martin *et al.*, 2002, Tanser *et al.*, 2006, Ray and Ebener, 2008). Martin *et al.* (2002) derived a cost surface using travel time to public and private health providers in Cornwall, in England. Friction surfaces have also been shown to perform better where different modes of transport are used such as vehicular as well as walking and travel speeds for these categorized modes can be varied (Tanser *et al.*, 2006, Rodrigue *et al.*, 2009, Alegana *et al.*, 2012).

1.4.4 Analysing utilisation of health facilities

A classical measure of utilization is by observing true rates of use at a health facility and comparing the rate against expected population (Densen, 1972, Aday and Andersen, 1974a). This

approach is referred to as realized or revealed access (Cromley and McLafferty, 2002). This method relies on recording the number of visits on a monthly basis and comparing to the expected utilisation based on the catchment population. Aday and Andersen (1974) observe that such usage rates can be analysed based on facility type, site, purpose (preventive or curative) and frequency of visits. Although this method is attractive to use since it is based on actual patient data, generalization is not straight forward. For instance, in low income countries, the treatment seeking pattern is not uniform, there are multiple sources of drugs and there is the problem of defining expected population (Agyepong and Manderson, 1994, McCombie, 1996, Goodman *et al.*, 2007).

A different approach of measuring utilisation is by using potential access (Khan, 1992, Khan and Bhardwaj, 1994, Ensor and Cooper, 2002, Guagliardo, 2004). The terms *potential spatial access* and *potential aspatial access* have been used previously, with the first derived based on spatial metrics such as distance while the second approach uses non-spatial metrics such as cost, socio-economic status, gender, age and cultural factors (Khan, 1992, Khan and Bhardwaj, 1994). Implementations of both approaches vary substantially from simple distance models (NoorAli *et al.*, 1999, Noor *et al.*, 2003, Jordan *et al.*, 2004), gravity models (Wang, 2007, De Vries *et al.*, 2009) to statistical and qualitative assessment (Rutebemberwa *et al.*, 2009, Comber *et al.*, 2011, Hadley, 2011) or using small area estimation (Joseph and Phillips, 1984, Cromley and McLafferty, 2002).

In summary, the classifications are based on the healthcare systems as well as a description of the denominator population. Thiessen polygons and provider-to-population ratios assume

uniform use and fail to account for interaction between different service providers. Statistical and gravity models may be more accurate if population and provider characteristics are accounted for. They may also be used subsequently to derive health facility service areas.

1.4.5 Various types of health facility catchments

A catchment represents a zone around an entity (such as a health facility) that draws the majority of users (patients) (Cromley and McLafferty, 2002, Schuurman *et al.*, 2006). The natural size of the catchment area may vary depending on factors such as the underlying population distribution, population utilisation pattern and facility attractiveness (Ensor and Cooper, 2002, Gething *et al.*, 2006, Noor *et al.*, 2006). In some cases, the size of the catchment is fixed (a *mandated* catchment) based on government regulation (Jenkins and Campbell, 1996). However, catchments may overlap where facilities are close to each other in space, a phenomenon that depicts competition between providers (Schuurman *et al.*, 2006). In health related studies, knowledge of catchment areas is useful in understanding access, utilisation and in estimating disease burden.

1.4.5.1 Natural catchments

Natural catchments delineate regions from which patients are drawn given enabling factors such as distance, travel time, cost, quality of service or cultural factors (Cromley and McLafferty, 2002). Natural catchments have been discussed in a school context (Parsons *et al.*, 2000, Martin and Atkinson, 2001) but far less in health research with complexities in representing spatial and non-spatial dimensions. In sparsely populated regions where only one facility may exist, the natural catchment may draw the whole population of the region. Natural catchments may also

overlap especially where there is a dense distribution of facilities (Cromley and McLafferty, 2002). Tanser *et al.* (2001) represented natural catchments for facilities in various rural districts in South Africa by plotting the households where patients originated. The natural catchments in the study areas overlapped where two facilities were located in close proximity as well as at the border of the study district (Tanser *et al.*, 2001). At such boundaries, utilisation was effectively evaluated using household surveys (Tanser *et al.*, 2001). A similar approach was carried out in Kenya but using a fuzzy classification based on reported use (Gething *et al.*, 2004).

1.4.5.2 Mandated catchments

Mandated catchments are common in systems where patients are assigned to various healthcare providers through a registration system. An example of the mandated catchment could be drawn from the implementation of the National Health Service (NHS) in Great Britain where general practitioners (GPs) were required to serve certain population groups (Bullen *et al.*, 1996, Jenkins and Campbell, 1996). The GPs were generally required to provide the geographic extent of areas served although no regulation was provided to control size of the service areas (Jenkins and Campbell, 1996). In modern practice, such service areas can easily be allocated using Thiessen (Voronoi) polygons with the aim of having an equal area share (Burrough and McDonnell, 1998) or using buffers based on distance. Mandated catchment populations are usually affected by changes in residential addresses (e.g. postcode) or population movement (Cromley and McLafferty, 2002). Moreover, methods such as Voronoi polygons may also fail to account for competition between different providers (Jenkins and Campbell, 1996, Cromley and McLafferty, 2002).

1.4.5.3 Empirical catchments

Empirical catchments are based on observational relationships between a patient's location and use of health facilities (Cromley and McLafferty, 2002). Although the form of geographic analysis deployed is usually based on constraining of a spatial interaction model (Bailey and Gatrell, 1995) based on distance or travel time, they are generally not very different from natural catchments especially in sparsely spatially distributed facility locations (Bullen *et al.*, 1996). Where information on patient flow is unavailable, theoretical analysis based on Thiessen polygons may be adopted (Noor *et al.*, 2006). In a developing country context, such as Kenya and Namibia, empirical catchments were derived from household surveys that define patient location, triangulated with facility location information (Noor *et al.*, 2006, Tanser *et al.*, 2006). In developed countries, such as Canada and Great Britain, such catchment areas have in the past been derived based on population movement between place of residence and facility location (Roos, 1993, Bullen *et al.*, 1996, Schuurman *et al.*, 2006, Zinszer *et al.*, 2010).

1.4.5.4 Two-Step floating catchment area (2SFCA)

The 2SFCA method was proposed by Luo and Wang (2003) to measure relative accessibility between different service areas in Chicago. The method as summarized by Cromley and Mclafferty (2010) requires input of service provider locations, population centres as well as some measure of facility capacity. First, a threshold distance or travel time is determined and used to calculate service areas around facilities. Provider-to-populations ratios are subsequently derived for each provider. Secondly, the population centres are used to search for number of providers within a pre-defined threshold distance. Provider-to-population ratios calculated in the first step are then summed for the number of providers within the population centroid, thereby, highlighting regions with greater access (Luo and Wang, 2003, Luo, 2004, Cromley and

McLafferty, 2010). Luo (2004) used this method to measure relative physician need in north Illinois by first drawing catchments (buffers) based on the maximum distance that an individual is willing to travel and measuring relative accessibility based on census tract centroids. The method as outlined by Luo and Qi (2009) has two main deficiencies related to the use of Euclidean distance measures or the use of circular buffers while calculating the provider ratio leaves out providers outside the catchment. An enhanced approach to this method involves incorporating distance decay functions by introducing a weighted distance from provider to population (Luo and Qi, 2009) while other methods have incorporated demand as well as provider characteristics for optimization (Ngui and Apparicio, 2011).

1.4.6 Factors affecting the measurement of healthcare access and utilisation

Several factors affect utilization of healthcare services. These can be grouped largely into three categories classed as *need* or *individual* (perception of illness, age, gender, cultural); *enabling* (cost and socio-economic) and *provider specific* (Distance, quality of care, size, attitude of physician, waiting time) (Aday and Andersen, 1974b, Joseph and Phillips, 1984).

Distance or travel time has already been studied at great length with most studies demonstrating it as the most important factor before an individual decides to seek medical care (Airey, 1989, 1992, Buor, 2003, Moisi *et al.*, 2011). Thus, patients are unlikely to travel a larger distance to a provider due to increasing cost in addition to other factors such as cultural identification or lack of familiarity with farther providers (Kloos, 1990). Studies by Airey (1989 and (1992) showed that improving road condition may improve utilisation rates observed at health facilities. Another example by Miosi *et al.* (2012) showed that increasing travel time to health facility was

associated with greater disease severity. However, some studies have also identified the phenomenon of by-passing of facilities (Akin and Hutchinson, 1999, Leonard *et al.*, 2002). The majority of such cases are usually either due to referral to a higher level facility with specialized treatment or seeking better quality of care (Girt, 1973, Roghmann and Zastowny, 1979, NoorAli *et al.*, 1999). Kloos (1990) illustrated that women are unlikely to seek medical treatment when certain services are administered by male practitioners. Other behavioural factors relate to staff attitudes towards clients, opening hours and availability of medicine (Jayawardene, 1993, Williams and Jones, 2004).

1.5 Disease Mapping

1.5.1 Review of disease mapping approaches in Africa

Disease maps are being used increasingly as tools for decision making in many malaria control programmes. They are useful tools for assessing the impact of various interventions and understanding populations at risk. Historical disease maps were based on expert opinion with simple geographic representation and lacked modern spatiotemporal analysis (Lysenko and Semashko, 1968). Recently, there has been a remarkable improvement in the assembly of malaria data as well as in mapping risk (Snow *et al.*, 2005, Hay and Snow, 2006, Guerra *et al.*, 2008, Gething *et al.*, 2011b) at continental and global level (Guerra *et al.*, 2007, Snow *et al.*, 2012). At county level, household surveys such as the MIS are useful (Roll Back Malaria Monitoring and Evaluation Reference Group *et al.*, 2005) in providing disease data and complement the national level surveillance system data from health facilities.

Remote sensing and GIS tools emerged in the late 1990s and early 2000s in the production of disease maps (Omumbo *et al.*, 1998, Craig *et al.*, 1999, Kleinschmidt *et al.*, 2000). Remote sensing is a scientific tool that broadly involves the study of phenomena at a distance, but focus on Earth observation based on reflected or emitted electromagnetic energy (Campbell and Wynne, 2011). GIS on the other hand is an information system involving collection, assembly, storage, analysis, interpretation, output and dissemination of spatially referenced data (Burrough and McDonnell, 1998). Omumbo *et al.* (1998) used GIS to map malaria risk in Kenya while Craig *et al.* (1999) produced a climate suitability map of malaria transmission at continental level that involved remotely sensed data. The MARA/ARMA project was the first attempt at modelling malaria seasonality at continental level (MARA/ARMA, 2004). Kleinschmidt *et al.* (2000) used a combination of regression and spatial statistical approaches (kriging on residuals) to predict malaria risk in Mali. The use of remote sensing techniques was further demonstrated in several studies (Hay *et al.*, 1998, Hay *et al.*, 2000a, Omumbo *et al.*, 2000, Omumbo *et al.*, 2004, Omumbo *et al.*, 2005) in mapping vector distributions (Coetzee *et al.*, 2000) and parasitic disease (Brooker *et al.*, 2001, Brooker *et al.*, 2002, Rinaldi *et al.*, 2004). What was common to these studies was the integration of GIS and remote sensing techniques and in some cases incorporating external statistical approaches, since the majority of standard GIS software packages have limited statistical modelling capability and are not able to analyse statistically the relationship between environmental covariates and disease.

Model-based geostatistical approaches are able to analyse geocoded data in space and time as well as relate these data to environmental variables (Christakos, 2000, Barnerjee *et al.*, 2004). This approach goes beyond the normal assumption of independence between observations by

quantifying spatial autocorrelation, usually modelled as a function of distance (Cressie, 1993). Covariance between spatial points is quantified as a function of distance, for example, in the analysis of malaria in Kenya and Somalia (Noor *et al.*, 2008, Noor *et al.*, 2009c) and childhood malaria in the Gambia (Diggle *et al.*, 2002). Data may also be referenced to areas or administrative provinces or districts. Areal data models (lattice methods) are commonly used with such data sets by relating two spatial regions using neighbourhood matrices (Barnerjee *et al.*, 2004). Common approaches include the simultaneous autoregressive (SAR) (Whittle, 1954) and the conditional autoregressive (CAR) models (Clayton and Kaldor, 1987). Kleinschmidt *et al.* (2002) used the CAR modelling approach while analysing the incidence of malaria in Kwa-Zulu Natal in South Africa (Kleinschmidt *et al.*, 2002). Point patterns may also be used to evaluate disease clustering as well as determine risk factors associated with disease events (Brooker *et al.*, 2004, Mirghani *et al.*, 2010) but are not discussed in this context.

Post-2005 has seen the development of Bayesian Hierarchical Models (BHM) where inference is based on a posterior distribution $[\theta|Z]$ which requires the likelihood and a *prior* $[\theta]$(with Z as the data and θ are parameters) (Gething *et al.*, 2008, Vounatsou *et al.*, 2009, Schrödle and Held, 2010, Duncan, 2011, Reid *et al.*, 2012). The likelihood based approach relies on marginal probabilities of the unknown quantities given the data (Barnerjee *et al.*, 2004, Cressie and Wikle, 2011), for example, in Craig *et al.* (1999). BHM models partition the mapping process into data models $[Z|Y,\theta]$ (involving distribution of observations), an underlying biological process model $[Y|\theta]$ (for example, in the case of disease) that leads to the observed phenomenon, and the unknown quantities $[\theta]$ (parameters) associated with the process (Barnerjee *et al.*, 2004, Banerjee and Fuentes, 2011). In this way, the process is represented separately and the uncertainties are

quantified systematically in terms of *conditional distributions* (Barnerjee *et al.*, 2004). The important differences here relate to uncertainties recognized in the BHM compared to the frequentist approach where a model is simply fitted to data based on likelihood. Bayes theorem (Bayes, 1763) is used to provide a *posterior distribution* (a conditional distribution of both the process and unknown quantities given the data). Thus;

$$(Y,\theta \mid Z) = \frac{[Z \mid Y,\theta][Y \mid \theta][\theta]}{\iint [Z \mid Y,\theta][Y \mid \theta][\theta] dy d\theta}$$

where $[Z \mid Y,\theta]$ is the data model and $[Y \mid \theta]$ is the underlying process model given the unknown quantities θ. The numerator is a direct product of these quantities but the normalizing quantity requires numerical analysis usually via a simulation approach such as Markov chain Monte Carlo (MCMC). Other approaches include the Laplace approximations, rejection sampling, slice sampling and importance sampling (Rue and Martino, 2007, Cressie and Wikle, 2011). Examples using MCMC at a national level are in analysing malaria transmission in Mali (Gemperli *et al.*, 2006) in mapping the risk of malaria infection in Somalia and Kenya (Noor *et al.*, 2008, Noor *et al.*, 2009c). These methods have also been applied at continental and global scales (Hay *et al.*, 2009b, Gething *et al.*, 2011b, Noor *et al.*, 2014). Lattice methods using a Bayesian framework were used in Malawi and South Africa in analysing fever treatment and malaria incidence (Kazembe, 2007, Kazembe *et al.*, 2007) and in South Africa (Kleinschmidt *et al.*, 2002). The development of numerical statistical analysis using Laplace approximations (Rue and Martino, 2007) may well increase the use of Bayesian approaches in disease mapping (Schrödle and Held, 2010, Ramiro Ruiz-Cárdenas *et al.*, 2012).

1.5.2 Introduction to frequentist methods

Early geostatistical applications were applied in geology and mining, although there are other applications in other disciplines (Zhou *et al.*, 2007, Hengl *et al.*, 2009). Zhou *et al.* (2007) and Hengl *et al.* (2009) review various research studies involving geostatistical applications and suggest that the majority of published articles are applied in the geosciences. These classical methods have rapidly evolved since the 1960s in line with the emergence of statistical computer packages that can easily implement models. They are also useful exploratory data analysis techniques for formulating complex hierarchical modelling approaches (Barnerjee *et al.*, 2004).

The concept underlying geostatistical application is that each observation s in a two dimensional (2D) space, $D \subset R^2$, is a drawn from a distribution (usually Gaussian). Thus, the Random Variable (RV) Z_u at a point u can have a series of outcomes (realizations) in space and relate to another point at a different location based on a function of distance (generally Euclidean distance) (Cressie, 1985a, 1986). The collection of random variables and realizations has strict stationarity if for any set $n \geq 1$, the distribution of $(Z(u_1),.....,Z(u_n)$ is equal to that of $(Z(u_1+h),.....,Z(u_n+h)$ where h is the lag vector in $D \subset R^2$ (Cressie, 1985a, Isaacks and Srivastava, 1989, Cressie, 1990). Second order stationarity is implied if the process has a constant mean, thus, $E(Z(u) \equiv \mu$ and $Cov\{Z(u), Z(u+h)\} = C(h)$ where $s \in D$, $s+h \in D$. Second order stationarity is not strictly required since the desired property is that the mean and variance are homogeneous within a distance h (*intrinsic stationarity*). Thus (Cressie, 1985a, Barnerjee *et al.*, 2004):

$$E[(Z(u+h)-Z(u)]^2 = Var((Z(u+h)-Z(u)) = 2\gamma(h)$$

where $2\gamma(h)$ defines a *variogram* (Figure 1.6 Page 59). The *semivariogram* $\gamma(h)$ is a graphical representation of autocorrelation with the lag distance (Cressie and Zimmerman, 1992, Cressie, 1993, Hudson and Wackernagel, 1994) and given weak second-order stationarity relates to the covariance as:

$$2\gamma(h) = Var(Z(u+h) - Z(u))$$

$$= Var(Z(u+h)] + Var(Z(u)) - 2Cov(Z(u+h), Z(u))$$

$$2\gamma(h) = C(0) + C(0) - 2C(h)$$

thus,

$$\gamma(h) = C(0) - C(h)$$

Cressie (1993) discusses methods of estimating valid variograms using the covariance. One way of dealing with non-stationarity is assuming that at large values of h then covariance $C(h) \to 0$, thus, treating the model as having second order stationarity or constant mean if the drift $d(t) = E(z_t)$ (Cressie, 1985a, Barnerjee *et al.*, 2004). Another approach of dealing with non-stationarity includes the use of low order polynomials and stratification by dividing the area of interest into sub-regions (Cressie, 1985a).

A plot of γ against lag distance results in a *variogram cloud* which is usually diffuse and difficult to interpret scientifically (Isaacks and Srivastava, 1989, Goovaerts, 1997). The semivariogram is a preferred visual plot compared to the variogram cloud because it is averaged at specific distances for $N(h)$ data pairs (Cressie, 1985b, Cressie, 1993). It represents a summary of autocorrelation with a specific distance (radius) (Cressie, 1990) and is valid as long as there is no trend (*isotropy*), along a certain direction. Otherwise, the trend has to be removed before

estimation of a semivariogram (Isaacks and Srivastava, 1989). In addition some discontinuities (measurement errors) may be observed at the origin (where lag distance is zero), usually known as a *nugget effect* (Isaacks and Srivastava, 1989). The term originates from differences observed between two sample locations in gold mining now attributed to spatial, sampling or random errors. Some models advocate for a full nugget effect, thus, a nugget incorporated at zero distance, in some cases zero nugget effect is included at zero distance (Goovaerts, 1997, 2006, 2010).

Zonal anisotropy is usually rare in practice but it can be assessed by plotting variograms in different directions and assessing their similarities (Goovaerts, 1997). Zimmerman (1993) describes different forms of anisotropy related to the range, nugget and sill. For *range anisotropy* a general suggestion is to increase the range along the axis of variation to reduce the effect of covariance structure perpendicular to the minor axis while dealing with (Goovaerts, 1997) or incorporating a nested model (combination of two different models) (Zimmerman, 1993). The other empirical variogram is calculated by:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Zu_i + Z(u_i + h)]^2 \text{ with covariance computed as :}$$

$$C(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} Z(u_i).Z(u+h) - m_{-h}.m_{+h}$$

where $m_{-h} = \frac{1}{N(h)} \sum_{i=1}^{N(h)} Z(u_i)$ and $m_{+h} = \frac{1}{N(h)} \sum_{i=1}^{N(h)} Z(u+h_i)$

The values of m_{-h} and m_{+h} correspond to lag means of the tail and head values, while ordered values of $C(h_1),.....,C(h_n)$ are referred to as an auto-covariance function or spatial covariance (Goovaerts, 1997). The correlation between various data values can be summarized using a correlogram (Isaacks and Srivastava, 1989, Goovaerts, 1998). Gooverts (1997) provides a generalized function for the above equation representing the variogram, which is simply the first moment of inertia according to Isaaks and Srivastava (1989), by changing the power from 2 (the classical form of estimation) to a value $\omega = 1$(madogram) or $\omega = 0.5$(rodogram) (Goovaerts, 1997). Thus,

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Zu_i + Z(u_i + h)]^{\omega}$$

Permissible semivariogram models can be fitted such as the linear, spherical, exponential, power, Gaussian, and dampened hole amongst other model forms. Mathematical formulations of these models are illustrated in Isaaks and Srivastava (1989) and Banerjee *et al.* (2004). For example the exponential model takes the form:

$$\gamma(h) = \begin{cases} \tau^2 + \sigma^2(1 - \exp(-\phi h) & for \quad t > 0 \\ 0 & otherwise \end{cases}$$

Where τ is the nugget, σ is the sill while ϕ is the range, the distance above which there is no spatial autocorrelation between pairs (Cressie, 1985a).

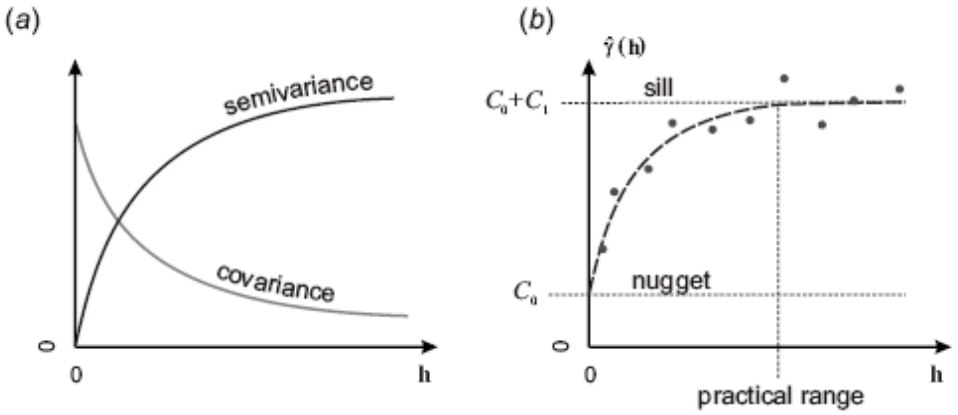

Figure 1.6: Semi-variogram and Covariance
Figure (a) Illustrates the relationship between semi-variance and covariance (*y*-axis) with lag distance (*x*-axis) while (b) shows the variogram parameters.

1.5.3 Geostatistical models

Geostatistical modelling has two broad objectives, the first relating to characterisation of spatial structure in data relating to the mean and variability between observations. The second aim of these models is to carry out predictions at unobserved locations (Cressie and Zimmerman, 1992). Interpolation is assumed if prediction is carried out within the data range $\{ Y(s) : s \in D \}$) for sites S_n and *Y(s)* observables or extrapolation if predictions are carried out beyond the data range. The latter is often discouraged due to uncertainties associated with prediction (Isaacks and Srivastava, 1989). The spatial linear predictor is expressed as $\sum \ell_i Y(s_i) + \delta_0$ (Barnerjee *et al.*, 2004). In general, $\delta_0(.)$ corresponds to errors with $E\{\delta_0(.)\} = 0$ and $E\{\delta_0(.)^2\} < \infty$ for all $s \in D$ (Cressie and Zimmerman, 1992). Thus estimating $\gamma(h)$ (the variogram) without any covariates, is often termed as ordinary kriging.

Universal kriging (Cressie, 1993) applies when covariates are incorporated into the model. Thus the linear predictor, maintaining the same notation for $\sum$ (covariance) as above with nugget effect can be expressed as:

$Y = X\beta + \varepsilon$ where $\varepsilon \sim N(0, \sum)$ further discussions in (Barnerjee *et al.*, 2004) (section 2.4)

Many sets of models can be used in practice for spatial prediction. These include the generalized linear models (GLMs), general additive models (GAM), semi-parametric regression approaches and geo-additive models (such as Bayesian Maximum Entropy) (Shepard, 1968, Christakos, 2000, Hengl *et al.*, 2009). Model choice largely depends on the problem at hand, and validation can often be assessed using an independent control dataset or by cross-validation with a subset of the original data. For ordinary kriging, for example, the mean prediction error $E\left[(Y(S_0) - f(y))^2 \mid y\right]$ is often assessed along with the root mean square error (RMSE) (Isaacks and Srivastava, 1989, Barnerjee *et al.*, 2004). Cross-validation can be carried out by (a) dividing the data set into *k* parts (validation and prediction set), (b) leave-one-out approach where each point is used iteratively and (c) jack-knife procedure that is similar to leave-one-out approach but estimates biases in the statistical method rather than the data points (Cressie and Zimmerman, 1992).

1.5.4 Areal-data models

Spatial data may often be referenced to irregular polygons such as districts, regions or census units or regularized grid cells. Spatial patterns in such measurements are modelled via area models, typically the conditional autoregressive (CAR) or the simultaneous autoregressive

approach (SAR) (Brook, 1964, Besag, 1974). This section will briefly discuss the CAR models which are more popular. Indeed literature considers SAR as a special case of CAR (see the discussion by Cressie, (1993). The CAR analytical framework involves spatial smoothing where observations in neighbouring spatial units are often pushed up (in small observation and large neighbours) or pushed down (for large observations and small neighbours) (Cressie, 1993, Stern and Cressie, 1999). The level of smoothing applied depends on the modelling framework as well as the physical phenomenon under study. Benerjee *et al.* (2004) observes that maximal smoothing yields common value estimates between the spatial units under study and a more suitable smoothing approach should take into consideration the arrangement of spatial units to yield optimal spatial variation. A general problem common to this approach, however, relates to change of statistical outputs as a result of change in shape or size of the geographic unit, the *modifiable areal unit problem* (MAUP) (Robinson, 1950) although hierarchical models have in the past been proposed to mitigate outcomes related to MAUP (Barnerjee *et al.*, 2004) (pg 182 to pg 205).

The general formulation of areal data models is the introduction of a spatial structure via a neighbourhood matrix w_{ij} assuming that $Y_1,.....,Y_n$ corresponds to a set of observations for spatial areal units $1,.....,n$ (Barnerjee *et al.*, 2004). The neighbourhood matrix w_{ij} represents weights that have been introduced based on different functions, such as, $w_{ij}=1$ for i and j with a common boundary or $w_{ij}=0$ (otherwise). Examples of such formulations were illustrated in mapping rates of cancer (Bernardinelli and Montomoli, 1992) and in incidence of malaria (Kleinschmidt *et al.*, 2002). Other forms of the weight matrix can be based on distance between centroids of various geographic regions (Barnerjee *et al.*, 2004).

The basic idea of the CAR model lies in the joint dependencies provided by a neighbourhood matrix as formulated by Besag (1974). The conditional dependencies among a set of RVs result in a Markov Random Field (MRF) such that the probability $P(x_i \mid x_1,....,x_{i-1},x_{i+1},.....,x_n)$ depends on x_j for $j \neq j$ site and $j(\neq i)$ is a neighbour of i. An imposed condition is that of positive definiteness where probability $P(0) > 0$ (Besag, 1974).

1.5.5 Review of Spatiotemporal methods

Spatiotemporal modelling involves measuring processes that occur both in spatial and temporal domains (Christakos, 2000, Barnerjee *et al.*, 2004). Earlier attempts at modelling spatiotemporal fields resulted in static maps based on different time (t) instances. Tobias and Salas (1985), for example, compared different mechanical and statistical interpolation techniques for precipitation data spanning 30 years with the aim of comparing outcomes at different times (Tabios and Salas, 1985). Hudson and Wackernagel (1994) also modelled temperature for the month of January in Scotland using kriging with external drift while Goovearts and Chiang (1993) investigated soil-nitrogen mineralization over the winter period (Goovaerts and Chiang, 1993). Other earlier studies used spatial time series (Cliff *et al.*, 1975, Bennett, 1979) which could not be interpolated at unobserved locations and required external computation. Furthermore, time-series approaches could not sufficiently relate the spatial aspect of data (ordered or random) with time which is usually ordered (past to present to future) (Cressie, 1993).

In the late 1980s, the spatiotemporal geostatistical models incorporated time as an additional domain to existing spatial statistical numerical methods by assuming a separable correlation

structure of residuals (Egbert and Lettenmaier, 1986, Rohuani and Hall (1989), for example, used this approach in summing two variograms. A similar method was used by Egbert and Lettenmaier (1986) when modelling atmospheric variables in the U.S., Rouhani and Myers (1990) later identified caveats when multiplying to space and time variograms due to dimensions (2D for space and 1D in time domain) as well as due to scale or units of measurements. Secondly, earlier experiments indicated that geostatistical properties such as isotropy were easier implemented when in the spatial domain compared to the temporal domain (Rouhani and Myers, 1990).

A slight deviation from these models used a product of the correlation structure in space and time. For example, in (Guttorp *et al.*, 1994) on ozone-monitoring applications and on rainfall acidity levels (Loader and Switzer, 1992). Dimitrakopoulos (1994) proposed a product-sum covariance structure that was non-separable and where units of measurements were converted to a common measure (Dimitrakopoulos and Lou, 1994). The conversion of units (space and time) to a common measure, however, meant that interpretation of autocorrelation was lost. Several permissible product-sum covariance functions were proposed subsequently in the 1990s (Cressie and Huang, 1999, Kyriakidis and Journel, 1999) and later on with trend modelling using polynomials, Fourier transformations and a mixture of the two approaches (Kyriakidis and Journel, 1999). Cressie and Huang (1999) reviewed space-time covariance functions starting with the limiting case of the product separable models, without space and time interaction, to full product-sum models that support interaction. It was not until the early 2000s, given the contributions by De Iaco *et al.* (2001), that these stationary non-separable models (generalized product-sum covariance) became easily implementable (De Cesare *et al.*, 2001, De Iaco *et al.*,

2001, Kyriakidis and Journel, 2001, De Cesare *et al.*, 2002). The main constraint imposed on coefficients of the product-sum model involved the requirement for positive definiteness (global sill ≥ 0) (De Iaco *et al.*, 2001). De Cesare (2002) published several FOTRAN related programs for implementing product-sum covariance structures modified from earlier GSLIB application (Deutsch and Journel, 1998). Examples of such model formulations were also provided in Kyriakidis and Journel (2001). There has been a recent improvement to the original programs published by De Iaco in 2001 by incorporation of multivariate variables or covariates (De Iaco *et al.*, 2005) and through subsequent published code (De Iaco *et al.*, 2010, De Iaco *et al.*, 2011, De Iaco and Posa, 2011). The increasing availability of software in recent times (post 2000s) has also seen more applications involving space-time interactions as well as hierarchical model based approach.

1.5.6 Spatiotemporal geostatistical and areal models

The introduction of time into spatial process models introduces model complexity due to specifications regarding spatial autocorrelation and temporal autocorrelation (Egbert and Lettenmaier, 1986, Rouhani and Myers, 1990). In this context, as is similar to spatial methods, distinctions are made based on the type of data where Gaussian process models are typically used for point referenced data while CAR specifications apply to the areal data types. Another complexity related to modelling spatiotemporal data is that of missing data (Christakos, 2000, Barnerjee *et al.*, 2004). Banerjee *et al.* (2004) discuss the problem of missing data as (a) that of spatial positions where predictions are performed to points with no observations as in kriging, (b) missing time points and (c) based on both space and time. The latter may be treated as cases of both interpolation and extrapolation and a hierarchical modelling approach adopted.

Joint space-time formulation requires simultaneous observations in space and time, based on RF $Z(s,t),(s,t) \in D \times T$, separated by lag vector (h,τ) where $h = s - s'$ and $\tau = t - t'$ refer to spatial and temporal lags respectively. Kyriakidis *et al.* (1999) reviewed the single and multiple separable spatiotemporal RF models along with limitations attached to these modelling frameworks such as lack of interaction in separable structures (Kyriakidis and Journel, 1999). In general, assuming stationarity, the Gaussian spatiotemporal process model is decomposed into a global mean $\mu(s,t)$ and a residual component $\varepsilon(s,t)$ based on a linear combination of residuals (Loader and Switzer, 1992, Kyriakidis and Journel, 1999, Gelfand *et al.*, 2003). Thus:

$$Z(s,t) = \mu(s,t) + \varepsilon(s,t) \qquad \forall (s,t) \in D \times T$$

The hierarchical models may include a vector of covariates $x(s,t)$ such that $\mu(s,t) = x(s,t)^T \beta$ with β coefficients. Further, $\varepsilon(s,t)$ may be decomposed into a Gaussian white noise component $e(s,t)$ and a mean-zero Gaussian parameter $\omega(s,t)$ (Barnerjee *et al.*, 2004, Banerjee *et al.*, 2008).

Often the assumption of stationarity is violated in many space-time models. Examples of these include disease mapping applications where data may often exhibit non-stationarity due to ecological or external factors such as the impact of interventions (Diggle *et al.*, 1998, Diggle *et al.*, 2002, Gemperli, 2003, Gemperli *et al.*, 2006). For non-stationary models, the mean part can be decomposed such that $E\{\mu(s,t) = m(s,t)$ is a spatially and temporally varying function (Kyriakidis and Journel, 1999, Barnerjee *et al.*, 2004). Deterministic models may also be included and written in regression form as (Goovaerts, 1997):

$$\mu(s,t) = \sum_{i=0}^{I}\sum_{j=0}^{J} b_{ij}\, f_{ij}(s,t) \qquad \forall (s,t) \in D \times T$$

Where the b_{ij} are unknown coefficients and $f_{ij}(s,t)$ are basis functions suitably selected to model the mean for data spanning $i = 1,.....,I$ in space and $j = 1,.....,J$ in the temporal domain.

Inference on non-stationary cases based on the above deterministic model is usually problematic because the covariance structure of the residual component is usually not readily available. Rouhani and Hall (1989) proposed the use of Intrinsic Random Functions (IRF) to determine the generalized covariance by considering the linear combination of differences in space and time of the data. Another possible remedy, although not simple, is to base inference based on data points that do not exhibit drift (Kyriakidis and Journel, 1999).

A similar model is adopted for areal data, where $Z_{it} = \mu_{it} + e_{it}$ for the i^{th} polygon at time t (Barnerjee *et al.*, 2004). Using the same notation for covariates and decomposing the residual part:

$$Z_{it} = x_{it}^{T}\beta + \varepsilon_{it}$$

With the ω_{it} representing the spatiotemporal random effects often modelled via CAR while the ε_{it} represent the unstructured unobserved effects. Area unit data may be based on counts. Thus, the Gaussian specifications are usually replaced with a Poisson model in such cases (Bernardinelli *et al.*, 2007).

1.6 Purpose and scope of the study

1.6.1 Justification of the study

There has been an increase in funds targeted towards healthcare programmes in many sub-Saharan African countries (Snow *et al.*, 2010b, World Health Organization, 2014d). This has been aimed at improving the delivery of interventions and extending primary care to the poor population to achieve the MDGs on reducing under-five child mortality by two thirds (MDG 4) and combating malaria and other infectious diseases by 2015 (MDG 6) (United Nations Development Programme, 2003, United Nations, 2010). In areas of low malaria transmission, such as the selected case studies (Namibia, Eritrea and Afghanistan), parasite prevalence estimates may be inefficient due to: (a) a requirement for large sample sizes to detect the low infection rates, and (b) the high seasonal variability of infections in low transmission settings (Hay *et al.*, 2008, Gething *et al.*, 2011b). Passive and active case detection is recommended by WHO in such settings (World Health Organizastion, 2007).

Most low transmission countries do not have functioning active case detection systems. These countries rely on passive case detection. Ideally, the use of such data for accurate estimation of disease incidence requires that all cases are parasitological diagnosed at health facility level and are reported through the HMIS; the denominator catchment population of the health facilities is known; and knowledge of the overall burden of fever within the community can be quantified (Breman and Holloway, 2007, Mueller *et al.*, 2011). However, as discussed in Section 1.3, countries report a mixture of confirmed and suspected cases, data is usually reported through the public health sector, the reporting rates are spatially and temporally incomplete and a considerable proportion of fevers are treated outside of the public health sector at home or in

private health facilities. In addition, few countries have mapped their health facilities making it difficult to link cases to catchment population to quantify incidence.

Therefore the aim of this study was to tap into the vast array of available national household survey data to model health facility catchment population and model disease incidence based on HMIS data in selected low malaria transmission countries. Some of the challenges in using the HMIS data are subsequently addressed in this thesis. For example, the adjustment for health facility utilisation by the population, imputation of missing data and adjustment for reported suspected malaria cases using slide positivity rates. The results were shared with country-specific ministries of health and national malaria control programmes to support planning and malaria elimination efforts.

1.6.2 Objectives

1.6.2.1 Main objective

The main objective of this thesis was to estimate the spatio-temporal distribution of malaria incidence for Namibia, Afghanistan, and Eritrea based on the treatment of fever by developing models of public health facility utilisation and deriving the catchment population.

1.6.2.2 Specific objectives

The specific objectives were:

1. To assemble geospatial health facility databases for Namibia, Eritrea and Afghanistan, malaria case data within the health facilities, household survey data on their use and other spatial ancillary data.

2. To model health facility catchments and catchment populations using treatment-seeking patterns from household surveys, travel times, facility characteristics, and population distributions.
3. To evaluate the occurrence and distribution of malaria cases in public health facilities over time and space.
4. To estimate malaria clinical burden at decision making units for the national malaria control programmes in the study countries.

1.7 Thesis outline

This thesis focuses on three low malaria transmission case studies: Namibia (Southern Africa), Afghanistan (South West Asia) and Eritrea (Horn of Africa). Table 1.1 provides summary characteristics of the three study countries. These countries were selected because they are all at a similar level in terms of malaria transmission. For example, *P. falciparum* prevalence in Namibia was 1.3% in 2010, in Eritrea 1% and <1% in Afghanistan. Therefore, the study countries provided relevant examples where research on mapping malaria in low transmission settings at a national level, could be explored using data on the number of presumed and confirmed malaria cases from HMIS. The national malaria control programmes were also keen in facilitating assembly of the HMIS data. In addition, the household data useful in modelling the utilization of health facilities were available in the public domain from the wealth of national household surveys of MIS, DHS or MICS surveys. For each country, geographic access to health services via probabilistic approaches was analysed followed by estimation of malaria burden.

Table 1.1: Summary of study sites characteristics

Country	Malaria parasites	Main malaria vector(s)	Mean parasite prevalence (percentage) [2010]		Overall percentage persons who slept under ITN[1] last night [year]	Estimated percentage public health sector use for fever treatment
			P*f*	P*v*	MIS	
Eritrea	*P. falciparum; P. vivax*	Anopheles arabiensis	1	No data	55.1 [2012]	60.9
Namibia	*P. falciparum*	An. arabiensis; An. gambiae; An. funestus	1.3	NA	22.9 [2009]	65.3
Afghanistan	*P. falciparum; P. vivax*	An. superpictus; An. culicifacies; An. hycranus; An. pulcherimus; An. fluviatilis; An. stephensi.	0.1	0.6	15.0 [2011]	44.3

1. Insecticide Treated Net (ITN) is (a) a permanent net that does not require any treatment or (b) a pre-treated net obtained within the last 6-12 months or (c) a net that has been soaked with insecticide within the past 6-12 months.
2. NA – Not applicable

Chapter 2 of this thesis is a case study of Namibia. Namibia declared an elimination ambition, being part of the elimination-eight (E8) initiative (Southern Africa Roll Back Malaria Network (SARN), 2010). The elimination-eight comprised four first-tier countries aiming for elimination (Botswana, Namibia, South Africa and Swaziland) and four second-tier countries (Angola, Mozambique, Zambia and Zimbabwe) in a control phase (Malaria Elimination 8 Ministerial Meeting, 2009, Ministry of Health and Social Services, 2010c). In 2010, Namibia launched a national malaria strategy for the period 2010 to 2016 with the goal of achieving pre-elimination by 2016, thus reducing malaria case incidence to less than 1 per 1000 population (Ministry of Health and Social Services, 2010c, d). The aim of the study in namibia was to assess the baseline incidence in 2009 upon which future disease trends can be compared. The chapter focused mainly on modelling of *Plasmodium falciparum* malaria incidence in northern Namibia using Bayesian approaches to assess feasibility of pre-elimination by 2016. The pre-requisite of this analysis was the analsis of public health facility utilisation described in Master's thesis and

subsequently published in 2012[1]. Therefore, the section on modelling public healthcare utilisation was only mentioned, briefly, to provide context and uniformity with other country case studies. *P. falciparum*, the main malaria species in Namibia, incidence was modelled using HMIS data for 2009 at the constiuency level and summaries provided at health district level useful for NMCP planning and decision making. The findings[2] on incidence were discussed in the broad context of feasilibility of achieving pre-elimination targets by 2016 highlighting regions where concerted control is required in Namibia.

Chapter 3 is a case study of Afghanistan in Asia where, despite instability and poor infrastructure, substantial resources have been invested in malaria control in Afghanistan since 2000 with financial support from external agencies, notably the Global Fund to fight AIDS, Tuberculosis and Malaria and the United States Agency for International Development (USAID) (Ministry of Public Health, 2008b). An immediate aim by the National Malaria and Leishmaniasis Control Programme was to reduce case incidence by 60% by 2013 in addition to improving case management and vector control. The focus here was to track progress towards the national target between 2006 and 2009 and provide estimates of clinical burden of *P. falciparum* and *P. vivax*. A Similar analytical framework to Namibia was adopted except that there were two main malaria parasites in Afghanistan, the *P. falciparum* and the *P. vivax*. The analysis of incidence aimed at identifying the co-distribution of the two parasites and

[1] Alegana VA, Wright JA, Uusiku P, Noor AM, Snow RW, Atkinson PM (2012). Spatial modeling of healthcare utilization for treatment of fever in Namibia. *International Journal of Health Geographics,* 11: e6.

[2] Alegana VA, Atkinson PM, Wright JA, Kamwi R, Uusiku P,Katokel S, Snow RW, Noor AM (2013). Estimation of malaria incidence in northern Namibia in 2009 using Bayesian conditional-autoregressive spatial-temporal models. *Spatial and Spatio-temporal* Epidemiology 7: 25-36.

implications for malaria control and elimination. The findings in from of spatio-temporal incidence maps were discussed at the end of the chapter.

Chapter 4 focuses on Eritrea in the horn of Africa. Eritrea is one of few countries in Africa that have recently reduced malaria burden by >75%. Eritrea has been in a consolidation phase since the mid 2000s and the NMCP is aiming for pre-elimination. The national malaria control programme has been in consolidation phase, but has set ambitions for pre-elimination in current malaria strategy (East Africa Roll Back Malaria Network (EARN), 2013). Eritrea was similar to Afghanistan in terms of prevalence of both *P. falciparum* and *P. vivax*. However, majority of the burden is from *P. falciparum*. This chapter, therefore, explored implications for moving to pre-elimination given the trends incidence.

The discussion in chapter 5, summarises findings from the three case studies and assesses the possibilities for sustained control on pre-elimination. The policy implications for each case study based on household and HMIS data, modelling healthcare access as well as disease incidence are also discussed. The future potential of the methods, particularly on the use of HMIS data to describe malaria incidence in low endemic regions rather than prevalence from community surveys is also discussed. The last section of the chapter points out some limitations of the modelling approaches used in this thesis and outlines recommendations for future studies.

1.8 Modelling approach used in the thesis

1.8.1 A mini review on methods of identifying high risk populations in low transmission settings

A literature search was conducted from Pubmed and the Web of Science to identify research related to low malaria transmission settings from 2005 to 2014. The search generated 2,335 articles including journal publications, book sections, and conference proceedings. These were reduced to 88 after screening the titles and abstracts to include only research related to malaria risk in low transmission settings. Commonly excluded themes were malaria interventions, clinical trials, travel health, drug efficacy and resistance, biological and physiological studies as well as diagnostics.

Of the 88 studies, 70 had been published during or after 2010. The majority involved parasite prevalence either directly estimating prevalence or comparing the performances of different approaches (RDTs, microscopy, PCR and serology). Three studies in Sri Lanka (Rajakaruna *et al.*, 2010), Zambia (Davis *et al.*, 2011) and Swaziland (Sturrock *et al.*, 2013a) demonstrated the usefulness of combining passive case surveillance with reactive case detection to identify asymptomatic infections. Except for the study in Swaziland, most studies were conducted over small areas focusing on one district or province and in some cases involving only one health facility. This illustrates the challenge of conducting (sero)-prevalence or molecular-based studies at the national level. In summary, there was an improvement in techniques (such as using PCR and advanced statistical methods) with time. The studies contributed by this thesis highlight advances in modelling (including in modelling spatial decay in health facility utilisation to estimate denominators) for estimating malaria incidence from routine data.

1.8.2 Issues in statistical modelling disease data

There are several issues related to data and methods when modelling disease data. First, there has been an increase in the availability of spatio-temporal datasets in last ten years from widely available household surveys at a national level and independent studies that are becoming increasingly available to the public domain. These data exist at varying spatial and temporal scales which makes it challenging to compare directly between countries and sites as well as imputing the missing data points. Secondly, the increasing availability of remotely sensed data has given rise to large set of spatio-temporal climatic or ecological covariates. Modelling these datasets using kriging, for example, requires assembly of large databases and requires inversion of large variance-covariance matrices generally of order $O(n^3)$ (Barnerjee *et al.*, 2004). Numerical algorithms such as MCMC are slow computationally and may result in poor mixing of chains as well as convergence issues.

Several approaches have been proposed in the literature to handle large datasets including reducing the dimension of the covariance matrix to a sparse nature using methods such as covariance tapering (Furrer *et al.*, 2006, Kaufman *et al.*, 2008). Vecchia (1988) proposed the partitioning of the density matrix into sub-vectors and computing the likelihood as a joint conditional product (Vecchia, 1988, Stein *et al.*, 2004). Slightly different approaches involved the use of reduced rank Kriging using basis functions (Cressie and Johannesson, 2008), modelling in the spectral domain using the Whittle likelihood (Fuentes, 2002) and the use of lattice methods (Whittle, 1963, Rue and Held, 2005). The sparseness property of the covariance was an earlier popular approach (see methods such as rejection sampling under MCMC (Gelman

and Rubin, 1992, Gelman and Rubin, 1996)). The lattice approaches used in this thesis involve expressing a random field model as a solution to Stochastic Partial Differential Equations (SPDEs). The use of Gaussian Markov Random Fields (GMRF) represented via linear basis functions has been proposed as a replacement to the Gaussian Field (GF) (Lindgren *et al.*, 2011). The implication is that the GMRF enjoys Markovian properties, such as the sparseness of variance-covariance matrixes that are computationally efficient compared to the dense covariances associated with a GF.

The assumptions of stationarity can be violated for large geographic domains (Fuentes, 2002, Stein, 2005, Gosoniu *et al.*, 2006). Stationary refers to a constant parameter mean ($\mu(s) \equiv \mu$) for a spatial process $Z(s) \quad s \in D \subset \Re^d$ (i.e. the $E(Z(s)) = \mu(s)$). Weak stationarity is implied if for a Euclidean distance $h = \|s_1 - s_2\|$, the covariance $C(\cdot) = C(h)$. This implies that the covariance between two locations depends only on the Euclidean distance. For disease mapping (such as malaria), non-stationarity (by allowing spatial structure or spectral density to vary by location) may be useful property of the model given that processes such as malaria interventions that affect the disease vary spatially (Gemperli *et al.*, 2006, Gosoniu *et al.*, 2006). There are several forms of investigating the requirements for stationarity, for example, using nonlinear approaches (Fuentes, 2005). Gemperli (2003) divided the study area into tiles in an attempt to introduce space-varying parameters. This, however, may result in boundary or edge effects when stitching the tiles back together and the independence assumptions of tiles may not be appropriate (Gosoniu *et al.*, 2006). With advances in computer programming and software, it is easier to introduce non-stationarity by supplying a vector of coordinates to the model of space varying parameters. Such models have been proposed with separable or non-separable covariance

structures (Gneiting, 2002, Stein, 2005, Cressie and Johannesson, 2008, Gething *et al.*, 2011b) with separable structure most favoured because it easily handles the two domains (space and time) as Kronecker products. Kronecker products are easier to handle computationally due to their good mathematical properties for computing quantities such as determinants or matrix inversion.

Lastly, it is not always the case that the relationship between the covariates and disease data is linear. This, however, could be investigated using nonlinear approaches that relax the linearity assumptions and fitting using a modelling strategy without linear assumptions. Examples of such algorithms include polynomial regressions (Cleveland, 1979, Cleveland, 1981), kernel smoothing (Silverman, 1981, 1986) and splines (Eilers and Marx, 1996). High-order curves such as B-splines are also proposed to represent the functional relationship between disease prevalence and climatic covariates (Lindgren and Rue, 2008). Thus, the nonlinear approaches provide an additional computational aid that is better than the usual reliance on categorizing environmental covariates. These higher order curves can also be applied to modelling seasonality that is driven largely by climatic conditions (such as low prevalence in extremely arid environments).

1.8.3 Integrated Nested Laplace Approximation (INLA) and Gaussian Markov Random Fields (GMRFs)

The goal of Bayesian inference is to learn about the posterior mean and perhaps variance of some unknown parameters given some observed data. A distributional model is, thus, based on likelihood $f(y|\theta)$ where y is the observed data with θ as unknown parameters and prior

distribution $\pi(\theta \mid \lambda)$ with additional hyperparameters λ. Thus, from Bayes theorem the posterior distribution is product of prior and the likelihood given the observed data as:

$$P(\theta \mid y, \lambda) = \frac{p(y, \theta \mid \lambda)}{p(y \mid \lambda)} = \frac{p(y, \theta \mid \lambda)}{\int p(y, \theta \mid \lambda) d\theta} = \frac{p(y \mid \theta)\pi(\theta \mid \lambda)}{\int p(y \mid \theta)\pi(\theta \mid \lambda) d\theta}$$

$$P(\theta \mid y) = \frac{p(y, \theta)}{p(y)} = \frac{p(y \mid \theta)\pi(\theta \mid \lambda)h(\lambda) d\lambda}{\int p(y \mid \theta)\pi(\theta \mid \lambda)h(\lambda) d\theta d\lambda}$$

where h and π can be conjugate prior information (see the appendix section for examples of conjugate priors). The attractive properties of the posterior are the mean, mode and median of the posterior $p(\theta \mid y)$. The mode

$$\hat{\theta} : P(\hat{\theta} \mid y) = Sup_{\theta} p(\hat{\theta} \mid y)$$

does not require an integration algorithm and is central in Bayesian analysis using INLA where the marginal is Gaussian. A popular approach to computing the above integrants is to use MCMC (Kloek and Dijk, 1978) which is non-deterministic by drawing samples (often correlated) from the posterior to arrive at closed values. A histogram distribution of the posterior is usually sufficient for inference. Common recursive algorithms are the Gibbs sampler and Metropolis-Hastings (Barnerjee *et al.*, 2004). The INLA method, used in this thesis, arrives at a closed solution using curvature of the mode evaluated at suitable sampling points (Rue *et al.*, 2009). The difference between these two approaches lies in computational efficiency and only applies for a class of Latent Gaussian Models (LGM) such as spatial and spatio-temporal models using GMRFs (Martins *et al.*, 2013).

GMRFs are widely used in Bayesian hierarchical models for lattice data due to their relative ease of implementation and they also possesses desirable Markov properties (Rue and Held, 2005). For example, they are useful at representing dependence of an unobserved process (the latent effect) at the second stage of a hierarchical Bayesian model. A random vector $x=(x_1,.....x_n)^T$ can be defined as a GMRF with mean μ and positive definite precision matrix Q with density:

$$\pi(x) = 2\pi^{-n/2} \, |Q|^{1/2} \, \exp[-\frac{1}{2}(x-\mu)^T Q(x-\mu)]$$

where Q is the precision matrix with covariance matrix $\Sigma = Q^{-1}$. The computational advantage of Markov properties arise from the sparcity of Q. If two random variables (RV) x_i and x_j $(i \neq j)$ are conditionally independent with conditional density $\pi(.|.)$ for $\pi(x_i, x_j \mid x_{-ij}) = \pi(x_i \mid x_{-ij}) \cdot \pi(x_j \mid x_{-ij})$ then, it also follows that, for the two RVs, $Q_{ij} \neq 0$ if $j \in (i, N_i)$ for a neighbourhood N_i and matrix operations reduced to order of $O(n^2)$ for spatio-temporal applications (Rue and Held, 2005). For geostatistical applications a Gaussian Field (GF) with a Matérn covariance is represented as a GMRF through a Stochastic Partial Differential Equation (SPDE) approach based on a basis function representation (Lindgren *et al.*, 2011). Thus, the dense structure of the covariance matrix in a GF is reduced when using a GMRF with a neighbourhood structure. The representation using a neighbourhood structure makes it possible and efficient to use INLA (Lindgren *et al.*, 2011, Cameletti *et al.*, 2012, Lindgren, 2013).

In INLA, the linear predictor η_i can generally be modelled with covariate effects in an additive manner (Rue and Martino, 2007, Rue *et al.*, 2009, Schrödle and Held, 2010) with the response

coming from a selected family linking to the predictor through some link function $g(x,\theta)$, thus, $g(\mu_i)=\eta_i$ with

$$\eta_i = \alpha + \sum_{j=1}^{nf} f^{(j)}(u_{ji}) + \sum_{k=1}^{n\beta} \beta_k z_{ki} + \varepsilon_i$$

where $f^{(j)}$ is a some function (could be linear or non-linear such as penalized splines or random walks (Fahrmeir and Kneib, 2009)) on a variable u, β_k are the coefficients for the covariates z ,and ε represents the residual error effects. The variables (η, α, f, β) together define the latent field associated with some hyperparameters θ. The joint posterior density $\pi(y \mid x,\theta)$ given a set of parameters is given as,

$$\pi(x, y \mid \theta) \;\; \alpha \;\; \pi(\theta)\,\pi(x \mid \theta) \prod \pi(y_i \mid x_i, \theta)$$

$$\alpha \;\; \pi(\theta) \;\; |Q|^{1/2} \exp[-\frac{1}{2} x^T \;\; Q(\theta) \;\; x + \sum \;\; \log\{\pi(y_i \mid x_i, \theta\}$$

where y are the observations, x latent Gaussian variables with θ hyperparameters. Bayesian implementation, in INLA for small θ (< 12) typically computes the conditional distribution of a latent field given the hyperparameters (Fahrmeir and Lang, 2001, Rue and Martino, 2007, Rue *et al.*, 2009). The desired posterior marginal distribution given observations is then calculated as:

$$\tilde{\pi}(x_i \mid y_i) \;\; = \;\; \int \tilde{\pi}(x_i \mid \theta, y)\, \tilde{\pi}(\theta \mid y)\, d\theta$$

$$\tilde{\pi}(\theta_j \mid y) \;\; = \;\; \int \tilde{\pi}(\theta \mid y) \;\; d\theta_{-j}$$

with the integral evaluated via a finite sum in INLA as (Rue *et al.*, 2009)

$$\tilde{\pi}(x_i \mid y) \;\; = \;\; \sum_k \tilde{\pi}(x_i \mid \theta_k, y) \;\; \tilde{\pi}(\theta_k \mid y) \;\; \Delta_k$$

with Δ_k as weights calculated at appropriate values of θ_k computed iteratively. The first procedure computes $\tilde{\pi}(\theta \mid y)$ as an approximation to $\pi(\theta \mid y)$ followed by $\tilde{\pi}(x_i \mid \theta, y)$ as an approximation to the conditional marginal distribution to $\pi(x_i \mid \theta, y)$. Lastly, given full conditional distributions, θ is integrated out. The first approximation is obtained as;

$$\tilde{\pi}(\theta \mid y) \quad = \frac{\pi(x, \theta, \quad y)}{\tilde{\pi}G(x \mid \theta, y)}\Bigg|_{x = \; x^*(\theta)}$$

where $\tilde{\pi}G(x \mid \theta, y)$ is the Gaussian approximation to the conditional of x evaluated at the mode $x^*(\theta)$ obtained by an optimization algorithm using a quasi-Newton approach (Fahrmeir and Lang, 2001, Rue *et al.*, 2009). This is the Laplace approximation of the marginal posterior distribution (Martins *et al.*, 2013) and is usually accurate as long as the expected posterior is dominated by one mode. Thus, locating the mode (using an optimization algorithm) is key in the successful application of the INLA method. The error, proposed to be of order $O(n_d^{-3/2})$ (Rue *et al.*, 2009), is minimal especially where the full Laplace option is selected when using the INLA algorithm (Tierney and Kadane, 1986, Martins *et al.*, 2013). The quantity $\pi(x_i \mid \theta, y)$ is then evaluated using a Laplace approximation (Rue *et al.*, 2009). Further details of the INLA numerical algorithms can be found in Martins *et al.* (2013).

$$\tilde{\pi}(x_i \mid \theta, y) \quad = \frac{\pi(x, \theta, \quad y)}{\tilde{\pi}GG(x_{-i} \mid x_i, \theta, y)}\Bigg|_{x_{-i} = \; x^*_{-i}(x_i, \theta)}$$

1.8.4 Analytical protocol and the Bayesian modelling approach adopted in the thesis

The overall analysis flow is presented schematically in Figure 1.7 which shows the data input and processing leading to a Bayesian framework. This includes adjusting for rates of attendance

or health facility utilisation, rate of reporting within the HMIS based on number of returned reports at each health facility and an adjustment of the suspected cases using a slide positivity rate. The focus of the rest of this section is to outline the Bayesian modelling framework.

CAR and geostatistical models were used to assess the burden of Malaria in the selected case studies. The INLA methodology was adopted for each study with actual models slightly modified to better fit the data and underlying disease dynamics to answer specific research questions, and also increase understanding of different modelling frameworks. The model specification for each study is included in the relevant chapters of the thesis and includes information on type and magnitude of priors used as well as assumptions attached to the parameters (for example, linear assumptions for Namibia and Afghanistan compared to non-linear assumptions for Eritrea). To understand the effect of including extra random effects, one level was specified in the Namibia CAR model compared to three different levels in Afghanistan and Eritrea (where facility, district and province independent effects were used). In Eritrea, non-linear effects included a seasonal model with spatio-temporal covariates (rainfall, and temperature). For Namibia, random effects were modelled at constituency and facility level only.

For modelling slide positivity rate, the model parameters such as range were estimated using the size of the geographic domain (country boundary). Covariates in each case study were selected based on previous knowledge from other studies and statistically via a best fitting generalized linear mixed model with lowest Bayesian Information Criterion (BIC). This resulted in different covariates in each model setup. Intervention effects such as ITN distribution were not included as part of the covariate set to reduce circularity and avoid over-fitting (Illian *et al.*, 2012). Edge

effects were minimized by expanding the study domain by at least 100 km. The Kronecker product feature $(Q = Q_s \otimes Q_t)$ was used to construct a separable covariance structure

$$Cov\{Z(h,u)\} = C_S(h)C_T(u)$$

and the non-separable covariance function (De Iaco *et al.*, 2001, Gneiting, 2002):

$$C(h,u) = \alpha_0 C_S(h)C_T(u) + \alpha_1 C_S(h) + \alpha_2 C_T(u) \qquad (h,u) \in \Re^d \times T$$

where α_0, α_1 and α_2 are non-negative coefficients while C_S and C_T are spatial and temporal covariance functions respectively. Thus, non-separable covariance functions were constructed to improve the mixing of two domains (Knorr-Held, 2000). However, the results for the separable covariance function are given in all the three case studies. The mean component in the modelling specification follows an independent autoregressive component $x \sim N(0,(\tau(1-\rho^2))^{-1})$ with two hyperparameters on τ (gamma prior) and $\rho < 1$ with normal prior.

Where several models were fitted, model selection was based on the significance of the parameters and Deviance Information Criterion (DIC) (Spiegelhalter *et al.*, 2002). The effective number of parameters for a model is defined by Spiegelhalter *et al.* (2002) as $p_D = \bar{D}(\theta) - D(\bar{\theta})$ where $\bar{D}$ is the posterior mean model deviance, $E(\theta \mid y) = \bar{\theta}$ is the posterior mean of parameters, $D(\theta)$ is the Bayesian deviance at the posterior estimate of parameters θ. Other quantities such as the Root Mean Square Error (RMSE), the mean error and the cross validation statistics were used to assess the selected model.

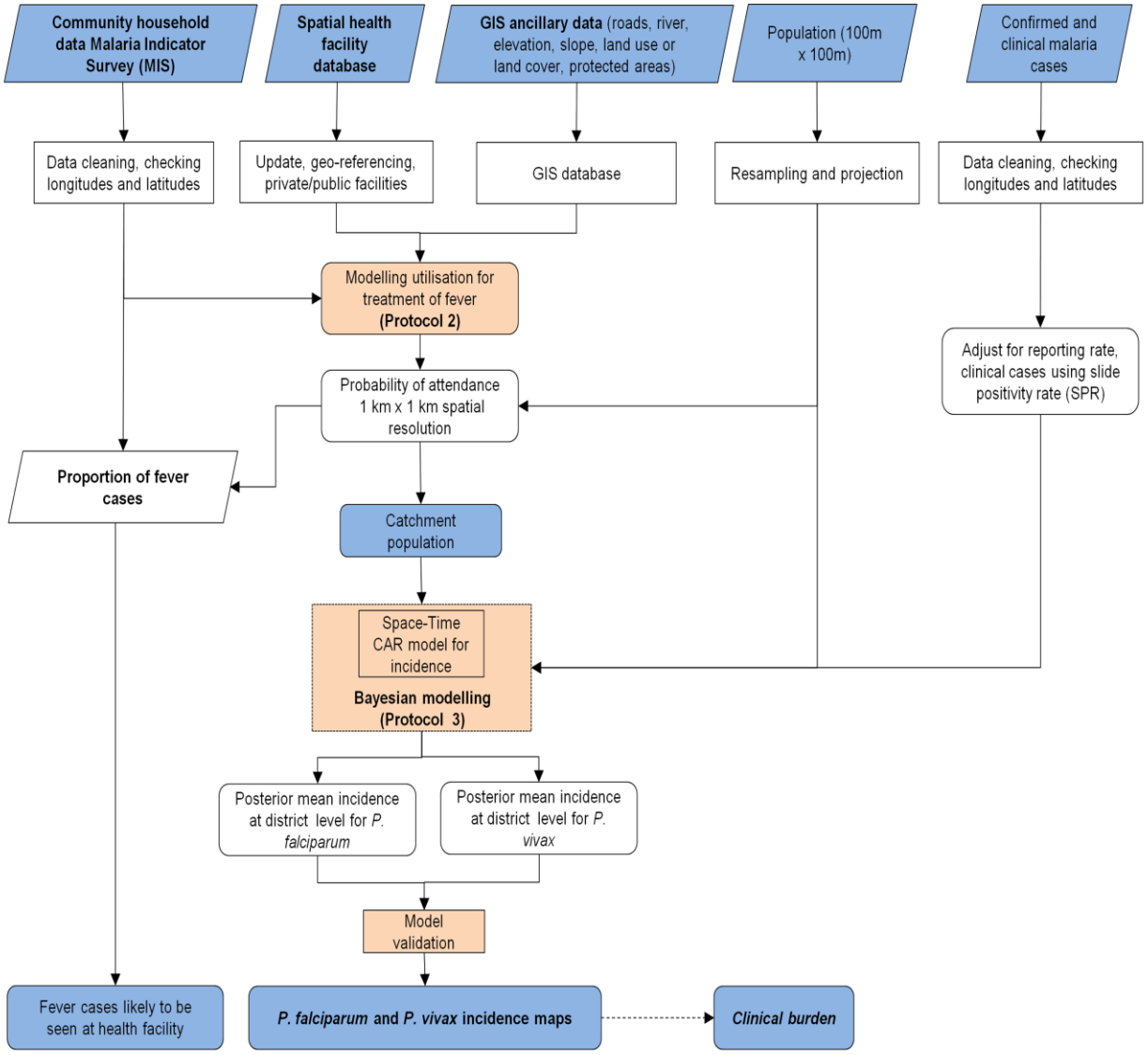

Figure 1.7 Schematic analysis flow

The overall methodology used in the thesis where main data inputs are at the top and the major sections are those in brown in terms of modelling probability of attendance for fever treatment, the Bayesian modelling framework as well as validation Bayesian model outputs.

CHAPTER 2: Case Study 1

Estimation of malaria incidence in northern Namibia in 2009 using Bayesian Conditional-Autoregressive (CAR) spatio-temporal models

2.1 Background

Namibia declared ambition to eliminate malaria by 2020 after reducing the burden significantly with sustained control (Southern African Development Community, 2012, Noor *et al.*, 2013a, Noor *et al.*, 2013b). In 2009, the Elimination Eight (E8) initiative was launched, under which eight southern African countries decided to collaborate to eliminate malaria in Namibia, Botswana, South Africa, and Swaziland. These ambitions were not only motivated both by reported substantial declines in malaria burden in the four countries but also by the renewed interest in malaria elimination following the global call in 2008 (World health Organizastion, 2008). In 2010, Namibia launched a national malaria strategy for the period 2010 to 2016 with the goal of reducing malaria case incidence to 10 persons per 1000 population by 2013 and to move the country to pre-elimination status by 2016 with a case incidence of no more than 1 person per 1000 population (Ministry of Health and Social Services, 2010c, d). Evidence from the malaria indicator survey (MIS) conducted in 2009 showed a mean community *Plasmodium falciparum* prevalence of less than 3% nationally (Ministry of Health and Social Services, 2010b) which is also a threshold at which countries are advised to consider pre-elimination as well strengthen surveillance. At this threshold, one of the proposed appropriate approaches is the use of case incidence data for measuring malaria risk since parasite prevalence surveys maybe inadequate (Yekutiel, 1960, Hay *et al.*, 2008).

Namibia, like other malaria eliminating countries in SSA, is yet to adopt active case-detection (ACD) systems (World Health Organization, 2012b) and the main surveillance data are from passive case detection (PCD), assembled through the public health sector. HMIS data, as outlined previously in the introduction chapter, have deficiencies that limit their use for

estimating malaria incidence. For example, a proportion of malaria cases in Namibia may be treated outside of the public health sector (Cibulskis *et al.*, 2007, Cibulskis *et al.*, 2011) while a proportion of health facilities in the public sector may fail to report. Those health facilities that report to HMIS may not do this consistently, thus, making data spatially and temporally incomplete (Murray *et al.*, 2004, Stansfield, 2005, Gething *et al.*, 2006, Gething *et al.*, 2008). The use of HMIS data requires approaches that adjust for non-utilisation of the public health sector; incomplete data reporting which underestimate burden and the presumptive diagnosis of malaria which inflate incidence (Cibulskis *et al.*, 2011, Alegana *et al.*, 2012). In addition, these approaches must harness the spatial and temporal autocorrelation of the available data to areas and periods where data are missing and estimate robustly the uncertainties of these estimates (Loha and Lindtjorn, 2010, Reid *et al.*, 2012).

Bayesian hierarchical conditional auto-regressive (CAR) models were used to smooth incidence using HMIS data while incorporating a set of environmental or ecological variables (Gelfand and Vounatsou, 2003, Barnerjee *et al.*, 2004, Gething *et al.*, 2006). This approach has been used previously in modelling spatial-temporal variation of disease risk in Yunnan province in China (Clements *et al.*, 2009) and in identifying social and ecological factors driving malaria risk in Vietnam (Manh *et al.*, 2011). The main advantage of this approach is that it can handle uncertainty in a coherent manner, is able to smooth risk in areas where data are not recorded and smooth variability where the denominator (population catchment) is small (Gelfand and Vounatsou, 2003, Reid *et al.*, 2012). These approaches are used in this study with the primary aim of smoothing the incidence of *P. falciparum* at second administrative unit level (constituencies) in northern Namibia where malaria is considered endemic.

The chapter provides an overview of Namibia in terms of geography and health goals and the healthcare system. A review of healthcare utilisation in section 2.5.5 provides a platform for analysis of *P. falciparum* incidence in Northern Namibia based on cases reported in 2009. The smoothing of incidence is based on a Bayesian conditional-autoregressive (CAR) zero-inflated Poisson model. The rationale for using zero-inflated model is provided prior to model specification. The last section discusses findings in context of malaria pre-elimination in Namibia and highlights regions where malaria control should be focused.

2.2 The Namibia context

2.2.1 Geography

Namibia is located in southern Africa at approximately latitudes 17^{o} S and 29^{o} S and longitudes 11^{o} E and 26^{o} E. It is bordered by Angola and Zambia to the north, Botswana and Zimbabwe to the east and Atlantic Ocean to the west (Figure 2.1). The country's total surface area is approximately 824,116 km^2 (ranked as 34^{th} in the world in terms of size) and is divided into 13 regions (administrative level 1 boundaries), 34 health districts and 108 constituencies (Zere *et al.*, 2006, Ministry of Health and Social Services, 2010c). Namibia is one of least densely populated country in the world after Mongolia, with approximate density of 2.5 inhabitants per square km. Population is estimated at just over 2.2 million. Majority live in the northern regions of the country (National Planning Commission, 2012). The common geographic features are the Namib Desert in the west and the Kalahari Desert in the East (with Botswana and South Africa) extending to Orange River in the south. Aridity constrains malaria risk to the northern regions (Ministry of Health and Social Services, 2010c, Snow *et al.*, 2010a). The central plateau extends

from north to south in the western part of the country. The West and coastal parts receive approximately 5 mm and 20 mm of rain respectively. The average daily temperatures range from 20 °C to 34 °C in the summer and 18 °C to 22 °C in the in winter. Temperatures at the coast are driven largely by the cold Benguela currents from the Atlantic Ocean.

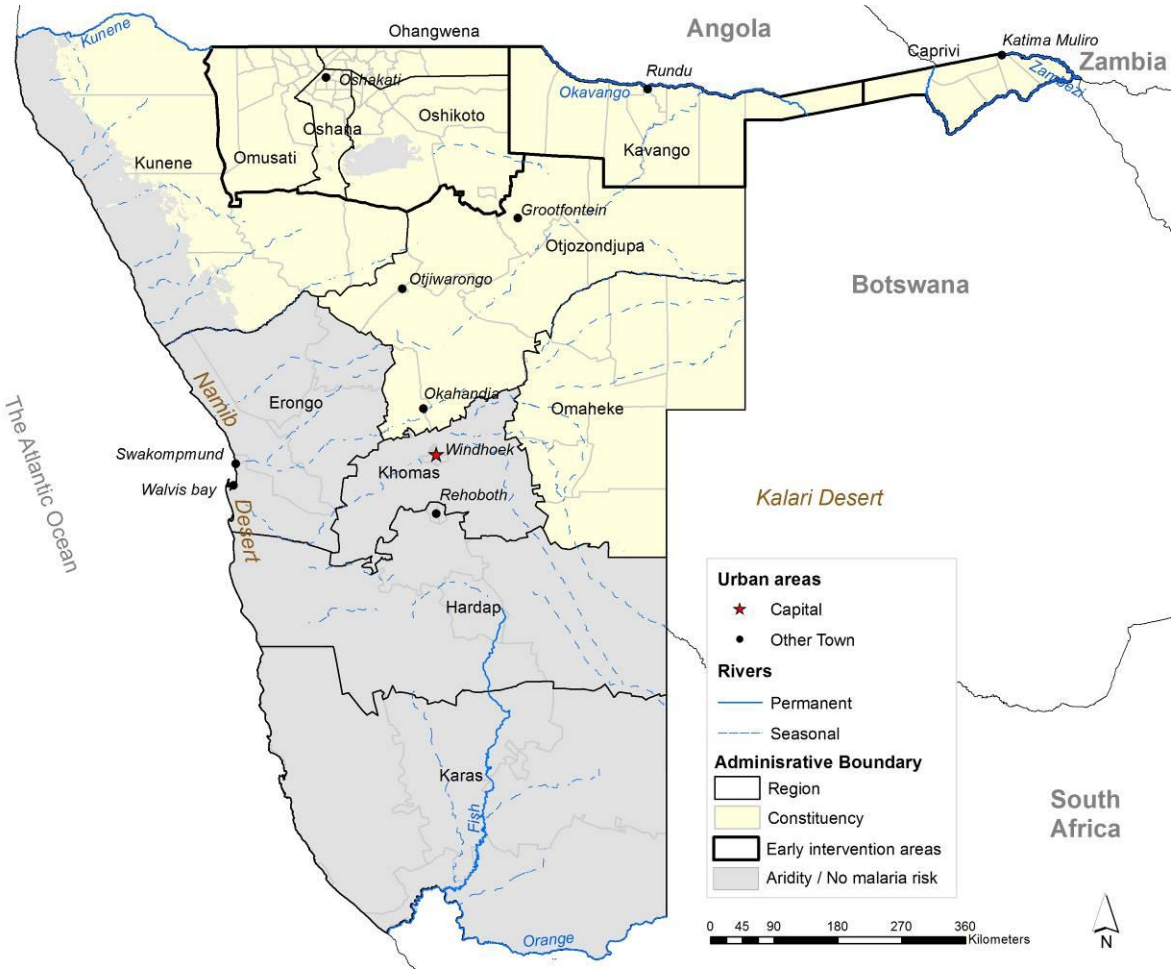

Figure 2.1: Map of Namibia showing administrative boundaries and limits of *P. falciparum* risk
Map of Namibia showing administrative 1 boundary (regions) and constituencies and the locations of the major urban areas. The zoned area in the north (bold line) show regions where IRS was conducted (1965 to 1980s) while the grey regions have no malaria risk mainly due to extreme aridity (Alegana *et al.*, 2013, Noor *et al.*, 2013b).

2.2.2 Namibia progress on MDGs and health targets

Namibia has made steady economic progress since independence in 1990 with GDP per capital estimated at 4.9 in 2012 (National Planning Commission Secretariat, 2013, The World Bank,

2013). It is ranked 128 out of 187 countries on the Human Development Index (HDI) (UNDP, 2013b). Despite recent economic progress, there still exist large inequalities in income (estimated Gini coefficient of 0.6 in 2010 (CIA, 2013)) and the majority of rural population remain poor. The overall life expectancy has declined to 48 years. Namibia is unlikely to achieve MDG 4 targets on child mortality and maternal mortality (MDG 5) (National Planning Commission and UNDP, 2010, Mbeeli *et al.*, 2011). For example, according to the 2006-07 DHS, the maternal mortality ratio (MMR) was 449 per 100,000 live births compared to 271 per 100,000 live births estimated in the 2000 DHS. The high MMR has been attributed to various factors including lack of skilled birth attendance and HIV/AIDS (Ministry of Health and Social Services, 2008b, National Planning Commission and UNDP, 2010, Mbeeli *et al.*, 2011). In 2008, the infant mortality rate was 46 deaths per 1000 live births while the under-five mortality rate was 69 deaths per 1000 live births (Ministry of Health and Social Services, 2008b, National Planning Commission, 2008). The reported rate of child mortality for SSA in 2010 (121 per 1000 live births) (United Nations, 2012) and a target of 28 deaths per 1000 live births for Namibia is less likely to be achieved by 2015.

However, progress has been made on reducing extreme poverty (estimated at 28%) and on malaria where the burden has been reduced significantly (National Planning Commission and UNDP, 2010). The reported incidence of malaria fell from 207 cases per 1000 population in 1996 compared to 63 cases per 1000 population in 2008 (National Planning Commission and UNDP, 2010). The rate of stunting (height-for-age) in the 2006-07 DHS was 29% and wasting (weight-for-height) was 7.5%. The level of malnutrition decreased with increasing wealth index (Ministry of Health and Social Services, 2008b). An improvement in the agricultural sector has

contributed to reduced poverty rates (Ministry of Health and Social Services, 2008b, National Planning Commission and UNDP, 2010) while aggressive malaria control strategies has reduced the malaria burden (Ministry of Health and Social Services, 2010b). Universal coverage of LLIN is yet to be achieved. The ownership of ITNs in the 2009 MIS was 54.8% with utilisation rates low at 34.2% amongst children under the age of five years.

The Namibia health goals are outlined in vision 2030 document targeting equity of access to quality healthcare services. The National health policy framework outlines the key health targets for Namibia for period 2010 to 2020 (Ministry of Health and Social Services, 2010d). It also outlines the management and strategic plans for malaria and other diseases under vision 2030 (National planning Commission, 2004). Examples of the targets in vision 2030 policy include: reducing the prevalence of HIV/AIDS and infectious diseases (malaria, TB and STIs); reducing vaccine preventable diseases; increasing family planning uptake and contraceptive use; and improving the provision of clean water and sanitation (National Planning Commission Secretariat, 2013).

With a substantial decline in malaria cases in recent years, the ministry of health and social services targets a pre-elimination status by 2015 and elimination by 2020 (Southern African Development Community, 2012, Noor *et al.*, 2013a, Noor *et al.*, 2013b). Current malaria strategies include the strengthening of the diagnosis of malaria, effective case management, increasing the coverage of insecticide treated bed nets (ITNs) in malaria endemic border areas, strengthening of community level surveillance and maintaining a malaria-free buffer extending across the border with neighbour countries (Ministry of Health and Social Services, 2010c, d,

Trans-Zambezi Malaria Initiative (TZMI), 2012, Noor *et al.*, 2013b). The 2008 DHS, for example, estimated 25% ownership of mosquito nets in a national representative sample of about 9,200 households. These are distributed mainly through mass campaigns, in hospitals and during immunizations.

2.2.3 History of malaria control in Namibia

Namibia introduced indoor spraying using DDT in 1965 with chloroquine used for treatment upon infection (Hansford, 1972). DDT was used in Ovambo land (present regions: Ohangwena, Oshana, Oshikoko and Omusati), Kavango and Caprivi. By 1979, IRS had contributed to a substantial decrease of malaria vectors namely the *An. Gambiae complex* and the *An. Funestus*. The period between 1975 and 1990 was marked by instability during the war of independence and malaria control activities reduced significantly. Chloroquine resistance was also reported during this period and a combination of these factors (instability and reduced malaria control) led to an increase in malaria cases in the late 1980s (Noor *et al.*, 2013b, Noor *et al.*, 2013c).

After the 1990 independence, the National Vector-borne and Disease Prevention (NVDCP) increased malaria control activities. An increase in funding particularly after the launch of the Global fund to Fight AIDS, Tuberculosis and Malaria (GFATM), increased the coverage of interventions. To date, MoHSS has received over USD 26 million in funding since January, 2005 (The Global Fund, 2013c). ITN ownership, for example, has increased and was estimated as 54.8% in the 2009 MIS. The first line treatment for malaria using ACTs was adopted in 2006 and AL is now used as treatment for confirmed *P. falciparum* cases with SP or Fansidar used during pregnancy (Ministry of Health and Social Services, 2010c).

2.2.4 Organisation of Namibia's healthcare system and delivery

The Ministry of health and social services is responsible for provision of all forms of healthcare in Namibia including rehabilitative, preventive and curative services. The main target is to provide primary care to the population via clinics, health centres, district and regional (referral) hospitals (Ministry of Health and Social Services, 1998, 2008a). Financing of healthcare is mainly by government (approximately 12%) and donors although the general public pay for some services (out of pocket payments) (Ministry of Health and Social Services, 2010c).

The regional and district administration manage referral facilities at the regional level as well as support management at the district level (34 health districts) (Ministry of Health and Social Services, 1995). The role of the national-level administration include planning, formulating healthcare policies including legislation and regulation (Ministry of Health and Social Services, 1995). There are 13 regional administrative areas headed by a deputy permanent secretary and responsible for policy implementation in the 34 health districts (Ministry of Health and Social Services, 1995, El Obeid *et al.*, 2001). Majority of in-patient services are provided by the tertiary higher-level facilities with outpatient care is mostly provided at the first-tier facilities such as clinics, VCTs and sick-bays.

Clinics and health centres constitute the first-tier facilities providing basic services to the population. There is also an involvement of community health workers in some regions. These basic facilities, however, lack skilled labour and provide limited services on neonatal care, emergency obstetrics, infant and maternal nutrition, Integrated Management of Neonatal and Childhood Illnesses (IMNCI) and breastfeeding programmes. Few outreach centres are linked to

clinics in rural areas (Ministry of Health and Social Services, 1995, 2007). Private facilities and private for non-profit (mission facilities) are run predominantly by private individuals. There exists church or faith based health facilities that provide healthcare services to general public (El Obeid *et al.*, 2001, Ministry of Health and Social Services, 2007).

2.2.4.1 Description of health facilities census in Namibia in 2009

Namibia conducted a national health facility census in 2009 in an effort to understand service provision and improve resource allocation (Ministry of Health and Social Services, 2010a). In total, there were 46 hospitals (10.3%), 49 (11%) health centres, 327 (73.3%) clinics, 15 (3.3%) stand-alone voluntary counselling centres (VCTs) and 9 (2.1%) sick-bays that provide various child and HIV/AIDS related services. Majority were managed by the MoHSS (70.4%), 48 (10.8) by missions and NGOs, 14 (3.1%) belonged to the ministry of defence (MoD) and 70 (15.7%) were private. Only two health facilities in Kavango region out of the 446 facilities listed were not visited. The geographic location of each facility was recorded using a global positioning system (GPS). For health facilities, where the GPS coordinates were not taken, the longitude and latitude was established through a geographic place name or village name from a geo-database of place names in Namibia (Geonames, 2010).

Majority of tertiary facilities were located in urban centres and on main access roads (Figure 2.2 below). The spatial distribution of health facilities is similar to population distribution with majority of facilities located in the north. The analysis of utilisation and malaria incidence did not include the specialised facilities managed by the MOD/police, the private facilities and VCT since the focus was on public healthcare utilisation and disease burden estimation within these health facilities.

2.2.5 A review of public health facility coverage and utilisation in Namibia in 2009

The analysis of coverage and utilisation in 2009 was based on public health facility data described in the section 2.2.4 and the rate of fever treatment for children under the age of five years based on the 2008-09 Namibia MIS. Relevant analytical procedures and discussion on public healthcare utilisation are outlined in Alegana *et al.* (2012). In brief, travel times were derived between population centres (households) and health facilities. Utilisation pattern was then modelled based on reported patterns of attendance (assumed at the nearest facility) for fever treatment and the theoretical derived patterns. The analysis focused on universe of all facilities and there was no stratification by facility type.

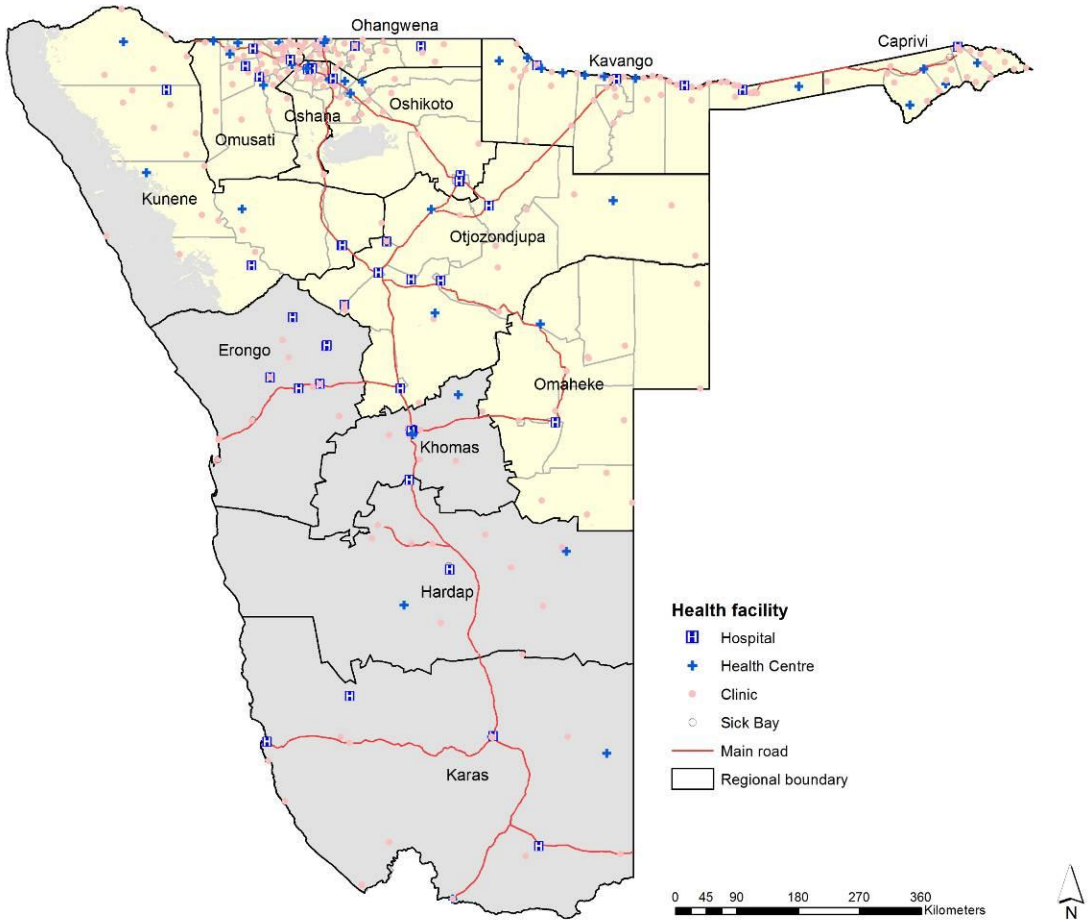

Figure 2.2: Distribution of Health facilities in Namibia in 2009
Public health facilities in Namibia in 2009 showing the majority of major hospitals are located in main urban centres and on the main roads. The grey areas are non-malaria due to extreme aridity (Alegana *et al.*, 2013, Noor *et al.*, 2013b)

The analysis of public health facility coverage showed that a greater number of facilities were located in northern Namibia (Alegana *et al.*, 2012). Assessment of health facility coverage using a 3-hour travel time cut off suggested high coverage rates of greater than 80% for most regions. However, public healthcare utilisation varied for each region (Table 2.1 below). Of the estimated 162,286 children under the age of five years in the northern Namibia, 160,294 (98.8%) were estimated to be within a public health facility catchment (Figure 2.3 Page97) and the burden of fever was approximately 24,830 cases in 2009. The proportion of fevers within the catchments

was 90.8% (22,553) and 16,195 (65.3%) were likely to seek treatment in the public health sector. This suggested that approximately 8,616 (34.7%) febrile children were unlikely to use at a public health facility in northern Namibia including the 4,030 (47%) fever cases that lived outside of a health facility catchment (Table 2.1). The lowest utilisation rates were in Kunene (41.3%) while the highest were in Oshana (75.4%).

Table 2.1: Estimated coverage and utilisation of public health facilities in northern Namibia
The table is for children under the age of five years based on derived catchment population and fever treatment pattern

	Estimated number of children under five years of age in 2009	Estimated number of children under five years of age within a PHF[1] catchment	Estimated number of fever cases among children under five years of age based on MIS prevalence	Number(Percentage) of children under five years of age with fever likely to attend a PHF[1]
Region				
Caprivi	8,881	8,741	2,433	1,637(67.3)
Kavango	20,244	20,374	4,825	3,264(67.6)
Kunene	8,192	7,363	1,425	588(41.3)
Ohangwena	32,167	30,863	3,793	2,695(71)
Omaheke	11,550	11,051	1,974	1,060(53.7)
Omusati	27,386	26,993	3,478	2,522(72.5)
Oshana	14,973	13,088	2,186	1,648(75.4)
Oshikoto	19,918	236,61[2]	1,395	932(66.8)
Otjozondjupa	18,977	18,160	3,321	1,868(56.3)
Travel time				
< 30 minutes	51,791	51,791	8,021	6,056(75.5)
> 30 minutes -< 1 hour	98,620	98,620	14,902	11,218(75.3)
>1 -< 2 hours	138,219	138,219	19,035	14,136(74.3)
>2 -< 3 hours	160,294	160,294	20,799	15,279(73.5)
>3 hours	1,992	-	4,031	934(23.2)
Probability of attendance				
< 0.50	13,698	11,808	2,277	19(0.8)
>0.50 - < 0.60	7,820	7,908	1,356	717(52.9)
>0.60- < 0.70	18,944	18,963	2,928	1,925(65.7)
>0.70- < 0.75	120,862	121,615	18,269	13,553(74.2)
Total	**162,286**	**1,602,94[3]**	**24,830**	**16,214 (65.3)**

1. PHF is an abbreviation for 'Public Health Facility', which in this case does not include private facilities or privates for profit.
2. For Oshikoto region, the estimated number of children (0-4 years) slightly exceeds the overall population estimate for the region. This is because the catchment boundaries in some cases overlap the regional boundaries.
3. The total number of children 0-4 years old in catchment boundaries was lower than the total estimated under fives population because of (a) not all children within the catchment were assumed to use the public health facility (b) the catchment boundaries did not covering 100% of the entire population by limiting maximum travel time to 3 hours from the decay model.

Figure 2.3 shows the modelled facility catchments nationally, although, for incidence analysis only northern based facilities were used.

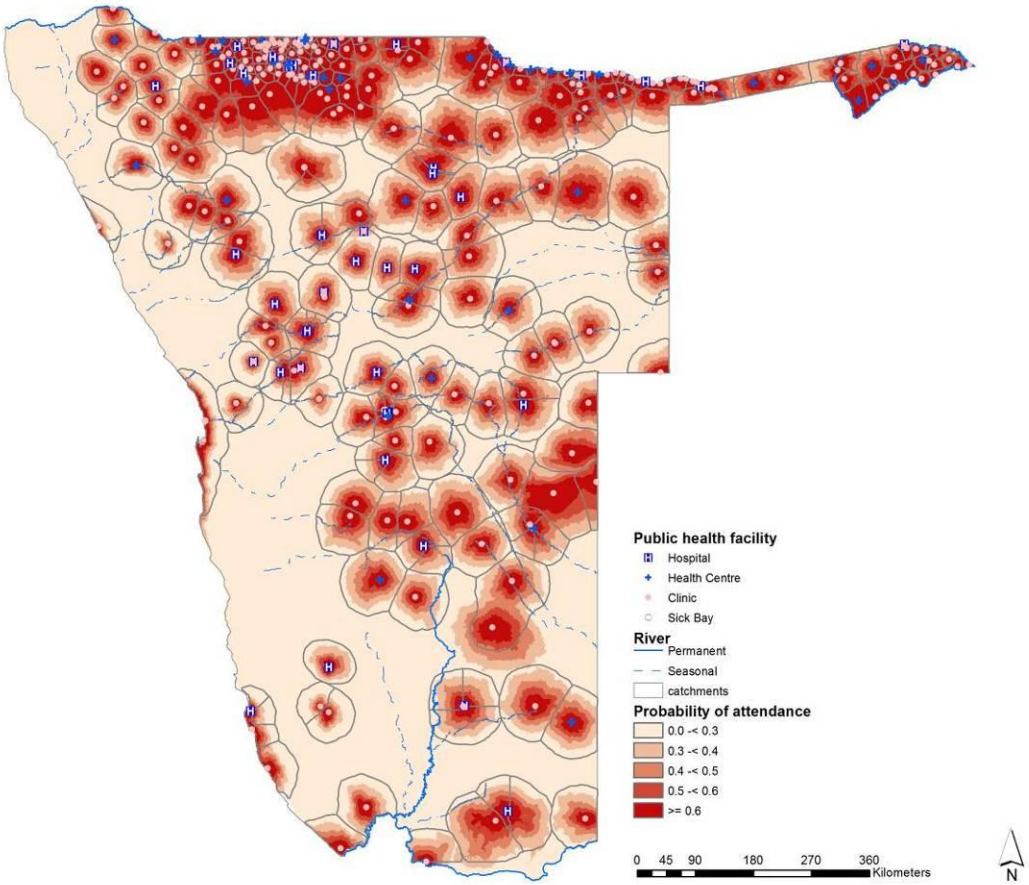

Figure 2.3: Probability of fever treatment and health facility catchments in Namibia
The probability of using public health facility for fever treatment at 1 x 1 km based on the MIS data superimposed with the delieanated facility catchments.

2.3 Analysis of malaria in 2009 in northern Namibia

2.3.1 Description of HMIS data

HMIS data for 2009 was obtained from MoHSS for 273 public health facilities in the northern Namibia after the national Service Provision Assessment (SPA) survey (Ministry of Health and Social Services (MoHSS) and ICF Macro, 2010). Of the 273 facilities in the north, 13 were private health facilities located in the urban centres and these were not included in the subsequent analysis of incidence. Three constituencies had no data (public health facilities) and were used as missing data after linking each facility to administrative areas. Missing data were

imputed as *NAs*. Data represented all age population of suspected and blood-slide confirmed malaria cases for 2009.

A monthly aggregate of all cases had been recorded for each health facility resulting in 3,120 facility-month records excluding the 13 private facilities. For majority of primary facilities, RDTs were used routinely to examine blood samples from most patients although a few, mostly at tertiary facilities, were examined using microscopy (Ministry of Health and Social Services, 2010a). Thus, since it was not possible to distinguish cases that had been confirmed using an RDT or via microscopy, there was no stratification based on diagnosis. A reporting rate, calculated as a proportion of received reports over the expected number was applied to facility catchments while the slide positivity rate at each health facility was used to adjust the suspected cases. The latter was necessary to avoid overestimating incidence (where true cases are treated as a summation of clinical and confirmed case while ignoring the SPR at the facility). The opposite is also true, thus:

$$SPR = \frac{Confirmed \quad cases}{Total \quad number \quad exa\min ed}$$

Where, $cases = Confirmed \quad Case \quad + \quad (Suspected \quad cases \quad \times \quad SPR)$

In total, malaria reports were available for most but the 17 facilities in 2009 (missing data). Thus, the data was considerably complete (over 90%), in terms of reporting rates, for the majority of facilities with a zero recorded where there was no confirmed or a suspected case. Figure 2.4 (below) shows a temporal plot of the malaria cases in 2009 by month.

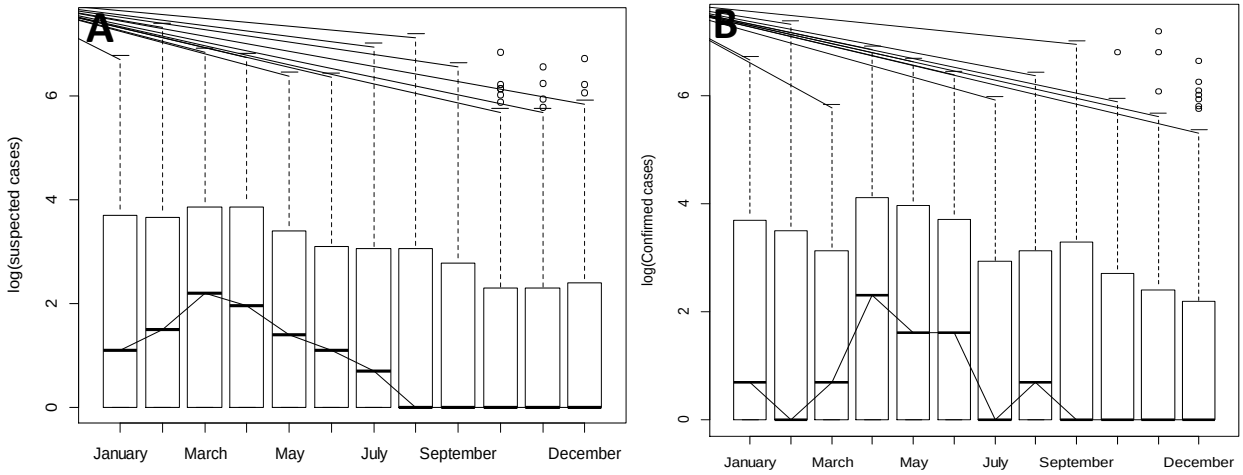

Figure 2.4: Box plot of malaria cases in 2009 in Namibia

Figure 2.4 (a) shows the suspected cases of malaria by months in Namibia in 2009 with a Gaussian fit (line) through the median cases. A peak occurred in March and was lowest towards the end of the year while (b) is based on confirmed malaria cases.

2.3.2 Development of population denominator for analysis of incidence

The Namibia population surface from Worldpop (Worldpop, 2010) (Figure 2.5) was derived from a combination of census, population settlements and land cover data using dasymetric approaches (Briggs *et al.*, 2007). Dasymetric techniques involve the disaggregation of census data to improve their spatial resolution (Bhaduri *et al.*, 2007). A Land cover surface for Namibia was obtained from the MEdium Resolution Imaging Spectrometer (MERIS) GlobCover product and combined with fine spatial resolution data on settlements to produce an improved surface that represented where people lived (Linard *et al.*, 2010). GlobCover was originally provided at a spatial resolution of 300 m and its land cover classification is compatible with the UN land cover classification system (LCCS) (FAO, 2000). Settlements data for northern regions of Caprivi and Kavango were obtained from the environmental atlas project (Mendelsohn and Roberts, 1997, Mendelsohn *et al.*, 2000, Mendelsohn and El Obeid, 2001) while an estimation of urban population in Windhoek was based on the census of 1991 and water demand report from water

resources management review (Water Resource Programme, 2007). A finer land cover or land use layer was created to include detailed information on roads, rivers and settlements extents. The land cover classes were then assigned a population weight based on density which was used to re-distribute population polygon data in 1 x 1 km spatial pixels (Linard *et al.*, 2010). The resulting national population map was then projected forward to 2009 using the United Nations' (UN) urban and rural inter-censual growth rates (http://esa.un.org/unup/).

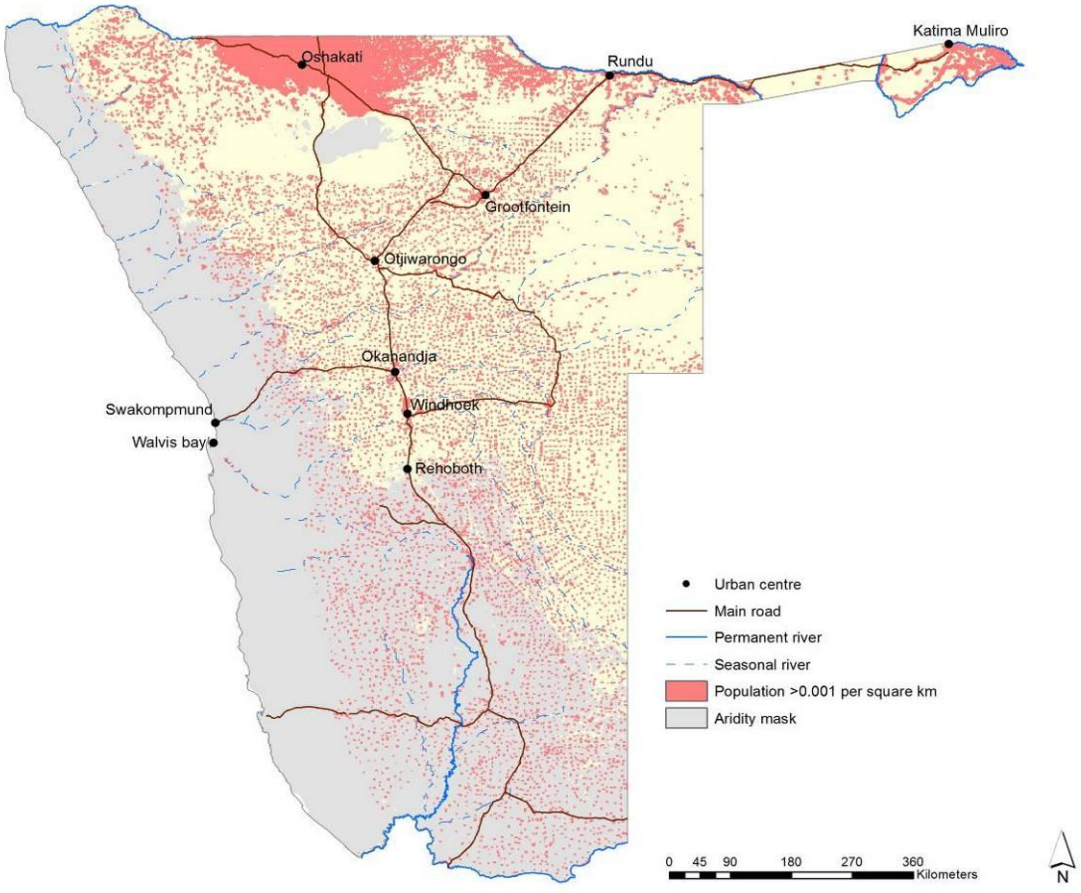

Figure 2.5 Map of population density in Namibia
Map of Namibia showing areas where population is greater than 0.001 per 1 km^2. The grey mask of aridity corresponds to regions where MODIS EVI>0.1.

2.3.3 Assembly of environmental or ecological covariates for malaria risk in Namibia

Malaria is driven by environmental and ecological factors such as rainfall, vegetation, altitude, humidity, temperature and human habitation that affect the development and survival of the malaria parasite and vector (Molineaux, 1988). Thus inclusion of these covariates is an important step in modelling malaria incidence.

The mean enhanced vegetation index (EVI) in Namibia for 2009 was derived from the MODerate-resolution Imaging Spectroradiometer (MODIS) sensor imagery as a measure of vegetation (Hay *et al.*, 2006, Scharlemann *et al.*, 2008) while monthly 2009 precipitation data were obtained from the Tropical Rainfall Measuring Mission (TRMM 3B43) (Huffman and Bolvin, 2011, NASA, 2011) (Figure 2.6). TRMM 3B43, a gridded mean monthly average precipitation product in $mmhr^{-1}$ at 0.25^{o} x 0.25^{o} spatial resolution (Huffman, 1997), was used. It is produced from multi-satellite precipitation analysis (TMPA) approach (Huffman and Bolvin, 2011) that combines satellite-sensor data and rain gauge (ground measurements) observations. A temperature suitability index (TSI) was obtained from the Malaria Atlas Project (http://www.map.ox.ac.uk) was used as temperature covariate. The TSI ranged from 0 (not suitable) to 1 (most suitable) and showed areas where temperature support parasite sporogony in Namibia (Gething *et al.*, 2011a). The average values of EVI, precipitation and TSI were then computed for each constituency. Finally, urban areas were based on the Global Rural Urban Mapping Project (GRUMP) (Balk *et al.*, 2004, Center for International Earth Science Information Network (CIESIN), 2004). Proportion of urban population was calculated by intersecting the urban grid with the population grid. Processing of this environmental grids

involve resampling to 1 x 1 km spatial resolution and extracting a value at each facility location in ArcGIS 10 (ESRI, Redlands, CA, USA).

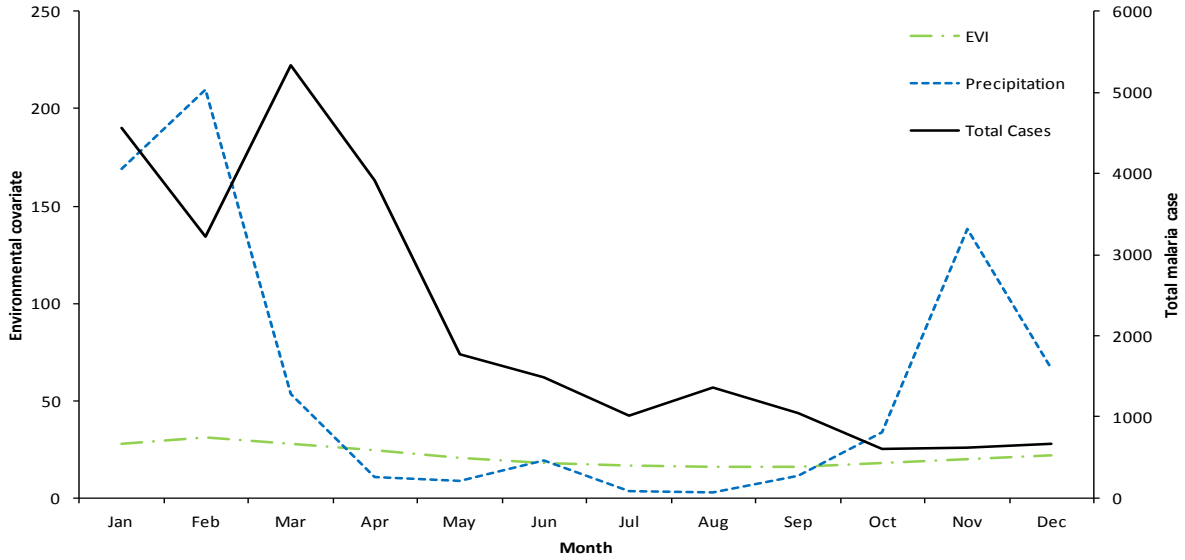

Figure 2.6: Plot of environmental covariates with malaria cases for Namibia
Plot showing the assembled total malaria cases in 2009 (————) and the extracted mean monthly environmental covariates: EVI (— · — · — · —) and precipitation (- - - - - - - - - - -).

2.3.4 Preliminary analysis of environmental covariates and crude incidence

Preliminary model selection of covariates that best describes the response (incidence) is a widely accepted exercise in statistical modelling (Murtaugh, 2009). The choice of covariates should be guided by the principle of parsimony (identifying few and easily interpretable covariates) (Murtaugh, 2009). Secondly, covariates improve statistical model fit and increase the precision of predicted estimates. Their inclusion increases the model parameter space (complexity) and, if not carefully selected, risk over-fitting (Babyak, 2004). For example, too many covariates may introduce artificial relations (due to interactions) with the outcome variable that are not easy to tease out. In addition, covariates often increase the R^2 value of regression models, especially if the number of observations compared to predictors is small, without significantly increasing prediction accuracy. This problem can be pronounced when data assembled are from

observational studies based on different study designs, sampling considerations and sample sizes which are then combined to describe a random process (Babyak, 2004).

There are several approaches in ecology reviewed by Murtaugh (2009) including widely criticised stepwise procedures see (Whittingham *et al.*, 2006, Mundry and Nunn, 2009, Murtaugh, 2009) and references therein. Subset selection based on a statistical criterion, such as the Akaike information criterion (AIC), is the most commonly used in statistical modelling. These criterion methods aim to penalize model complexity (McLeod and Xu, 2008).

In this study, preliminary non-spatial generalised linear regression analysis with response variable (crude incidence) was conducted. Figure 2.7 show the association of assembled environmental covariates with crude incidence using a scatter plot analysis in R version 14 [http://www.r-project.org/]. The extracted values, for each covariate at each health facility, were used in continuous form in a generalized (multivariate) linear regression model with the response variable being the observed crude incidence rates.

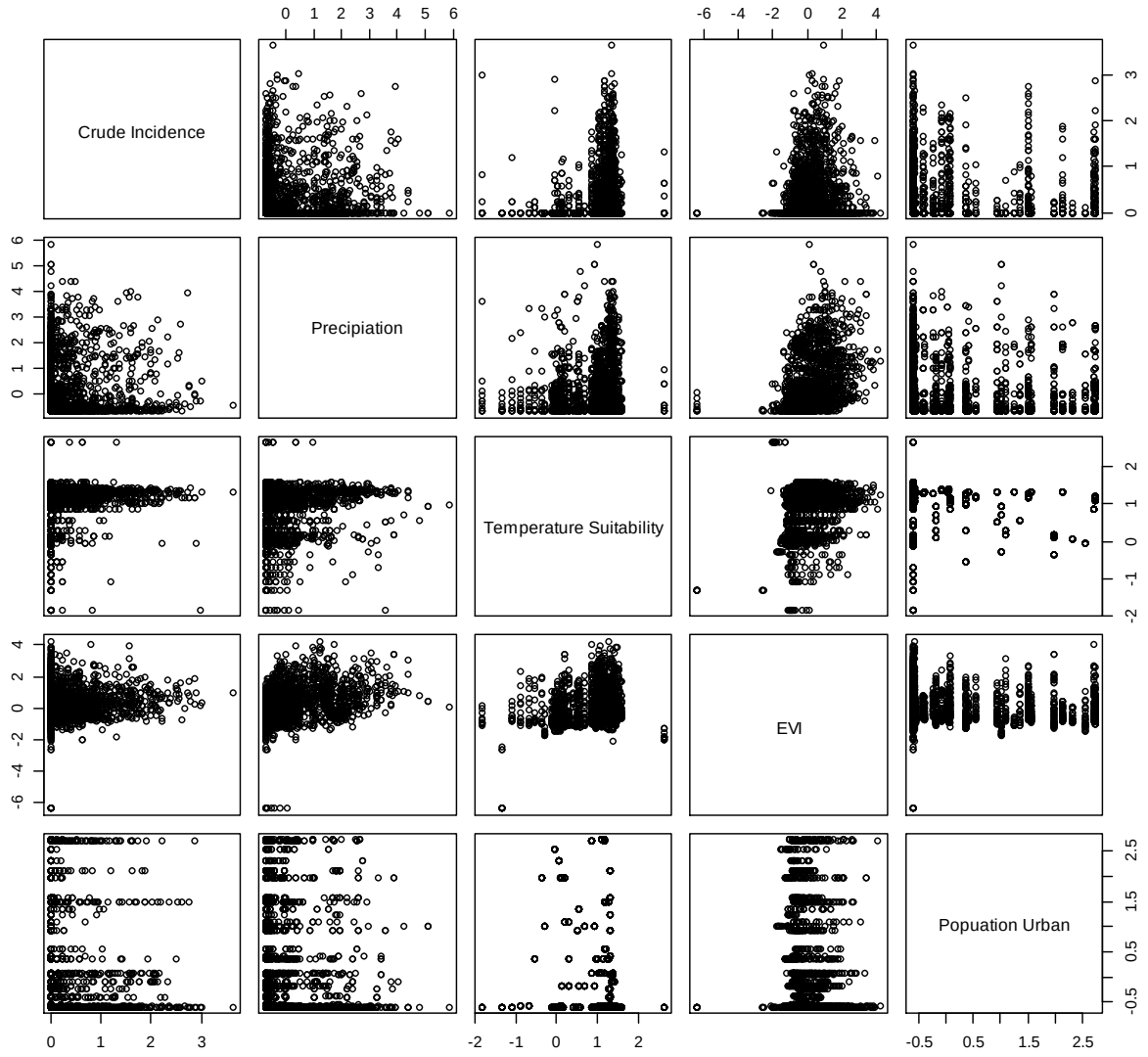

Figure 2.7: Scatterplot matrix showing association between covariates and crude incidence
Association between crude incidence per 1000 population (log-transformed) with the standardised environmental covariates. The first row shows the crude incidence (response) against each explanatory variable. None of the predictor variable is categorical.

A set of covariates were selected using the *bestglm* package in R (McLeod and Xu, 2008, Xu, 2010) using the Bayesian information criterion (BIC). The final model was based on prediction error of the response and from candidate models based on covariates. For a response variable $y = y_1 \ldots\ldots\ldots\ldots y_n$ with $X = X_1 \ldots\ldots\ldots X_k$ matrix of covariates and letting S_k, $k = 1 \ldots\ldots d$ be subset of models of size *k*:

$$BIC = -2 \cdot \log lik + k \cdot \log N$$

For a given set of candidate models, M_m, with parameters θ_m, the posterior probability can be given by:

$$\Pr(M_m \mid Z) \alpha \Pr(M_m) \cdot \int \Pr(Z \mid \theta_m, M_m) \Pr(\theta_m \mid M_m) d\theta_m$$

where $\Pr(\theta_m \mid M_m)$ is the prior distribution. Xu (2010) use the BIC*q* model with an imposed Bernoulli prior and compared with normal BIC model with flat normal prior. The BIC*q* is written as:

$$BIC_q = -2 \log L[\hat{\theta}(S_k)] + k \log n - 2k \log[q/(1-q)]$$

where $L[\hat{\theta}(S_k)]$ is the likelihood, q is the prior with probability lying between 0 and 1 $p = q(1-q)$.

The best parsimonious model was selected based on *K*-fold cross validation approach where the data were split into *k*-fold partitions and performance evaluated using one fold validation set (Hastie *et al.*, 2009). The lowest validation sum of squares S_k, were obtained after minimizing the prediction error with $E\{K(S_k)\} = qK$. The cross-validation was evaluated as:

$$CV = \frac{1}{n} \sum_{1}^{k} S_k$$

Univariate non-spatial regression analysis showed that the EVI (coefficient of regression, 95% CI: 0.37, 0.31 – 0.44, p<0.001), TSI (0.77, 0.59 – 0.96, p<0.001) and precipitation (0.15, 0.08 – 0.21, p=0.002) were important explanatory variables of crude incidence. The percentage of urban resident population produced a negative association with incidence (–0.02, -0.10 – 0.50, p<0.001). From multivariate models, only EVI (coefficient of regression: 0.0867; *p*<0.001) and TSI (coefficient of regression: 0.0959; *p*<0.001) was selected as best combination set of

covariates. However, EVI (coefficient of regression: 0.0969; p<0.001) was selected in the final parsimonious model after cross-validation and minimizing the cross-validation error (Figure 2.8). Thus, only this covariate was used in subsequent modelling work.

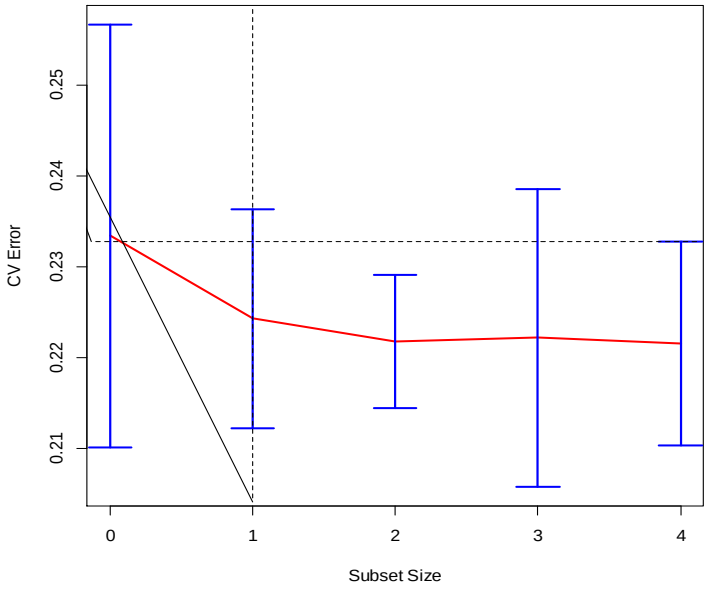

Figure 2.8: Covariate selection based on cross-validation error
Estimated prediction error curve (red) and cross validation error (bars). The model complexity increased on the x-axis while the estimate of standard error and prediction error was based on fivefold cross-validation. The best fit model within one standard error (standard deviation rule $(CV \pm s\sqrt{K})$) was indicated by broken line.

2.3.5 Bayesian model specification for malaria incidence

Preliminary exploratory analysis revealed that zero cases were reported in 65.3% of the facility level monthly returns (Figure 2.4 Page 99). Overall, 43% of facilities reported no cases in March and over 60% from May to December. Following other studies in low transmission settings (Lambert, 1992, Manh *et al.*, 2011), a Zero-inflated Poisson (ZIP) model was used to handle the excess zero counts. The ZIP models have also been applied previously in mapping the malaria vector sporozoite rate (Nobre *et al.*, 2005, Amek *et al.*, 2011) and in schistosomiasis (Vounatsou *et al.*, 2009). They are formulated as two component mixture model given as (Lambert, 1992, Angers and Biswas, 2003, Ghosh *et al.*, 2006);

$$f(y,\varphi,\mu) \sim \begin{cases} \varphi_i + (1-\varphi_i)\ e^{-\lambda} & y_i = 0 \\ (1-\varphi_i)\ Po(y,\ \mu) & y_i > 0 \end{cases}$$

for observation i and $(1-e^{\lambda})^{-1} < \varphi_i < 1$. The probability parameter φ_i is modelled with a 'structural' zero or defaulting to a general Poisson model ($\Pr(X = k) = \lambda^k e^{-\lambda} / k!$) while $\text{logit}(p_{ij}) = \log[\ p_{ij} /(1-p_{ij})]$.

In general φ_i and μ_i can depend on a set of covariates such that:

$$f(y, p, \mu) = (1-p_{ij})\, Po(y,\ 0) + p_{ij} Po(y,\ \exp[\alpha + \beta^T X_{ij}])$$

with α forming an intercept modified by a $q \times 1$ vector of X_j covariates with unknown coefficients β and $\log(\mu_i) = \sum_{j=1}^{k} x_{ij}\beta_j$. For $y_i > 0$, the second part of the equation can be generalized as:

$$(1-\phi)\frac{(1-\alpha k)^{k-1}}{k!}\frac{(\lambda e^{-\alpha\lambda})^k}{e^{\lambda}} \qquad k = 1,2....,n$$

where (α, ϕ, λ) are the parameters of interest requiring prior distributions to estimate the posterior.

A conditional auto-regressive (CAR) spatio-temporal model was implemented using EVI as covariate at constituency level. Bayesian hierarchical conditional auto-regressive (CAR) models was used to smooth incidence by constituency in northern Namibia where routine surveillance is inefficient and incorporating a set of environmental or ecological factors and random effects (Gelfand and Vounatsou, 2003, Barnerjee *et al.*, 2004, Gething *et al.*, 2006). Random effects were included at facility and constituency level to improve spatial variability in smoothing such that the linear predictor was written as:

$$\log \mu_i = \log(E_i) + \alpha + X_{ij}^T \beta_{ij} + f_1(s_u) + f_2(s_u) + \psi_i + f(t)$$

Where E_i is the expected number of cases adjusted for utilisation at facility i, α is the intercept and the unstructured components at facility $f_1(s_u)$ and $f_2(s_u)$ and facility and constituency level respectively. The $f(t)$ represents the temporal mean component modelled as an auto-regressive process with first term $x_0 = \rho \ x_{0-1} + \xi_0 \quad \xi_i \sim N(0, \tau^2)$ obtained from a stationary distribution $N(0, \sigma_w^2 \Sigma)$. The fixed effects were assigned flat (non-informative) priors $P\alpha 1$. The random effects were assigned zero-mean Gaussian priors $\theta_{unstr} \sim N(0, \tau^2)$ with the hyperparameters assigned Inverse Gamma $IG(a,b)$ priors $a = 0, b = 0.00005$ in line with other standard studies (Kazembe, 2007). The conditional spatial prior was used to model contiguous areas where the risk will be similar. Thus:

$$S_i \mid S_{-i}; \tau^2 \sim N\left(\frac{1}{k_i}\sum_{\in\delta} s_{-i}, \frac{\tau_s^2}{k_i}\right)$$

where S_i and S_{-i} are adjacent constituencies with k_i number of neighbours and τ_s^2 is the variance. Conditional adjacency matrix of weights was modelled as $W_{ij} = 1$ for two neighbouring regions or $W_{ij} = 0$ otherwise following Bernardinelli *et al.* (1997) (Bernardinelli *et al.*, 1997). The likelihood with inclusion of the terms and covariates is:

$$L_i(\beta_i, \tau_i, \varphi, \rho, Z \mid y, \alpha) = \prod_{j=1}^{Ei} [p_{ij} \Pr(Y_{ij} = 0 \mid Z_{ij} = 1, \theta)]^{Z_{ij}} \times [(1 - p_{ij}) \Pr(Y_{ij} = y_{ij} \mid Z_{ij} = 0, \ \alpha, \ \theta)]^{1-Z_{ij}}$$

2.3.6 Bayesian model specification for slide positivity rates at health facilities

2.3.6.1 Overview of slide positivity analysis

A slide positivity rate based on the cases diagnosed parasitologically at each health facility was analysed using a Bayesian hierarchical geostatistical model. The modelling challenge involved

interpolating point referenced spatio-temporal data to 1 x 1 km fine spatial resolution estimates of *P. falciparum* positivity rates in 2009. The data on slide positivity rates comprised 260 public health facilities of which 26 (10%) were selected as a validation set. This resulted in approximately 2,808 spatio-temporal unique points. EVI was used a covariate. Modelling involved the use of Gaussian Markov Random Fields (GMRF) via Stochastic Partial Differential Equation (SPDE) approach outlined in section 1.8, chapter 1.

2.3.6.1 Model specification using the SPDE approach

Let $Y(s_i, t)$ denote the response (slide positivity) at facility s_i, $i = 1,2,.....,N$ in month t, $t = 1,.....,T$. The $Y(s_i, t)$ arise from a binomial likelihood such that the probability that a case is positive for a single blood test for *P. falciparum* was; $y_i \mid p_i \sim Binomial(n_i, p_i)$ with a logit link function $p(\eta) = \{\exp(\eta)/1+\exp(\eta)\}$.

Model specification followed the approach proposed by Lindgren *et al.* (2011) where an SPDE is an approximation to a continuous domain Gaussian field using a GMRF with a matérn covariance function (section 1.8), to produce a continuous map of slide positivity rate at 1 x 1 km spatial resolution. GMRFs result in sparse covariance matrices that are computationally faster. The model is a realisation of a spatio-temporal process of the outcome variable (incidence) at each facility location defined by longitude and latitude, month, covariates and a measurement error defined by Gaussian white noise. The resulting space-time covariance matrix from the spatial and temporal domains informs the spatial range and temporal lag of the prediction model. The SPDE/GMRF for spatial-temporal field is given as (Lindgren, 2013):

$$\frac{\partial}{\partial t}(k(s)^2 - \Delta)^{\alpha/2}(\tau(s) \cdot x(s,t)) = W(s,t), \qquad (s,t) \in \Psi$$

where k is spatial scale adjustment parameter of the field, $\alpha = v + d/2$ controls the amount of smoothing and τ the variance. Non-stationarity was introduced by allowing for spatial varying parameter on $\tau(s)$ (i.e. using the coordinates of health facilities). To obtain the desired Markov structure representation the SPDE is projected to the domain using the spatial and temporal basis function $x(s,t)$,

$$x(s,t) = \sum_k \psi_k(s,t) x_k$$

where the distribution of $x(s,t)$ approximates the distribution of solution to SPDE in the spatio-temporal domain with $\psi(s,t) = \psi_i^s(s) \cdot \psi_j^t(t)$. 2-D piece-wise linear basis functions were used in the spatial domain and one degree B-splines for temporal domain.

The general Bayesian model has the linear mixed model form;

$$Y(s_i,t) = \eta(s_i,t) + e(s_i,t)$$

Where $\eta(s_i,t)$ denotes the overall mean structure (or linear predictor) and $e(s_i,t)$ the residual error term. This model constituted a general form of regression models of the exponential family. The mean component $\eta(s_i,t)$ was set to $\eta(s_i,t) = x(s,t)^T \beta$ to allow for spatial-temporal varying covariate (EVI). The additional component $e(s_i,t)$ were further decomposed into $e(s_i,t) = \alpha_s(t) + w(s_i,t) + \varepsilon(s_i,t)$ following Barnerjee *et al.* (2004), where $w(s_i,t)$ is a mean-zero spatio-temporal process modelled from matérn covariance function and $\varepsilon(s_i,t)$ is Gaussian white noise process. This specification separates the spatial-temporal effects from pure error effects.

The space-time covariance was modelled as a Kronecker product $Q = Q_s \otimes Q_t$. The projection of the space-time SPDE to the basis representation approach ensures that the GMRF maintains markovian properties with its evaluation resulting in a sparse precision matrix (Q). The realizations of above domain have spatial Matérn covariance function which is flexible within the many families of covariance functions (Barnerjee *et al.*, 2004). For example, the marginal variance parameter $\sigma^2 = (\Gamma(\nu)/\Gamma(\alpha)(4\pi)^{d/2} k^{2\nu} \tau^2$ is related to the smoothing parameter via $\nu = \alpha - d/2$. Furthermore, the popular exponential function can be identified by setting $\nu = 0.5$ and $\alpha = 1.5$ (Lindgren, 2013). The following covariance specification was used:

$$Cov\{w(s_i,t), w(s_j,t)\} = \begin{cases} 0 & if \quad t \neq t' \\ \sigma_w^2 C(h) & if \quad t = t' \end{cases}$$

where $h = \|s_i - s_j\|$ is the Euclidean spatial distance and $Var(w) = \sigma_w^2$ (the marginal variance). The matérn covariance function for $h > 0$:

$$C(\cdot) = \frac{\sigma_w^2}{2^{\nu-1}\Gamma(\nu)} (kh)^{\nu} K_{\nu}(kh)$$

where K_ν is Bessel function of second kind and of order ν (the smoothing parameter), k as a scaling parameter and σ_w^2 is the marginal variance. Values of $\nu = 1/2$ corresponds to exponential covariance function $C(u) = \sigma^2 \exp(-\phi u)$ (Barnerjee *et al.*, 2004, Zhang, 2004, Sahu *et al.*, 2013). INLA implementation takes values of $0 \leq \alpha < 2$ for $\nu = \alpha - d/2$ with marginal variance parameter as $\sigma^2 = \Gamma(\nu)/(\Gamma(\alpha)(4\pi)^{h/2} k^{2\nu} \tau^2)$ and spatial range defined as $\phi \approx \sqrt{8\nu}/k$ (Lindgren *et al.*, 2011).

The σ_{ε}^2 prior specification was $\sigma_{\varepsilon}^2 \sim IG(1, 0.0005)$. The $\alpha_s(t)$ term in $e(s_i,t) = \alpha_s(t) + w(s_i,t) + \varepsilon(s_i,t)$ was modelled as first order auto-regressive terms ($\rho < 1$; $\alpha \sim N(0, \sigma_{\alpha}^2 A(\rho))$) with a time interval of 12 months (Lindgren and Rue, 2008). Thus, indexed time points $t = 1,.....,T$ were associated with a dummy time variable such that $\alpha(t+1) = \rho\alpha(t) + \eta(s_i,t)$.

Priors for $W \sim N(0, \sigma_w^2 Q(k,\tau))$ with $Q = Q_t \otimes Q_s$ was a sparse precision matrix (Lindgren, 2013) with two parameters $\log(\tau) = \theta_1$ and $\log(k) = \theta_2$ with (θ_1, θ_2) having joint independent normal default priors. Value for σ was set to one ($\sigma = 1$, i.e. $\log \sigma = 0$) and an approximate range as 1/5 of the domain used as initial parameters. Thus from (Lindgren and Rue, 2013), parameterisation was based on approximate prior information for field standard deviation (σ) and the range (r).

$$\log \sigma(s) = b_0^{\sigma}(s) + \sum_{k=1}^{p} b_k^{\sigma}(s)\theta_k$$

$$\log r(s) = b_0^{r}(s) + \sum_{k=1}^{p} b_k^{r}(s)\theta_k$$

The above specification the joint distribution can be written as:

$$Y \mid \beta, \sigma_{\alpha}^2, \sigma_w^2, \sigma_{\varepsilon}^2, \rho \sim N(\mu,\ \sigma_{\alpha}^2 A(\rho) \otimes I_{n\times 1} I_{n\times 1}^{'} + \sigma_w^2 I_{T\times 1} I_{T\times 1}^{'} \otimes \Sigma + \sigma_{\varepsilon}^2 I_{Tn\times Tn}$$

where I_n is an identity matrix of order n and $\otimes$ is the Kronecker product.

2.3.7 Model validation and scoring rules

Model comparison was done using deviance information criterion (DIC) (Spiegelhalter *et al.*, 2002). A linear Pearson correlation coefficient was calculated for selected model to compare the estimates to the observed values. This was based on 26 health facilities selected randomly as

validation set. Model calibration (statistical consistency) and sharpness (concentration) was conducted using the probability integral transform (PIT) and the conditional predictive ordinate (CPO), a leave-one-out cross-validation approach in which a estimate is validated based on the fitted model and the remaining data only (Spiegelhalter *et al.*, 2002, Czado *et al.*, 2009). The CPO, the probability of observing a value given all other data, was examined for all observations in a full Laplace model (Martins *et al.*, 2013). Model scoring rules such as the square error score (SES) and the ranked probability score (Gneiting and Raftery, 2007) were also computed. Gneiting and Raftery (2007) discuss model scoring procedures such as the standard error score (SES) and the ranked probability score (RPS). To predict a value y_p given other values, the predictive posterior density $P(y_p \mid y_{-p})$ is given by:

$$P(y_p \mid y_{-p}) = \int \pi(y \mid \theta, y_{-p}) \pi(\theta \mid y_{-p}) d\theta_p$$

and obtained via a finite sum with weights Δ_j:

$$P(y_p \mid y_{-p}) \approx \sum_{1}^{j} \pi(y_p \mid \theta_j, y_{-p}) \pi(\theta_j \mid y_{-p}) \Delta_j$$

A score is then said to be proper if there is consistency between the news estimates and the observations (model is correctly calibrated). The SES and RPS are computed as:

$$SES(P, y) = (y - \mu_P)^2$$

$$RPS(P, y) = \sum_{k=0}^{\infty} \left[P(Y \le k) - 1(y \le k) \right]^2$$

where P is the predictive posterior distribution with mean μ_p and standard deviation σ_p and y is the observed count (Gneiting and Raftery, 2007).

The SES is comparable to square mean error $(y-\mu)^2$ except that it applies to a predictive distribution. Lastly, sensitivity analyses for the slide positivity rate were conducted using the root mean square error (RMSE), the absolute mean error (MAE) that summarised the closeness of validation set data to observed values. These two quantities also estimate model bias and accuracy. Additionally, nominal model coverage of 95% credible intervals was assessed based on the validation set. The MAE and the RMSE is given by:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Z^*(x) - Z(x)\right|$$

$$RMSE = \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}(Z^*(x) - Z(x))^2\right)}$$

2.4 Results

2.4.1 Summary of assembled data

Table 2.2 provides a summary of assembled health facility data and mean slide positivity rates by region in northern Namibia. 17 health facilities had no malaria reports in 2009, but, the remaining health outlets returned complete reports every month with no cases reported as a zero. In total, 134,851 cases were diagnosed clinically while 90,835 individuals were examined for malaria parasites of which, 9,893 were positive. Higher case loads were reported in regions bordering Angola compared to the southern regions. The unadjusted mean slide positivity rate was 11.2 [95% CI 6.7 - 15.7] (Table 2.2) and crude annual incidence was 16 cases per 1000 population.

Table 2.2 Summary of number of suspected and confirmed malaria cases in northern Namibia by region. The last column show the percentage of population likely to use a health facility for fever treatment based on the reported pattern form household survey (MIS)

Region	Health district	Number of health facilities (number with missing data)	Number of constituencies	Confirmed malaria cases	Suspected malaria cases	Mean slide positivity rate (95% CI)	Population 2009	Percent of population attending a PHF[1] modelled
Caprivi	Katima	27(2)	6	954	10,605	21.1 (17.9-24.3)	87,088	68.0
Kavango	Andara	10(0)	1	309	4,293	9.2 (7.0-11.3)	26,677	71.1
	Nankudu	11(1)	2	244	7,662	8.4 (6.0-10.8)	48,715	64.2
	Nyangana	8(0)	1	665	3,063	25 (20.1-29.9)	19,815	71.9
	Rundu	23(1)	5	1,176	34,608	16.4 (13.4-19.4)	119,855	71.1
Kunene	Khorixas	8(0)	1	1	89	2.7 (-0.5-6.1)	12,469	61.4
	Opuwo	14(0)	3	539	856	47.3 (40.7-53.8)	52,485	52.5
	Outjo	4(0)	2	1	53	1.1 (-0.2-2.5)	20,395	53.4
Ohangwena	Eenhana	10(1)	4	379	3,956	7.1 (4.8-9.4)	80,419	68.2
	Engela	16(0)	6	916	13,774	9.8 (7.8-11.9)	131,744	74.2
	Kongo	4(1)	1	529	1,788	24.3 (15.8-32.8)	24,744	61.5
Omaheke	Gobabis	14(2)	7	11	96	13.8 (9.5-18.1)	68,433	62.1
Omusati	Okahao	9(1)	2	384	9,066	4.1 (2.1-6.1)	29,964	73.6
	Oshikuku	19(0)	5	436	10,315	3.6 (2.6-4.7)	101,587	75.2
	Outapi	10(0)	2	1,970	9,846	8.4 (6.6-10.2)	48,812	70.8
	Tsandi	10(1)	3	617	5,339	8.4 (6.2-10.6)	54,418	70.1
Oshana[2]	Oshakati[2]	19(4)	10	353	9,133	3.1 (1.9-4.3)	169,053	75.4
Oshikoto	Onandjokwe	16(0)	8	266	8,516	2.3 (1.6-3.1)	146,436	69.8
	Tsumeb	5(1)	2	28	628	5.1 (1.8-8.4)	29,094	67.4
Otjozondjupa	Grootfontein	6(0)	2	59	547	13.7 (7.1-20.3)	33,347	61.3
	Okahandja	2(1)	2	3	110	5.2 (0.2-10.1)	40,209	64.2
	Okakarara	5(0)	1	17	189	11.6 (5.0-18.2)	21,748	56.6
	Otjiwarongo	10(1)	2	36	319	5.4 (2.6-8.1)	42,336	67.3
Total		**260(17)**	**78**	**9,893**	**134,851**	**11.2 (6.7-15.7)**	**1,409,841**	**65.3[3]**

1. PHF is an abbreviation for 'Public Health Facility', which in this case does not include private facilities or privates for profit
2. Two constituencies in Oshana region (Okatyali and Ompundja) did not have any health facilities, thus, the polygons where treated as missing data.
3. Public health facility attendance for treatment of fever based on probability of attendance and the distance decay effect. Description outlined in Alegana *et al.*, 2012.

2.4.2 The malaria incidence model results and validation

Table 2.3 shows comparison of two models Model 1(without covariates) and Model 2 (with covariates) based on the DIC, SES and RPS. Model 2 was marginally better compared to Model 1. The standard deviation of predictive distribution for Model 2 was also lower compared to Model 1. The lower the predictive score the better the model. The conditional predictive ordinate (CPO) for both models was 0.22 (Table 2.3) and since a smaller CPO value usually indicates

greater predictive accuracy (Schrödle and Held, 2010), the result suggested a small difference between the two fitted models. Model 2 (with EVI) was, however, used as the basis for presenting subsequent model outputs since covariate could be included when smoothing for unknown areas.

Table 2.3: Comparison of the implemented Bayesian models
Comparison is based on the deviance information criterion (DIC), the rank probability score (RPS) and the squared error score (SES). M2 had lower score of the DIC, RPS and SES

Model	DIC	Mean deviance	Number of effective parameters	CPO	RPS	SES	Standard deviation of predictive distribution	Mean of predictive distribution
Model 1 (without covariate)	3123.90	3113.22	9.79	0.23	0.6922	1.7039	1.3053	1.1526
Model 2 (with covariate)	3123.80	3112.08	10.68	0.23	0.6662	1.6093	1.2686	1.1343

Figure 2.9 shows a scatter plot of the estimates compared to the observed cases based on Model 2. This Pearson correlation coefficient for the model was 0.56. The correlation was based on the 26 health facilities selected randomly in northern Namibia. Table 2.4 lists parameters for the two CAR models. The seasonal effect parameter (2.02 with Crl 0.16 – 5.79), health facility random effect (6.95, Crl 2.65 – 13.22) and the constituency effect (0.20, CrI 0.02 – 0.57) were all significant at 95% Crl (Bayesian credible interval). There was marginal difference in the mean intercept: -1.80 Crl (-1.98 - -1.64) and -1.76 Crl (-1.93 - -1.58) for model with and without covariate information, respectively. The spatial effect parameter (ϕ) was significant in both models.

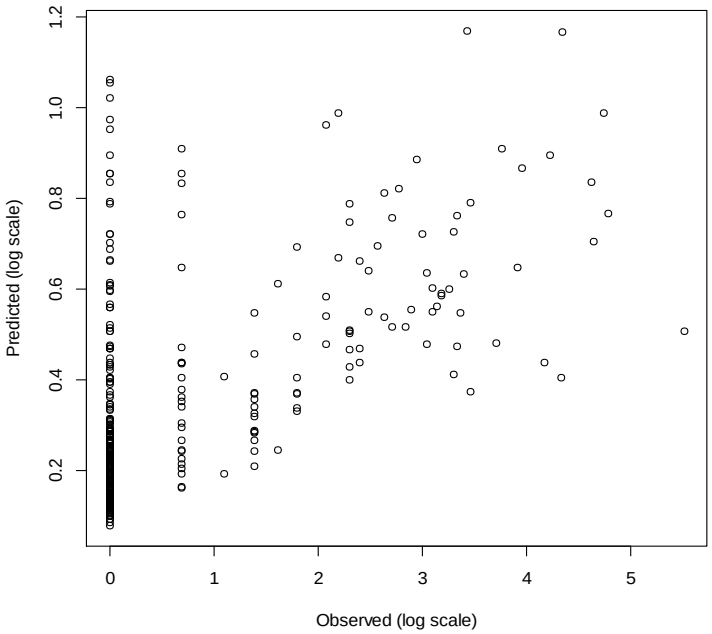

Figure 2.9: Comparison of observed and the estimated incidence in Namibia
Scatter plot of the observed cases on a log scale compared to the estimated case for the Bayesian model with environmental covariate (EVI).The Pearson correlation was 0.56.

Table 2.4: Posterior distribution of parameters for incidence model.
Posterior means and the 95% Bayesian credible intervals (Crl) of the parameters for the two zero-inflated CAR models of malaria incidence in northern Namibia on a log scale

Parameter	**Model 1**	**Model 2**
	Without covariates: Posterior mean, median, (95% CrI)	With environmental covariate: Posterior mean, median, (95% CrI)
μ (Intercept)	-1.763, -1.760 (-1.932 - -1.581)	-1.803,-1.800 (-1.980 - -1.639)
Enhanced Vegetation Index (EVI)	-	0.093, 0.093 (-0.028 - 0.211)
ϑ (parameter for Zero-inflation)	0.843, 0.843 (0.833 - 0.856)	0.843, 0.843 (0.833 - 0.854)
τ_m (seasonal random effect)	1.546, 1.023 (0.137 - 4.692)	2.015, 1.427 (0.161 - 5.789)
τ_f (facility random effect)	6.912, 5.836 (2.605 - 14.830)	6.952, 6.388 (2.641 - 13.220)
ϒ (unstructural random effect)	0.190, 0.136 (0.020 - 0.542)	0.200, 0.144 (0.019 - 0.568)
ϕ (structural random effect)	0.081, 0.045 (0.003 - 0.278)	0.080, 0.004 (0.030 - 0.276)

2.4.3 *Plasmodium falciparum* incidence in northern Namibia in 2009

Figure 2.10 (below) compares the estimated monthly incidence per 1000 population of *P. falciparum* with the calculated crude incidence based on the reported cases by month in 2009. Table 2.5 shows a comparison between the crude incidence and estimates based on the Bayesian approach. Incidence peaked in March and April period compared to later months of the year. The

Bayesian estimates for September to December period was higher, possibly due to the inclusion of spatial interaction effects and the environmental covariate. Figure 2.9 showed that the highest incidence was in Kunene, Kavango, Caprivi and in a few constituencies in Ohangwena region that borders Angola.

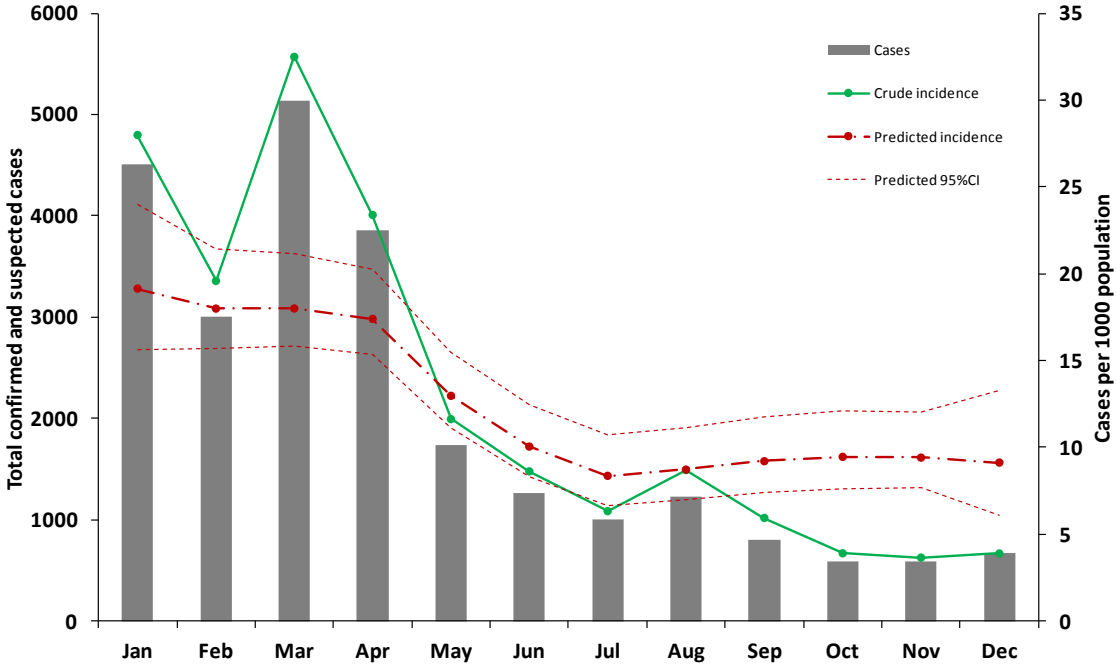

Figure 2.10: Temporal plot of estimated incidence
Plot showing the observed cases by month (dark grey vertical bars), the calculated crude incidence (green line) based on the reported cases and estimated population and the predicted incidence per 1000 population of P. falciparum malaria in northern Namibia in 2009 (dashed red line with upper and lower limits at 95% Crl).

Crude incidence estimates varied widely across the regions compared to a smoothed incidence from the Bayesian model that incorporated environment covariate. Crude estimates can over-estimate incidence where the denominator population is small, and do not account for spatial and temporal dependencies in the data. For example, crude annual incidence in Katima district in Caprivi region was 133 cases per 1000 population compared to a smoothed estimate of 12 cases per 1000 population when using a Bayesian model. The predicted annual mean incidence from

the Bayesian CAR model was 13 cases per 1000 population in the 78 constituencies in northern Namibia.

Figure 2.11 shows the predicted mean spatio-temporal maps of malaria incidence in northern Namibia in 2009. The annualised incidence for 2009 is shown in Figure 2.12 with corresponding standard deviations in Figure 2.13. The highest estimated incidence was between 15 and 20 cases per 1000 population during January to April period and in constituencies bordering Angola. A similar pattern of incidence was depicted in the annual mean maps. The lowest risk was in southern most constituencies of Omaheke. The estimates in the southern constituencies also had higher standard deviations compared to northern constituencies.

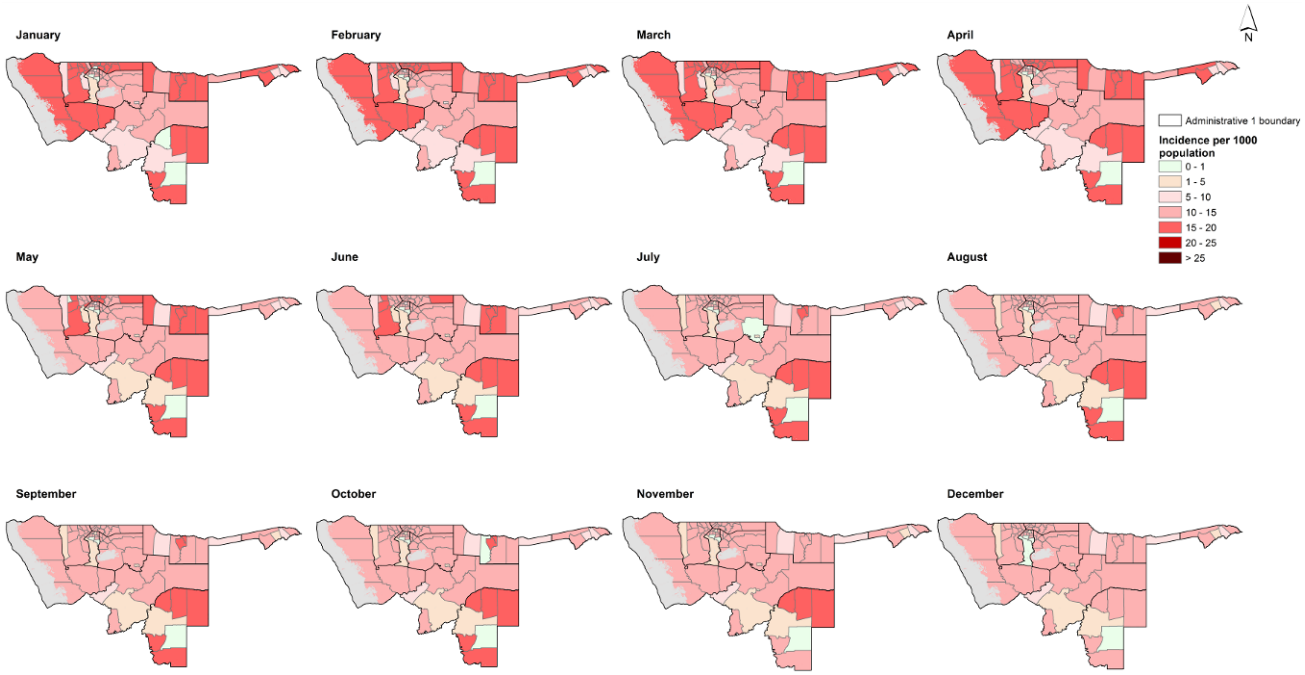

Figure 2.11: Monthly maps of *P. falciparum* incidence in 2009 in northern Namibia
estimated incidence of *P. falciparum* malaria at constituency level using Bayesian spatio-temporal CAR zero-inflated models with environmental covariates (Model 2)

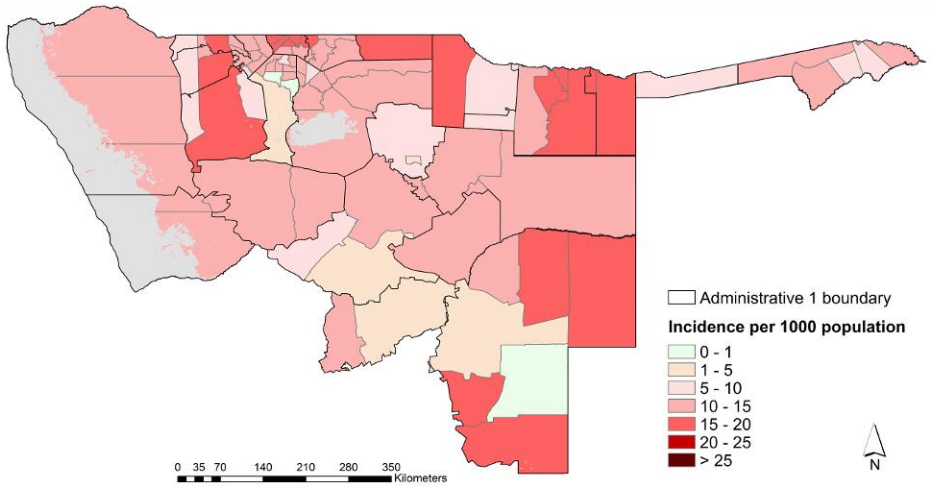

Figure 2.12: Mean incidence of *P. falciparum* in northern Namibia in 2009
The mean annual incidence of *P. falciparum* malaria based on Bayesian CAR with environmental covariates (Model 2).

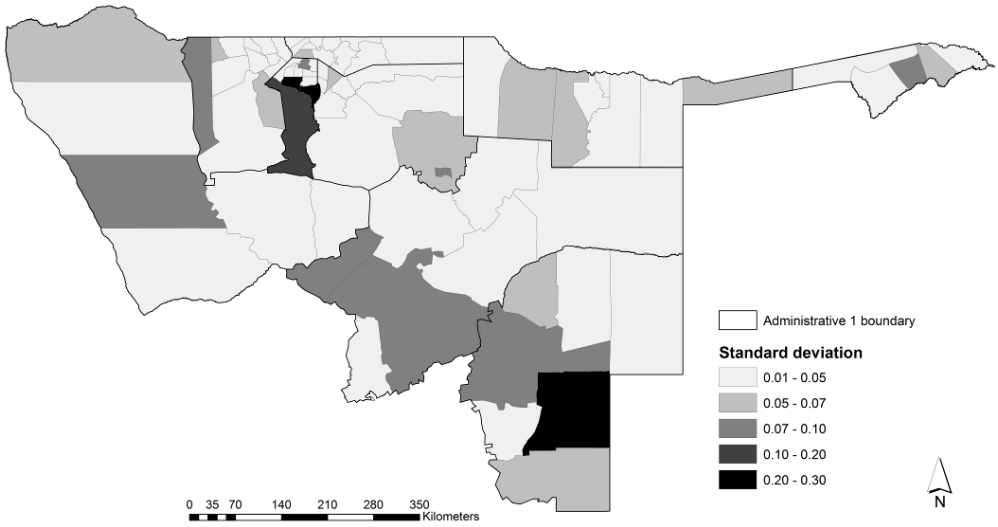

Figure 2.13: Standard deviation map of annualised incidence for 2009
Figure 2.11: The standard deviation map based on Bayesian model. The largest values were in regions where no or less data was reported.

Table 2.5: Comparison of crude malaria incidence and estimated incidence from a Bayesian approach
Crude incidence was derived based on total malaria cases and the population likely to attend treatment while the Bayesian spatio-temporal approach was used to smooth this incidence across regions in northern Namibia.

Region	Health district	Number of health facilities (number with missing data)	Number of constituencies	Crude incidence per 1000 population	Mean estimated malaria incidence per 1000 population via Bayesian approach
Caprivi	Katima	29(2)	6	132.7	11.5 (9.8-13.9)
Kavango	Andara	10(0)	1	24.5	9.3 (7.5-11.9)
	Nankudu	13(1)	2	15.5	12.5 (10.5-15.4)
	Nyangana	8(0)	1	78.7	15.1 (13.2-17.7)
	Rundu	23(1)	5	51.1	12.7 (10.8-15.2)
Kunene	Khorixas	8(0)	1	0.2	14.6 (12.5-17.6)
	Opuwo	16(0)	3	20.6	11.8 (9.7-15)
	Outjo	4(0)	2	0.1	14.5 (12.6-17.3)
Ohangwena	Eenhana	10(1)	4	9.7	14.8 (12.9-17.4)
	Engela	17(0)	6	16.8	14.3 (12.5-16.9)
	Kongo	4(1)	1	48.5	15.6 (13.5-18.8)
Omaheke	Gobabis	16(2)	7	0.3	11.5 (9.6-14.4)
Omusati	Okahao	9(1)	2	31.9	12.3 (10.4-14.9)
	Oshikuku	19(0)	5	10.8	14.5 (12.8-16.9)
	Outapi	10(0)	2	65.6	14.8 (12.9-17.5)
	Tsandi	10(1)	3	21.7	10.3 (8.5-12.7)
Oshana[2]	Oshakati[2]	20(4)	10	4.2	8.8 (7.6-10.6)
Oshikoto	Onandjokwe	16(0)	8	4.4	12.8 (11.1-15.5)
	Tsumeb	6(1)	2	1.9	6.7 (5.0-9.3)
Otjozondjupa	Grootfontein	7(0)	2	3.6	13.4 (11.6-16.3)
	Okahandja	3(1)	2	0.1	9.2 (7.5-12.1)
	Okakarara	5(0)	1	1.6	13.6 (11.7-16.5)
	Otjiwarongo	10(1)	2	1.3	9.7 (7.9-12.6)
Total		**273(17)**	**78**	**17.4**	**12.5 (10.3-14.6)**

2.4.4 Assessing the population at risk of malaria in 2009 in northern Namibia

Table 2.6 (below) shows the estimated population at risk of *P. falciparum* malaria by region in 2009. Based on the Bayesian model, 383,632 people (27.2%) lived in areas where case incidence

was greater than 15 cases per 1000 population; slightly more than half 745,903 (52.9%) lived in areas where case incidence was between 10 to 15 cases per 1000 population; approximately 216,512 (15.4%) resided in regions with an average of 5 to 10 cases per 1000 population; 49,005 (3.5%) in areas with greater than 1 case, but less than 5 cases per 1000 population and 1% of population lived in regions with less than 1 case per 1000 population.

Table 2.6: Population at risk in Namibia based on estimated incidence.
Estimated population in northern Namibia at risk of *P. falciparum* in 2009 by region. Each classification class represent estimated incidence per 1000 population

Region	< 1.0	>1.0 -< 5.0	>5.0 -< 10.0	>10.0 -< 15.0	> 15.0	Total
Caprivi	0	0	25,614	61,474	0	87,088
Kavango	0	0	75,393	26,097	113,572	215,062
Kunene	0	0	0	85,348	0	85,348
Ohangwena	0	0	0	91,438	145,469	236,907
Omaheke	7,249	9,376	0	7,153	44,656	68,433
Omusati	0	0	27,741	127,104	79,935	234,780
Oshana	7,539	3,866	32,573	125,074	0	169,053
Oshikoto	0	19,041	24,938	131,550	0	175,529
Otjozondjupa	0	16,722	30,254	90,664	0	137,640
Total	**14,789**	**49,005**	**216,512**	**745,903**	**383,632**	**1,409,841**

2.4.5 The predicted slide positivity rate in 2009

2.4.5.1 Model validation results for slide positivity rate

Table 2.7 shows parameters of the slide positivity rates model. The model had an absolute mean error of 1.1 which measured the overall magnitude of predictions and a RMSE of 2.1. The Pearson correlation coefficient which measured the linear association between the predicted and the observed values was 0.7 as shown by the scatter plot (Figure 2.14 (A)). The absolute error of 1.1 indicates a small variation (in terms of magnitude) between observed and predictions. The analysis of standardized residuals using a semi-variogram showed minimum spatial structure (Figure 2.14 (B)) i.e. the residual spatial autocorrelation unexplained by the model after accounting for spatial effects.

Table 2.7: Bayesian geostatistical model parameters of Slide positivity rates for Namibia in 2009

	DIC	P_D[1]	MAE	R.M.S.E	Probability of prediction interval (%)[2]	Pearson correlation (%)
SPR Model	372.98	115.03	1.06	2.10	93.82	0.6985

1. P_D represent the effective number of parameters
2. The nominal probability of prediction is 95%

The actual coverage probability of a prediction interval was used on a validation set to test how well the posterior distributions captured uncertainty at nominal 95% credible interval. Thus for a perfect model, 95% of value should fall within 95% of credible interval predicted at each location. The actual coverage probability was 93.8% indicating a slight underestimation of uncertainty. The RMSE was 2.1 % and corresponding AR(1) coefficient for structure time component was 0.8430. The model spatial range was 54.8 Km. Further parameters from the model such as the marginal variances σ_w^2 are tabulated (Table 2.8).

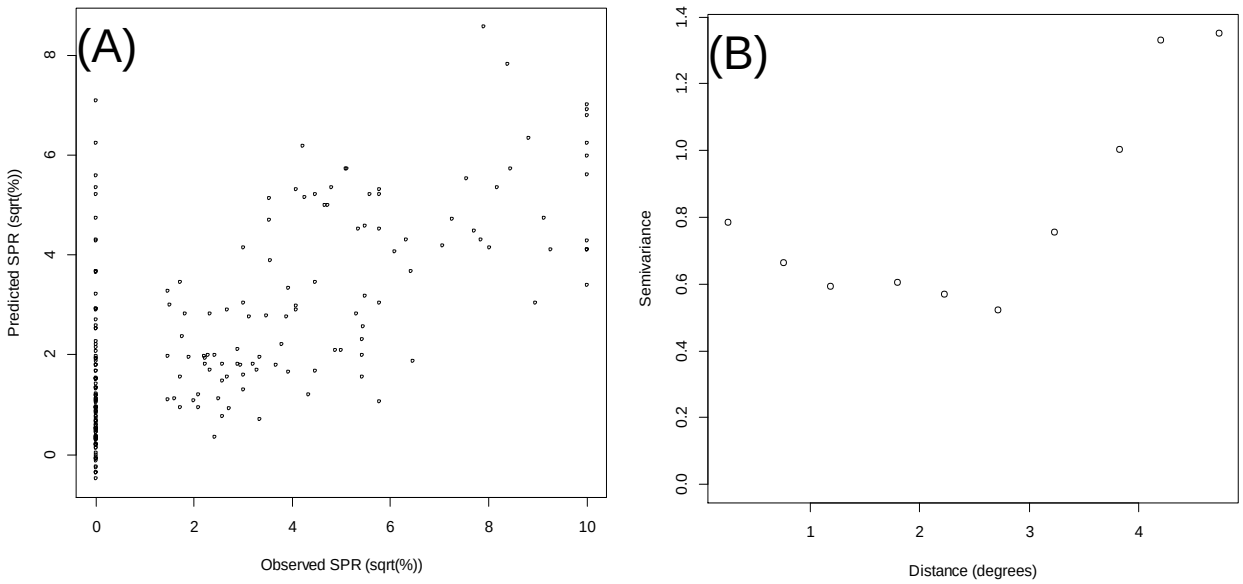

Figure 2.14: Scatter plot of predicted and observed slide positivity rate and semi-variogram of residuals
(A) Scatter plot of the predictions and the actual values (B) standardized residual semi-variogram

Table 2.8: Posterior distribution of parameters for slide positivity rate
Posterior estimates (mean, standard deviation and quantiles) of the fixed components (intercept and EVI), the matérn marginal variance component), the AR(1) coefficient and the model range

Parameter	Mean	Std. dev	5%	50%	95%
β_0	0.2648	0.0612	0.0164	0.2647	0.3660
β_1	0.0404	0.0090	0.0255	0.0404	0.0551
σ_w^3	0.1576	0.0476	0.0945	0.1496	0.2475
ρ	0.843	0.0981	0.6555	0.8626	0.9627
ϕ	0.3093	0.0750	0.2095	0.2965	0.4510

2.4.5.2 Predicted slide positivity rate

Figure 2.15 shows the predicted continuous and binned maps of slide positivity rate weighted by the probability of seeking treatment at the nearest public health facility when sick with fever. Slide positivity was higher in regions bordering Angola, consistent with predicted incidence prediction. The positivity rates in Omaheke regions were low. Table 2.9 shows the average positivity by region. The mean slide positivity rate for northern Namibia was 4.5% (minimum 0.2%, maximum 19.2%). The constituency with highest slide positivity rate was Eenhana in Ohangwena region. Figure 2.16 show the associated standard deviations of the predictions which were higher in areas with no data. Overall, majority of regions had slide positivity of less than 5%, which is a threshold for pre-elimination. Three regions, Capirivi, Ohangwena and Kavango had more than 5% positivity threshold.

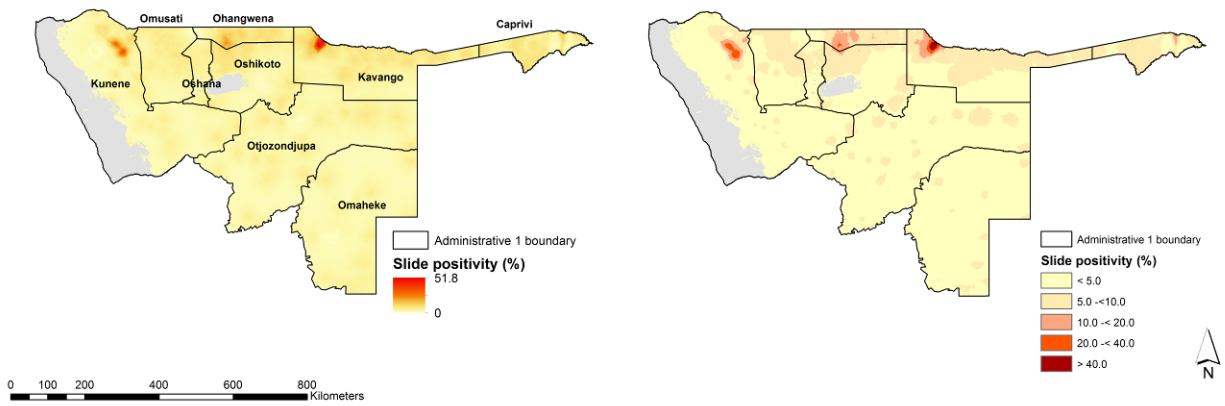

Figure 2.15: Posterior mean predictions of slide positivity rate at 1 km x 1 km
Posterior mean estimates of the slide positivity rate in northern Namibia in 2009 on a continuous scale (the first map) and categorized map (5 classes of : <5%; 5%-<10%; 10% -< 20%; 20% -< 40% and ≥40%).The grey mask is the aridity corresponding to regions were EVI<0.1

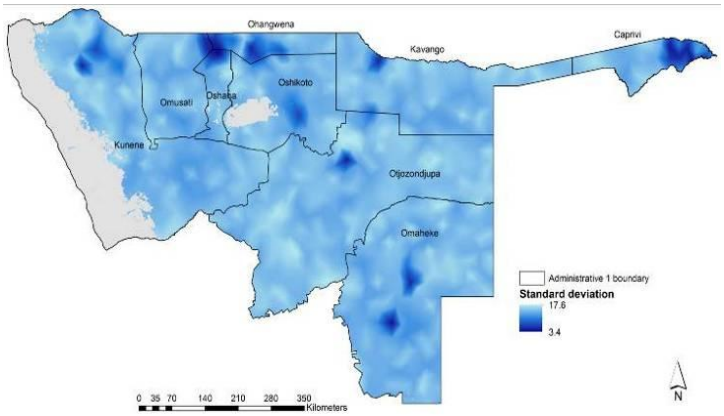

Figure 2.16: Standard deviation map of slide positivity rate

Table 2.9 Slide Positivity Rate by region in Namibia.
Comparison of predicted slide positivity rates (%) by region in northern Namibia in 2009

Region	Mean SPR	Minimum (SPR)	Maximum (SPR)	Range (SPR)
Omusati	4.4	0.0	9.1	9.1
Oshikoto	4.0	0.1	15.4	15.3
Oshana	4.3	0.4	7.9	7.4
Ohangwena	8.2	0.0	22.8	22.8
Otjozondjupa	3.0	0.0	7.7	7.7
Kunene	2.3	0.0	36.1	36.1
Kavango	6.1	1.2	51.8	50.6
Caprivi	6.0	0.0	14.6	14.6
Omaheke	2.4	0.0	7.4	7.4
Average	**4.5**	**0.2**	**19.2**	**19.0**

2.5 Discussion

2.5.1 Introduction

The evaluation of malaria pre-elimination or elimination programme requires a detailed description of local epidemiology. These descriptions are required as baseline upon which evaluation and transition of risk is estimated. This study investigated two important aspects on baseline malaria incidence in 2009 and comparison of this incidence with the slide positivity rates. The smoothing of incidence and prediction of slide positivity rates was carried out using Bayesian model based approaches with the former including a CAR prior to smooth risk at constituency level. A higher incidence of malaria was observed between January and April in the constituencies bordering Angola and Zambia while lower values were estimated between the months of July and December. The mean malaria incidence estimated for northern Namibia was 13 cases per 1000 population for 2009. For the same northern regions a 4.5% mean slide positivity rate was predicted using a model based geostatistical approach. Both modelling approaches included environmental covariates. According to the WHO, the target threshold for pre-elimination based on incidence is less than 1 case per 1000 population or in areas where slide positivity rates are less that 5% (World Health Organization, 2007a). Incidence was estimated based on reported clinical and confirmed cases in 2009 in northern Namibia while slide positivity rate only looked at cases examined at the facility. The two indices are different but can be useful indicators to establish baseline makers of malaria elimination in Namibia.

2.5.2 Implications for malaria control and elimination in Namibia based on estimated incidence

From the monthly maps of Namibia (Figure 2.9), a higher incidence of malaria was observed between January and April, while, lower values were observed for the July and December period. Thus, malaria control efforts should be lagged with the peak observed in early months of the year. These include the distribution of ITNs, parasitological diagnosis and treatment (World Health Organization, 2014d). Universal ITN coverage should target areas with high malaria incidence such as Caprivi, Omusati and Kavango.

Malaria transmission in Namibia is likely to be highly seasonal with precipitation months between November and March. The findings also suggest that malaria cases peak early in the year in March and April. However, precipitation patterns could vary from year to year and this may lead to low number of malaria cases in drought years or epidemics in a rain year in the northern Namibia (De Meillon, 1951, Ministry of Health and Social Services and World Health Organization, 1996). Aridity limits transmission on the western coast with the Atlantic ocean and in the southern desert fringe regions. These have been masked in the maps presented (Figure 2.9 and Figure 2.10). The southern regions of Hardap, Karas, Erongo and Khomas are masked because they are defined as malaria free, although, Khomas and Erongo may support unstable transmission (Snow *et al.*, 2010a).

The mean incidence observed for 2009 was highest in regions bordering Angola, Zambia and Botswana. Historical *Plasmodium falciparum* data for Namibia between 1969 and 1992 (Noor *et*

al., 2013a, Noor *et al.*, 2013b) suggest a parasite prevalence of greater than 5% in Kavango and other northern regions along the border with Angola. In addition, Craig and others showed that in Botswana, the area along the north-western border areas with Namibia had relatively high prevalence (Craig *et al.*, 2007). For these border constituencies concerted efforts with neighbouring countries have to be put in place to realize the pre-elimination targets (Noor *et al.*, 2013b, Noor *et al.*, 2013c). It is important to note that the malaria parasite can be imported as a result of residence returning after travelling to other endemic regions (Angola, for example), by visitors from across the border, permanent migration or by infected mosquitoes moving to the area (Cohen *et al.*, 2012). A higher Incidence in the border regions could well be driven by cross border population movement (Cosner *et al.*, 2009, Noor *et al.*, 2013b). Similar suggestions were made for two districts in South Africa close to the Mozambique border (Kleinschmidt *et al.*, 2002) and in Yunnan province in China that borders Myanmar, Laos and Vietnam (Clements *et al.*, 2009).

2.5.3 Comparison of estimated incidence, parasite prevalence and slide positivity rates

Figure 2.17 (below) shows comparisons between the estimates (incidence and slide positivity rate) with parasite prevalence at community level. Parasite prevalence maps were produced by interpolating community prevalence data (Noor *et al.*, 2014). There few points to consider. First, the parasite prevalence is representative of specific age group for children (2-10 years) but modelled using Bayesian approaches. Secondly, data for the parasite prevalence was assembled from independent surveys randomly within the population, the denominator being number of people examined within the community. There was a positive correlation (Pearson correlation coefficient 0.5) between the estimated incidence per 1000 population and the age specific (2-10

years) parasite prevalence. This suggested that areas with higher incidence of *P. falciparum* in Namibia were associated positively with parasite prevalence. Similar results were observed for the slide positivity rate (Pearson correlation coefficient 0.5).

Namibia NMCP has already deployed the use of RDTS and microscopy at health facilities for diagnosis of suspected cases prior to treatment (Ministry of Health and Social Services, 2010c). In low transmission settings, where malaria is highly seasonal and unstable, fevers due to malaria are less common and low parasite density is not detectable easily by routine diagnostic tools such as RDTs. The mean slide positivity rate for Namibia in 2009 was 4.5%. In northern Namibia, only Kavango, Caprivi and Ohangwena had slide positivity exceeding the 5% threshold for pre-elimination after adjusting for utilisation of facilities for fever treatment. The rest of regions in the north appear to be within pre-elimination targets based on baseline rate of slide positivity. At low parasite densities, the slide positivity rate may well be an indicator of endemicity in population using the health facility and has been illustrated in São Tomé and Príncipe (Lee *et al.*, 2010) and in Yuunan province in China (Bi *et al.*, 2012). Namibia was also one of the countries in SSA where only small proportion of cases is treated in the private sector (Cohen *et al.*, 2012). Another characteristic of low transmission settings is that malaria may be in marginalised population e.g. at the borders (World health Organisation, 2012a). For Namibia, this was also evident from the posterior mean of the slide positivity for 2009 (Figure 2.13) where constituencies in Kunene, Ohangwena and Kabe constituency in Caprivi region had the highest slide positivity rates.

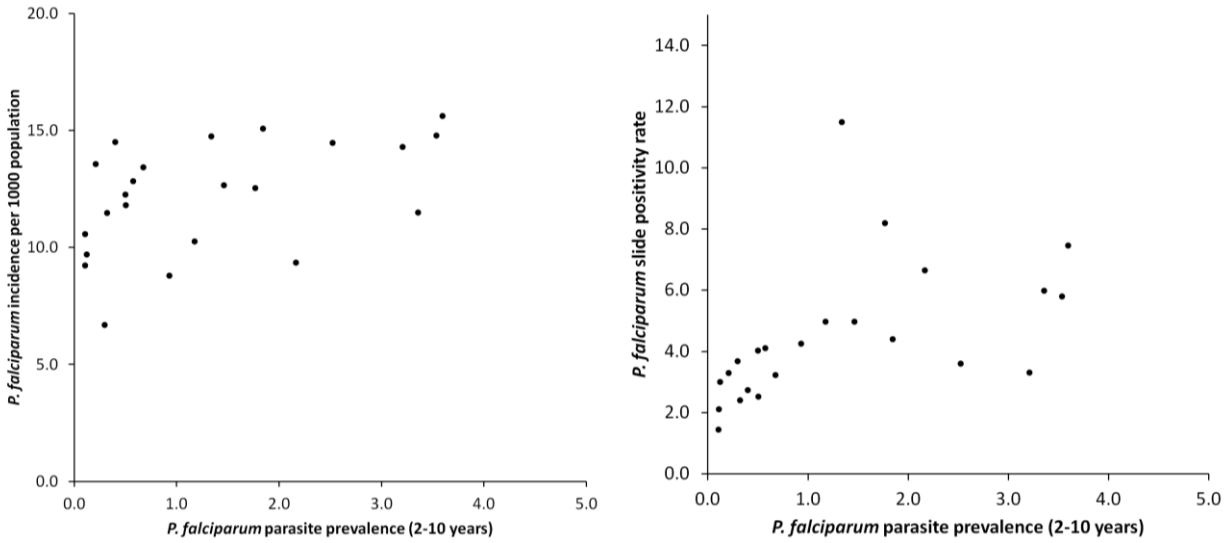

Figure 2.17: Comparison of incidence, slide positivity and parasite prevalence
Scatter plot of community measured parasite rate with the modelled slide positivity rate at health facility level.

2.5.4 The model based approach for *P. falciparum* incidence compared to crude incidence

The spatio-temporal model included the random effects at facility and constituency level as well as introduced dependencies spatially and temporally using the CAR prior affect and auto-regressive component respectively. Thus the modelled risk was not only as a function of smoothing but also taking into account of spatial autocorrelation via the GMRF (Banerjee and Carlin, 2003, Rue and Held, 2005). In general there was smoothing of incidence towards the overall mean. This contributed to wider differences between crude incidence and the smooth estimates in some regions. This is likely to be a factor of using 0-1 weights as neighbour matrix during the model set up (Section 2.3.5). This means that neighbour areas are correlated and treated as independent if not a neighbour of the region of interest. Other studies have found differences in level of smoothing based on specification of the neighbourhood matrix (Earnest *et al.*, 2007). It is possible to mitigate this effect by introducing higher level random effects while maintaining a local CAR prior effect to control local-mean smoothing (Lee, 2011), in addition to

the facility level effects. This alternative methodology is explored in next chapter on Afghanistan.

Crude incidence estimates are overestimated where the denominator population is small and malaria cases are few. The hierarchical Bayesian zero-inflated CAR model addressed several sources of uncertainty in comparison to the crude incidence estimates. First, the model was applied at facility level and, therefore, the method not only takes into account the nature of the facility, but also season and environmental factors (at facility level) in adjusting for under-reporting. Secondly, incidence was smoothed across the facility reports, thereby addressing the potential impact of model instability resulting from small numbers of reported cases, apparent in the facility data. Smoothing incidence also reduces the potential impacts of under-reporting of cases by facilities. Third, incorporating the environmental covariate (EVI) explained spatial variation where data were absent in addition to providing information on the climatic suitability of malaria transmission, for example, in Omaheke region (Craig *et al.*, 1999, Guerra *et al.*, 2008). This suggested that the inclusion of environmental covariates improved the model estimates for a few constituencies (in Kunene and Omaheke), but only marginally as suggested by the DIC between the two models implemented.

The cross-validation approach used in this case study draws from predictive distribution of the model. This deviates from large literatures that mainly use statistics such as the mean error (Clements *et al.*, 2009, Raso *et al.*, 2012, Noor *et al.*, 2013c, Sahu *et al.*, 2013). The mean error, for instance, summarise the overall model performance which could deviate from individual estimates. The analysis in this thesis largely employs the use of model scoring rules to validate

the predictive performance (Gelfand and Ghosh, 1998, Gneiting and Raftery, 2007, Czado *et al.*, 2009). Since the goal is estimation and at unknown locations, the validation procedures used here are likely to assess the predictive performance in a better way. Moreover, the R-square value is likely to improve with increasing number of covariates without improving model performance (Gething *et al.*, 2011b). The Namibia approach used a single covariate and with relatively 90% complete reporting.

2.5.5 Limitations

The approach presented here drew upon a comparatively data rich setting based mainly of RDT and microscopy diagnosis. First these diagnostic tools are documented to have different sensitivities (WHO-FIND, 2009). RDT, for example, may record false positives resulting in overestimation of cases (Bell *et al.*, 2005). Improvement in diagnosis in low transmission areas may involve the use of polymerase chain reaction (PCR) (Zakeri *et al.*, 2010, Noor *et al.*, 2011, Gething *et al.*, 2012). The recent improvements in case management in Namibia in which all suspected malaria fevers are diagnosed parasitologically before treatment will reduce the need for adjustment for test positivity rate. In addition, planned improvements in HMIS reporting and quality and transition to active case detection mean smaller adjustments for treatment seeking and reporting will be required in future. This is may also be useful for external validation of the zero-inflated Bayesian models. Precise incidence estimates should provide a basis for targeting active case detection efforts at specific locations and in specific months, potentially making such resource-intensive efforts more cost-effective. A comparison of resulting estimates with the standard non-spatial WHO approach (Cibulskis *et al.*, 2011) shows that the latter estimates a higher annual malaria incidence of 23 per 1000 population and is likely to be less imprecise. The

main difference between the WHO approach and the one used in this thesis is inclusion of extra parametrisation in Bayesian modelling in addition to dealing with spatial (and temporal) auto-correlation.

One drawback of many studies analysing areal data, and one common to the Bayesian approach used in this study, is the modified areal unit problem (MAUP), a well-known analytical problem in geography that could affect the observed statistical results with a change in shape or size of spatial polygons used in the analysis (Robinson, 1950, Wakefield, 2003, Barnerjee *et al.*, 2004). In this study constituencies were selected as the basis for presenting estimates, with the aim of providing information at this level to health authorities, though the model was fitted at facility level. There is therefore a potential impact of MAUP both in terms of the shape of constituencies and in smoothing at constituency level from facility level data. Secondly, the data used for this study were obtained from the Namibia HMIS which covers the majority of public health facilities in the north. This means that the findings are relevant only for the 12-month time-series in 2009. The results could be improved by inclusion of more data and at different time points to draw more stable long-term spatio-temporal patterns (Zhou *et al.*, 2005). In addition, the modelling approach excluded the effects of population movements between regions, especially across borders, while the relations between the environmental variables could change across space and at(WHO-FIND, 2009) shorter time periods than those considered (Hay *et al.*, 2008). Finally, some sources of uncertainty remain. In particular, the underlying the care-seeking behaviour data used to adjust denominator populations relate to children under 5 years, not the whole population. Utilisation rates were estimated from cross-sectional surveys and therefore may not capture temporal changes in care-seeking behaviour. The underlying utilisation data also relate to fever rather than malaria *per se*.

2.6 Conclusion

The current efforts of the NVDCP to focus aggressive malaria control activities around the border regions with Angola, Zambia and Botswana are likely to play an important role in achieving malaria elimination by 2020 in Namibia (Ministry of Health and Social Services, 2010c, Trans-Zambezi Malaria Initiative (TZMI), 2012, Noor *et al.*, 2013b). Current results suggest that the country may be within the pre-elimination targets in most parts of the northern region. This study provides additional information to identify the highest malaria risk areas in Namibia and when used together with evidence from modelled community parasite prevalence surveys on receptive and contemporary malaria risk, should support malaria control and elimination initiatives in the country. However, long term spatio-temporal trends in incidence will be useful to assess progress over time.

CHAPTER 3: Case Study 2

Examining coverage and utilisation of healthcare in Afghanistan to estimate the incidence of *Plasmodium vivax* and *Plasmodium falciparum* malaria

3.1 The Afghanistan context

3.1.1 Background

In this chapter, passive case data from HMIS was used to estimate the burden of *Plasmodium vivax* and *Plasmodium falciparum* in Afghanistan. Of the 10 countries in World Health Organisation Eastern Mediterranean Region (WHO/EMRO) with ongoing malaria transmission, Afghanistan has the second highest malaria burden after South Sudan, predominantly due to *P. vivax* species with only a small proportion of cases due to *P. falciparum* (Safi *et al.*, 2009a, Safi *et al.*, 2009b). Being a mountainous country, previous studies suggested that malaria risk is higher in regions of altitude ≤ 2000m and in river valleys where irrigation and rice cultivation is practiced (Kolaczinski *et al.*, 2005, Safi *et al.*, 2009b). *P. vivax* is predominant in Asia and only patchy in Sub-Saharan Africa due to absence of Duffy antigen that is usually required for red blood cell invasion stage (Gething *et al.*, 2012). *P. vivax* is also known as relapsing malaria because it can stay dormant for long periods (months or years) in the liver (hypnozoite stage) (Douglas *et al.*, 2011, White, 2011).

Southeast Asia has the most malaria vectors compared to SSA. The main malaria vectors in Afghanistan are about six, namely: *Anopheles superpictus*, *An. culicifacies*, *An. stephensi*, *An. hycranus*, *An. pulcherimus* and *An. Fluviatilis*; although *An. stephensi* and *An. culicifacies* are predominantly found in the east of the country (Rowland *et al.*, 2002, Safi *et al.*, 2009b). In Afghanistan, earlier vector control using dichlorodiphenyltrichloroethane (DDT) during the eradication programme (1957-1969) had reduced vector abundance (Ramachandra, 1951, Dy, 1954, Kolaczinski *et al.*, 2005). However, malaria vectors re-emerged as a result of resistance to insecticides (Eshghy and Nushin, 1978, Delfini, 1989) and a decline in malaria control due to

instability after the Soviet invasion in 1979. The alternative use of malathion and larvivorous fish (*Gambusia affinis*) had limited success (Eshghy and Nushin, 1978, WHO/EMRO, 2003). In addition to re-emergence of malaria vectors, chloroquine resistance and population movement mainly from returning refugees contributed to an increase in malaria burden in Afghanistan (Delfini, 1989, Rowland *et al.*, 1997, Shah *et al.*, 1997).

The national malaria and Leishmaniasis control programme (NMLCP) was re-constituted in 2002 and formulated a national malaria strategic plan (2006 to 2010). Artemisinin-based combination therapies (ACTs), mainly using AS+SP for *P. falciparum*, were adopted in 2004 to replace the SP monotherapy. Chloroquine remains effective for treatment of diagnosed cases of *P. vivax*. Funding from the Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM) (The Global Fund, 2013a) (round 8) has increased procurement and social marketing of ITNs. By 2012, the malaria programme had benefited from approximately US $ 50.2 million in disbursements from the GFATM since 2004.

To track progress towards national targets, the National Malaria and Leishmaniasis Control Programme (NMLCP) and partners established a routine information system to report monthly malaria cases by health facility (Ministry of Public Health, 2006). The health information system, however, captured passively detected case data from only public-based health facilities which contained both suspected and parasitologically confirmed cases. Although healthcare delivery has improved recently in Afghanistan after introduction of the Basic Package for Health Services (BPHS) (Edward *et al.*, 2011), a significant challenge facing Afghanistan is effective delivery of interventions to population at risk. Access and utilisation of public health facilities

are not well understood at a national level especially for remote and conflict-affected provinces. Past and recent service provision assessments were constrained to only a few provinces and cannot be generalized for entire country (Reilley *et al.*, 2004, Edward *et al.*, 2011). Lastly, HMIS is hindered by challenges of low parasite confirmation rates which inflate malaria cases. In addition, low reporting rates tend to underestimate disease burdens because of the spatially and temporally incomplete data (Gething *et al.*, 2006).

The aim of this study is to provide reliable estimates of *P. vivax* and *P. falciparum* at a sub-national level in Afghanistan to enhance decision making. The first section of the chapter provides an overview of healthcare in Afghanistan, including geography, the health system and a review of healthcare access in Afghanistan (section 3.1). Geographic access is subject of section 3.2 that looks into aspects of utilisation and coverage of health services. In terms of modelling utilisation of public health facilities, there is a slight difference from Namibia analysis. Here, analysis was stratified by facility type. The purpose of stratification was to investigate whether non-stratification resulted in a mean distance-decay model compared to health facility specific patterns. Section 3.4 outlines incidence and slide positivity analysis based on Bayesian model based approaches. The slide positivity analysis here was similar to Namibia analysis. A different model parametrisation, for malaria incidence, was used compared to Namibia. The Bayesian hierarchical model here included random effects at facility, district and province levels. In Namibia, random effects were specified only at facility and constituency levels. The aim of extra parametrisation was to improve spatial smoothing of incidence. The results of both healthcare access and incidence modelling are subsequently discussed in section 3.5.

3.1.2 Geography

Afghanistan is a landlocked country located in south central Asia with an estimated land surface area of 640,000 km^2. Its estimated population is 30 million with majority living in the northern and south eastern regions of the country (World Health Organization, 2013a). It is boarded by the commonwealth states of Tajikistan, Turkmenistan and Uzbekistan in the north, Iran to the west, Pakistan to the south east and China in the north east (Figure 3.1). Afghanistan is divided into 34 administrative provinces and 398 districts. It is a highly mountainous country dominated by Hindu Kush range from the central to the north eastern regions. The major rivers are the Helmand in the south, the Kabul river system that passes through the capital to the east, the Hari Rud to the west and the Amu river to the north. The south west regions of Afghanistan bordering Pakistan and Iran are predominantly desert compared to the south east regions that experience some low rainfall. Population is highly dependent on irrigation for Agricultural productivity with the main cultivated crop as rice, in the valley regions. Irrigation in the valleys is aided by melting snow at high mountain ranges from the central regions which in turn provide breeding areas for *Anopheles*. Temperature in the summer can reach as high as 45 °C and fall below 0°C in the winter months.

Figure 3.1: Google map of Afghanistan
Source: Google Maps (Google, 2013)

3.1.3 Organisation of Afghanistan's Healthcare system and financing

Healthcare system in Afghanistan exhibits a hierarchical structure categorised as primary level, the secondary level and the tertiary level. The primary level forms the Basic Package for Health Services (BPHS) constituting clinics, health posts and Maternal Child Health (MCH) centres, basic and Comprehensive Health Centres (CHC). District hospitals form the secondary level, while, provincial and regional facilities form the tertiary level. Secondary and tertiary level facilities constitute the Essential Package for Hospital Services (EPHS) and serves as referrals for the BPHS-level facilities (Ministry of Public Health, 2008b). The BPHS and the EPHS were constituted in 2002 by the Ministry of Public Health (MoPH) (Ministry of Public Health, 2003, Waldman *et al.*, 2006, Ministry of Public Health, 2010a). Prior to the formation of the BPHS and

the EPHS, community outreach programmes served at the village level and were supported by sub-health centres and MCHs. District hospitals were previous known as rural-level facilities and were intermediary to the lower-tier health centres.

The BPHS provides primary care to the rural population and was expanded by contracting services to NGOs and MoPH partners (Waldman *et al.*, 2006, Steinhardt *et al.*, 2011). Priority areas included maternal and new born health (e.g family planning, ANC, PNC); child health (expanded program for immunization (EPI) and integrated management of childhood illness (IMCI)), general nutrition (malnutrition); communicable diseases (malaria, TB and HIV/AIDS); mental health; disability and supply of essential medicine (Ministry of Public Health, 2010a). Health service delivery in Afghanistan is hampered by poor infrastructure following years of conflict, unequal distribution of health facilities, costs of providing medical care and lack of personnel (Acerra *et al.*, 2009, Ministry of Public Health, 2010a). At village level community health workers manage the health posts which provide basic preventive and curative services. In remote areas Mobile Health Teams (MHTs) are used, especially for immunization programmes such as EPI (Belay, 2010, Ministry of Public Health, 2010a). The basic health centres link basic service providers at the community level with the next service tier (the CHC) that are in turn linked to higher level hospitals providing both inpatient and outpatient services. Thus, where no regional or tertiary facility exists, district hospitals are the main referral centres. Inpatient facilities are provided mainly at the tertiary level (Ministry of Public Health, 2010a). At health posts, only clinical diagnosis of malaria is provided along with other basic services. Severe and complicated illness is referred to higher-level district hospitals (Ministry of Public Health, 2010a).

Financing of healthcare is shared partly by government, private and external sources (Palmer *et al.*, 2006, Ministry of Public Health, 2009b, Belay, 2010). The public financing schemes include social insurance offered mainly to government employees or through public based organisations. Other forms of insurance is also provided by private companies or organisations (Ministry of Public Health, 2011). Private financing is through direct out of pocket payments at the peripheral health facilities or through charity or private insurance. Government funds in Afghanistan are channelled through the provincial and regional directorates that manage and coordinate activities at the district level. The general per capita expenditure on healthcare in 2008 was estimated to be US $ 42 (Ministry of Public Health, 2011) which was lower than the proposed minimum amount of US $ 44 per annum on an individual for basic and life saving health services (World Health Organisation, 2012b). Majority (75%) of direct healthcare costs is provided by household through OOP payments (Ministry of Public Health, 2011).

3.1.4 Population health and the MDGs in Afghanistan

Afghanistan has been plagued by conflict since the soviet invasion in (1979 -1996) and during the Taliban reign (1996-2001). This has led to destruction of basic infrastructure, rise in poverty and contributed to low economic development (The PLoS Medicine Editors, 2011, Jacobs *et al.*, 2012, The World Bank, 2012, Singh *et al.*, 2013). Afghanistan endorsed the Millennium declaration in 2004 which extended the country deadline to 2020 rather than the 2015. It also included an extra ninth goal '*to enhance security*' which remains a major problem (Ministry of Finance, 2010).

Afghanistan was ranked as one of the bottom countries (175 out of 186 countries) according to human development index (0.374) in 2012 (UNDP, 2013a) and has only made little progress in addressing the problems of poverty, illiteracy, infrastructure and healthcare provision. Conflict and poor health has contributed to low life expectancy at birth which is estimated currently at 47 years for males and 50 years for the female population; the infant mortality rate is estimated at 129 per 1000 live births; while the under-five mortality rate stands at 191 per 1000 children not surviving age 1 (Afghan Public Health Institute *et al.*, 2011). A mortality survey conducted in 2010 showed higher all-cause mortality in adult population aged over 50 years. Pregnancy-related mortality was estimated to be 327 (95% confidence interval: 260-394) and was four times higher in rural areas compared to urban areas (Afghan Public Health Institute *et al.*, 2011). It is less likely that progress will be made to reduce the current MMR which is compounded by factors such as illiteracy and lack of autonomy in the households amongst the female population as well as poor access to quality healthcare (Acerra *et al.*, 2009, Ministry of Finance, 2010, Trani *et al.*, 2010).

The burden and deaths due to malaria have declined as a result of recent intervention strategies as a result of increased funding from the Global Fund (The Global Fund, 2013a). This includes increasing use of ITNs as protection against malaria. The burden of Tuberculosis however remains one the highest in the Eastern Mediterranean Region (Mauch *et al.*, 2010, Delawer *et al.*, 2013). These challenges may be improved by better targeting of external aid, reducing poverty, improving primary care and quality of healthcare.

3.1.5 Review of healthcare access in Afghanistan

Access and utilisation of health services continues to be a subject of debate in high (Department of Health, 2009, Laudicella *et al.*, 2012) and low income countries (Jacobs *et al.*, 2012, Singh *et al.*, 2013). The situation is even more severe in Afghanistan following decades of conflict that has led to destruction of basic infrastructure (Ameli and Newbrander, 2008, The PLoS Medicine Editors, 2011, The World Bank, 2012). The Basic Package for Health Services (BPHS) was introduced in 2004 to expand coverage (Edward *et al.*, 2011) using contracting mechanisms through NGOs. Many NGO based health facilities in Afghanistan are supported by the World Bank, the USAID and the European Union (Sabri *et al.*, 2007, Ameli and Newbrander, 2008). Contracting has been used previously to improve healthcare delivery in other settings such as: Asia (Pakistan, India, Cambodia and Bangladesh); Africa (Senegal and Madagascar) and Central and South America (Bolivia, Haiti and Guatemala) (Loevinsohn and Harding, 2005, Liu *et al.*, 2008).

Previous studies on healthcare access in Afghanistan were based on *demand* factors such as the perceived need, cost and quality of care (Edward *et al.*, 2011). Examples of *demand*-based studies in Afghanistan include an investigation on the use of preventive and curative services (Newbrander *et al.*, 2007, Ameli and Newbrander, 2008, Steinhardt *et al.*, 2011), which suggested presence of female health worker as an attractive factor influencing use. A survey conducted in 2004 after the introduction of BPHS in 2004 found low utilisation of health services (25%) amongst vulnerable population (the disabled) (Trani *et al.*, 2010, Trani and Barbou-des-Courieres, 2012) and identified cultural perception , availability and cost as important factors. A different study examining affordability showed an increase use of health

facilities where user fee had been abolished in 2008 (Steinhardt *et al.*, 2011). These studies although were important in identifying other forms of access did not measure spatial access at a national level. A general health survey conducted in few regions in 2006 estimated that 60% of the population were within 2-hour walking time to the nearest facility (Ministry of Public Health, 2008a).

3.2 Estimating the spatial coverage and utilisation of public healthcare facilities for treatment of fever in Afghanistan

3.2.1 Data

3.2.1.1 Assembly of spatial health facility database in Afghanistan

A health facility database was obtained from the national health management information system (HMIS) through the Afghanistan National Malaria and Leishmaniasis Control Programme (NMLCP). Routine health facility assessments are carried out by the Ministry of Public Health (MoPH) through the subsequent National Health Service Performance Assessment (NHSPA) surveys (Waldman *et al.*, 2006, Peters *et al.*, 2007). The surveys aim to assess performance of the BPHS. Health facilities were classified into three broad categories that combined the BPHS and EPHS as: basic facilities made up of HPs, clinics and MCHs; health centres (the comprehensive health centres) and a third category representing tertiary facilities, the hospitals (Figure 3.2 below). The spatial coordinates were established using a non-differential handheld global positioning systems (GPS) receiver during the assessment surveys or in some cases using a database of placenames.

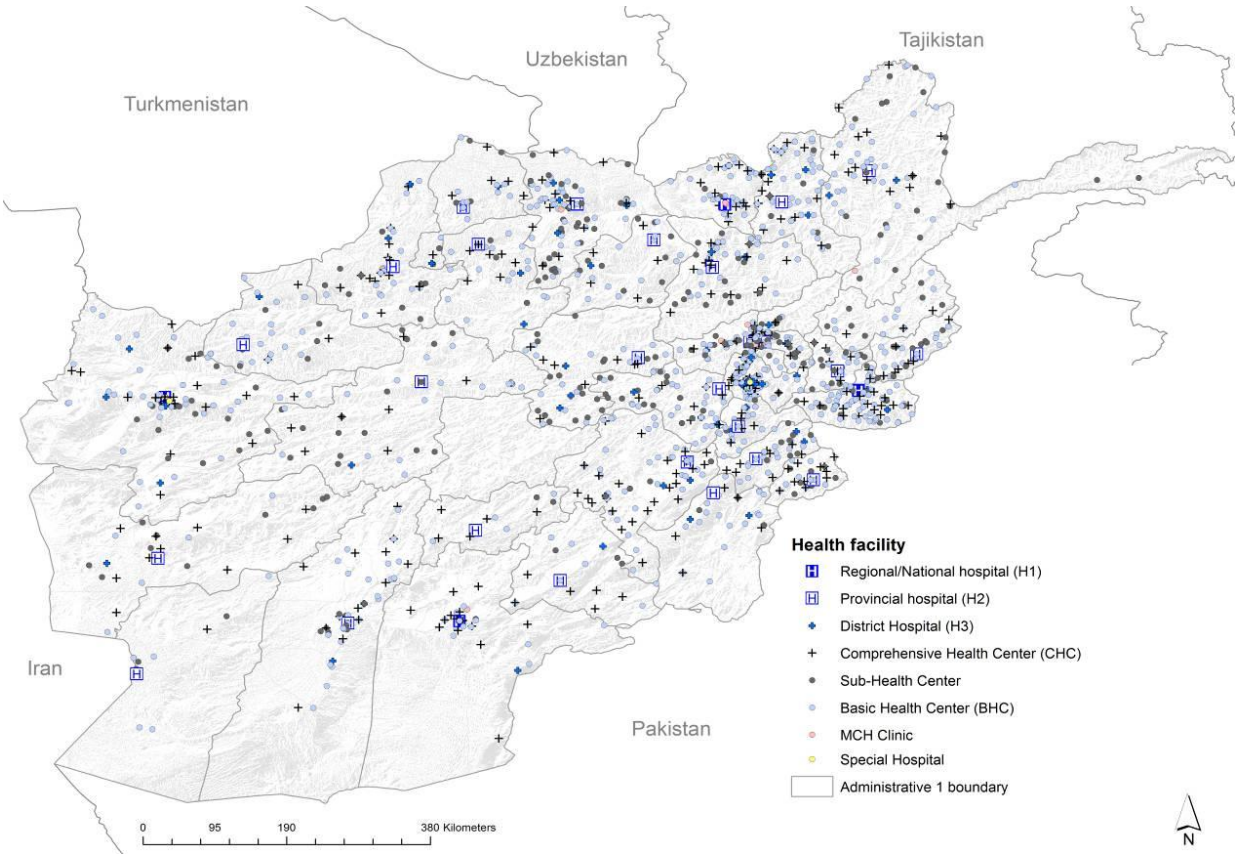

Figure 3.2: Distribution of public-based health facilities in Afghanistan
Distribution of active health facilities in Afghanistan by comprising of: basic health centres (n=773), comprehensive health centres (n=354), sub-health centres (n=344), district hospitals (n=95), special emergency hospitals (n=30), provincial hospitals (n=29) and regional or national referral hospitals (n= 5).

3.2.1.2 Population

An Afghan gridded population surface was obtained from Asiapop at 0.000833^{o} x 0.000833^{o} spatial resolution (Gaughan *et al.*, 2013) (Figure 3.3). In brief, a fine spatial resolution population map was produced from a combination of settlement, land use or land cover and basic infrastructure data (Gaughan *et al.*, 2013). The method of mapping population distribution involved disaggregating census counts in areal units via weights derived from land cover and land use data (Bhaduri *et al.*, 2007, Tatem *et al.*, 2007, Linard *et al.*, 2012). For each region, population density was calculated based on land cover class, rurality, census counts and an

adjustment factor that ensures the total population sums to a known value. Urban extents were derived from the Global Rural and Urban Mapping (GRUMP) project (Center for International Earth Science Information Network (CIESIN), 2004, Tatem *et al.*, 2007), but adjusted for urban extent based on the MDA Geocover (MDA, 2013), and reclassified to conform to the UN Land Cover Classification system (UNLCC) (FAO, 2000, Tatem *et al.*, 2012). An updated finer spatial resolution land cover map was created to include detailed information on roads, settlements and inhabitable areas. The resulting classes were then used to disaggregate count data based on a weight value assigned to each land cover pixel. Population densities were derived as a ratio of the known population count and the total habitable land area normalized by an adjustment factor (Linard *et al.*, 2012, Gaughan *et al.*, 2013). The resulting 2010 national population map was projected using the United Nations' (UN) Population Division national inter-censual growth rates (UN Population Division, 2013). For measuring utilisation rates, 2011 population estimates were used and these were back projected to 2009 when estimating population at risk of *P. falciparum* and *P. vivax*.

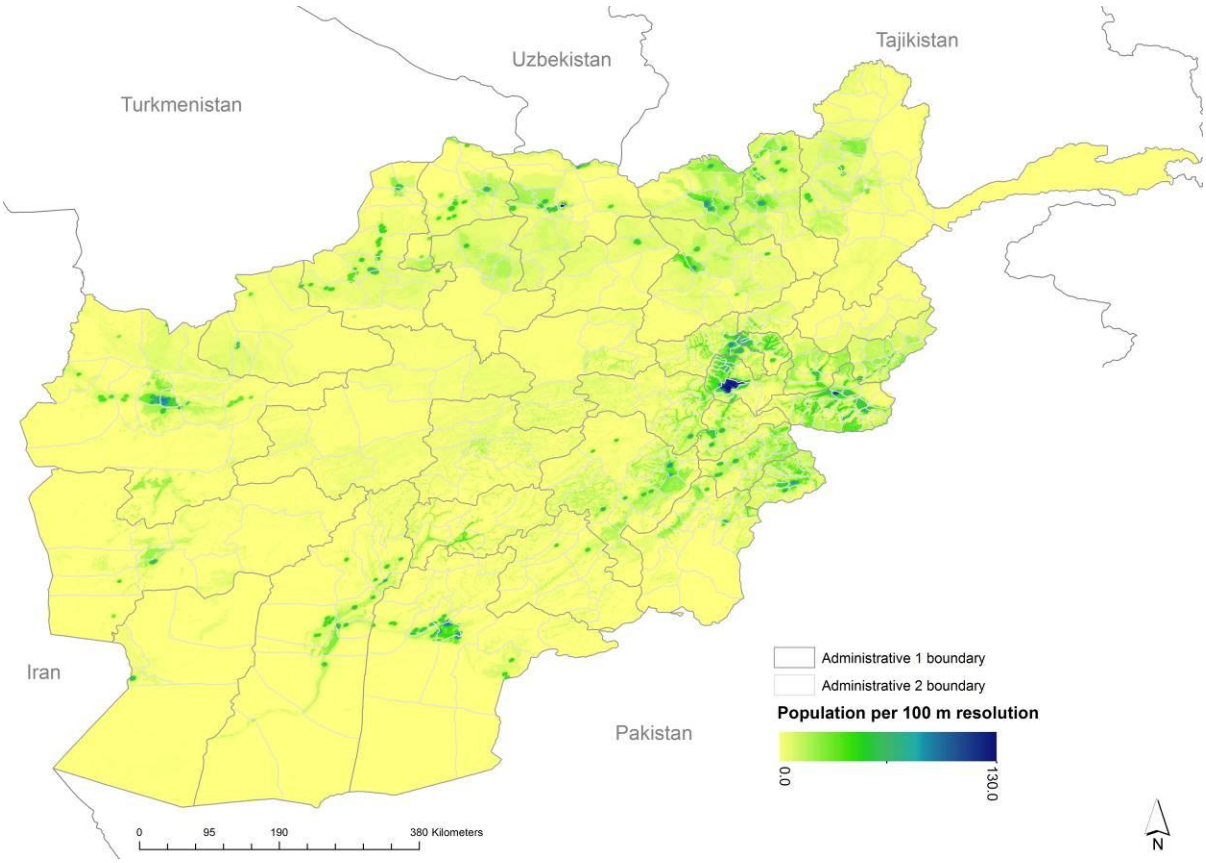

Figure 3.3: Population density map of Afghanistan for 2010
Population distribution in Afghanistan from Asia pop at approximately 100 m x 100 m spatial resolution.

3.2.1.3 National household survey data for analysing malaria treatment seeking behaviour

Data for modelling healthcare utilisation for treatment of fever was obtained from the MIS carried out between September and October 2011 (n = 15,442 individuals) (Ministry of Public Health, 2012). The MIS was conducted in 21 provinces, across different malaria strata in Afghanistan, but excluded the southern regions for security reasons. A multi-stage probability sampling design was adopted in line with other MIS surveys conducted in other low income countries (RBM-MERG, 2008) where clusters or villages were first selected randomly in a district via probability sampling and then households within the selected clusters were sampled randomly (Ministry of Public Health, 2012). Through self-reporting (except for children below 15 years), two-week fever prevalence and treatment seeking behaviour were recorded for all the

respondents that had been invited to participate. The Geographic location of all clusters was recorded using a hand-held GPS receiver (Garmin International Inc., Olathe, KS, USA).

3.2.1.4 National level GIS data

Ancillary GIS data used for healthcare access and utilisation analysis included roads and rivers obtained from an online resource (MapCruzin, 2013) distributed through the GNU General Public License (GPL) (Free Software Foundation, 2007) and constituted initially by AIMS (http://www.aims.org.af/). Elevation data were obtained from the Advanced Spaceborne Thermal Emission and Reflection Radiometer-Global Digital Elevation Model (ASTER-GDEM) (Huggel *et al.*, 2008) while the land use-land cover layer was downloaded from the MEdium Resolution Imaging Spectrometer (MERIS) GlobCover product (ESA, 2010).

3.2.2 Developing surface of travel time and probability of attendance

A combination of land cover, elevation, road and river data layers was used to generate a cost surface of travel times between public health facilities and population locations in AccessMod (version 3.0) (Ray and Ebener, 2008). A gridded (raster) surface was generated based on cumulative travel speed between patient origins (households) and destinations (public health facilities) at 1 km by 1 km spatial resolution. Travel speeds (Table 3.1) were assigned to each land cover pixel based on recommendations from previous studies, for example, by using motorized transport on tarmac roads and walking on bare land (Ray and Ebener, 2008, Alegana *et al.*, 2012, Huerta Munoz and Kallestal, 2012). The derived friction raster was used to extract travel times between cluster locations and health facilities using ArcGIS (ESRI, Redlands, CA). Thus, all the individuals in a cluster were assigned an average community travel time to the nearest public health facility. Since all age-cohort data was available, a preliminary analysis that

included age (regression coefficient -0.448, 95%CI: -0.904 – 0.003, p=0.07), gender (regression coefficient -0.003, 95%CI: -0.015 – 0.007, p=0.5) and derived travel time (regression coefficient -0.170, 95%CI: -0.3000 – -0.043, p<0.001) suggested that the latter had a larger effect on the use of public health facilities. Thus, age and sex were not considered subsequently when modelling the distance decay curve. A three parameter logistic regression models of the form $Y = \kappa / 1 + e^{(\alpha - x)/\beta}$ (Pinheiro and Bates, 2002) were then fitted to estimate health facility attendance based on the extracted theoretical travel times, assuming utilisation was for the nearest health facility. Four models representing the universe of all public health facilities, hospitals (District, regional and national referral), health centres (sub-health centres and CHC) and basic facilities (MCHs, clinics and HPs) were fitted separately to the survey data in the R statistical software (R Development Core Team, 2010). The model coefficients: α an asymptote factor at an inflection point β, a distance decay parameter and κ, a limiting function on the y-axis that measured the probability of attendance when distance was zero, were recorded along with the goodness-of-fit statistic, *t*, and the *p*-values. A gridded surface of probability of attendance was derived by applying the logistic model to the gridded cost surface in ArcGIS (ESRI, Redlands, CA, version 10).

Table 3.1: Input data for analysis of utilisation of public health facilities in Afghanistan
Description of various data and their sources used as inputs in calculating travel time to the active public health facilities in Afghanistan. The assumed travel speeds for each input feature are also shown

Map Layer	Description	Classification	Speed (km/h)	Mode[1]
Land use/ land cover	Spatial representation of all different land use and land cover types. Two land cover grids were processed (1) a basic land cover grid (2) a combined grid that incorporates roads and rivers with the same resolution as the DEM	Irrigated, rain fed, mosaic or vegetated croplands	3.0	Walking
		Open or closed broadleaved, needle leaved, deciduous or evergreen tree cover	3.0	Walking
		open or closed mixed broadleaved forest/tree cover	1.5	Walking
		Mosaic, closed to open grassland/shrubland	1.5	Walking
		Sparse Vegetation	1.5	Walking
		Open or closed broadleaved regularly flooded	0.5	Walking
		Artificial/urban areas	30.0	None
		Bare areas/desert	1.0	Walking
		Ice/ permanent snow	0	None

Roads	Classified into five categories; class A (highways), class B (secondary roads), tertiary Class C and Class D roads as well as street level urban roads. Each road class was assigned a different speed limit.	Class A roads	60.0	Motorised
		Class B roads	30.0	Motorised
		Class C roads	10.0	Cycling
		Class D roads	4.0	Walking
		Street level roads in urban areas	20.0	None
Rivers	GIS layer representing barrier to movement. Only major rivers were used to reduce the complexity of running the algorithms	NA[3]	0	NA[2]
Digital elevation model	Altitude values that are used in anisotropic calculation; Original DEM 30 m ASTER grid; resampled to 1 km pixel size	Degree of Slope (< 0.5^o)	4.88	Walking
		Degree of Slope (5.0^o)	3.71	Walking
		Degree of Slope (10.0^o)	2.71	Walking
		Degree of Slope (20.0^o)	1.41	Walking
		Degree of Slope (30.0^o)	0.66	Walking

1. Assumed mode of travel to health facility, as either walking on foot, cycling, using motorise transport as on roads or a combination of the different modes. Anisotropic movement for walking based on Tobler's equation, (V=6*exp(-3.5abs[Tan(slope in degrees/57.296) + 0.05]) (Tobler, 1993) where V is the speed with slope derived from DEM or for cycling (Walter, 2008), was applied for traversing across a pixel. For example, on a flat terrain, the walking speed is 5.0 km hr^{-1}.
2. NA is an abbreviation for ' Not Applicable'

3.2.3 Developing health facility catchments used in the analysis of malaria incidence and assessing spatial coverage

Catchment areas were derived for various public health facilities using the '*cost allocation*' Spatial Analysis tool in ArcGIS (ESRI, Redlands, CA). The distance decay model based on the universal facility list was selected to zone catchment areas while limiting the maximum travel time to 2 hours. This time limit has been used to measure physical accessibility in Afghan BPHS reports and in related research (Loevinsohn B, 2008, Acerra *et al.*, 2009, Ministry of Public Health, 2010a). Population counts in various catchments were extracted using the hard catchment boundaries and multiplied by the probability of attendance for fever treatment to derive the proportion of population likely to attend a public health facility. These counts were subsequently used in the analysis of incidence. Similarly, the population outside the two hour threshold was estimated. The rate of fever reported from the MIS at province level was multiplied by the population to generate a fever burden map. The number of fever cases within each catchment was estimated based on the fever burden. Lastly, the number of fever cases likely to attend a

public health facility was calculated by multiplying the estimated number of cases by the probability of attendance.

3.2.4 Results of analysis of spatial coverage and utilisation of public health facilities in Afghanistan

3.2.4.1 Fever prevalence in 2011 and treatment seeking behaviour

In total 1,629 public health facilities were assembled (Figure 3.2 Page 147) and these were distributed across the 34 provinces in Afghanistan. Only Kabul and Nangahar provinces had more than 100 public health facilities while no facilities were recorded for Day Kundi province in the central highlands. The number of basic health facilities (n=754) was similar to the number of health centres (n=698). 11,307 people in 185 clusters were interviewed in the 2008 MIS. Overall, fever prevalence was estimated as 3.4% (95%CI: 3.1 – 3.7) in all age populations in 2008, of which 59.8% (95%CI: 54.9 – 64.7) sought treatment. Of those who sought treatment, 29.0% (95%CI: 24.4 – 33.6) used the public sector compared to the 30.8% (95%CI: 26.2 – 35.5) who used other sectors including the private health facilities. Thus, fever treatment in the public and other sectors in 2008 was low generally. From the estimated fever cases, Hirat and Nangahar provinces had a burden greater than 30.0%.

3.2.4.2 Distance decay model fitting results

Figure 3.4 (below) shows the distance decay curves by facility type for utilization of public health facilities with increasing travel time on the x-axis for Afghanistan. Table 3.2 lists the various parameters of the fitted distance-decay models along with their respective p-values. In all the four models, treatment-seeking behaviour decayed rapidly after 90 minutes. The coefficients of the various decay curves were all significant with $p<0.001$ and the sum of squared residuals

indicating a good model fit with the observed fever treatment patterns. Treatment seeking patterns for the basic facilities tended to be similar to hospital utilisation patterns. Health centres pattern was close to the universal "all" model (Figure 3.4). In addition, the maximum probability of use was higher for HCs (0.789) compared to the other facility types. The universal model was used subsequently to delineate catchment areas to estimate coverage. Secondly, this model had lower values of sum of squared residuals and standard error as shown in Table 3.2.

Table 3.2: Logistic Model parameters by hospital type and for the universe of all facilities in Afghanistan

		Model Parameter					
	Number of facilities	α	β	κ	**p-value (all parameters)**	**Residual standard error**	**sum of squared residuals**
All	1,581	3.1906	-0.2908	0.8681	<0.000	0.0021	0.0015
Hospitals	129	3.8439	-0.3139	0.8702	<0.000	0.0218	0.0176
Health Centres	698	2.8896	-0.2863	0.8768	<0.000	0.0182	0.0121
Basic	754	2.8066	-0.2843	0.8805	<0.000	0.0175	0.0113

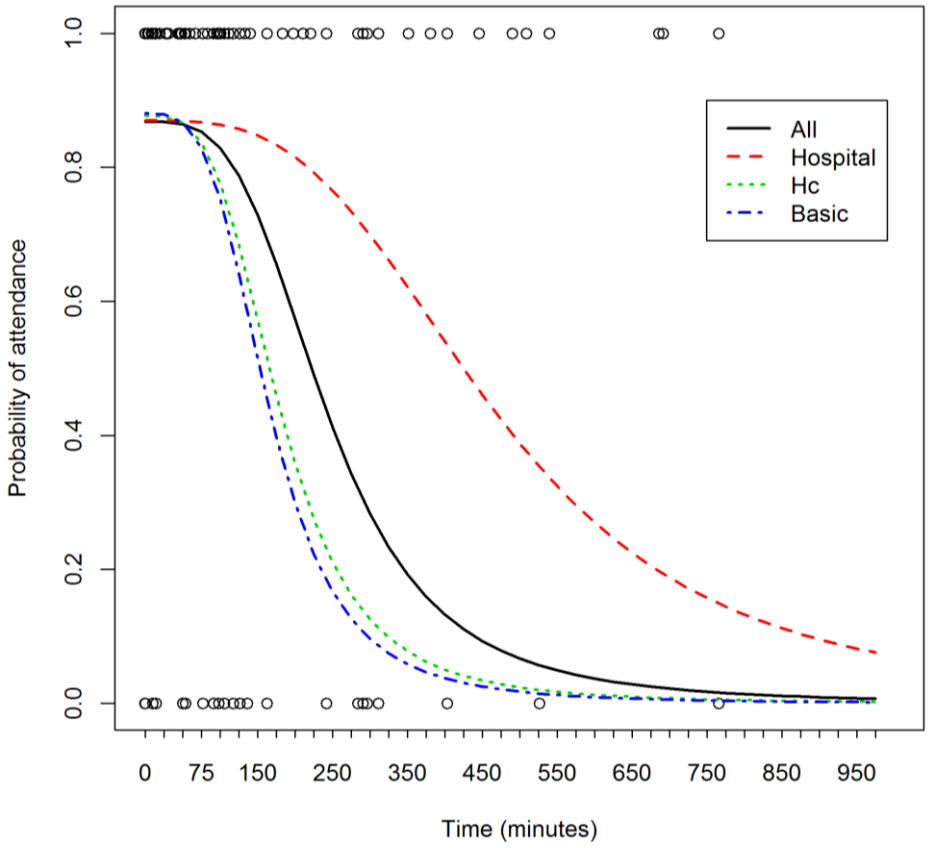

Figure 3.4: Distance decay curves for public health facility use in Afghanistan

Distance decay curves for the MIS survey (2011) showing probability of public health facility use for fever treatment in Afghanistan (y-axis) against increasing travel times (x-axis). The model was run using log-transformed travel time (x-axis) then back-transformed for presentation purposes. The attendance pattern (1 = attendance and 0 = non-attendance) is also superimposed on the decay curve.

3.2.4.3 Probability of attendance for fever treatment

Figure 3.5 shows the gridded probabilistic surface of attendance for fever treatment to all health facilities at 1 x 1 km spatial resolution. This gridded surface had been used to delineate the health facility catchments as shown in Figure 3.6. Of the estimated population (29.8 million) in 2008, 25,574,396 (85.8%) were estimated to be within a public health facility catchment (Table 3.3 on Page 161 and Figure 3.6). Further, 15,987,842 (62.5%) of those within a public health facility catchment were within 30 minutes and 8,772,663 (34.3%) were within distances where the probability of attendance was ≥60%. 11,119,031 (43.5%) had much lower probabilities of attendance (≤0.20). From the modelled fever burden based on the distance decay curves, the estimated national fever burden was 3,440 cases assuming a single episode of fever in October 2008 of which 3,146 (91.5%) were within 2 hours' travel time to the nearest public health facility (Table 3.3). Finally, 832 (24.1%) of these fever cases were likely to have been treated in the public sector based on the distance decay model.

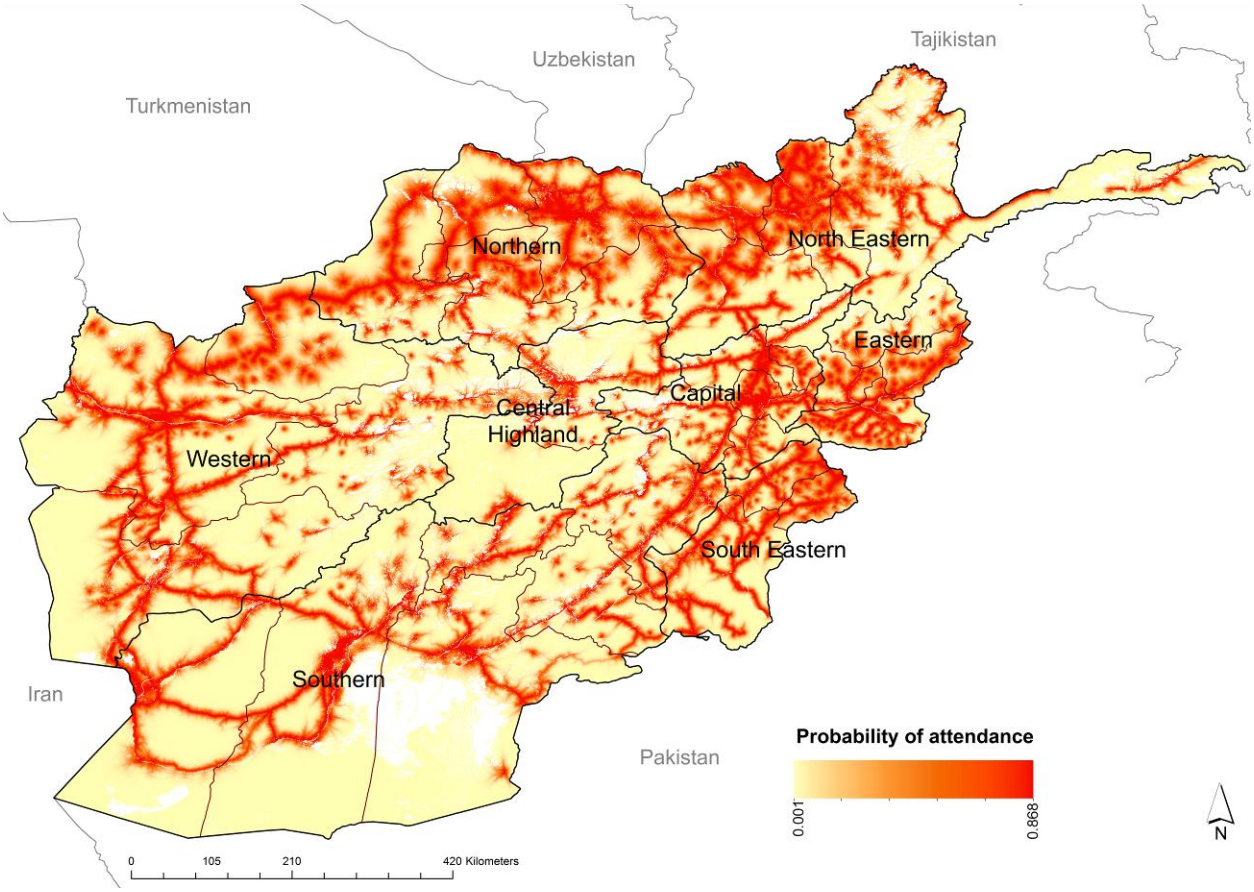

Figure 3.5 Probability of health facility use for fever treatment in Afghanistan
Map of probability of attendance for treatment of fever for all age population in Afghanistan at the nearest public health facility based on the 2008 MIS survey. The lowest probability was 0.001 and the highest probability was 0.768.

3.3 Discussion on coverage and utilisation of healthcare facilities for fever treatment

This section assessed the coverage and utilisation of the existing active public health facility in Afghanistan. First, health facility catchments based on a 2-hour cut-off to nearest health facility were derived and used to assess coverage in Afghanistan (Table 3.3). Secondly, a travel time metric was derived and used to derive a surface of probability of attendance for fever treatment (Figure 3.5). The main findings suggest that majority of existing active health facilities were located close to the population. 1,538 (94.4%) of health facilities were within 30 minutes of travel and 1,292 (79.3%) located in distances where probability of utilisation was ≥0.60.

Findings from the Afghanistan National Health Resource Assessment (NHRA) survey carried out in 2002 showed that most health facilities were located close to roads suggesting a shorter travel time if motorised transport is used (Ministry of public Health, 2002). A large proportion of population (85.8%) was also estimated to live within the derived health facility catchment boundaries, thus, suggesting that a substantial majority of population experience good coverage. The results were also similar to those reported by the national Afghanistan Health Survey (AHS) of 2006 which indicated that over 60% of the rural population was within 2 hours of a health facility (Ministry of Public Health, 2008a). This study included urban areas and modelled travel by including motorized and non-motorized modes.

From the MIS survey, the reported two-week-period fever prevalence was low at 6.4% (95%CI: 6.0 – 6.8) and reported rates of treatment in the public sector were 44.7% (95%CI: 38.8 – 50.6), but not different from overall use of the private sector which was 42.2% (95%CI: 36.3 – 48.1). In this study, the distance decay model predicted 44.3% of the estimated fever cases in 2011 were likely to have been treated in the public sector. Thus, the predicted rate of utilisation in the public sector was similar to the observed utilisation rates. The small differences could be attributed to factors such as socio-economic status, gender, over-reporting and sampling in the MIS or the propagation of error from different data sources. Moreover, the severity of fever condition is likely to influence the decision to seek medical treatment in addition to above factors. For example, in the BPHS community health workers at community level treat mild conditions including fever (Acerra *et al.*, 2009, Ministry of Public Health, 2010a). In the 2006 AHS survey, illness severity and distance were reported as the top two factors explaining failure to seek medical treatment. A study among the nomadic Fulani population in Nigeria reported

similar findings where treatment was either delayed or sought based on severity (Akogun *et al.*, 2012).

From the universal distance decay model, utilisation of public health facilities declined rapidly after about 120 minutes regardless of the facility type (Figure 3.4). The 2-hour cut-off (taken from policy documents (Ministry of public Health, 2002, 2008a)) used previously in assessing coverage appears to correspond roughly to a mean rate of distance decay observed in Figure 3.4 and may, therefore, be a reasonable choice. In addition, the rate of decay based on the derived travel time, was rapid for basic health facilities compared to that from previous studies (Noor *et al.*, 2006, Alegana *et al.*, 2012) which could reflect a reluctance to travel longer distances for security reasons (Acerra *et al.*, 2009), cost as highlighted by Ameli and Newbrander (2008) or due to gender and cultural norms. A study conducted in Kabul by Mashal *et al.*, (2008) identify factors such as mothers' lack of autonomy in the household and level of education contributed to a poor health seeking behaviour resulting in poor health outcomes (Mashal *et al.*, 2008).

From the fitted logistic models, the use of basic facilities was very similar to the health centres but different for the hospitals. The modelled pattern suggested that population is likely to travel greater distances for hospital based services and shorter distances (travel times) for basic services. In addition, slightly higher probabilities of use at zero distance were observed for basic health facilities κ(0.881, p<0.001) and for the health centres κ(0.877, p<0.001) (Figure 3.4 and Table 3.2). This phenomenon could be attributed to proximity of the basic health facilities to population and tendency to use basic facilities for uncomplicated illness (Belay, 2010). Variation in travel modes could also affect pattern of use as well as the unobserved effects such as

perception of quality of services (Trani *et al.*, 2010, Trani and Barbou-des-Courieres, 2012). Empirical data to test these assumptions are not available readily at the national level and analysis was restricted to interaction with the closest public health facility.

Other sources of errors remain. These include the exclusion of factors explaining health facility utilisation such as household income, healthcare costs, wealth and cultural preferences (Joseph and Phillips, 1984, Akin and Hutchinson, 1999, Tanser *et al.*, 2001, Leonard *et al.*, 2002, Noor *et al.*, 2006, Gething *et al.*, 2007, Das *et al.*, 2013). Access, as a multidimensional concept, is affected by these factors. Although the household data on reported rates of fever and treatment seeking behaviour were representative for all ages, the inclusion of the above factors could alter the probability of using the nearest health facility. Since this study has focused on patients' interactions with the public health sector, the inclusion of private sector facilities may well alter the patterns of use observed here. Additionally, the different modes of transport such as walking or use of motorized transport could differ from those assumed in the model. Data on actual mode of transport used while travelling to a health facility are rarely available. Further, the calculated fever burden was based on a survey-derived regional period prevalence rate, when point prevalence may vary at community or facility catchment level (Youssef *et al.*, 2010, Elmardi *et al.*, 2011). This fine resolution point prevalence may be different significantly from the regional mean. The study did not account for the effects of conflict while modelling utilisation. It was assumed that conflict effects were inherent in respondent answers at the survey stage. Such effects required a time series of data on incidences of conflict to identify stable hotspot areas that was beyond the scope of this study. Future studies could investigate the probability of conflict as an adjustment variable in determining probability of health facility use in fragile provinces. A by

product of this study was an estimation of population in health facility catchment areas to enable estimation of incidence outlined in the next section.

Table 3.3: Population within health facility catchments in Afghanistan.
Estimated population data for 2011 by province and modelled treatment seeking for fever at the nearest public health facility

	Hospitals (Provincial/Regional/District)	Health Centers (Comprehensive and sub health Centers)	Basic Health facility (HPs/Clinics/ MCH)	Other Facilities	Total health facilities	Estimated Population in 2008 (All ages)	Population (percentage) in PHF[1] catchments	Estimated fever burden for 2011 from MIS prevalence	Number of fever cases likely to attend a PHF[1] (%)
Probability of attendance									
≤0.20	15	93	117	2	227	13,649,346	9,071,360	124,967	5,931 (4.7)
>0.20 -< 0.50	5	26	45	4	80	3,267,810	3,267,810	39,725	13,124 (33)
>0.50 -< 0.60	5	9	15	1	30	2,287,390	2,287,390	28,231	16,628 (58.9)
≥ 0.60	104	570	577	41	1292	13,166,100	13,158,313	134,595	109,402 (81.3)
Travel time									
≤ 30.00 minutes	121	661	708	48	1538	17,890,000	17,890,000	192,425	136,890 (71.1)
>30.00 min - <1.00 hours	7	27	28	0	62	23,306,950	23,306,950	66,141	7,495 (11.3)
>1.00 hours -< 2.00 hours	1	9	16	0	26	27,792,610	27,792,610	41,906	672 (1.6)
≥ 2.00 hours	0	1	2	0	3	4,578,036	-	27,045	28 (0.1)
Province									
Badakhshan	3	35	34	1	73	1,351,920	974,008	11,324	1,749 (15.4)
Badghis	2	10	23	0	35	584,251	488,069	10,897	2,222 (20.3)
Baghlan	3	30	21	0	54	1,101,920	925,927	5,399	1,754 (32.4)
Balkh	7	41	46	3	97	1,493,720	1,456,970	20,686	12,427 (60.0)
Bamyan	4	27	18	1	50	534,916	288,573	34	3 (9.8)
Day Kundi	-	-	-	-	-	578,854	88,678	20,689	595 (2.8)
Farah	2	21	6	1	30	615,616	402,153	0	0
Faryab	3	26	22	0	51	1,153,790	1,048,770	25	4 (17.4)
Ghazni	4	31	35	1	71	1,527,840	1,244,250	19	1 (6.2)
Ghor	3	25	21	0	49	814,963	375,385	27	0
Hilmand	4	25	27	1	57	1,100,790	805,764	6	0
Hirat	5	42	40	1	88	2,221,470	1,973,640	12	1 (4.2)
Jawzjan	4	13	16	0	33	665,411	644,905	8,974	4,036 (44.9)
Kabul	41	39	61	25	166	4,872,250	4,840,000	20,601	16,365 (79.4)
Kandahar	2	18	18	2	40	1,460,940	1,168,280	0	0
Kapisa	2	17	16	2	37	488,739	466,414	33	8 (24.6)
Khost	1	19	9	1	30	697,456	696,841	3,535	1,750 (49.4)

Kunar	1	18	13	0	32	547,826	532,095	42,121	14,645 (34.7)
Kunduz	2	17	32	2	53	1,203,720	1,177,470	13,974	7,306 (52.2)
Laghman	1	21	16	0	38	535,115	504,745	9,610	3,776 (39.2)
Logar	3	11	20	0	34	479,402	432,285	21	2 (8.8)
Nangarhar	6	33	63	1	103	1,828,820	1,751,860	119,758	63,249 (52.8)
Nimroz	1	9	5	1	16	198,911	139,575	0	0
Nuristan	0	12	12	0	24	187,256	136,591	77	8 (10.0)
Paktika	3	10	18	0	31	538,622	475,151	6	1 (15.2)
Paktya	3	17	17	0	37	683,023	675,139	4,851	2,116 (43.6)
Panjshir	1	6	4	0	11	150,659	88,775	4	0
Parwan	2	31	32	5	70	884,168	774,160	6,792	3,545 (52.1)
Samangan	2	15	13	0	30	467,796	391,992	16	2 (15.3)
Sari Pul	3	14	16	0	33	723,273	632,692	8,584	2,288 (26.6)
Takhar	4	21	38	0	63	1,194,430	1,149,240	7,276	4,031 (55.4)
Uruzgan	1	7	6	0	14	428,158	296,071	11	1 (6.0)
Wardak	4	28	27	0	59	683,159	518,201	12,157	3,200 (26.3)
Zabul	2	9	9	0	20	371,462	220,204	0	0
Total	**129**	**698**	**754**	**48**	**1629**	**32,370,646**	**27,784,873[2]**	**327,517**	**145085(44.3)**

1. PHF is abbreviation for Public Health Facility
2. The total number of people in the catchment was lower than the overall estimated population because some population were outside the catchment boundary, thus not covering 100% population and not entire population is likely to use a PHF.

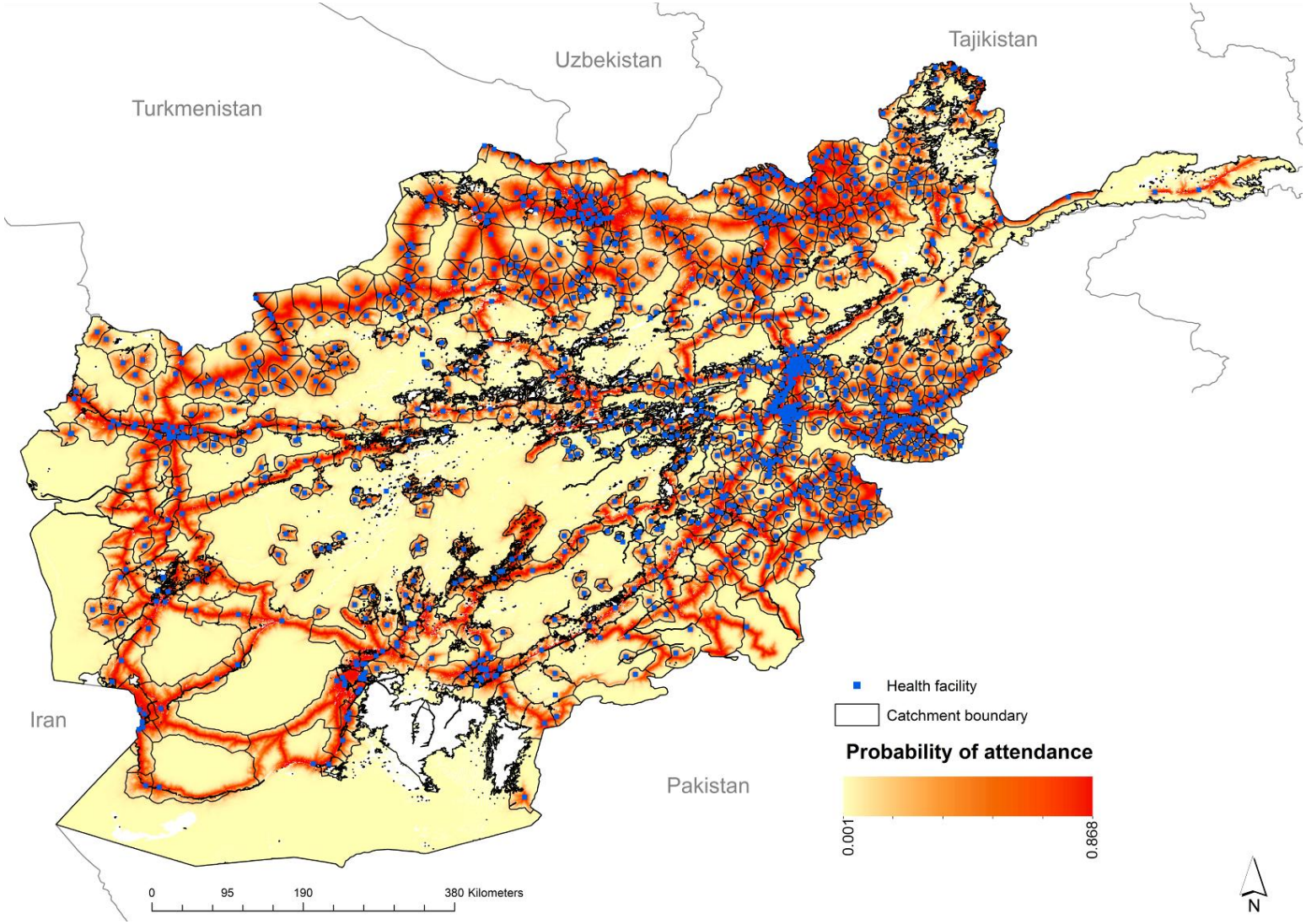

Figure 3.6: Delineated public-based health facility catchments

Map of public health facility catchment areas for Afghanistan derived from modelled travel time superimposed on the probability of attendance of nearest public health facility when sick with fever

3.4 Analysis of incidence of *P. vivax* and *P. falciparum* in Afghanistan

3.4.1 Assembly of HMIS data

The Afghanistan Health Management Information Systems (HMIS) records the number and types of illness including malaria. All health facilities are required to report the number of malaria cases on a monthly basis including the suspected or cases confirmed through a laboratory test (RDT or Microscopy). In general, the use of parasitological diagnosis at basic facilities was low during the study period (2006-2009) (Ministry of Public Health, 2010a). Complicated or severe malaria cases are referred to higher-level facilities (comprehensive health centres, district, provincial and regional hospitals) where parasitological diagnosis is available.

Table 3.4 shows a summary of the assembled malaria case data for *P. falciparum* and *P. vivax* for the 48 months from January 2006 to December 2009. Data were based on outpatient cases observed at each facility. The slide positivity rate was used to adjust the suspected cases similar to Namibia analysis (Section 2.3.1). Health facility utilisation rates and the rate of reporting (i.e. the number of received reports divided by the total expected) was used to adjust the catchment population (section 3.2.3). This was important since not all individuals use the public sector and there was sporadic reporting by health facility to the HMIS. Adjustment for slide positivity was necessary to avoid underestimating incidence (if suspected cases are ignored) or overestimating incidence (where true cases are treated as a summation of clinical and confirmed case while ignoring the SPR at the facility). Parasitological diagnosis (microscopy or RDTs) was conducted at higher-tier facilities (hospitals and health centres) where laboratory facilities exist while clinical diagnosis was predominantly used at lower-level facilities such as health posts (Supplementary Information (SI)). No cases were examined or reported for 228 facilities which were treated as missing data while data for mobile units (*n=93*) were omitted from the final

analysis since they serve as outreach centres from major facilities. The missing spatial and temporal structures of data were imputed as '*NAs*' and predictions made at missing locations

$$SPR = \frac{Confirmed \quad cases}{Total \quad number \quad exa\,\min ed}$$

$$TMC = Confirmed \quad Case \; + \; \left(Suspected \quad cases \; \times \; SPR\right)$$

Overall, Table 3.4 suggests a decline in slide positivity rates from 2006 to 2009. For example slide positivity for hospitals for P*f* was 1.7% in 2006 compared to 0.9% in 2009 and similarly for P*v* (10.1% and 5.0% respectively). A large proportion of reported cases are based on clinical diagnosis.

Table 3.4: Summary of assembled malaria case data
Number of malaria cases (*Plasmodium falciparum* (P*f*), *Plasmodium vivax* (P*v*) and clinical) assembled by year and average positivity rates

	Malaria cases (Slide positivity Rate %)											
	2006			**2007**			**2008**			**2009**		
Type of Facility	P*f*	P*v*	Clinical	P*f*	P*v*	Clinical	P*f*	P*v*	Clinical	P*f*	P*v*	Clinical
Provincial/Regional Hospitals	1,122 (1.7)	6,800 (10.1)	10,866	1,318 (1.5)	8,679 (10.0)	13,636	177 (0.4)	3,658 (8.0)	5,964	950 (0.9)	5,437 (5.0)	8,597
District Hospital	562 (0.7)	10,270 (13.4)	19,471	590 (0.7)	10,258 (13.0)	23,038	112 (0.3)	2,273 (6.6)	15,504	508 (0.6)	8,727 (10.8)	18,783
Comprehensive Health Center (CHC)	2,481 (1.3)	38,040 (19.6)	129,101	2,474 (1.1)	37,134 (16.8)	141,066	571 (0.4)	17,820 (13.5)	107,277	1,527 (0.8)	26,295 (13.8)	108,238
Sub Health Center (SC)	3 (1.2)	46 (19.01)	407	1 (0.1)	65 (8.7)	1,432	627 (1.0)	9,541 (16.7)	46,675	17 (1.8)	76 (8.0)	22,292
Basic Health Center (BHC) (clinics/HPs/MCH)	1216 (1.3)	17,744 (18.3)	169,272	978 (0.8)	22,823 (19.8)	200,197	2,391 (1.0)	36,451 (14.4)	208,949	683 (0.5)	18,898 (14.5)	165,858
Total	**5,384 (1.2)**	**72,900 (16.7)**	**329,117**	**5361 (1.1)**	**78,959 (15.7)**	**379,369**	**3,878 (0.7)**	**69,743 (13.4)**	**384,369**	**3,685 (0.7)**	**59,433 (11.7)**	**323,768**

3.4.2 Assembly of environmental or ecological covariates for malaria risk

Malaria transmission in Afghanistan is constrained by altitude, temperature (Gething *et al.*, 2011a) and aridity (Guerra *et al.*, 2008, Guerra *et al.*, 2010) which affect parasite sporogony and vector development (Safi *et al.*, 2009b). Environmental covariates were assembled from remotely sensed data and extracted for each health facility. For the districts where no health facilities existed, a mean value of the covariate was used. This district-level average was also

used for health facilities where no geographic coordinates had been established (*n=108*). All the grid surfaces were resampled to a common spatial resolution (cell size 0.008333^{o} x 0.008333^{o}).

A temperature suitability index (TSI) (Gething *et al.*, 2011a) rather than the actual temperature values were used since TSI was modelled from long-term mean monthly temperature data from global climate data (WoldClim, http://www.worldclim.org/) (Hijmans *et al.*, 2005). TSI represented the optimum temperature suitability (from 0 (unsuitable) to 1 (most suitable)) for *P. falciparum* and *P. vivax* transmission based on the survival of malaria vectors and on the duration of sporogony (effect on the malaria parasite). An average monthly enhanced vegetation index (EVI) for the four year period was downloaded from the MODerate-resolution Imaging Spectroradiometer (MODIS) sensor imagery (available at http://modis.gsfc.nasa.gov/data/) as measure of vegetation cover (Hay *et al.*, 2006, Scharlemann *et al.*, 2008). The rate of precipitation was obtained from the Tropical Rainfall Measuring Mission (TRMM 3B43) (Huffman and Bolvin, 2011, NASA, 2011). TRMM 3B43 is archived at 0.25^{o} x 0.25^{o} spatial resolution and represent average rate of precipitation in $mmhr^{-1}$ produced after combining satellite data and information from ground stations (rain gauges) (Huffman, 1997, Huffman and Bolvin, 2011). The hourly rate was converted to a monthly average based on number of days per calendar month (i.e. by multiplying the gridded rate by 24 hours and by 30 days in a month).

3.4.3 Bayesian model specification for analysis of malaria incidence in Afghanistan

A zero-inflated Poisson model, similar to Namibia analysis, was implemented in Afghanistan based on observed cases for *P. falciparum* and *P. vivax*.Thus,

$$Y_{ij} \sim \begin{cases} 0 & with\ probability\ P_{ij} \\ Poisson(\mu_{ij}) & with\ probability\ (1-P_{ij}) \end{cases}$$

where $Poisson(\mu_{ij}) = \exp(-\mu_{ij})\,\mu_{ij}^{y_{ij}} / y_{ij}!$ with the $\log it(p_{ij}) = \log[\,p_{ij}/(1-p_{ij})]$.

The general log linear mixture model was:

$$\log \mu_i = \log(E_i) + \alpha + X_{ij}^T \beta_{ij} + f(s_u) + \varphi_i + f(t)$$

where E_i was the expected number of cases adjusted for utilisation at each facility i, α as the intercept, with the $f_{unstr}(\cdot)$ terms representing the unstructured spatial effects at facility, district and province levels different to Namibia where random effects were only at district and facility level. The extra randomisation aimed to improve the spatial smoothing of incidence. The $f_t(.)$ represented the seasonal or temporal effects. Thus, the likelihood of the data assuming similar covariates for the zero-state and the Poisson state as:

$$L_i(\theta, Z \mid y, \alpha) = \prod_{j=1}^{Ei} [\,p_{ij}\ \Pr(Y_{ij} = 0 \mid Z_{ij} = 1, \theta)]^{Z_{ij}} \times\ [(1-p_{ij})\Pr(Y_{ij} = y_{ij} \mid Z_{ij} = 0,\ \alpha,\ \theta)]^{1-Z_{ij}}$$

Inverse Gamma priors $IG(a,b)\,a = 1.0, b = 0.005$ were assigned to precision hyperparameters τ^2 for the unstructured effects components $\theta_{unstr} \sim N(0, \tau^2)$ at facility and district and province level. For the temporal trend, a first-order auto-regressive process, $\rho Y(s_i, t_{i-1})$ with the first term coming from a stationary distribution $N(0, \sigma_w^2 \sum)$ that depends on past values $x_i = \rho\ \ x_{i-1} + \xi_i \quad \xi_i \sim N(0, \tau^2)$ for $0 > \rho < 1$ was assigned (Sahu and Bakar, 2012). The conditional-autoregressive prior was used as a spatial effect at the district level. The conditional prior for neighbouring districts $(\varphi_j,\ j \neq i)$ was specified in a similar way to Namibia study as $(\varphi_i \sim N(\mu_{\phi i}, \sigma_{\phi i}^2)$ where $\mu_{\phi i} = \Sigma_{j \neq i} W_{ij} \varphi_j / \Sigma_{j \neq i} W_{ij})$; $\sigma_{\phi i}^2 = 1/\gamma_\varphi \Sigma_{j \neq i} W_{ij})$ (Bernardinelli *et al.*, 1997). The W_{ij} represented an adjacency matrix of weights assigned as $W_{ij} = 1$ for two neighbouring regions or $W_{ij} = 0$ otherwise. Flat priors $\pi(\theta)\ \ \alpha\ 1$ were assigned on the fixed covariate effects. The posterior taking into account of the priors $\pi\{\theta(\beta_i, \tau_i, \varphi, \rho)\}$,

$$f(\theta,\alpha,Z\mid y)=\prod_{i=1}^{m}\left\{\prod_{j=1}^{n_i}[P_{ij}^{Z_{ij}}[(1-P_{ij})\Pr(Y_{ij}=y_{ij}\mid Z_{ij}=0,\theta]^{1-Z_{ij}}\right.$$
$$\times\left|\Psi_\beta\right|^{-\frac{1}{2}}\exp\left((-\frac{1}{2}(\beta-\beta_0)'\Psi_\beta^{-1}(\beta-\beta_0)\right)$$
$$\times\sigma^{-a}\exp\left((\frac{-b}{2\sigma^2}\right)$$
$$\times\Pi(\theta)$$

3.4.4 Bayesian model specification of slide positivity rates at health facilities in Afghanistan

Here, the interest was to investigate the distribution of malaria species based on the slide positivity rates. A hierarchical Bayesian geostatistical model was subsequently used to predict SPR at fine spatial resolution (1 x 1 km). Thus, let $Z(s_i,t)$ denote the response (the SPR) for particular malaria species at facility S_i in a particular month, $i=1,.....,n;t=1,.....,T$. $Z(s_i,t)$ is a realisation from a binomial process i.e. the probability that a case is positive for a single blood test for either *P. vivax* or *P. falciparum* modelled separately, $y_i \mid p_i \sim Binomial\ (n_i\ ,\ p_i)$ with a logit link function $p(\eta)=\{\exp(\eta)/1+\exp(\eta)\}$. The probability of $Z(s_i,t)$ was taken to be independently distributed samples, $Z=(z_1,.....,z_N)$ as;

$$P(Z\mid p)=\prod_{1}^{k}P_k^{N_k}\qquad N_k=\sum_i\delta(x_i=k)$$

and a likelihood function as

$$f(x\mid p)=\frac{n}{x!(n-x)!}p^x(1-p)^{n-x}$$

The hierarchical model was decomposed into the observations with measurement error term,

$$Z(s_i,t)=\eta(s_i,t)+\varepsilon(s_i,t)\qquad i=1,...n,\ \ t=1,....T$$

where $\eta(s_i,t)$ represented the underlying spatio-temporal biological process with an error term $\varepsilon(s_i,t)\overset{i.i.d}{\sim}N(0,\sigma_\varepsilon^2\ I_n)$ with prior $\sigma_\varepsilon^2\sim IG(1,0.0005)$. The mean component was modelled as a

combination of the first order auto-regressive process $\rho Y(s_i, t_{i-1})$ and the covariates $x(s_i, t) = \{x_1(s_i, t).....x_p(s_i, t)\}^T$ with the first term coming from a stationary distribution $N(0, \sigma_w^2 \sum)$ that depends on past values for $0 > \rho < 1$ with a non stationary matérn covariance function (Sahu and Bakar, 2012). Thus,

$$\eta(s_i, t) = \rho Y_{t-1} + x(s_i, t)\ \beta + w(s_i, t)$$

An approximate range of 1/5 of the spatial domain was used while the initial values for the marginal variance parameter and the scaling parameter in the matérn model were set to 1 and 0.1 respectively. Flat priors were used for fixed parameters $\pi(\beta)\alpha 1$. The posterior is evaluated as the product of likelihood given all the model parameters.

$$\eta(s_i, t) \mid \beta, \tau, \phi_s, \rho_t \sim GMRF(\mu, \Sigma)$$

With Gaussian Markov Random Field (GMRF) used as a representation of the Gaussian Field evaluated using finite element methods (Lindgren *et al.*, 2011). The region of study was expanded by 100km at the border to reduce edge effects associated with Neumann boundaries in SPDE (Cameletti *et al.*, 2012, Lindgren, 2013).

3.4.5 Model choice and validation for incidence and SPR analysis

Four spatio-temporal models were compared to assess effect of the introduced random effects at province, district and facility level as well as the inclusion of the covariates. Model choice was based on the Deviance Information Criterion (DIC) (Spiegelhalter *et al.*, 2002) and marginal likelihood. Both modelling approaches included EVI, temperature suitability index and precipitation as covariates. Posterior mean predictions were carried out at 1 x 1 km spatial resolution with associated standard error maps.

Sensitivity analyses were conducted using the root mean square error (RMSE), the mean and the absolute mean error (MAE) that summarised the closeness of validation set data to observed values as well model scoring rules based on the probabilistic values from predictive distribution of the model compared to actual observations (Gneiting and Raftery, 2007). Model measures of uncertainty included the standard error score (SES), the Dawid-Sebastiani score (DSS) reviewed by Gneiting and Raftery (2007). These two parameters are calculated as:

$$SES(P, y) = (y - \mu P)^2$$

$$DSS(P, y) = \frac{1}{2} \cdot \left(\log\left(\sigma^2\right) + \left(\left(y - \mu_p\right) / \sigma_p \right)^2 \right)$$

where P is the predictive posterior distribution with a mean μ and standard deviation σ_p (Gneiting and Raftery, 2007). SES is similar to the mean square error (MSE) but applies to the predictive posterior distribution.

For geostatistical model, the nominal model coverage of 95% credible intervals was assessed based on the validation set (see section 2.3.6). The MAE and the R.M.S.E were also calculated as to estimate bias and accuracy of the model. Thus,

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left| Z^*(x) - Z(x) \right|$$

$$RMSE = \sqrt{\left(\frac{1}{n} \sum_{i=1}^{n} (Z^*(x) - Z(x))^2 \right)}$$

Lastly, the spatial structure in the residuals was assessed using semivariogram plots of in the spatial and temporal domain.

3.4.6 Results of estimating the incidence of *P. falciparum* and *P. vivax* in Afghanistan

Table 3.5 compares the four spatio-temporal models implemented along with associated model parameters for both *P. vivax* and *P. falciparum*. Comparison using the DIC showed that the fourth model (M4) provided the best trade-off between model fit and parsimony compared to the other three models. This model, however, had more number of effective parameters (P_D). For both *P. vivax* and *P. falciparum*, the standard error in M4 of the predictive distribution was also lower and was used for further analysis both for *P. vivax* and *P. falciparum*.

The mean error based on a 10% validation set was -0.30 and -0.44 for *P. vivax* and *P. falciparum*, respectively showing an overall tendency to under-estimate by less than 0.5 incidence cases per 1000 population. The Pearson correlation was 0.63 for *P. vivax* and 0.62 for *P. falciparum*.

Table 3.5: Bayesian model comparison for incidence
Models with and without random effects and covariates (M1 with no random effects or environmental covariates; M2: with random effects but no environmental covariates; M3: with environmental covariates but no random effects; M4 with random effects and environmental covariates)

	Model	DIC	P_D	Mlik (Integration)	Variance of predictive distribution	Std error of predictive distribution	Mean Error	R^2
P. falciparum	M1	3670.00	86.80	-1824.57	0.002	1.026	-	-
	M2	3596.90	95.60	-1824.94	0.005	1.042	-	-
	M3	3599.48	90.64	-1821.78	0.002	1.026	-	-
	M4	3570.76	96.85	-1804.94	0.002	1.022	-0.442	0.619
P. vivax	M1	20933.49	203.48	-10571.74	0.001	1.054	-	-
	M2	20781.31	301.97	-10538.10	0.001	1.049	-	-
	M3	20935.46	206.49	-10593.93	0.001	1.052	-	-
	M4	20780.64	301.46	-10554.87	0.001	1.047	-0.308	0.629

DIC: Deviance Information Criteria, **P_D**: effective number of parameters, **mlik**: maximum likelihood estimate

Table 3.6 lists the posterior distributions of the fixed effects, the unstructured components, and the temporal and spatial parameters for both the *P. vivax* and *P. falciparum* (for model M4).

None of the environmental covariate were significant at 95% Crl based on the *P. falciparum* model but temperature suitability (0.123, 95% Crl 0.046 – 0.202) was significant based on the *P. vivax* model. All other model parameters were significant at 95% Crl.

Table 3.6: Bayesian estimates of model parameters.
Parameters of the selected Bayesian models (M4) for both *P. falciparum* and *P. vivax* (sequentially as intercept β_0, EVI, TSI, Precipitation, random effects at (facility, district and province), temporal parameter and spatial effect ϕ)

	Parameter	Mean	Sd	5%	50%	95%
P. falciparum	β_0	-3.630	0.387	-4.244	-3.633	-3.008
	β_1	-0.031	0.079	-0.162	-0.031	0.099
	β_2	0.164	0.127	-0.042	0.163	0.334
	β_3	0.008	0.051	-0.077	0.008	0.091
	τ_1	1.940	1.903	0.192	1.380	5.534
	τ_2	2.484	0.829	1.355	2.369	4.010
	τ_3	3.668	1.164	2.040	3.521	5.838
	ρ	0.849	0.117	0.617	0.881	0.969
	ϕ	5.492	4.535	0.698	2.376	20.970
P. vivax	β_0	-2.065	0.240	-2.451	-2.069	-1.662
	β_1	-0.026	0.019	-0.058	-0.026	0.005
	β_2	0.124	0.048	0.046	0.124	0.202
	β_3	0.013	0.011	-0.005	0.013	0.031
	τ_1	8.383	1.778	6.095	8.057	11.750
	τ_2	2.081	1.976	0.181	1.500	5.888
	τ_3	7.972	3.953	3.897	6.922	15.530
	ρ	0.728	0.098	0.551	0.737	0.872
	ϕ	3.141	0.983	1.759	3.024	4.933

Figure 3.7 shows monthly (n = 48) variation of incidence for *P. vivax* and *P. falciparum*. The incidence of *P. vivax* was highest in August (7.611 95% Crl 4.849 – 11.721) compared to *P. falciparum* which was highest in November (mean incidence per 1,000 population 2.403 95% Crl 0.929 – 5.276) and lowest in May (0.830 95% Crl 0.303 – 1.783).

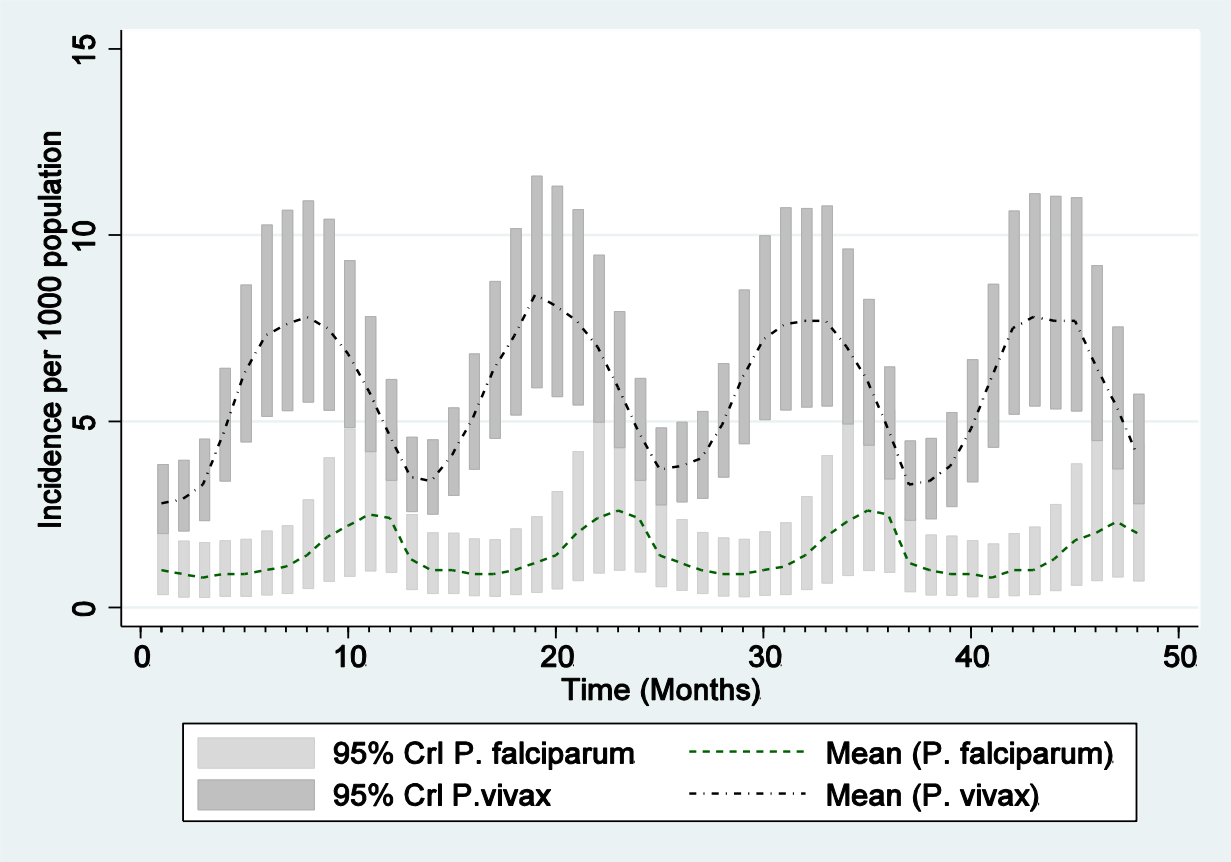

Figure 3.7: Time series plot of estimated incidence for 48 months in Afghanistan
Monthly predicted cases per 1000 population for P. falciparum and P. vivax with error bars showing 95% Bayesian credible interval. P. vivax cases have a peak in July and August compared to P. falciparum that peaks in November.

Figure 3.8 (below) and Figure 3.9 (Page 175) shows maps of monthly mean incidence of *P. vivax* and *P. falciparum*, respectively, at district level. The incidence of *P. falciparum* was very low generally compared to *P. vivax*. The results also showed that districts in south-east and eastern provinces with a high incidence of *P. vivax* also tended to have higher incidence of *P. falciparum* or *vice versa*.

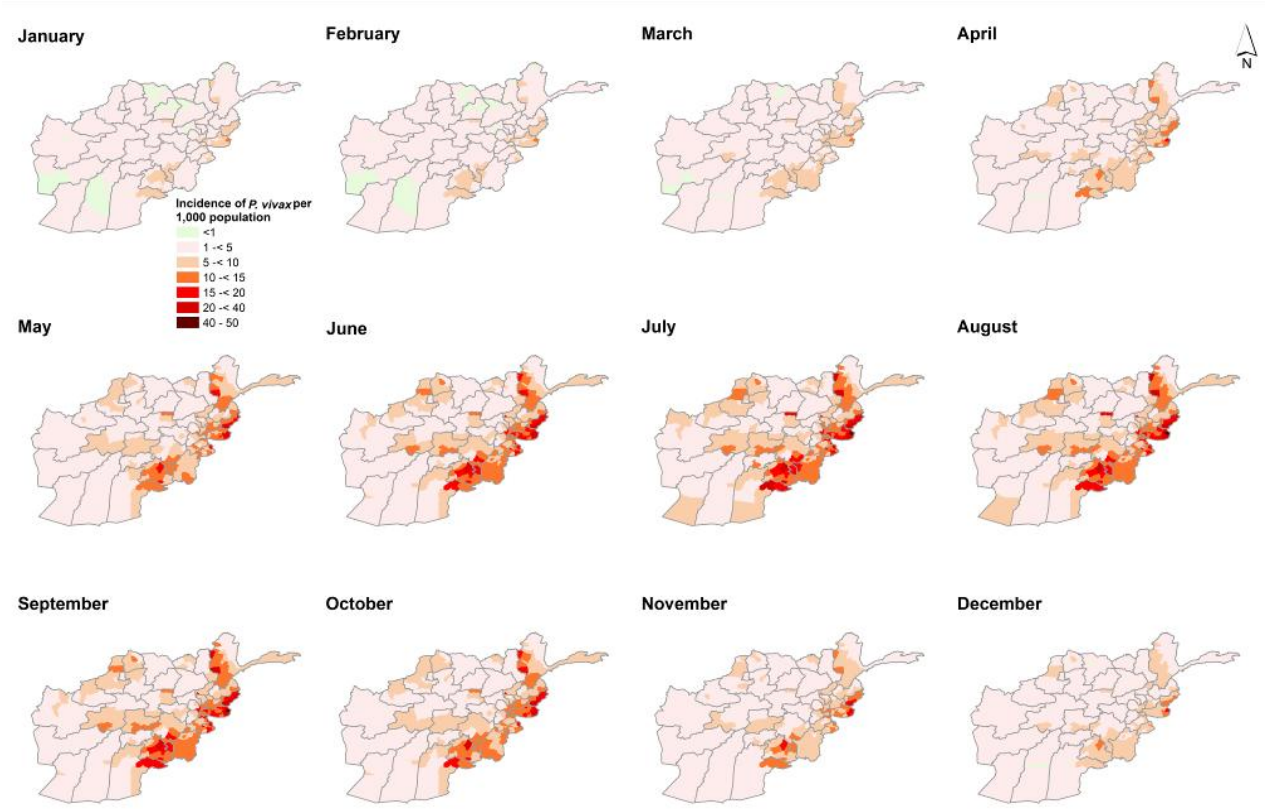

Figure 3.8: Posterior mean monthly incidence of *P. vivax* per 1000 population
Mean monthly maps of *P. vivax* incidence per 1000 population for Afghanistan using a Bayesian CAR model formulated at the facility level and included environmental covariates (Rainfall, TSI and EVI) and spatial random effects to account for regional heterogeneity. P. vivax constitutes a major burden in Afghanistan and experiences a peak in the July-August period

The estimated mean annual incidence for *P. vivax* was 5.1 cases per 1,000 population and 1.2 cases per 1,000 population for *P. falciparum*. Figure 3.10 (Page 176) shows the mean annual maps at the district level for *P. vivax* and for *P. falciparum*, respectively. Less than 1 case per 1000 population of *P. falciparum* was estimated for most districts on annual basis compared to *P. vivax*. Similarly, annual estimates showed that incidence was highest in the southern, south-eastern and the eastern regions for both parasites. The estimated mean incidence in the most recent data year (2009) for *P. vivax* was 5.4 (95% Crl 3.2 – 9.2) cases per 1,000 population and 1.2 (95% Crl 0.4 – 2.9) cases per 1,000 population for *P. falciparum*. Comparison between the

baseline in 2006 and in 2009 showed small change in incidence (4.9, 95% Crl 3.0 - 7.8 and 5.1, 95% Crl 3.2 – 8.1 respectively for *P. vivax*; 1.1, 95% Crl 0.3 – 2.4 and 1.1, 95% Crl 0.3 – 2.5 respectively for *P. falciparum*). However, there was a slight increase in malaria incidence in 2008 for both *P. vivax* and *P. falciparum* as estimated by the model, but, dropped subsequently to the 2006 level in 2009. The mean percentage change in incidence in the 34 provinces between the baseline year and 2009 for *P. vivax* was 3.0 and 5.9 for *P. falciparum* (Table 3). *P. vivax* reduced in 17 of the 34 provinces in Afghanistan while *P. falciparum* reduced in 13 provinces

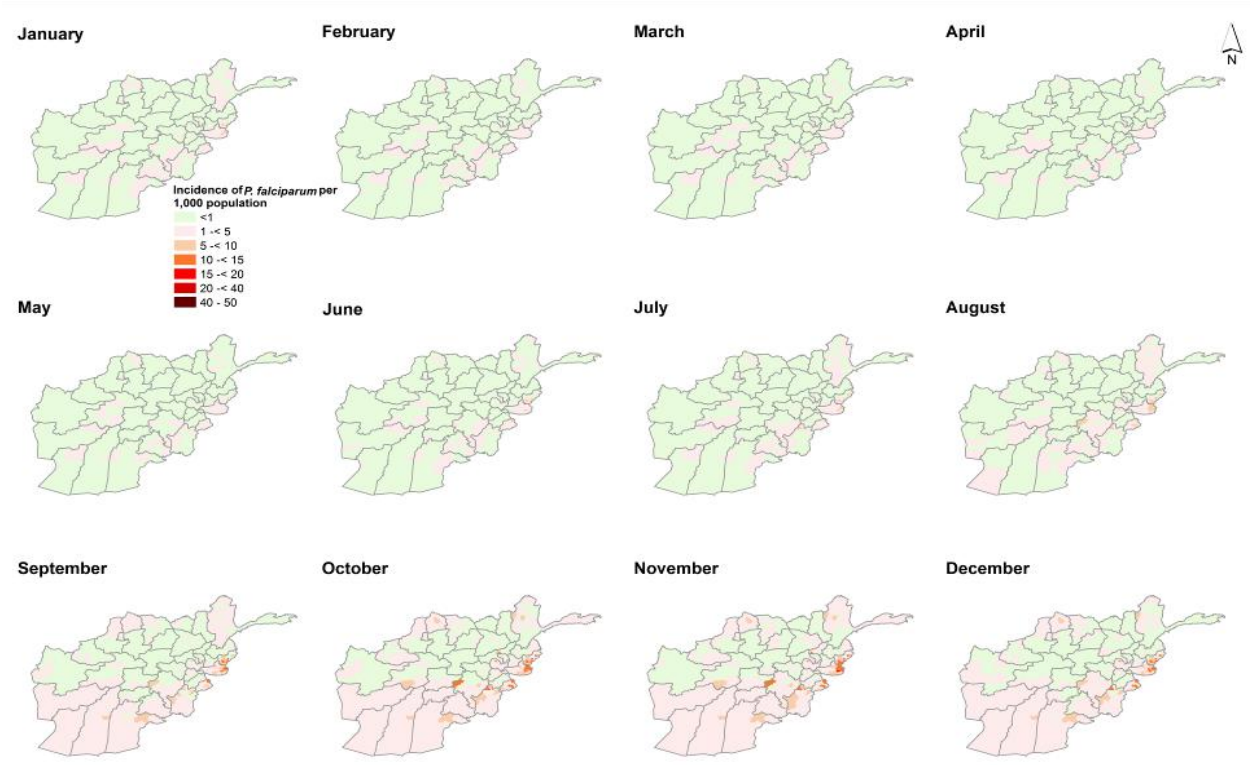

Figure 3.9 Posterior mean monthly incidence of *P. falciparum* per 1000 population
Mean monthly maps of *P. falciparum* incidence per 1000 population for Afghanistan using a Bayesian CAR model formulated at the facility level and included environmental covariates (Rainfall, TSI and EVI) and spatial random effects to account for regional heterogeneity.

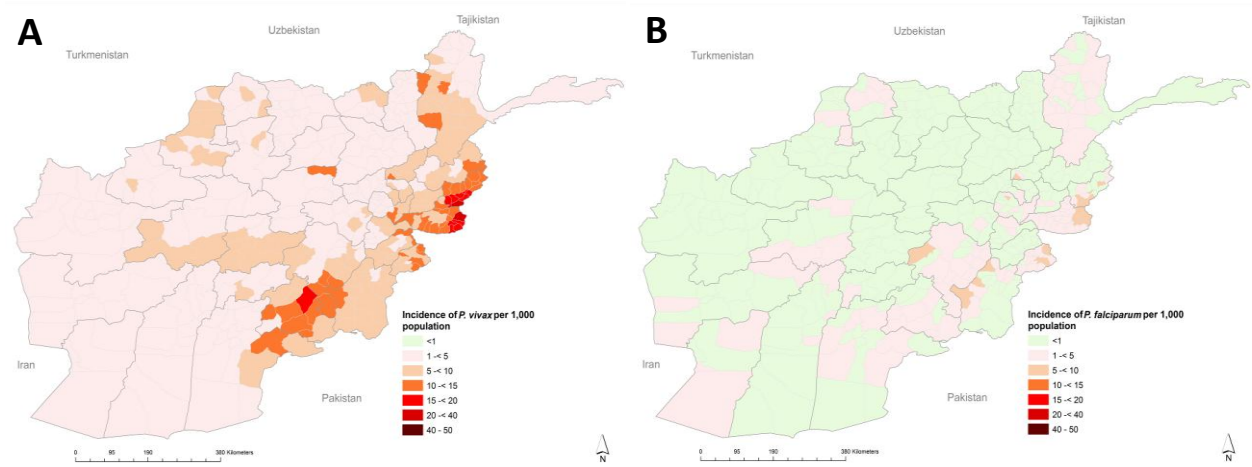

Figure 3.10: Posterior mean annual incidence maps
The mean annual incidence of (a) *P. vivax* and (b) *P. falciparum* by district in Afghanistan

3.4.7 Results of slide positivity prediction for *P. falciparum* and *P. vivax* in Afghanistan

Table 3.7 lists model parameters of the slide positivity rate. The MAE and the RMSE for the *P. vivax* was 0.42 and 0.81 and for *P. falciparum* 0.83 and 1.05. The 95% nominal coverage for *P. vivax* was 96.75% showing tendency to over predict at 95% nominal. The model range for *P. vivax* and *P. falciparum* was different (Table 3.7). Semivariogram plots of the small scale variation (residuals) in the spatial and temporal domain based on the validation set indicated minimum spatial and temporal autocorrelation (Figure 3.11). The semivariogram for residuals in the spatial domain for *P. falciparum* (Figure 3.11(a) (i)) data had shorter range but larger sill compared to *P. vivax* (Figure 3.11(b) (i)). The structures were also different marginally in the temporal domain Figure 3.11(a) (ii) and Figure 3.11(b) (ii) respectively.

Table 3.7: Models for slide positivity rate in Afghanistan.
Bayesian model comparison based on separable covariance function (Product) (M1) and no-separable form (Product-sum) (M2) for *P. falciparum* (P*f*) and *P. vivax* (P*v*).

Species	DIC	P_D[1]	MAE	R.M.S.E	Probability of prediction interval (%)[2]	Model Range (m)
P*f*	27393.81	200.03	0.4218	0.8063	95.37	3,395.73
P*v*	34996.00	269.08	0.8368	1.0489	96.75	45,198.81

1. P_D represent the effective number of parameters that represent model complexity
2. The nominal probability of prediction is 95%

Table 3.8 shows the posterior summaries of the fixed effects along and the random effects. Of the three selected environmental covariates, only TSI (β_3) was an important predictor of *P. falciparum* (0.05 95% Crl 0.02 – 0.08) while for *P. vivax* EVI (β_1) (-0.04 95% Crl -0.07 - -0.02) and precipitation (β_2) (0.02 95% Crl 0.00 – 0.11) were significant for *P. vivax*. The nominal range for *P. falciparum* was also shorter (3,395.7 m 95% Crl 1,521.78 – 6,196.59) compared to the model range for *P. vivax* (45,198.8 m 95% Crl 37,591.30 – 52,711.04).

Table 3.8: Posterior estimates of parameter for slide positivity rate.
Distribution of posterior estimates (mean, standard deviation and quantiles) of the fixed components (intercept (β_0), EVI (β_1), precipitation (β_2) and TSI (β_3), the matérn variance component (σ_w^2)) and the model range (ϕ)

	Parameter	Mean	Std. dev	5%	50%	95%
P*v*	β_0	1.3466	0.3299	0.7979	1.3663	1.8223
	β_1	-0.0448	0.016	-0.0711	-0.0448	-0.0185
	β_2	0.0213	0.0127	0.0004	0.0275	0.1059
	β_3	0.0274	0.0477	-0.0511	0.0275	0.1059
	σ_w^2	1.4204	0.0584	1.3319	1.4156	1.523
	ϕ	45,198.81	4,605.16	37,591.30	45,235.97	52,711.04
P*f*	β_0	0.323	0.0326	0.2696	0.3229	0.3768
	β_1	-0.004	0.0109	-0.0219	-0.004	0.014
	β_2	0.0076	0.009	-0.0072	0.0076	0.0224
	β_3	0.051	0.0202	0.0179	0.0509	0.0843
	σ_w^2	18.8035	18.6075	3.271	12.954	53.6073
	ϕ	3,395.73	1,471.85	1,521.78	3,120.85	6,196.59

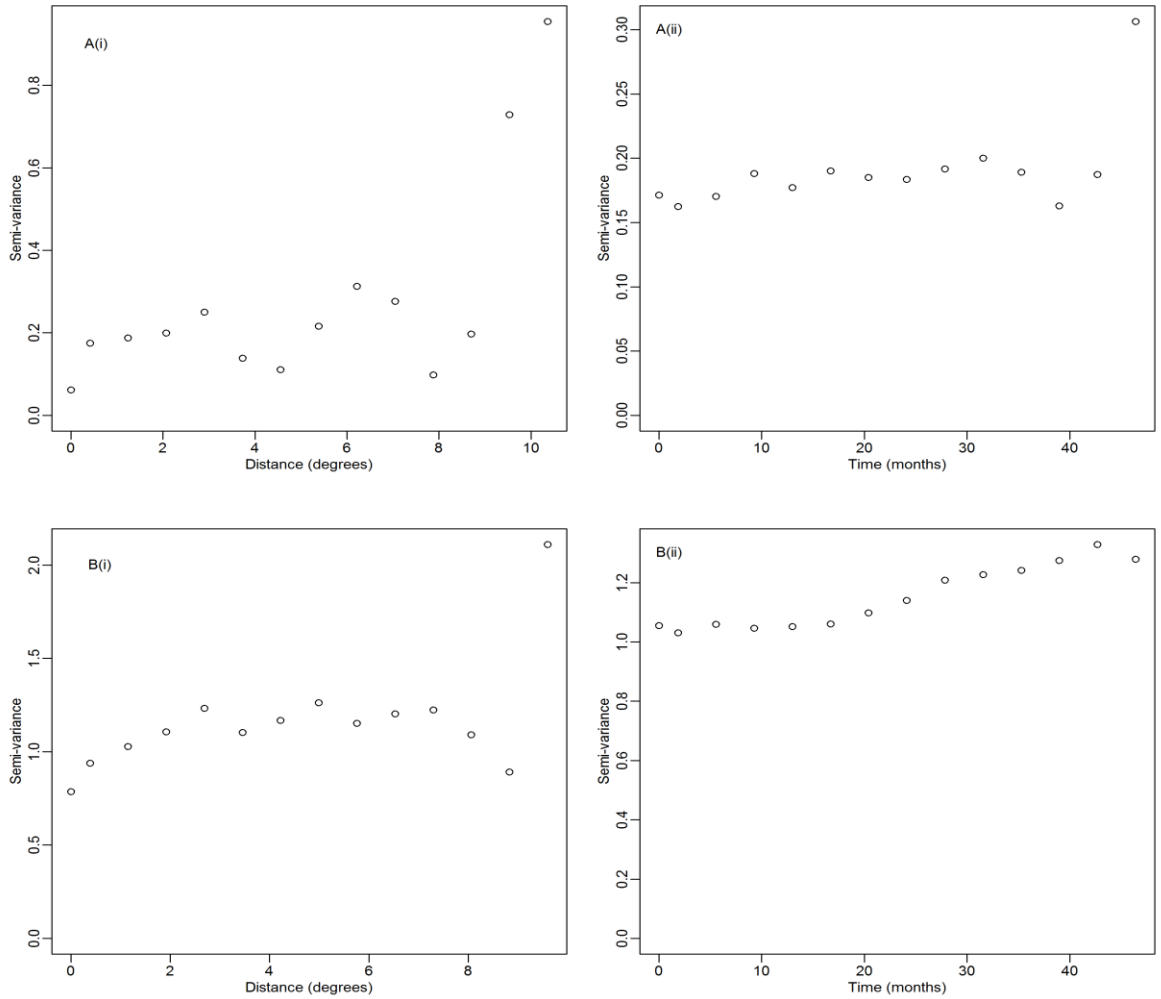

Figure 3.11: Semi-variograms of residuals in spatial and temporal domain
Semivariogram plots of the residuals in spatial (i) and temporal (ii) domain for *P. falciparum* **(a)** and *P. vivax* **(b)** based on the predictions on the hold out set. The x-axis shows distance in degrees latitude and longitude (decimal degrees) whiles the y-axis shows semi-variance. There was only minimal spatial structure shown in both temporal domains compared to the spatial domain in both malaria species.

Figure 3.12 shows the continuous and binned predictions at 1 x 1 km of Slide positivity rate for 2009 along with standard errors for *P. falciparum* and *P. vivax* respectively.

Table 3.9 provides summaries of slide positivity rate by province weighted by probability of health facility utilisation. For *P. vivax*, the mean slide positivity was >1% (mean 1.3%; minimum 0.0%; maximum 14.3%) while for *P. vivax*, slide positivity was <1% (mean 0.01%; minimum

0.0%; maximum 0.3%). The slide positivity rate was highest in southern and south-east provinces. The highest uncertainties were in districts with no facilities and where the rate of reporting was poor. For *P. falciparum* majority of provinces had a positivity rate of 0% -< 0.1% compared to *P. vivax* where most provinces (28 out of 36) had mean estimate between 0.5% -< 5.0% (Table 3.9).

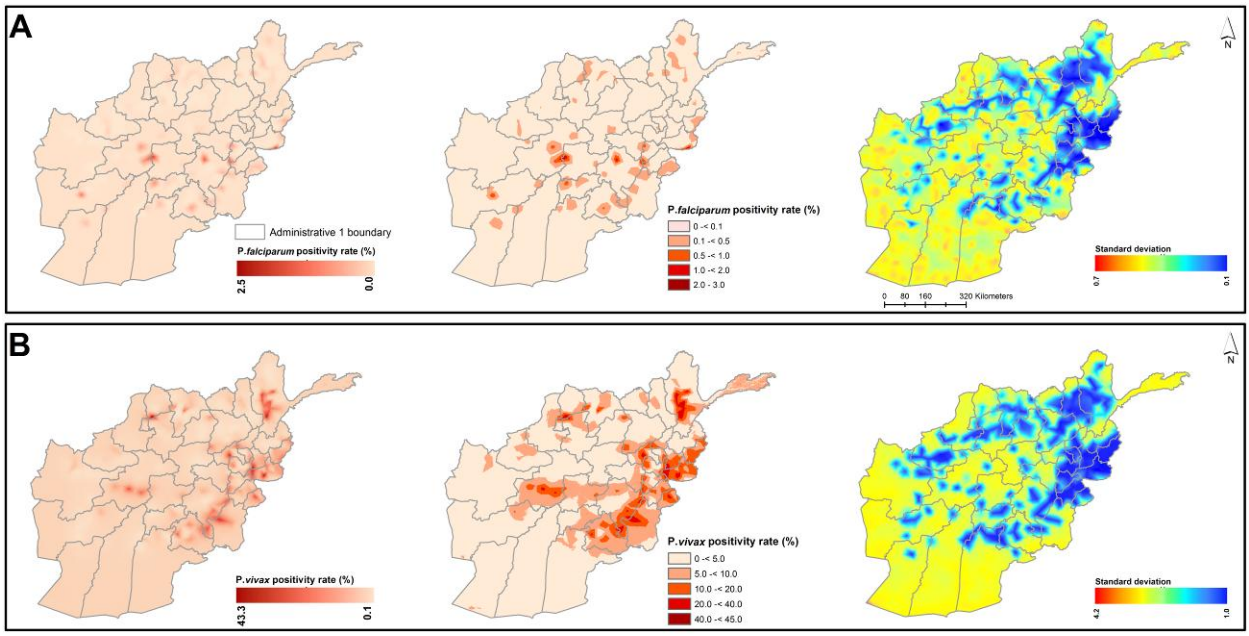

Figure 3.12: Posterior mean predictions of slide positivity rate
Posterior mean predictions and standard error maps. **Panel A** the posterior mean for *P. falciparum* with first map on continuous scale, the second map classified in 5 endemicity classes of <0.1%; 0.1 -<0.5; 0.5 -< 1.0; 1.0 -<2.0 and ≥2.0 and the third map being the standard error map. **Panel B** the posterior mean for *P. vivax* on continuous scale, in 5 classes of <5.0; 5.0 -< 10.0; 10.0 -< 20.0; 20.0 -< 40.0 and ≥40.0 and a standard error map. Higher standard error was in regions with sparse or no data points.

Figure 3.13 shows a comparison between incidence, community prevalence and slide positivity rate. Parasite prevalence estimates for Afghanistan were obtained from the global endemicity maps (Gething *et al.*, 2011b, Gething *et al.*, 2012). There was positive association of incidence and slide positivity comparison with parasite prevalence (Pearson correlation coefficient 0.6 and 0.4 for *P. vivax* and 0.3 for *P. falciparum*, for incidence).

Table 3.9: Mean slide positivity rate by region in Afghanistan

Summary (in percentage) of the mean predicted slide positivity rate (SPR) at health facilities weighted by probability of use for *P. falciparum* and *P. vivax* by Province in Afghanistan in 2009

	P. vivax				*P. falciparum*			
Name	**Mean**	**Minimum**	**Maximum**	**Range**	**Mean**	**Minimum**	**Maximum**	**Range**
Badakhshan	0.65	0.00	22.71	22.71	0.01	0.00	0.21	0.21
Badghis	0.64	0.00	7.03	7.03	0.01	0.00	0.15	0.15
Baghlan	0.50	0.00	14.85	14.85	0.00	0.00	0.05	0.05
Balkh	1.49	0.00	15.27	15.27	0.02	0.00	0.21	0.21
Bamyan	0.48	0.00	11.60	11.60	0.01	0.00	0.14	0.14
Day Kundi	0.13	0.00	5.00	5.00	0.00	0.00	0.07	0.07
Farah	0.35	0.00	7.83	7.83	0.01	0.00	0.77	0.77
Faryab	0.77	0.00	5.31	5.31	0.01	0.00	0.06	0.06
Ghazni	1.49	0.00	22.42	22.42	0.02	0.00	0.77	0.77
Ghor	0.42	0.00	19.86	19.86	0.01	0.00	0.41	0.41
Hilmand	0.39	0.00	4.18	4.18	0.00	0.00	0.19	0.19
Hirat	0.57	0.00	5.81	5.81	0.01	0.00	0.08	0.08
Jawzjan	1.42	0.00	22.93	22.93	0.01	0.00	0.18	0.18
Kabul	3.27	0.00	21.48	21.48	0.01	0.00	0.05	0.05
Kandahar	0.30	0.00	8.43	8.43	0.00	0.00	0.18	0.18
Kapisa	1.30	0.00	9.83	9.83	0.00	0.00	0.07	0.07
Khost	3.46	0.01	21.79	21.77	0.05	0.00	0.28	0.28
Kunar	3.42	0.00	17.66	17.66	0.03	0.00	0.39	0.39
Kunduz	1.08	0.00	4.88	4.88	0.01	0.00	0.05	0.05
Laghman	3.12	0.00	30.31	30.31	0.00	0.00	0.05	0.05
Logar	1.15	0.00	12.38	12.38	0.02	0.00	0.68	0.68
Nangarhar	4.10	0.00	27.26	27.26	0.02	0.00	0.96	0.96
Nimroz	0.38	0.00	4.20	4.20	0.01	0.00	0.25	0.25
Nuristan	0.35	0.00	9.47	9.47	0.00	0.00	0.09	0.09
Paktika	1.60	0.00	25.56	25.56	0.02	0.00	0.37	0.37
Paktya	2.29	0.01	11.55	11.54	0.02	0.00	0.36	0.36
Panjshir	0.93	0.00	16.27	16.27	0.00	0.00	0.05	0.05
Parwan	1.75	0.00	22.17	22.17	0.01	0.00	0.09	0.09
Samangan	0.67	0.00	13.44	13.44	0.01	0.00	0.10	0.10
Sari Pul	0.94	0.00	22.30	22.30	0.00	0.00	0.12	0.12
Takhar	1.46	0.00	8.04	8.04	0.01	0.00	0.21	0.21
Uruzgan	0.66	0.00	6.61	6.61	0.01	0.00	0.45	0.45
Wardak	0.84	0.00	10.68	10.68	0.01	0.00	0.44	0.44
Zabul	0.76	0.00	15.69	15.69	0.01	0.00	0.41	0.41
Mean	**1.27**	**0.00**	**14.26**	**14.26**	**0.01**	**0.00**	**0.26**	**0.26**

3.4.8 Assessing the population at risk of malaria based on estimated incidence

Of the 30.6 million people in 2009, approximately 32.0% of the population lived in regions where *P. vivax* was greater than 1 case per 1000 population compared to 23.7% for *P. falciparum*. Table 3.10 provides summaries of population at risk by region. Overall, 1.3% of the population in Balkh province, were estimated to live in districts with <1 case per 1,000 population, the majority (66.7%) in districts of 1 to < 5 vivax cases per 1,000 population, 23.3% in 5 to < 10 cases per 1,000 population, 8.4% in 10 to < 20 cases per 1,000 population and 0.3%

of the population, in eastern Afghanistan in Kunar and Nangarhar provinces, were classified as residing in districts with annual *P. vivax* case incidence of >20 cases per 1,000 population. For *P. falciparum*, 76.3% lived in districts where *P. falciparum* case incidence was <1 per 1,000 population, while 20.9% lived in areas were incidence of *P. falciparum* was 1 to < 5 cases per 1,000 population. A minority (2.8%) were classified to live in districts with an estimated annual incidence of 5 to < 10 *P. falciparum* cases per 1,000 population.

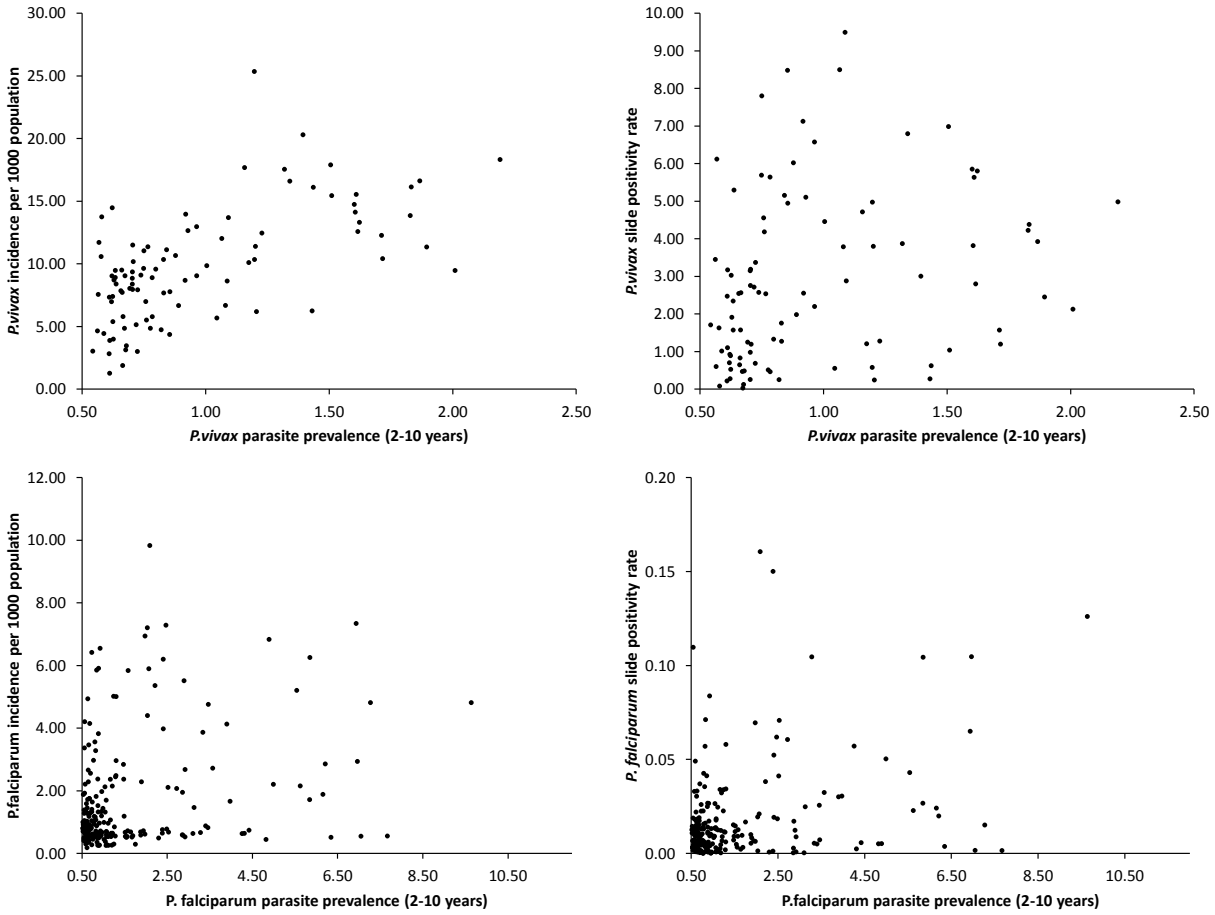

Figure 3.13 Comparison of incidence and slide positivity rate with parasite prevalence
Scatter plots comparing the incidence per 1000 population for *P. vivax* (top left) and slide positivity (top right) with community parasite prevalence for 2010; P. falciparum is shown at the bottom. There was positive correlation with incidence *P. vivax* and *P. falciparum*.

Table 3.10: Estimated population at risk in 2009 by Province for *Plasmodium falciparum* and *Plasmodium vivax*

Province	*Plasmodium vivax* incidence per 1,000 population									*Plasmodium falciparum* incidence per 1,000 population								Total Population
	Estimated Pv Clinical burden	crude incidence	Estimated mean incidence 2009	SD[1]	% change Baseline (2006 and 2009	< 1 (%)	1 -< 5 (%)	5 -< 10 (%)	≥ 10 (%)	Estimated Pf Clinical burden	crude incidence	Estimated mean incidence 2009	SD[1]	% change Baseline (2006 and 2009	< 1 (%)	1 -< 5 (%)	5 -< 10 (%)	
Kabul	19,788	4.7	4.3	1.1	0.59	0	4,513,039 (98.1)	72,584 (1.6)	16222 (0.4)	5,016	2.7	1.1	0.7	-0.52	4,132,170 (89.8)	425,758 (9.3)	43,917 (1)	4,601,845
Kapisa	2,497	5.3	5.4	0.8	0.59	0	215,825 (46.8)	245,789 (53.2)	0	268	1.7	0.6	0.5	-0.21	461,613 (100)	0	0	461,613
Logar	1,988	5.1	4.4	1.1	-2.01	0	308,882 (68.2)	143,913 (31.8)	0	317	2.2	0.7	0.6	-5.99	425,918 (94.1)	26,877 (5.9)	0	452,795
Panjshir	777	6.3	5.5	1.5	1.24	0	0	123,418 (86.7)	18879 (13.3)	203	5.1	1.4	0.8	1.74	99,628 (70.0)	23,790 (16.7)	18,879 (13.3)	142,298
Parwan	2,806	3.4	3.4	0.9	-4.04	0	778,082 (93.2)	57,015 (6.8)	0	576	2.2	0.7	0.6	-3.74	778,082 (93.2)	57,015 (6.8)	0	835,096
Wardak	2,439	4.9	3.8	1.1	-1.13	0	515,633 (79.9)	129,611 (20.1)	0	400	2.8	0.6	0.5	-3.68	573,533 (88.9)	71,711 (11.1)	0	645,244
Bamyan	2,172	4.4	4.3	1.3	-4.79	0	465,596 (92.2)	0	39633 (7.8)	273	1.3	0.5	0.5	-8.76	505,228 (100)	0	0	505,228
Day Kundi	2,280	6.1	4.2	1.5	0.19	0	143,333 (26.2	403,394 (73.8)	0	306	3.3	0.6	0.6	-0.02	546,727 (100)	0	0	546,727

							)											
Kunar	6,897	7.5	13.3	2.0	2.61	0	0	0	517,421 (100)	2,013	4.9	3.9	1.3	-0.03	175,647 (33.9)	182,013 (35.2)	159,762 (30.9)	517,421
Laghman	4,114	6.8	8.1	0.6	2.91	0	48,935 (9.7)	456,482 (90.3)	0	510	3.1	1.0	0.5	1.68	386,431 (76.5)	118,986 (23.5)	0	505,417
Nangarhar	23,043	7.5	13.3	0.9	5.48	0	0	733,865 (42.5)	993,460 (57.5)	6,391	4.5	3.7	1.2	3.94	0	1,453,897 (84.2)	273,428 (15.8)	1,727,324
Nuristan	1,043	5.8	5.9	1.1	3.20	0	35,830 (20.3)	112,523 (63.6)	28511 (16.1)	108	2.1	0.6	0.5	2.67	176,863 (100)	0	0	176,863
Badakhshan	6,895	4.7	5.4	1.4	0.59	0	247,984 (19.4)	785,904 (61.5)	243,005 (19)	1,302	2.4	1.0	0.8	0.93	618,838 (48.5)	658,054 (51.5)	0	1,276,892
Baghlan	3,039	4.5	2.9	1.0	2.05	0	1,040,766 (100)	0	0	385	0.9	0.4	0.4	5.79	1,040,766 (100)	0	0	1,040,766
Kunduz	4,639	5.4	4.1	0.7	-0.82	0	755,118 (66.4)	381,799 (33.6)	0	432	1.6	0.4	0.4	2.59	1,136,917 (100)	0	0	1,136,917
Takhar	3,452	4.6	3.1	0.7	1.73	0	1,128,142 (100)	0	0	463	2.0	0.4	0.3	3.65	1,128,142 (100)	0	0	1,128,142
Balkh	4,091	2.9	2.9	1.1	10.56	408,202 (28.9)	1,002,618 (71.1)	0	0	818	1.5	0.6	0.5	15.39	1,410,820 (100)	0	0	1,410,820
Faryab	5,209	5.4	4.8	1.3	-0.05	0	498,575 (45.8)	591,181 (54.2)	0	708	2.2	0.7	0.5	0.25	971,677 (89.2)	118,079 (10.8)	0	1,089,756

Jawzjan	2,583	3.8	4.1	1.3	-2.39	0	521,918 (83)	106,563 (17)	0	760	2.3	1.2	0.9	8.69	255,757 (40.7)	372,724 (59.3)	0	628,480
Samangan	1,219	2.6	2.8	1.0	-1.86	0	441,833 (100)	0	0	239	1.2	0.5	0.5	-4.16	441,833 (100)	0	0	441,833
Sari Pul	2,548	4.2	3.7	1.0	3.69	0	604,485 (88.5)	78,646 (11.5)	0	266	0.5	0.4	0.4	20.25	683,132 (100)	0	0	683,132
Khost	5,613	6.0	8.5	1.3	10.31	0	57,439 (8.7)	343,942 (52.2)	257,366 (39.1)	1,719	3.3	2.6	1.1	17.70	13,623 (2.1)	426,758 (64.8)	218,366 (33.1)	658,747
Paktika	3,037	5.3	6.0	1.9	-4.60	0	25,624 (5.0)	453,431 (89.1)	29,674 (5.8)	829	3.6	1.6	1.0	-13.29	171,714 (33.8)	216,884 (42.6)	120,131 (23.6)	508,729
Paktya	4,458	5.8	6.9	1.1	-0.44	0	115,091 (17.8)	432,258 (67)	97,766 (15.2)	387	2.5	0.6	0.5	1.12	645,114 (100)	0	0	645,114
Ghazni	9,033	5.1	6.3	1.2	2.46	0	530,263 (36.7)	739,121 (51.2)	173,664 (12)	2,165	2.7	1.5	1.0	16.31	321,390 (22.3)	1,084,947 (75.2)	36,711 (2.5)	1,443,048
Hilmand	2,859	3.4	2.8	1.1	-0.26	0	1,039,697 (100)	0	0	1,071	2.0	1.0	0.8	1.90	930,158 (89.5)	109,539 (10.5)	0	1,039,697
Kandahar	5,644	4.9	4.1	1.2	-0.08	0	1,161,551 (84.2)	178,308 (12.9)	40,003 (2.9)	1,421	2.7	1.0	0.8	-4.06	1,130,774 (81.9)	249,088 (18.1)	0	1,379,862
Nimroz	577	4.6	3.1	1.2	-2.38	0	187,872 (100)	0	0	178	1.9	1.0	0.8	-7.58	86,898 (46.3)	100,974 (53.7)	0	187,872
Uruzgan	1,601	4.2	4.0	1.1	-0.88	0	306,799 (75.9	97,595 (24.1)	0	311	1.9	0.8	0.7	0.87	332,451 (82.2)	71,944 (17.8)	0	404,395

							)												
Zabul	2,621	6.3	7.5	1.6	-6.88	0	20,586 (5.9)	144,845 (41.3)	185,415 (52.8)	379	2.7	1.1	0.8	-9.01	149,910 (42.7)	200,936 (57.3)	0	350,846
Badghis	2,257	4.7	4.1	0.9	1.92	0	534,183 (96.8)	17,642 (3.2)	0	353	2.0	0.6	0.5	7.60	551,825 (100)	0	0	551,825
Farah	1,459	2.8	2.5	1.0	-3.19	0	581,449 (100)	0	0	494	1.7	0.9	0.8	0.59	484,949 (83.4)	96,500 (16.6)	0	581,449
Ghor	3,141	4.5	4.1	1.4	-1.07	0	468,809 (60.9)	300,924 (39.1)	0	708	2.5	0.9	0.9	2.08	459,819 (59.7)	309,915 (40.3)	0	769,733
Hirat	7,532	3.3	3.6	1.2	16.37	0	2,098,175 (100)	0	0	1,133	1.2	0.5	0.5	23.14	209,817 5 (100)	0	0	2,098,175
	165,712	5.0	**5.4**	**1.2**	**1.62**	**408,202 (1.3)**	**20,394,129 (66.7)**	**7,130,753 (23.3)**	**2,641,017 (8.6)**	**36,077**	**2.7**	**1.2**	**0.7**	**2.26**	**23,326,522 (76.3)**	**6,376,386 (20.9)**	**871,194 (2.8)**	**30,574,102**

1. SD: Starndard Deviation

3.5 Discussion on the incidence and slide positivity rates of *P. vivax* and *P. falciparum* in Afghanistan

In this chapter, the distribution of *P. vivax* and *P. falciparum* malaria species in Afghanistan was modelled using HMIS data to estimate disease burden. The findings confirm *P. vivax* malaria morbidity in Afghanistan exceeds that for *P. falciparum*. The incidence of *P. vivax* and *P. falciparum* was estimated to be higher in the southern, south-eastern and eastern parts of Afghanistan. On average, the incidence of *P. falciparum* was low with majority of districts classified as <1 case per 1000 population. The crude estimate was 2.7 cases per 1000 population. For *P. vivax*, the estimated incidence was 5.4 cases per 1000 population compared to a crude estimate of 5.0 per 1000 population. The spatial distribution of both species was similar. Thus, the results suggested that the incidence of *P. vivax* was highest in the population highly endemic with *P. falciparum*. What was striking was the distribution of both malaria species in southern and south-eastern provinces based on both incidence and slide positivity rates. The mean slide positivity rate predicted was 1.27 % (minimum 0%; maximum 14.26%) and 0.01% (minimum 0%; maximum 0.26%) for *P. vivax* and *P. falciparum*, respectively. Given the additional spatial precision resulting from the facility, district and regional adjustments of incidence compared to a crude estimate, maps of both malaria species are useful for concerted planning. The smoothed incidence also incorporated environmental covariate to estimate incidence in districts where with no data.

3.5.1 Implications for malaria control and elimination in Afghanistan

Using the 2006 estimates as baseline, 17 and 13 provinces had already reduced *P. vivax* and *P. falciparum* incidence respectively by 2009. No reduction in incidence was estimated for

Nangahar, Balkh, Sari Pul, Khost and Hirat. Nangahar and Khost provinces in south-eastern regions of Afghanistan were amongst those with highest incidence for both parasites. From the MIS undertaken in 2008, Nangahar had an estimated long lasting insecticidal nets (LLINs) coverage of 19% while no LLINs use was observed in Hirat (Ministry of Public Health, 2009a). Sari Pul district, for example, had some of lowest rates of long lasting insecticidal nets (LLINs) coverage and access to treatment of care. In districts where indoor residual spraying (IRS) is used as the main vector control approach or to complement LLINs, the targeting of this intervention should be informed by the lag in the peak season of the two main malaria parasites. *P. vivax* peaks in August while *P. falciparum* peaks in November. IRS campaigns should therefore be planned in such away the insecticide are efficacious through the two peak seasons.

Of the 34 provinces of Afghanistan, five were considered to be malaria free based on altitude thresholds (Ministry of Public Health, 2010b). These provinces, however, accounted for 9.7% of all estimated cases in 2009 indicating a potential problem of importation of suspected cases due to human population movement in Afghanistan or foci transmission in valleys where climatic conditions are favourable. The available data, however, do not provide malaria case definitions and it is impossible to distinguish between imported and local cases. In the malaria free provinces, suspected imported infections should be documented and algorithms, based on travel history, could be used as the basis for case definitions. In addition, health advice and chemoprophylaxis for travellers from the malaria free to endemic provinces should be initiated as an additional package for malaria prevention. An incidence of less than 1 *P. falciparum* case per 1000 individuals is considered to be the threshold for pre-elimination by the WHO (World Health Organization, 2007a). By 2009, 21 provinces in Afghanistan had already achieved such a

threshold. However, the biggest challenge is likely to be operational and a comprehensive analysis of overall feasibility of *P. falciparum* elimination (Feachem *et al.*, 2010)

The analysis also showed that malaria in Afghanistan exhibits a seasonal peak between July and November. *P. vivax* tended to peak in August (mean incidence of 7.611 95% Crl 4.849 – 11.721) compared to *P. falciparum* which peaked in November (mean incidence 2.403 95% Crl 0.929 – 5.276). Incidence was lowest between January and May with variation resulting largely from climatic conditions in winter and spring. *P. vivax* hypnozoites are likely to survive the long winter season in Afghanistan, due to long latent periods, relapsing after the spring period in May. The result is a possible explanation of the early peak of vivax malaria for the July-August period. These maps may, therefore, provide a baseline for identifying areas where mixed infections are likely to occur.

The slide positivity varied between the two malaria species in the endemic provinces. For *P. falciparum*, SPR was less than 3% while for *P. vivax*, the mean SPR was 6.2%. Pre-elimination of malaria can be achieved at less than 5% positivity rate (World Health Organizastion, 2007). The spatial distribution observed in Figure 3.12 is driven partly by climatic conditions (e.g. temperature, rainfall, humidity) which affect parasite survival. The SPR is independent of population size and is a useful index in unstable areas where asymptomatic infections are not common. The Afghanistan analysis indicated that only TSI was an important covariate for estimating *P. falciparum* perhaps due for focal pattern observed compared to *P. vivax* where EVI and precipitation were more useful.

In terms of case management, the quantification of the co-distribution of *P. vivax* and *P. falciparum* in Figure 3.8 and Figure 3.9 may have useful implications for dual control approach for both species in endemic districts. Other studies elsewhere have shown an infection of *P. vivax* malaria subsequent to *P. falciparum* infection and cases of mixed infections may present a challenge for treatment (Looareesuwan *et al.*, 1987, Mehlotra *et al.*, 2000, Mayxay *et al.*, 2004, Douglas *et al.*, 2011). *P. vivax* infections tend to relapse more often because the hypnozoites can lie dormant in an infected liver for months (White, 2011), a factor that has important implications for its control. An additional characteristic of *P. vivax* is that it can induce fever at relatively low parasite densities (Price *et al.*, 2007, White, 2011), thus suspected cases require a parasitological diagnosis before treatment. The prevalence of glucose-6-phosphate dehydrogenase deficiency (G6PDd) is estimated to be 8% in Afghanistan (Howes *et al.*, 2012) which complicates the use of the reccomended 14 day regiment of primaquine (PQ) (World Health Organization, 2010a). The use of PQ in patients with G6PDd can cause severe haemolysis (Cappellini and Fiorelli, 2008, Leslie *et al.*, 2008). We suggest that improved maps of prevalence of G6PDd may be helpful in reducing disease burden. Chloroquine is used as first line treatment of *P. vivax* in Afghanistan as recommended for countries where it remains efficacious and where parasites can be isolated (World Health Organization, 2012b), while Artesunate with Sulfadoxine-Pyrimethamine (AS+SP) is used for *P. falciparum* (Ministry of Public Health, 2008b). However, where both species are endemic, the use of artemisinin-based combination therapies (ACTs) has been proposed (Douglas *et al.*, 2010, Sinclair *et al.*, 2011) and other clinical studies have shown a faster parasite clearance rate when ACTs were used (Nguyen *et al.*, 1993, Hamedi *et al.*, 2004, Dao *et al.*, 2007). Figure 3.10 indicates where such a case management approach could be beneficial.

3.5.2 Modelling gains for incidence analysis in Afghanistan

An independent linear approach was used when modelling both incidence and slide positivity rates for both parasites. While modelling large space-time data with gaps is a challenging task, the advantages of hierarchical model-based approaches lies in quantifying the mean process independently $\eta(s_i,t)$ while at same time including random effects (see also section 2.5.3). An autoregressive time varying factor was used in the model with an assumption that estimates evolves from previous values but modified by spatial and spatio-temporal set of covariates $X^{'}(s_i,t)\beta$ (Sahu and Bakar, 2012). Crude estimates of incidence was much close to the smooth values in Afghanistan compared to Namibia. The advantages of smoothed incidence over crude estimates are discussed in section 2.5.4. However, the analytical gain in the Afghanistan study was the extra random effects included at province level while maintaining the CAR prior at the district level. Thus;

$$f_\theta = \varsigma_i + (\varphi_k + \xi_k) + \psi_j$$

Where ς_i are the facility level effects, φ_k is the CAR prior at district level with random effects ξ_k (with $\varphi_k \mid \sigma^2 \sim N(0,\sigma^2)$ (Barnerjee *et al.*, 2004, Lee, 2011)) and ψ_j represent higher level effects at the province level. This specification improved the spatial smoothing toward a regional (province) mean. This evident in closeness of crude estimates at regional level compared to smooth estimates.

The analysis in this section however did not explore the alternative approach of multivariate space-time random effects where the joint covariance for *P. vivax* and *P. falciparum* maybe be

based on linear transformation of the independent processes (Wagner and Tüchler, 2010, De Iaco *et al.,* 2011). Such modelling approach is however not straight forward and as observed here. For example in the geostatistical approach, the spatial range ϕ for both species was different which may complicate a joint modelling framework of both species.

3.5.3 Limitations in Afghanistan context

An important factor to consider while interpreting the results is that the data were based on either microscopy or RDTs, both of which have varying sensitivities (WHO-FIND, 2009). With such low infection rates and an increased likelihood of mixed infection in districts showing patterns of co-infection (Imwong *et al.*, 2012). It was not possible to distinguish the proportion of observed cases at health facilities over the four year period that were a result of new infections or relapsing. Such analysis may require additional models of transmission, for example, incoporating the force of infection (Gemperli *et al.*, 2006, Yukich *et al.*, 2012). Another limitation of the maps presented here is that the effects of migration or travel between various regions were not incorporated into the modelling framework. A study in south-eastern Afghanistan showed higher asymptomatic infections in the migrant population (Nateghpour *et al.*, 2011). Modelling migration patterns at national level was beyond the scope of this study. It was assumed that individuals would seek treatment at the nearest facility or at least within a district or one of its neighbours.

3.6 Conclusion

This study demonstrates how HMIS data can be assembled, integrated and interpolated to identify district with high malaria burden spatially and temporally. Maps were produced at the

level of decision-making units, which are useful to the malaria control programme in assessing the changing burden of disease in Afghanistan, targeting malaria interventions at the population most at risk, and planning health resources. It is likely that Afghanistan's NMLCP faces a challenge in reducing the burden and management of *P. vivax* infections compared to *P. falciparum*. The districts identified with high burden can form the basis of targeting mass ITN distribution. For areas showing co-distribution of both species, mixed infections should be investigated and careful case management strategies adopted. Donor commitment to financing of the BPHS in Afghanistan since 2004 has had a positive effect on improving coverage of healthcare. The analysis of public health sector utilisation undertaken here suggests that the majority of the population is within two hours of a health facility, indicating an improvement in healthcare delivery and availability of services.

CHAPTER 4: Case study 3

Mapping the seasonal transmission of *P. falciparum* and *P. vivax* in Eritrea using HMIS

4.1 Eritrea context

4.1.1 Background

The burden of malaria in Eritrea has reduced significantly in the recent past. Both *Plasmodium falciparum* and *Plasmodium vivax* are found in Eritrea, although, *falciparum* is the major contributor for most malaria related deaths. *Plasmodium vivax* is estimated to constitute approximately 46% of the total burden in Eritrea (World Health Organization, 2013b). The 2012 health report indicates that malaria accounts for approximately 1.5% of inpatient morbidity (IPD) and 0.5% out-patient morbidity (OPD) in children under the age of five years (MoH, 2013). This is a ranking of about 10^{th} and 11^{th} respectively for all IPD and OPD cases, respectively. The estimates in ages above five years are slightly higher for OPD cases at 1.3% and 4.9% for IPD cases.

The 2014 WHO malaria report highlighted Eritrea as one of the countries in SSA that reduced malaria cases by over 75% between 2000 and 2013 (World Health Organization, 2014d). The 2012 MIS estimated a mean parasite prevalence of 2% nationally (Ministry of Health, 2012). Several factors have contributed to a decline in malaria burden in Eritrea. The RBM implemented aggressive malaria control between 1999 and 2004. During this period, there was an increase in coverage and use of ITNs and LLINs; larvae source reduction and the use of selective indoor residual spraying (IRS) of insecticides using dichloro-diphenyl-trichloroethane (DDT) and organophosphates (Malathion). By 2005, Eritrea had superseded the Abuja target of ITN coverage of >60%. There was a change in the antimalarial drug policy in 2007 from mono-therapies to use of ACTs and these are now available freely in the public sector. Compulsory use

of diagnostics at health facilities was also introduced to improve case management (Nyarango *et al.*, 2006, Mufunda *et al.*, 2007).

The national malaria control programme is targeting pre-elimination (MoH and RBM, 2005, Mufunda *et al.*, 2007) which requires case incidence to be less than 1 case per 1000 per year (World Health Organizastion, 2007). Routine HMIS could be used to target reactive Active Case Detection (ACD) usually deployed only during epidemics. Mass screening of population or ACD is yet to be adopted as a routine surveillance strategy (East Africa Roll Back Malaria Network (EARN), 2013). Therefore, there is a need to identify foci districts with high transmission to guide intervention and active case surveillance in order to achieve <1 case per 1000 population threshold.

This chapter assesses the spatial and temporal distribution of malaria transmission in Eritrea from 2010 to 2012 using routine data. It qualifies as a low malaria transmission country targeting pre-elimination. The analysis is aimed at identifying low and moderate risk areas in Eritrea to support the changing malaria strategies of pre-elimination in addition to identifying seasonal trends in transmission of both *P. falciparum* and *P. vivax*. In terms of methodology, analysis of healthcare utilisation is a prerequisite for incidence analysis. There is also an adjustment of denominator population using health seeking behaviour pattern (section 4.2) and adjustment at facility level for slide positivity as well as the rate of health facility reporting (section 4.4). The analysis of incidence incorporates the use of nonlinear functions for covariates to improve the smoothing in temporal domain as well as the prediction of slide positivity rates. This approach is aimed at improving estimates in the temporal domain while maintaining gains in spatial domain.

4.1.2 Geography

Eritrea lies in the horn of Africa with an estimated land surface area of approximately 123,200 km^2. It is divided into six administrative 1 units (*Zobas*) and 58 districts (*Sub-Zobas*). Eritrea is bordered to the East by Djibouti which is also in pre-elimination of malaria with low and unstable transmission (Noor *et al.*, 2011, Ollivier *et al.*, 2011). There is documented evidence of declining burden in Ethiopia in the south of Eritrea (Otten *et al.*, 2009, Jima *et al.*, 2012) while in Sudan, to the west, the regions of Kassala and Gedaref experience moderate transmission (Hay *et al.*, 2009a, Gething *et al.*, 2011b) (Figure 4.1).

Malaria transmission in Eritrea is highly seasonal and unstable, driven by the climatic conditions that vary from the hot and dry desert strip coastline of approximately 2,234 km along the red sea to the cooler and wetter highland areas (inland) with an average 60mm of rainfall annually. The main rainy season is between June and September. The extended Ethiopian highlands dominate the central regions descending to the east to coastal plain, hilly to the north reaching 3000m above sea level and rolling plains to the west (towards Sudan). Temperatures in the summer months vary between 40 ^{0}C to 50 ^{0}C at the coastal strip and 16 ^{0}C to 30 ^{0}C inland throughout the year. These climatic conditions are favourable for malaria transmission in the highland regions. The extreme low rainfall in the lowlands causes aridity which is unfavourable for both malaria transmission and for agricultural needs of the population.

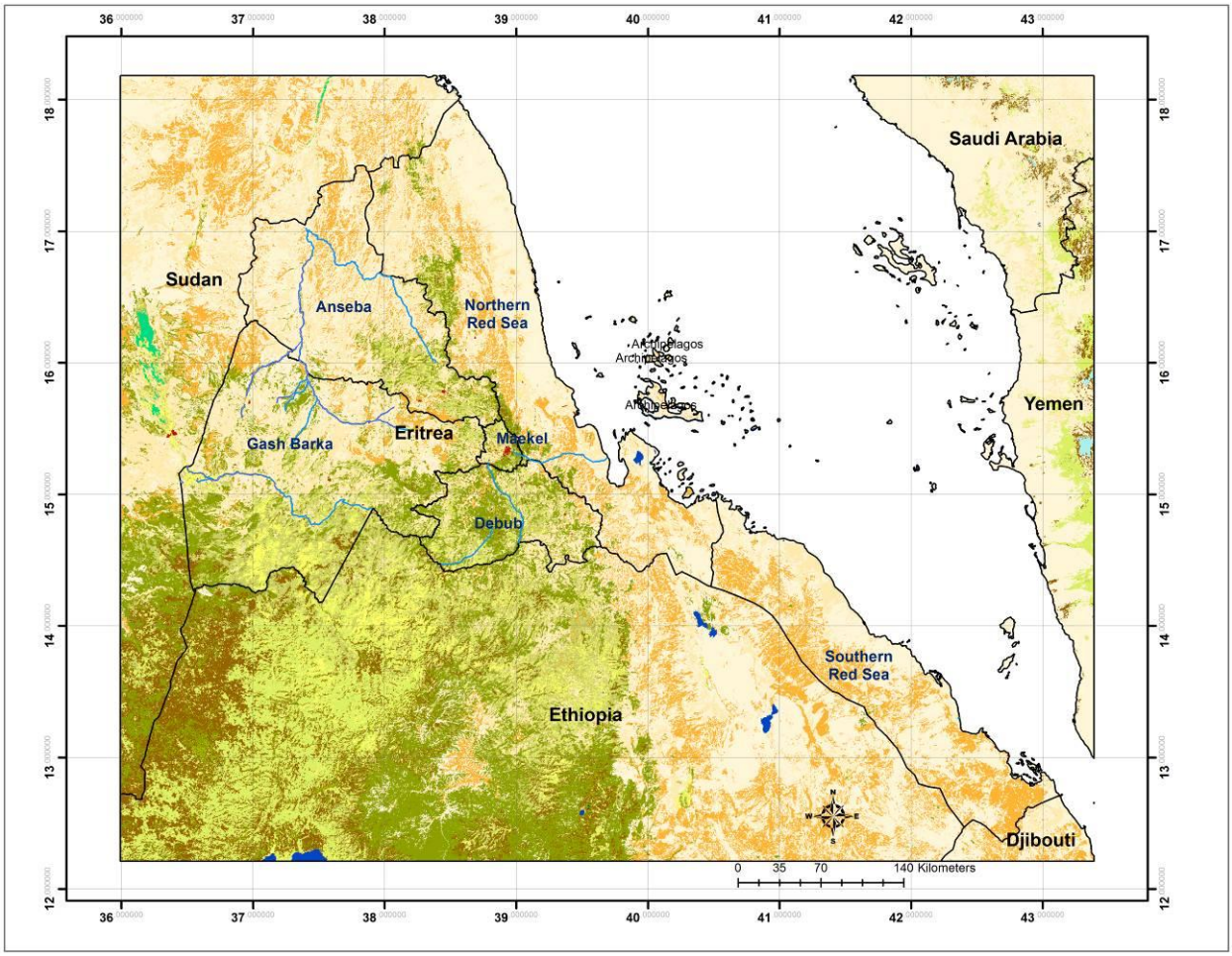

Figure 4.1: Map of Eritrea

4.1.3 Healthcare system in Eritrea: Organisation and delivery

The ministry of health is responsible for provision of preventive, curative and promotive healthcare services to the population. Like many other developing countries, the focus is on primary care to the rural populations in form of a Basic Healthcare Package (BHCP). This is outlined in the health policy which also aims at reducing disease burden, control communicable diseases, improve healthcare practices and implement a functional efficient healthcare system (MoH, 2010). In general, the healthcare system in Eritrea is hierarchical with management units set up at the national, regional and sub-regional level. The national level ensures the proposed

policies are implemented including: equitable distribution of health and social services; introducing national health insurance schemes; promoting healthcare services and good practices; and encourage participation of the private sector. Regional directories roles include the preparation of regulations, ensuring compliance with the national policy and encouraging regional level development. The coordination of development activities is carried out at district level including planning and implementation of policies.

Development and expansion of healthcare providers (hospitals, health centres and the health stations) was a priority of the ministry of health from 1995. Expanding the health system improved access and utilisation of healthcare facilities. For instance in 1996, 6% of population was estimated to be within 10km of a health facility which had increased to 70% by 2006 (MOH, 2008). By 2005 the number of health facilities was over 300 including both the public and private based facilities. There were approximately 21 major hospitals, four mini hospitals, 49 health centres and 178 health stations. The health stations and clinics form the primary-level facilities, while district (sub-*zoba*) level facilities and the major regional level facilities form the secondary and tertiary facilities respectively. The tertiary-level facilities are managed mainly by a *zoba* (regional) medical office and serve as referral centre for secondary and primary care facilities. They also form as focal point for the *zoba* conducting teaching and training, supporting operational research and *zoba* level as well as providing specialised services. National referral hospitals offer all services in a similar way to the *zoba* referral hospitals but with additional specialised services. In 2010, at least 340 facilities reported to the national HMIS. The MoH owns approximately 75.9% of the facilities, faith-based organisations (Eritrea Catholic

Secretariat (8.9%), Evangelical church (0.3%)), Private (2.4%), Industry facilities (9.2%) and other Non-governmental organisation constitute about 3.3%.

Primary care is provided through community-based health facilities with a catchment the size of a village (*kebabi*) (estimated population of 500 to 2000 people). Most of primary curative and preventive services are provided by CHA and skilled community health workers (CHW). Health station and clinics provide basic services for estimated 5000 to 10000 people. The health stations support the CHAs and CHWs and conduct regular training in addition to providing outreach services. Health centres and community based hospitals provide supervision to the health stations and clinics and are designed to serve a catchment population of approximately 30,000 to 50,000. They also form direct referrals to the clinics and health stations. The number of community health agents (CHA) has increased over the last few years. The CHAs provide basic curative and preventive services, for example, treatment of mild fever. Therefore a substantial number of people are treated at community level by the CHAs. It is estimated that between 2000 and 2004 CHA treated on average 50% of febrile events. There is improved training by year in practice as well as with change of health policy.

4.1.4 Health goals and progress on MDGS

Eritrea is a low income country with a Gross National Income (GNI) of 550 and low human development index (HDI) of 0.35 (World Bank, 2014). It was ranked 182 out of the 187 countries based on the HDI index. Current population is just under 7 million with a growth rate of approximately 2%. Health expenditure in 2011 formed approximately 2.6% of the total GDP (World Bank, 2014). Overall, life expectancy has increased to over 60 years which is higher than

the SSA average of 51 (MoH, 2010). The larges burden of disease comes from preventable and communicable diseases such as Acute Respiratory Infections (ARI) and maternal health-related problems. There is therefore an emphasis on health programs focusing on preventive activities in Eritrea. According to health assessment review in 2012, the top five cause of in-patient morbidity include Diarrhoea, HIV/AIDS, ARI, anaemia and heart diseases (MoH, 2013). Pneumonia, skin infections and ARIs are common in children under the age of five years. The under-five mortality rate was estimated to be 98 deaths per 1000 live births while infant mortality rate is 48 deaths per 1000 live births. This has declined significantly compared to 1992-1996 estimates of 121 deaths per 1000 live births and 67 deaths per 1000 live births respectively (National Statistics and Evaluation Office (NSEO) [Eritrea] and Macro, 2003).

Despite tremendous achievement in reducing malaria burden, the threat of resurgence remains due to combination of environmental or climatic factors and cross-border movement. While MDG targets remain on track in regard to infant and child mortality especially with declining burden of malaria, the maternal mortality ratio (MMR) remain a problem with estimate of 240 per 100,000 live births in 2011. Some of cultural factors and lack of skilled personnel contribute to maternal deaths (Sharan *et al.*, 2011). There are current attempts to train more skilled community health workers and to increase ANC coverage. Health worker population ratios indicate the ratio of number of doctors to population is approximately 0.48/10,000 which is close to the WHO limit of 1/10,000. The nurses ratio of 3.2/10000 and 6.5/10000 for associate nurses is within the advisory limits (East Africa Roll Back Malaria Network (EARN), 2013). This staffing challenges affect quality and competency of healthcare professionals.

In summary, Eritrea has made strides on some MDGs such as reducing the malaria burden, but remains behind on other MDGs such as poverty. Most population remains poor with heavy reliance on agriculture which is undermined by drought. The next section reviews history of malaria control.

4.1.5 History of malaria control in Eritrea

Malaria control in Eritrea is coordinated by the National malaria control programme which was established in 1999 following epidemics between 1997 and 1998. IRS started in 1965 during the GMEP to the late 1960s when it was discontinued due to instability during the occupancy of Ethiopia from 1952-1991. During the period of instability, much of country's infrastructure was destroyed and malaria control activities nearly stopped. A major re-building programme was initiated after independence declaration in 1993, which involve the re-construction of healthcare services and improving provision. Between 1997 and 1998 a severe epidemic was reported that resulted in over 200,000 cases. Previous reports indicate malaria cases in 1995 were less than 90,000. The number of impatient deaths reported during this epidemic surged to over 500. Subsequently, the National malaria control program (NMCP) was constituted with support from the World Health Organisation, The U.S. Agency for International development (USAID) and the World Bank.

A five year malaria attack phase was launched which commenced with DDT spraying in three *Zobas* worst hit during the epidemic (Debub, Garsh Barka and the Northern Red Sea). The RBM also supported the distribution of bed-nets through clinics to high risk groups (such as pregnant women) and through social marketing (Eisele *et al.*, 2006). Integrated vector management (IVM)

was rolled out to other regions (Anseba, Maekel and Southern Red Sea). As a strategy of managing resistance, DDT spraying in Debub was replaced partly by malathion towards the end of transmission season due to its short half-life (WHO/AFRO, 2007). As a result of these efforts, between 1999 and 2003, the malaria burden was halved. Funding from the GFTAM commenced in 2003. The first GFTAM disbursement period was from November 2003 to March 2004 and was instrumental in scaling up the coverage of ITNs such that by 2005, Eritrea had exceeded the Abuja targets of greater than 60% ITN coverage nationally (Eisele *et al.*, 2006). In 2004 RBM conducted an assessment survey to identify key priority areas and ways of consolidating malaria control gains. Some of the weaknesses identified were in case management, late presentation of cases at periphery health facilities for treatment, a need to increase the number of CHAs, sustaining ITN distribution and environmental management of IRS, improvement on awareness and integration of IMCI amongst other logistical issues.

The 2005 to 2009 malaria policy focused on consolidation and strengthening the health system on issues around the use of RDTs, training of CHAs and improving logistic supply system to guard against stock outs at health facilities. About USD 13 million was budgeted for case management, prevention, epidemic detection and prevention, operational research, program management and monitoring. Most of these funds were proposed during the sixth round of the GFTAM (The Global Fund, 2013b). About USD 11.05 million (94%) of proposed amount was disbursed commencing in November 2007 and that ended in 2012. This was deemed adequate for the malaria activities proposed during the consolidation phase. For instance, only 75 deaths were attributed to malaria in 2009 and over three million ITNs had been distributed by 2012. In 2007 there was a change in the antimalarial drug policy from the monotherapies. Chloroquine +

Sulfadoxine pyrimethamine (CQ+SP) was introduced as first line drug for uncomplicated malaria while Artesunate + Amodiaquine (AS + AQ) was introduced for first line treatment of confirmed falciparum malaria with quinine (QN) used if there was treatment failure. Chloroquine and Primaquine remained as first line treatment for confirmed vivax malaria. Progress has also been made in regard to the scale-up of LLINs. The 2012 MIS suggested availability of 1 LLIN for every 0.5 people. This, however, still falls short of the WHO recommendation of 1 LLIN for every 1.8 people (World Health Organization, 2012b). There was an 86% estimated ownership of LLINs and use at 67.4% for children under the age of five years in the same survey. The use of LLIN in all age population was 55% from the MIS (Ministry of Health, 2012). The use of diagnostics has also increase at health facilities nationally which has had an impact on case management by identifying and treating febrile cases (Nyarango *et al.*, 2006, Mufunda *et al.*, 2007). As a result of these activities malaria in Eritrea is ranked 10th and 11th in the IPD and OPD for the under-fives. Thus, the NMCP is re-orienting the programme to pre-elimination. The reduced burden has positive implications to indicators such as child mortality in Eritrea. The next section investigates access and utilisation of the public health sector in order to derive denominators of quantifying malaria incidence at facility level.

4.2 Analysis of coverage and utilisation of health services for treatment of fever

A national HMIS was first created in 1997 in Eritrea and this has strengthened over the years in terms of data collection and standardization. By 1998 a computerized information system had been set up for data storage and analysis. The health metric Network (HMN) continues to provide technical support to the MOH and in 2007 an assessment of data quality showed high accuracy (80.6%), timeliness (88.1%) and completeness (>90%) (MOH, 2008). There has been

no formal assessment of Eritrea healthcare service access and utilisation at a national level except for the designed population provider ratios. This section quantifies geographic access to health facilities in Eritrea in relation to physical distance and topography. The main objective is to assess health facility coverage and quantify utilisation. A second object is to delieanated catchments and estimate population within catchments for incidence analysis.

4.2.1 Data

4.2.1.1 Health facility database for Eritrea

A health facility database was obtained from the national malaria control programme. Health facilities in Eritrea were mapped using non-differential handheld GPS receivers as the same time with the computerized system. The process was supported by the health Metrics Network (MOH, 2008). There were approximately 287 public-based health facilities assembled majority owned by the MoH (85.7%). 2.4% were owned by the faith-based organisations (mainly the catholic and evangelical church) and a further 1.7% owned by the NGOs. The rest were industrial based facilities (Table 4.1).

Hospitals were ranked the highest with one as a referral facility, 19 regional based hospitals and five mini hospitals. Other levels included the health centres, health stations, clinics and specialised hospital. Providers included the government (majority), charity organizations, private individuals and other government agencies such as the police and ministry of defence (MOD). Each facility was linked to an administrative area (Administrative 1 boundary and district) (Figure 4.2).

Table 4.1: Public-based health facilities in each *Zoba* in Eritrea by type

Zoba	National Referral Hospital	Hospital	Mini Hospital	Health Centre	Health Station	Clinic	Special Hospital	Total
Anseba	0	1	0	9	26	3	0	39
Debub	0	3	2	11	47	2	0	65
Gash Barka	0	3	0	13	51	5	0	72
Maekel	1	7	1	10	24	2	4	49
Northern Red Sea	0	4	0	11	29	4	0	48
Southern Red Sea	0	1	2	0	11	0	0	14
Total	**1**	**19**	**5**	**54**	**188**	**16**	**4**	**287**

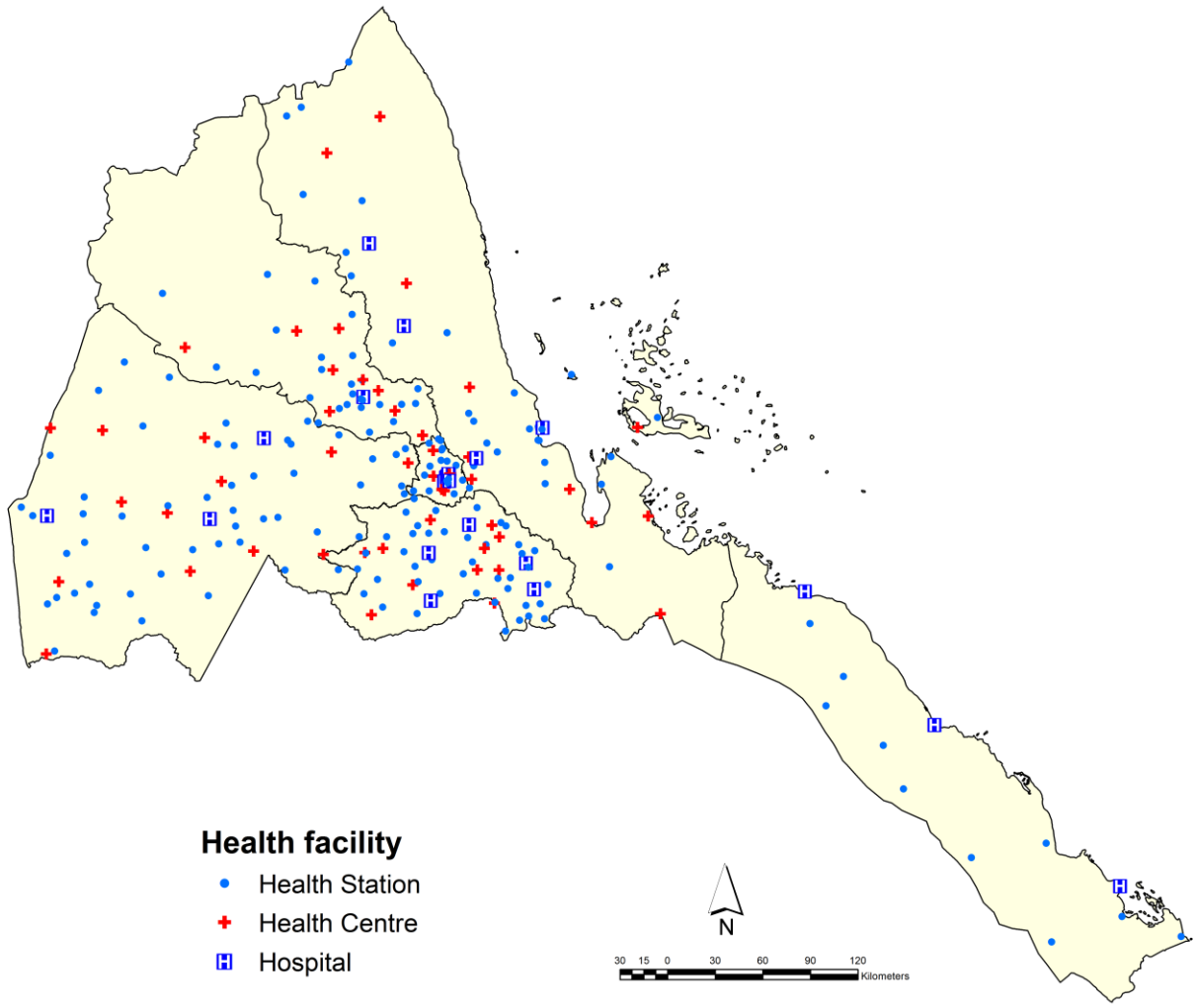

Figure 4.2: Distribution of health facilities in Eritrea
Map of health facilities in Eritrea by type. Hospitals (*n*=25), Health centres (*n*=54 and Health stations or clinics (*n*=204)

4.2.1.2 Eritrea Population map

An population surface (Figure 4.3) was obtained from Worldpop at 100 x 100 m spatial resolution (Worldpop, 2010). The methodology in modelling population distribution follows a

similar approach to the Namibia population surface and other countries in SSA (Linard *et al.*, 2012). In brief, a Dasymetric approach was used (Monmonier and Schnell, 1984, Briggs *et al.*, 2007), that involve the redistribution of census data (Tatem *et al.*, 2007). No population census that has ever been conducted in Eritrea; therefore, most population estimates were based on government estimates and publication from the United Nations population division (UN Population Division, 2013). The inclusion of land use and land cover data from GlobCover (Arino *et al.*, 2007) improves the representation of habitable land. The weights, calculated based on the density of habitable areas, were used to distribute census estimates at regional block level. The method assumes a direct relationship between the population density and land cover classes. Recent improvement of this technique using the random forest modelling approach suggest a linear relationship between the selected land cover covariates and population (Worldpop, 2010). The resulting national population map was then projected using the United Nations' (UN) inter-censual growth rates (UN Population Division, 2013).

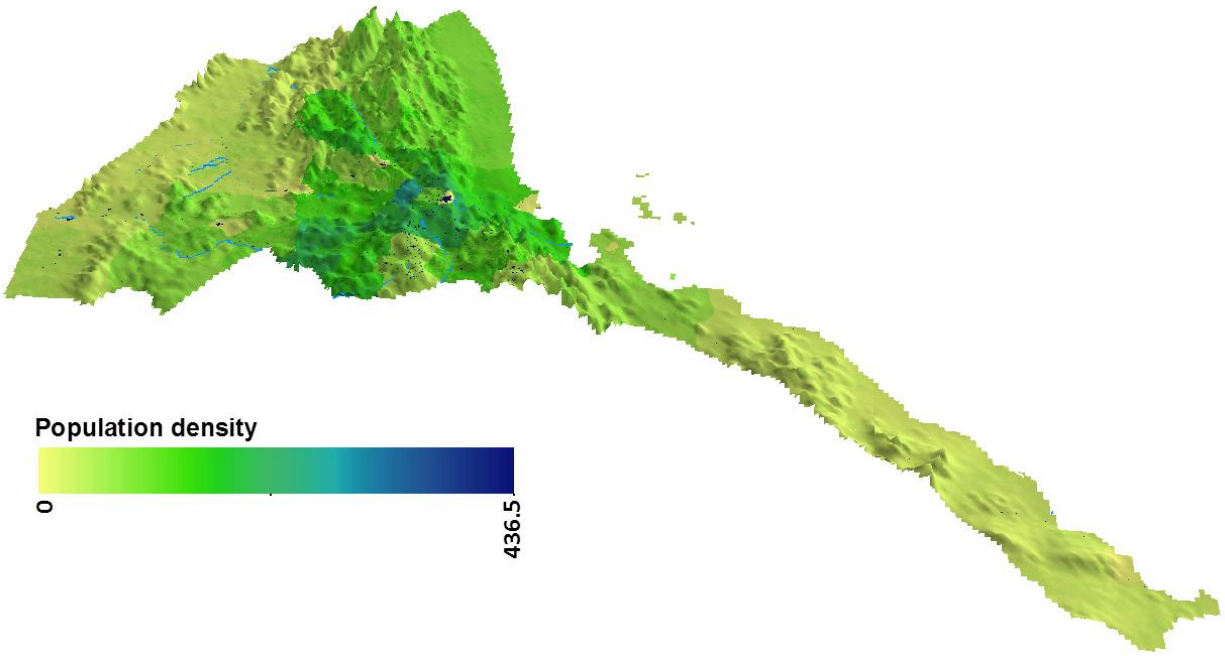

Figure 4.3: Population density map of Eritrea for 2010
Population density map superimposed to elevation map of Eritrea showing majority of population reside in central highland regions.

4.2.1.3 National representative surveys in Eritrea

Eritrea has to date conducted at least four nation representative household based surveys. Two DHS were conducted in 1995 and in 2002. The first MIS was conducted in 2008. The 2012 MIS was used in this study for analysis of fever treatment in the public sector. The MIS focused on malaria, and was powered primarily to measure the impact of interventions such as insecticides treated nets (ITNs) (Roll Back Malaria Monitoring and Evaluation Reference Group *et al.*, 2005).

3,845 households (8,533 individuals) in all age cohorts were sampled in the MIS carried out between September and October 2012 in four *zobas* namely: Anseba, North Red Sea, Debub and Garsh Barka. A two-stage sampling design was adopted (RBM-MERG, 2008) where villages (or clusters) were first selected based on the number of households in a probability-to-proportional-

size approach (PPS) and households were selected randomly at the second stage within the selected clusters (Ministry of Health, 2012). The geographic locations of the clusters were established using a handheld GPS receiver. A cluster comprised of approximately 19 households (variation of minimum 6 to a maximum of 21 households). The surveys provide information on fever prevalence amongst all-age cohorts and treatment seeking patterns in different sectors. The management of fever 2-weeks prior to the MIS was recorded for household members that reported a fever episode. The individual-level data were linked to administrative regions including information on type of residence (urban or rural status) and fever.

4.2.1.4 Ancillary national-level GIS data

Data on Land use and Land cover, elevation, road and rivers for Eritrea was downloaded from an online repository (DIVA-GIS, 2011) and from the data exchange platform for the Horn of Africa created by the UN (UNEP, 2002). Elevation (Figure 4.4) was obtained from the Advanced Spaceborne Thermal Emission and Reflection Radiometer-Global Digital Elevation Model (ASTER-GDEM) (NASA, 2012). ASTER-GDEM. The elevation data had a 30 m spatial archived using 1° by 1° tiles in GeoTIFF format. A land cover surface for 2009 was obtained from the Medium Resolution Imaging Spectrometer (MERIS) GlobCover product (ESA, 2010) at 300 m spatial resolution.

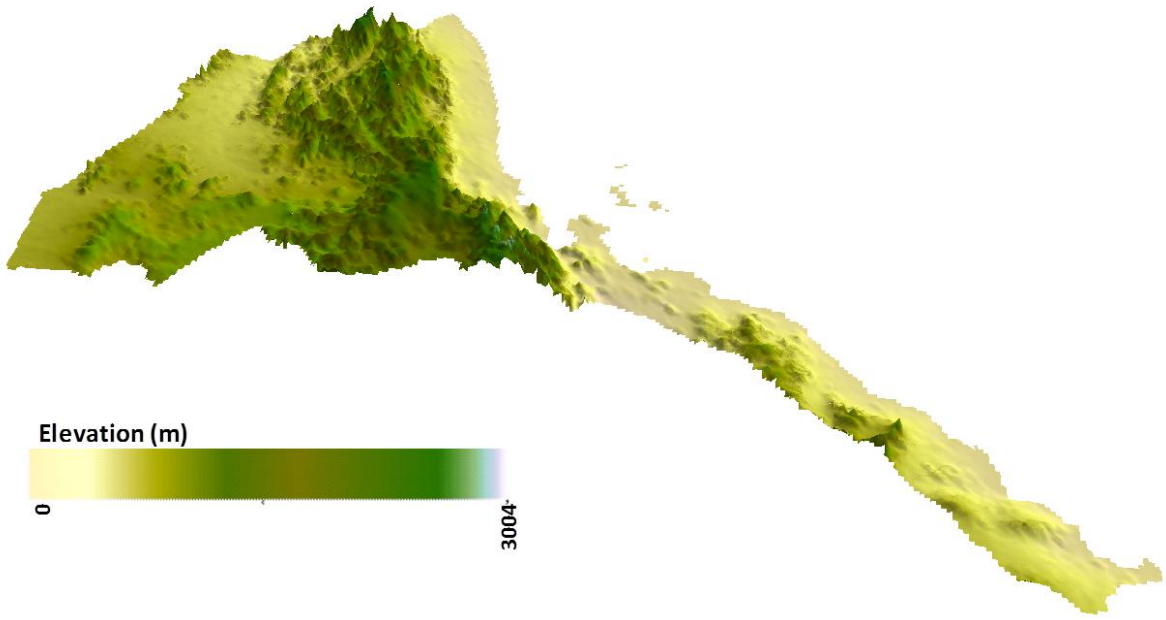

Figure 4.4: Digital elevation model (DEM) for Eritrea

4.2.2 Developing surface of travel time and probability of attendance for fever treatment in Eritrea

Ancillary data were combined to generate a cost surface of travel times between facilities and population locations in AccessMod (version 3.0) (Ray and Ebener, 2008). The cost surface was generated by assigning various travel speeds (Table 4.2) to land cover, slope (from elevation) and roads while rivers were used as barriers (Ray and Ebener, 2008). Travel speeds assigned to various land use or land cover types were similar to Namibia and Afghanistan analysis (Ray and Ebener, 2008, Alegana *et al.*, 2012, Huerta Munoz and Kallestal, 2012) and were extracted at cluster locations from the MIS for the 8,533 individuals. Overall, 1,247 individuals in all age cohorts reported at least a single fever episode two weeks prior to the survey while 60·8% reported to have sought treatment in public sector (Ministry of Health, 2012). Effects of residence (urban =1 or rural =0), gender (Male or female) and age were evaluated in a

generalized multiple regression model on the reported pattern of health facility attendance. Of which, only residence (regression coefficient -1·0, 95%CI: -1·4 – -0·8, p<0·001) was significant and was subsequently used in the fitted model. The effect of age (regression coefficient -0·1, 95%CI: -0·2 – 0·1, p=0·5) and gender (regression coefficient 0.02, 95%CI: -0·2 – 0·3, p=0·8) were not significant and were, therefore, dropped from the analysis. The travel times were used in the logistic regression model of the form $Y = C/(1+e^{(A-x)/B})$ (Pinheiro and Bates, 2002) to estimate probability of health facility attendance.

Table 4.2: Description data for modelling healthcare utilisation in Eritrea.
Various data and their sources used as inputs in calculating travel time to the active public health facilities in Eritrea.

Map Layer	Description	Classification	Speed (km/h)	Mode[1]
Land use/ land cover	Spatial representation of all different land use and land cover types. Two land cover grids were processed (1) a basic land cover grid (2) a combined grid that incorporates roads and rivers with the same resolution as the DEM	Irrigated, rain fed, mosaic or vegetated croplands	4.0	Walking
		Open or closed broadleaved, needle leaved, deciduous or evergreen tree cover	4.0	Walking
		open or closed mixed broadleaved forest/tree cover	2.5	Walking
		Mosaic, closed to open grassland/shrubland	2.5	Walking
		Sparse Vegetation	2.5	Walking
		Open or closed broadleaved regularly flooded	1.0	Walking
		Artificial/urban areas	30.0	None
		Bare areas/desert	1.0	Walking
		Ice/ permanent snow	0	None
Roads	Classified into five categories; class A (highways), class B (secondary roads), tertiary Class C and Class D roads as well as street level urban roads. Each road class was assigned a different speed limit.	Class A roads	60.0	Motorised
		Class B roads	40.0	Motorised
		Class C roads	20.0	Cycling
		Class D roads	5.0	Walking
		Street level roads in urban areas	30.0	None
Rivers	GIS layer representing barrier to movement. Only major rivers were used to reduce the complexity of running the algorithms	NA[3]	0	NA[2]
Digital elevation model	Altitude values that are used in anisotropic calculation; Original DEM 30 m ASTER grid; resampled to 1 km pixel size	Degree of Slope (< 0.5^o)	4.88	Walking
		Degree of Slope (5.0^o)	3.71	Walking
		Degree of Slope (10.0^o)	2.71	Walking
		Degree of Slope (20.0^o)	1.41	Walking
		Degree of Slope (30.0^o)	0.66	Walking

3. Assumed mode of travel to health facility, as either walking on foot, cycling, using motorise transport as on roads or a combination of the different modes. Anisotropic movement for walking based on Tobler's equation, (V=6*exp(-3.5abs[Tan(slope in degrees/57.296) + 0.05]) (Tobler, 1993) where V is the speed with slope derived from DEM or for cycling (Walter, 2008), was applied for traversing across a pixel. For example, on a flat terrain, the walking speed is 5.0 km hr^{-1}.
4. NA is an abbreviation for ' Not Applicable'

Figure 4.5 shows the universal (all facilities) distance-decay model of treatment seeking pattern based on the travel time and the reported fever treatment. Note that following the Afghanistan analysis, here, stratification by facility type was not used for public health sector coverage analysis. Treatment-seeking behaviour reduced rapidly after approximately 180 minutes. The residual standard error was 0.02 while the sum of squared residuals equal was 0.01 indicated a good model fit to the data. The coefficients of the distance-decay curve were all significant with $p<0.001$. The limiting coefficient was C (0.73) with the other two coefficients: the asymptote as A (3.83) and decay parameter B (-0.42).

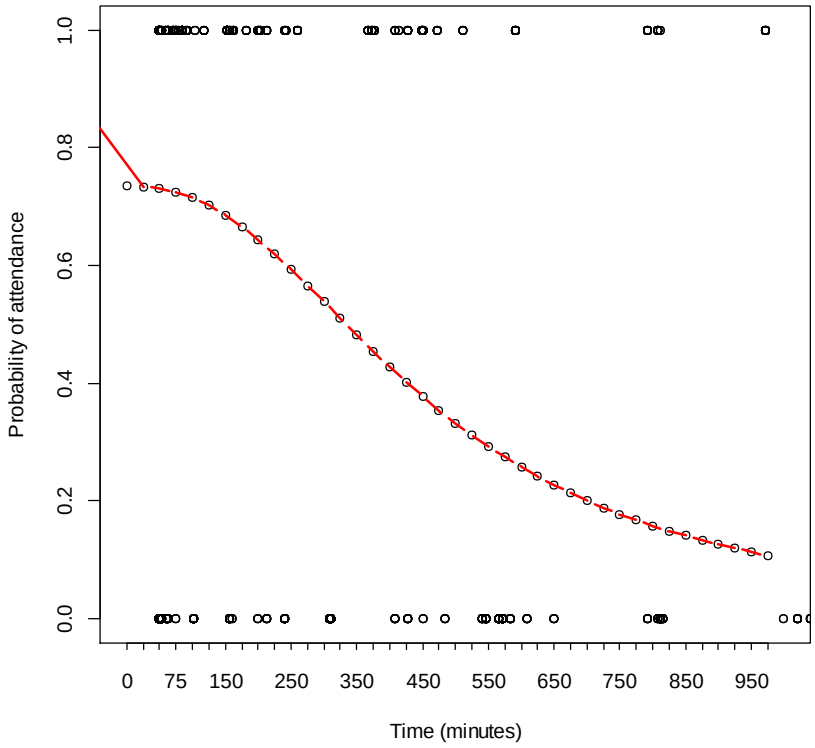

Figure 4.5: Distance decay curve of probability of fever treatment in Eritrea
Distance decay model based on the reported fever treatment in the 2012 MIS survey for Eritrea showing probability of treatment (*y*-axis) against travel times (*x*-axis). The model parameters were all significant at p<0.0001 with limiting factor C as 0·73 [95% CI 0·71 – 0·74], the asymptote 3·83[95% CI 3·80 – 3·86] and the decay parameter as -0·42 [95% CI -0·45 - -0·40]. The attendance pattern (1 = attendance and 0 = non-attendance) is also superimposed on the decay curve

4.2.3 Developing health facility catchments in GIS and assessing coverage of health facilities

A gridded surface of probability of health facility attendance was derived at 1 x 1 km spatial resolution based on the distance-decay curve. The distance decay model based on the universal facility list was selected to zone catchment areas while limiting the maximum travel time to 180 minutes. Population counts in various catchments were extracted using the catchment boundaries and multiplied by the probability of attendance for fever treatment to derive the population-weighted surface of health facility utilisation. Similarly, the population outside the threshold catchment boundary was estimated. A fever burden map was derived by combining population map with estimates of fever prevalence from the MIS. The number of fever cases in all age cohorts within each catchment was extracted based on the fever burden map. The number of fever cases likely to attend a public health facility was calculated by multiplying the estimated burden by the probability of attendance.

4.2.4 Results of analysis of coverage and utilisation of health facilities

3.2.4.1 Fever prevalence in 2012 and treatment seeking behaviour

From the MIS, approximately 8,533 individuals in 96 clusters were interviewed. Fever prevalence in all age-cohorts was 15.2% (95% CI 12.9 - 17.6). Of those with fever, the proportion that sought treatment in public sector was 56.2% (95% CI 53.5 – 59.0).

3.2.4.2 Probability of attendance for fever treatment

Figure 4.6(a) shows the gridded surface of attendance for fever treatment to all health facilities at 1 x 1 km spatial resolution. The gridded surface was used to delineate catchments shown in Figure 4.7 (b). Out of the estimated total population in 2012 (5,541,112) 67·8% were estimated

to be within 180 minutes of a health facility and therefore within the catchment. The estimated number of fever cases in 2012 based on the MIS prevalence and population map was 688,700 of which 515,936 were within the derived health facility catchments (Table 4.3 and Figure 4.6 (b)). Of the four regions covered by the MIS, the burden was higher in Debub and less in Anseba. Approximately 25% of the estimated fevers were outside the catchments of the health facilities and 61% of those in catchments were likely to seek treatment for fever. This means that approximately 39% of the population in catchments was not likely to seek treatment in the public sector. The modelled mean probability of attendance by region was highest in Maekel and lowest in Southern Red Sea (Table 4.3).

Table 4.3: Estimated population in Catchments by region
Estimated population in 2012 and the estimated counts in catchments by region as well as those in the catchment likely to seek treatment for fever based on the probabilistic measure of attendance. Maekel and Southern Red Sea were not sampled during the MIS survey.

Region	Estimated Population 2012	Estimated population 2012 (percentage) in health facility catchments	Fever prevalence (95% CI) in all age cohorts from the 2012 MIS	Estimated fever burden in health facility catchments based on the MIS prevalence	Fever cases (Percentage) in catchment likely to seek treatment
Anseba	734,948	525,349 (71.51)	12.4 (8.3 - 16.4)	60,343	35, 475 (58.8)
Debub	1,341,639	1,332,810 (99.3)	20.7 (16.4 - 13.0)	237,107	157,128 (66.3)
Gash Barka	1,297,092	1,177,388 (90.8)	12.7 (8.1 - 17.3)	133,913	78,132 (58.3)
Maekel	498,550	498,550 (100.0)	-	-	
Northern Red Sea	1,449,685	771,772 (53.2)	10.7 (7.5 - 14.0)	84,572	43,228 (51.1)
Southern Red Sea	219,198	99,379 (45.3)	-	-	
Total	**5,541,112**	**4,405,325 (79.5)**	**15.2 (12.9 - 17.6)**	**515,936**	**313,964 (60.9)**

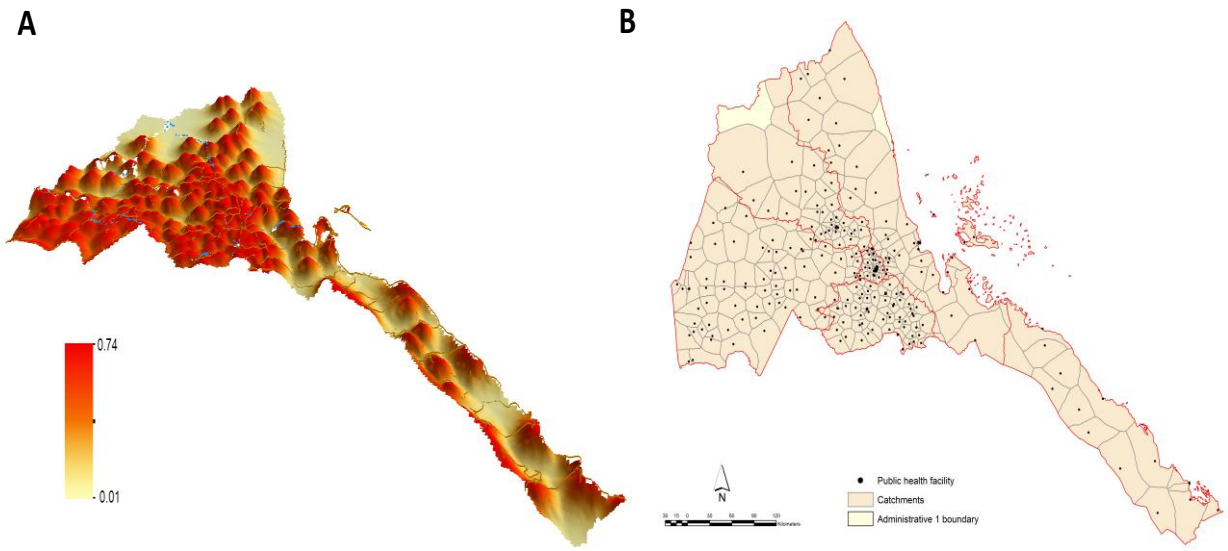

Figure 4.6 Probability of health facility use for fever treatment and delieanated catchments
Figure 4.7 **(a)** shows the modelled probability of attendance for fever at 1 km by 1 km resolution using logistic regression model based on the 2012 MIS. The lowest probability was 0.01 and the highest probability was 0.73 **(b)** the health facility catchments at a threshold travel time of 180 minutes

Table 4.4 provides an overall assessment of coverage of health facilities in relation to population distribution. In general there was 1 health facility (of any type) for every 20, 000 people. This ratio is likely to be more for hospitals given there was on average more than 2 hospitals in each region except in Anseba and the Southern Red Sea regions (Table 4.1, Page 205). Approximately 27% of the population was within 30 minutes of the nearest public health facility, while majority (73%) where within two hours of travel to the nearest health facility. There was poor geographic access in the Northern Red Sea and the Southern Red Sea regions where 59% and 66% of population, respectively, were at distance greater than two hours of the nearest health facility. In addition, approximately 9% of population in Northern Red Sea and 4% in Southern Red Sea were at distances within 30 minutes to the nearest health facility. In Maekel, 100% of the population was within two hours of the nearest health facility.

Table 4.4: An assessment of coverage of health facilities in relative to population and travel time in Eritrea in 2012

			Percentage of population to nearest health facility					
Region	**Population : Facility Ratio**	**Estimated Population 2012**	**30 minutes**	**1 hour**	**2 hours**	**> 2hours**	**3 hours**	**> 3hours**
Anseba	19863:1	734,948	21.4	39.5	64.1	35.9	71.5	28.5
Debub	20963:1	1,341,639	41.5	77.8	98.1	1.9	99.3	0.7
Gash Barka	18530:1	1,297,092	19.0	42.7	82.4	17.6	90.8	9.2
Maekel	12464:1	498,550	77.0	99.0	100.0	0.0	100.0	0.0
Northern Red Sea	32215:1	1,449,685	8.8	17.7	41.2	58.8	53.2	46.8
Southern Red Sea	15657:1	219,198	4.4	13.6	33.9	66.1	45.3	54.7
Total	**20523:1**	**5,541,112**	**26.7**	**48.2**	**72.7**	**27.3**	**79.5**	**20.5**

4.3 Discussion on coverage and utilisation of healthcare facilities for fever treatment

This section assessed coverage of healthcare facilities in Eritrea and their utilisation for fever treatment based on the MIS data. This study did not assess the overall expansion of healthcare system in terms of service provision over the years since there was no documented baseline survey to compare the results against. However, for effective service planning it is essential to outline systematically the coverage and utilisation of health facilities. Overall, 72% of the population were within 2 hours of the nearest health facility. Thus, the results suggest that the construction of health facilities since the mid 1990s has had a positive impact on increasing coverage. To assess utilisation, travel times were calculated between the facility location and population centres (clusters) and based on a 3 hour cut off (180 minutes). 60% of the estimated fever cases where likely to be treated in the public sector. The probability of seeking treatment was highest in Debub and Garsh Barka where there was a higher concentration of health facilities and population. The estimated number of fevers within the health facility catchments was 688,700 of which 61% were likely to be treated in the public sector. Overall, 172764 (25%) febrile cases were unlikely to use at a public health facility.

Coverage of health facilities was congruent with population distribution. The facility to population ratio was 1:20,000 at the national level. The ratio was better in Maekel with one health facility for every 12,000 people and the worst in the North Red Sea region (one facility for approximately 32,000 people). There was, in general, a balance between population distribution and the location of health facilities. 48% of the population was within 1 hour of a public-based health facility while approximately 7% were located in areas more than three hours from the nearest health facility. Majority of the population in Maekel, Debub and Garsh barka where located within 2 hours of the health facility.

Geographic access was poor in the Southern Red Sea and the Northern Red Sea regions where 47% and 55% of the population, respectively, were at distance greater than three hours of the nearest facility. In these two regions combined, only 8.2% of population was within 30 minutes of the nearest health facility. These two regions are sparsely populated generally with population density of 0.02 and 0.11 per square kilometre respectively. In Other regions in Maekel, Gash Barka, and Debub >80% of population were within two hours of the nearest health facility. Geographic access in these regions was higher compared to the regions bordering the Red sea. The subsequent analysis on modelling disease burden is dependent on a good distribution of health facilities in addition to availability of health reports.

Utilisation of health facilities was accessed using the 2012 MIS conducted during the malaria season. The reported fever prevalence (15.2% (95% CI 12.9 - 17.6)) was higher compared to the other low transmission cases studies in Namibia or Afghanistan (Ministry of Health, 2012). The MIS report suggested a high recognition of fever as a symptom for malaria during the

transmission (amongst the sampled population). More than 80% of the interviewed population associating fever with malaria. It is likely that the majority of cases are treated within the public health sector that includes a network of community health agents linked to lower tier facilities. The report also showed that 61% of children under the age of five years sought treatment for fever at public health facilities (Ministry of Health, 2012) and estimated use in all age cohorts was 55%. The CHAs play an important role in treatment seeking behaviour in Eritrea at community level since they can treat uncomplicated febrile cases. Training and expanding the network of the CHAs is entrenched in the national health strategy (MoH, 2010). Secondly, the size of the private sector is relatively small with private-based health facilities constituting about 3%. Majority of the public-based health facilities (76%) are managed and owned by the MoH (MOH, 2008).

The distance decay model estimated 61% of the estimated fever cases in 2012 were likely to have been treated in the public sector. Thus, the estimated rate of utilisation in the public sector was similar to the observed utilisation rate of the public sector with only a five percent difference based on all age cohorts. Differences in estimates could be as a result of sampling in the MIS or through error propagation due to combination of different data sets. Moreover, there could be fine scale variation at regional or local level compared to the generalized model in this study. Distance remains one of the factors affecting utilisation in Eritrea even for other services such as ANC (Sharan *et al.*, 2011). For example, findings from the study by Sharan and colleagues emphasised the disproportionate access to maternal services and highlighted distance as a major factor affecting utilisation. The relationship between distance and poor health outcomes has been demonstrated in many other studies (Al-Taiar *et al.*, 2008, O'Meara *et al.*, 2009, Moisi *et al.*,

2011), In Ethiopia there was an increased risk of death associated with an increased distance from a health facility (Okwaraji *et al.*, 2012). The recommended travel time from a health facility varies by location but longer distances may affect high risk groups, for example, the risk of miscarriage in pregnant women. Thus, CHAs play an important role in providing primary care to the marginalised population. CHAs are linked to lower-tier facilities (clinics and health stations) and access to these providers may probably alter the utilisation at health facility level.

In addition, the rate of distance-decay based on the derived travel time, was similar to the one modelled for Namibia (Alegana *et al.*, 2012). This decay was less rapid after seven and half hours. It is likely that some population in the regions along Red Sea travel a longer distance to facility compared to other regions. The study did not investigate the pattern of utilisation between various facility types. Given there at least two to three major hospitals in each *Zoba*, it is likely that these are referral centres for severe or complicated malaria within the *Zoba*. Perception of quality of services offered at the referral hospitals could also be a factor affecting utilisation (Trani *et al.*, 2010, Trani and Barbou-des-Courieres, 2012). Other limitations such as social-economic status, costs or cultural preferences were not evaluated. The advantage of this study was the use of all-age population cohorts increasing representation of general utilisation patterns in the population. Although, the private sector constitutes just a minority of providers in Eritrea, their inclusion in future studies may improve understanding of coverage and utilisation. Here, the quantification of catchment population at facility level served as pre-requisite in modelling malaria incidence.

4.4 Mapping malaria incidence and slide positivity rates in Eritrea

4.4.1 Assembly of malaria data (2010-2012)

Malaria data (Total Cases (TC)) for *P. falciparum* and *P. vivax* were obtained from the National Malaria Control programme (NMCP) for a three year period from 2010 to 2012. This data was extracted from the national HMIS database which reports on all illness and deaths recorded at a health facility. Facilities reported the number of malaria cases per month diagnosed clinically or confirmed through laboratory test (RDT or Microscopy). In total, 265 public based health facilities reported malaria data in all age groups from the 270 health facilities nationally excluding a few reproductive-based facilities (maternity and nursing homes) and special facilities (Psychiatric and Rehabilitation centres), stand alone voluntary counselling centres (VCT), dental clinics and educational or vocational centres. These facilities comprised of: two national referral centres; 21 hospitals (mini and regional hospitals); 53 health centres or community facilities; and 194 health stations. The overall reporting rate, calculated as a proportion of received reports over the expected number, was high (85.4%) (8301 reports of the expected 9720).

Although most cases in Eritrea were confirmed using a parasitology test prior to treatment, few were diagnosed clinically. Only 8.7% of possible monthly data points were missing and these were imputed as *NAs*. Adjustment of data at facility level was similar to Namibia and Afghanistan. This included the rate of utilisation, the rate of reporting at slide positivity rates. Table 4.5 shows a summary of the assembled case data for *P. falciparum* and *P. vivax* for the three-year period. The slide positivity rate (SPR) was used to adjust the suspected cases, where parasitology had been used and malaria species isolated. This was necessary to avoid underestimating incidence (if suspected unconfirmed cases are ignored) or overestimating

incidence (where true cases are treated as a summation of suspected and confirmed case while ignoring the SPR at the facility).

$$SPR = \frac{Confirmed \quad cases}{Total \quad number \quad exa\min ed}$$

$$TC = Confirmed \quad Case \quad + \quad (Suspected \quad cases \quad \times \quad SPR)$$

Overall, there were more cases in 2010 compared to 2011 to 2012. The highest number of cases was in Garsh Barka. Slide positivity rate was also highest in Gash Barka, for example: 64.4% in 2010 and 47.0% in 2012 in this region. SPR was low Northern Red Sea and the Southern Red Sea. The proportion of suspected cases was also small compared to the overall number of malaria cases in Eritrea.

Table 4.5: Summary of assemble HMIS data in Eritrea.
Assembled *P. falciparum* (P*f*) and *P. vivax* (P*v*) cases in Eritrea by region and year. The slide positivity rate is shown in brackets.

	Malaria cases (Slide positivity (%))								
	2010			2011			2012		
Region	**P*f***	**P*v***	**Suspected**	**P*f***	**P*v***	**Suspected**	**P*f***	**P*v***	**Suspected**
Anseba	1,415 (23.2)	578 (9.5)	159	1,094 (10.1)	837 (7.7)	472	1,903 (13.0)	827 (5.6)	28
Debub	9,928 (40.7)	3,205 (13.1)	447	3,923 (7.9)	3,959 (7.9)	2,561	6,624 (14)	3,476 (7.4)	1,068
Gash Barka	23,767 (64.4)	2,576 (7)	7,770	8,438 (25.7)	4,914 (15)	14,212	18,769 (47.0)	5,617 (14.1)	5,060
Maekel	489 (9.2)	254 (4.8)	560	516 (7.2)	493 (6.9)	807	596 (5.8)	526 (5.1)	843
Northern Red Sea	304 (7.9)	128 (3.3)	6	295 (11.8)	149 (5.9)	1	241 (5.4)	227 (5.1)	271
Southern Red Sea	6 (5.2)	0(0)	24	15 (23.4)	0(0)	71	6 (7.3)	5 (6.1)	160
Total	**35,909 (46.8)**	**6,741 (8.8)**	**8,636**	**14,281 (13.8)**	**10,352 (10)**	**18,122**	**28,139 (24.1)**	**10,678 (9.2)**	**5,238**

4.4.2 Assembly of environmental and ecological covariates in Eritrea

To model the risks of *P. falciparum* and *P. vivax* malaria transmission in time and space, a set of ecological and climatic covariates that affect the development and survival of the malaria parasite and malaria vector (*Anopheles* mosquito) therefore influencing transmission. The aim

was to select three to four covariates to avoid the problem of over-fitting where redundant covariates increase model complexity without changing model results significantly. These included precipitation, minimum temperature, maximum temperature and mean temperature, the normalized difference vegetation index (NDVI) and the enhanced vegetation index (EVI). The objective was to assemble time-series variables rather than the long term means. The methodological innovation in Eritrea was to fit nonlinear functions to the covariates over time variables rather using a fixed linear prior effect. The combined effect of these covariates is likely to affect the spatio-temporal heterogeneity of malaria. Nonlinear functions such as the first-order random walks may improve estimating the seasonal trends (Blangiardo *et al.*, 2013) and may also be useful in forecasting of cases (Sahu *et al.*, 2013).

The mean monthly temperature surfaces were downloaded from WorldClim at approximately 1 x 1 km spatial resolution (Hijmans *et al.*, 2005). The optimum temperature for the development of sporozoites, (the duration of *sporogony)* in mosquito, range between 25 ^{0}C to 30 ^{0}C. Higher temperatures above (>36 ^{0}C) result in mosquito mortality (Kirby and Lindsay, 2009) while colder temperatures (< 16 ^{0}C) impact on the parasite survival (Guerra *et al.*, 2008). Guerra *et al.* (2008) show *P. falciparum* can survive up to 16 ^{0}c and *P. vivax* ceases after 14 ^{0}C. Mean, maximum and minimum temperature grid surfaces were produced from long-term climate observations for the period 1950-2000, interpolated using smoothing spline algorithms (Hijmans *et al.*, 2005). A Temperature Suitability Index (TSI) was derived from the gridded surfaces of temperature (Gething *et al.*, 2011a). The development of TSI also included time series data on effect of temperature on vector survival and the duration of sporogony. A value of zero indicated inability of temperature within a localized area, pixel, to support vector survival.

Precipitation data were obtained from the Tropical Rainfall Measuring Mission sensor (TRMM 3B43 product) that combines ground observations and satellite sensor data to generate a gridded rainfall surface at approximately 0.25^{o} x 0.25^{o} spatial resolution (Huffman and Bolvin, 2011, NASA, 2011). The TRMM satellite orbits at approximately 401.5 km altitude with an inclination of about 35^{o} to the equator (Huffman and Bolvin, 2011). The temporal resolution is approximately 90 minutes enabling a global coverage on 24 hour basis at varying spatial scale. TRMM 3B43 is a gridded mean monthly average product of precipitation rate in $mmhr^{-1}$ (Huffman, 1997). It is produced after multi-satellite precipitation analysis that combines both satellite and ground observations (from rain gauges). Rigorous data checks are applied to both satellite and ground level data. Majority of the global products are provided and archive at 3-hourly interval with also monthly level products available after the application of TRMM multi-satellite precipitation analysis (TMPA) (Huffman and Bolvin, 2011)

Vegetation indices, namely the Normalised Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI) were obtained from the MODerate-resolution Imaging Spectroradiometer (MODIS) sensor (Scharlemann *et al.*, 2008) at 1 x 1 km spatial resolution. These indices are commonly used in mapping land cover changes (Martinez and Gilabert, 2009) and can quantifying changes in phenology. In disease mapping, the association between the change in environmental factors, disease vector and humans is important in measuring risk (Hay *et al.*, 1997). NDVI metric is affected by sensor calibration drift (Miura *et al.* 2006), atmospheric effects (Song *et al.* 2001), solar and satellite viewing angle, topography (Cuo *et al.* 2010) and shadow (Huemmrich 1996). EVI metric is usually preferred because of the reduced atmospheric

scatter since visible range spectrum is included in its construction (Wardlow and Egbert, 2010). NDVI is most sensitive to chlorophyll change before peak biomass cover in contrast to EVI which is more sensitive at seasonal peak (Huete *et al.*, 2002). The two indices are highly correlated (Wardlow *et al.*, 2007), complement each other and were both incorporated.

4.4.3 Preliminary analysis of covariates and matching to malaria cases at health facility

The covariates were matched to facility data in space and time. All the grid surfaces were resampled to a common resolution (cell size 0.008333^{o} x 0.008333^{o}). A standardization procedure was applied prior to the analysis by cantering on the mean and dividing by the standard deviation. A selection procedure was implemented to achieve a minimum set of covariates that have a plausible relationship with malaria transmission. In addition, the selection was used to remove highly co-correlated covariates. Pearson's linear correlation coefficient was estimated for the pair of covariates. A minimum set was selected using the *bestglm* package in R based on the smallest value of the Bayesian information criterion (BIC*q*) with a Bernoulli prior and smallest cross-validation error (McLeod and Xu, 2008, Hastie *et al.*, 2009, Xu, 2010). Similar analysis and specifications are in chapter 2 and chapter 3.

Preliminary analysis suggested a strong correlation amongst the temperature covariates (Pearson correlation > 0.75) and between NDVI and EVI (Pearson correlation = 0.97), which was expected. From regression analysis (Table 4.6 and Figure 4.7), precipitation, minimum temperature and maximum temperature were selected as best combination set of covariates for smoothing incidence of *P. falciparum* and *P. vivax*. EVI was included when modelling the slide

positivity rates. The parameter for the Bernoulli prior for *P. falciparum* was *q* (0, 7.31 and *q* (0, 3.71) for *P. vivax*.

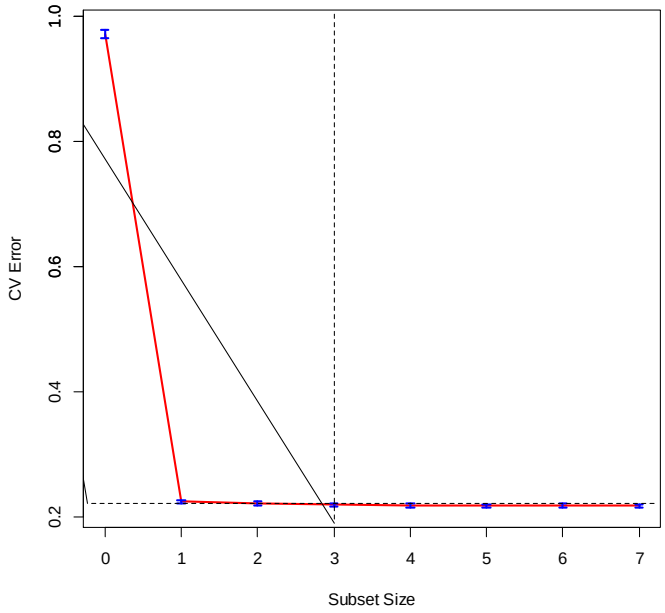

Figure 4.7 Covariate selection model based on cross validation error in Eritrea
Generalised linear model fit showing the prediction curve (red) and the cross validation error (bars). The estimate of standard error and prediction error was based on fivefold cross-validation (see chapter 2 section 2.3.4). Only one figure was displayed since the results were similar for P. falciparum and P. vivax.

Table 4.6: Regression coefficients for the best combination variables selected in a generalized linear model for *P. falciparum* and *P. vivax*

	P. falciparum			***P. vivax***		
Covariate	**Coefficient**	**Std Error**	***p*-value**	**Coefficient**	**Std Error**	***p*-value**
Precipitation	-0.09	0.01	< 0.001	-0.07	0.01	< 0.001
Maximum temperature	0.25	0.01	< 0.001	0.32	0.01	< 0.001
Minimum temperature	0.65	0.01	< 0.001	0.59	0.01	< 0.001

4.4.4 Bayesian model specification for incidence and slide positivity rate

4.4.4.1 Bayesian model specification for incidence

Two hierarchical Bayesian spatio-temporal Poisson conditional autoregressive (CAR) models for *P. falciparum* and *P. vivax* were fitted separately to smooth monthly malaria incidence at the district level. The number of cases at time j , (y_{ij}) was assumed to follow a Poisson distribution with expectation λ_{ij} .

$$y_{ij} \sim Poisson(\mu_{ij}); \quad \mu_{ij} = E(\lambda_{ij})$$

$$\Pr(Y = y_i) = \frac{\lambda^{y_i} e^{-\lambda}}{y_i!}; \quad \lambda > 0, \quad y_i = 0,1,2.....$$

The parameter was transformed such that:

$$\lambda = \log \mu \quad \sim N\left(\log a + \log b, a^{-1}\right)$$

for large a $\log gamma \approx Gaussian$. The projected denominator population varied by year based on the population growth rate, but not by month. The linear predictor was modelled additively as

$$\log \mu_{ij} = \log\left(E_{ij}\right) + \alpha + f_{ij}(t) + f_{ij}(geo) + f\left(\sum_{1}^{n} \beta_k z_{k(ij)}\right) + f(\varepsilon_i)$$

with α as the intercept and E_i is the expected number of cases adjusted for utilisation and rate of reporting. The likelihood given n *i.i.d* samples can be written as

$$f(y_1.....y_n \mid \mu) = \prod_{1}^{N} f(y_i \mid \mu)$$

$$\alpha \quad \mu^{\Sigma y_i} \exp\{-n\mu\}$$

$$\alpha \quad \prod_{i}^{N} \exp\left\{-\frac{1}{2} y_i (\log \lambda_i - \log y_i)^2\right\}$$

The posterior distribution was based on approximating *loggamma* distribution is evaluated as

$$P(\alpha, \beta \mid y_i) \quad \alpha \quad \exp\left\{-\frac{1}{2} y_i (\alpha + \beta x - \log y_i)^2\right\}$$

The additional $f(\cdot)$ terms in the linear predictor were used to relax the linear assumptions. For instance, a seasonal (time) term was included with length 36 months ($f(t) \sim LogGamma(1,0.0005)$). Nonlinear smoothing functions of first order random walk priors were used for the covariate effects with successive increments $\Delta x \sim N(0,1/\tau)$ with τ being the precision parameter (Fahrmeir and Kneib, 2009). The unstructured random effects at facility, district and province level were assumed to be independently distributed $(i.i.d)$ with zero mean and large variance (10,000). To capture the spatial effect $f_{ij}(geo)$, a conditional autoregressive prior was used (Besag *et al.*, 1991) with spatial dependence specified using an

adjacency matrix W. The weights were assigned as $W = 1$ for two neighbouring regions or $W = 0$ otherwise. Residuals were examined to assess for autocorrelation.

The model was assessed using a predictive posterior distribution of the missing data points i.e. imputed as NA. For estimation of an unknown data point y_p given other data, the predictive posterior density $P(y_p \mid y_{-p})$ is given by:

$$P(y_p \mid y_{-p}) = \int \pi(y \mid \theta, y_{-p})\, \pi(\theta \mid y_{-p})\, d\theta_p$$

This method was proposed by (Gneiting and Raftery, 2007) and can be used to obtain model scoring rules that can be used to assess model calibration (i.e. consistency between estimates and observations). The implemented scoring rules were the standard error score (SES), the ranked probabilities score (RPS) and the Dawid-Sebastiani score (DSS) computed as:

$$SES(P, y) = (y - \mu_P)^2$$

$$DSS(P, y) = \frac{1}{2} \cdot \left(\log\left(\sigma^2\right) + \left(\left(y - \mu_p\right) / \sigma_p\right)^2\right)$$

$$LogS(P, y) = -\log\left[P\left(Y - y_{observed}\right)\right]$$

$$RPS(P, y) = \sum_{k=0}^{\infty} \left[P(Y \le k) - 1(y \le k)\right]^2$$

where P is the predictive posterior distribution with mean μ_p and standard deviation σ_p and y is the observed count (Gneiting and Raftery, 2007). Note that the SES is analogous to mean square error $(y - \mu)^2$ with main difference being reference to the predictive posterior distribution. Thus, SES varies depending on the local mean while the other scores (e.g the RPS) are dependent on the whole predictive posterior distribution. The DSS is an alternative measure of predictive model choice criterion (Gelfand and Ghosh, 1998).

Pearson correlation was used to compare the predicted values to the observed and scatter plots produced for visual display. The correlation was based on 26 facilities selected randomly in Eritrea and residuals based on this hold-out set were checked for spatial and temporal autocorrelation. The leave-one-out cross validation score using the conditional predictive ordinate (CPO) was also evaluated.

4.4.4.2 Bayesian model specification for slide positivity rate

The model specification of slide positivity rate was similar to the one in chapter 3 section 3.5.4.A binomial outcome was fitted for *P. falciparum* and *P. vivax* with $y_i \mid p_i \sim Binomial\,(n_i\,,\,p_i)$ with shape dependent on $\tau^{y}(1-\tau)^{n-y}$ for

$$f(y \mid \tau) = \binom{n}{y} \tau^{y} (1-\tau)^{n-y}$$

with a logit link function $p(\eta) = \{\exp(\eta)/1+\exp(\eta)\}$. The hierarchical model was decomposed into the observations with measurement error term,

$$Z(s_i,t) = \eta(s_i,t) + \varepsilon(s_i,t) \qquad i = 1,...n, \;\; t = 1,....T$$

where $\eta(s_i,t)$ represented the underlying spatio-temporal biological process with an error term $\varepsilon(s_i,t) \overset{i.i.d}{\sim} N(0,\sigma_\varepsilon^2\, I_n)$ with prior $\sigma_\varepsilon^2 \sim IG(1,0.0005)$. A seasonal term was included in the model through the first order auto-regressive process $\rho Y(s_i,t_{i-1})$ for $|\rho|<1$ (Sahu and Bakar, 2012). The linear predictor was modelled as:

$$\eta(s_i,t) = \rho Y_{t-1} + x(s_i,t)\beta + w(s_i,t)$$

where $x(s_i,t) = \{x_1(s_i,t).....x_p(s_i,t)\}^T$ are covariates with β coefficients. The first term of $\rho Y(s_i,t_{i-1})$ is derived from stationary distribution $N\left(0,\sigma_w^2/\left(1-\rho^2\right)\right)$. The term $w(s_i,t)$ was modelled as a

separable covariance function using Kronecker product $(Q = Q_s \otimes Q_t)$. Prior range for matérn function was set to 1/5 of the spatial domain. Nonlinear smoothing functions of first order random walk priors were used for the covariate effects similar to incidence analysis. The posterior can be written as:

$$\eta(s_i, t) \mid \beta, \tau, \phi_s, \rho_t \sim GMRF(\mu, \Sigma)$$

The Gaussian Markov Random Field (GMRF) defined using Stochastic Partial Differential Equations (SPDE) evaluated using finite element methods (Lindgren *et al.*, 2011). The region of study was expanded by 100 km at the border to reduce edge effects associated with Neumann boundaries in SPDE (Cameletti *et al.*, 2012, Lindgren, 2013).

For the geostatistical model, sensitivity analyses were performed using the root mean square error (RMSE), the mean error and the absolute mean error (MAE) that summarised the closeness of validation set data to observed values. The nominal model coverage of 95% credible intervals and spatial structure of residuals was also assessed based on the validation set. MAE and the R.M.S.E are given by:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left| Z^*(x) - Z(x) \right|$$

$$RMSE = \sqrt{\left(\frac{1}{n} \sum_{i=1}^{n} (Z^*(x) - Z(x))^2 \right)}$$

4.4.5 Results

4.4.5.1 Bayesian CAR model for incidence results and validation

The mean error based on a validation set (*n*=120) in 10 districts selected randomly was 2.1 cases per 1000 population for the *P. falciparum* compared to 1.8 cases per 1000 population for *P. vivax*

which suggests a tendency to over-estimate by 1.9 case per 1000 population. Figure 4.8 shows scatter plots of crude and the estimated incidence per 1000 population. The Pearson correlation between the crude and the predicted incidence was 0.77 for *P. falciparum* and 0.68 for *P. vivax*. Residual analysis indicated that incorporation of environmental and spatial effects was useful in explaining most of spatial variation in the model with no amount of spatial or temporal autocorrelation evident in the residuals (Figure 4.9). The variance of the spatial random effect was reduced to 27.0% for *P. falciparum* and 23.3% for *P. vivax*.

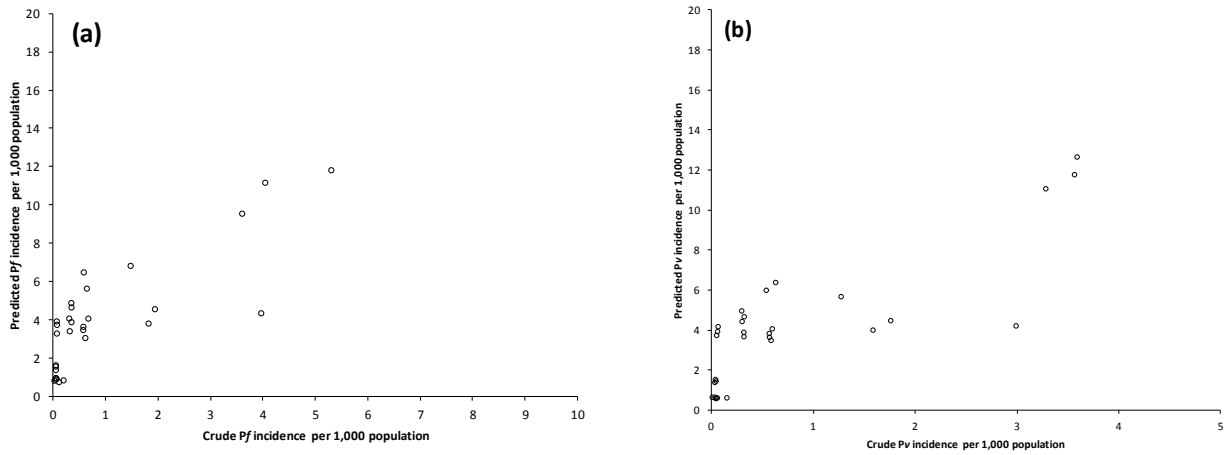

Figure 4.8: Comparison of observed and estimated incidence
Scatter plot of the observed crude incidence compared to the estimated incidence from the spatio-temporal Bayesian Poisson model for (**a**) *P. falciparum* and (**b**) *P. vivax*.

Table 4.7 shows internal model validation statics. Cross-validation was conducted based on CPO and none of the data points reported a CPO value of 1 (fail). No data values for the fitted model failed the CPO test which is likely to happen if the approximation of the latent Gaussian Field (GF) is less accurate (Czado *et al.*, 2009). The standard error score for the mean component for *P. vivax* model was smaller than that for *P. falciparum*. There were minimal difference in the DSS used to assess model calibration, perhaps since same model was applied to the two parasites with similar distribution in space and time. The same applied to the log-score and the ranked

probability score (RPS) which can be used for model comparison. A smaller log-score indicate a better model fit or calibration.

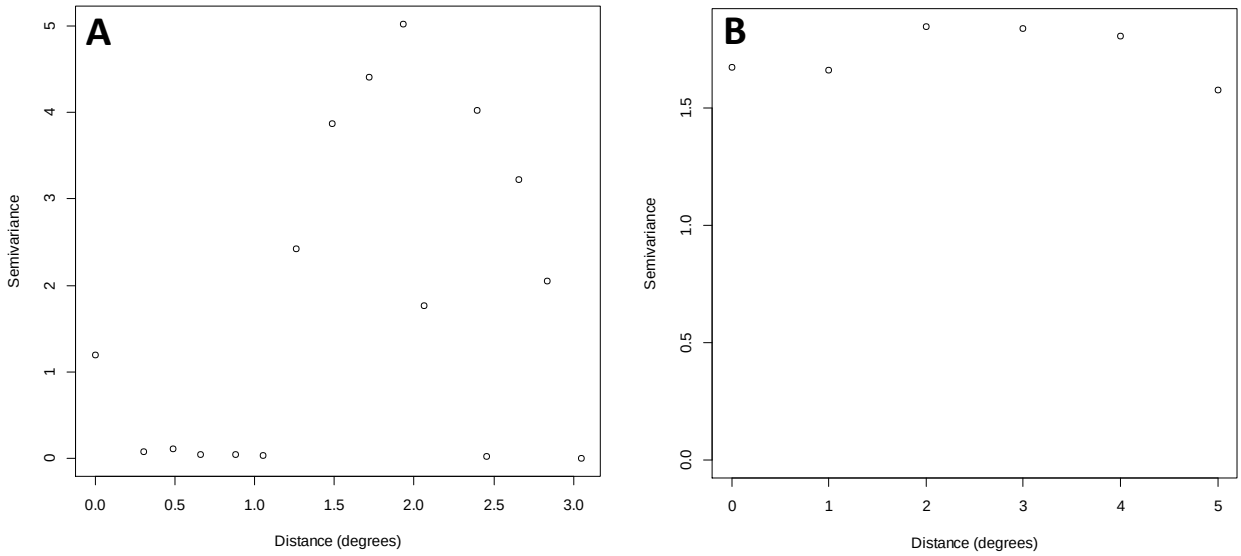

Figure 4.9: Semi-variogram plots of residuals for incidence analysis in Eritrea
Semi-variogram showing residual autocorrelation in (a) the spatial and (b) the temporal domain for *P. falciparum*. The plots for P. vivax were very similar to the *P. falciparum* plots and not diplayed. The plots suggest insignificant autocorrelation was left after running the model. The y-axis is the semi variance while the x-axis is distance in degree between pairs. The residuals were extracted for validation set data.

Table 4.7: Model validation results for incidence analysis.
The implemented scoring rules squared error score (SES), mean logarithmic sore (logS), the ranked probability score (RPS) and the Dawid-Sebastiani score (DSS) for *P. falciparum* and *P. vivax* malaria incidence. The scores are obtained by averaging (μ_{ij}) the scores of one-step-ahead predictions in six regions (n=60) as well as calculated based on the seasonal component

Model	Component	SES	DSS	Log score	RPS
P. falciparum	μ_{ij}	1.36	0.66	1.16	0.00
	τ	1.22	0.60	1.12	0.01
P. vivax	μ_{ij}	1.18	0.56	1.07	0.52
	τ	1.99	0.58	1.10	0.56

The DIC for the *P. falciparum* model was 13372.54 and for the *P. vivax* model was 11945.12, and both had a similar number of effective parameters (249.46 and 244.27, respectively) (Table

4.8). The posterior summaries of the parameters representing the fixed effects and random effects for *P. falciparum* and *P. vivax* are listed in Table 4.8.

Table 4.8: Posterior estimates of the parameters of Bayesian spatio-temporal Poisson model with 95% credible interval

Variable	Plasmodium falciparum Mean (95% Crl)	Plasmodium vivax Mean 95% Crl
District random effect	7.12 (5.53 - 8.58)	7.17 (5.38 - 8.62)
Province random effect	2.35 (1.15 - 4.16)	2.37 (1.36 - 3.85)
Facility random effect	1.05 (0.82 - 1.29)	0.90 (0.66 - 1.12)
Spatial CAR effect	0.96 (0.04 - 1.79)	1.02 (0.24 - 1.84)
Seasonal component	0.12 (0.02 - 0.36)	0.07 (0.01 - 0.20)
Maximum temperature	9.36 (7.80 - 10.83)	9.42 (7.69 - 10.89)
Minimum temperature	9.46 (7.81 - 10.89)	9.40 (7.78 - 10.86)
Precipitation	9.42 (7.83 - 10.87)	9.50 (7.82 - 10.92)
Variance of CAR effect	0.27	0.23
DIC	13372.54	11945.12
P_D	249.26	244.27
SES	0.53	0.65

Crl: Bayesian credible interval, CAR: Conditional autoregressive, DIC: Deviance Information Criterion, P_D: Effective parameters

4.4.5.1 The seasonal trends in covariates based on the fitted nonlinear functions

Covariates were associated positively with estimated incidence. The coefficients of temperature and precipitation were very similar compared to the coefficients between the random effects at province, district and facility level. *P. falciparum* constituted approximately 56.6% of the estimated malaria burden for the three years while 43.4% were *P. vivax*. Figure 4.10 and Figure 4.11 shows the non-linear effects on temperature, seasonality and rainfall when used in modelling incidence of *P. falciparum* and *P. vivax* respectively. Similar trends were observed in the slide positivity analysis using the same specification. There was a strong seasonal effect which peaked in October for *P. falciparum* and *P. vivax*. High temperature ($> 40^0$C) and rainfall (>300mm) had a negative effect on incidence which was more evident from the *P. vivax* model compared to *P. falciparum*. For *P. falciparum* the effect of rainfall seemed to plateau at levels exceeding 200mm. The non linear pattern is not surprising given that high temperature is associated with mosquito mortality and a high amount of rainfall may lead to larvae wash out (Paaijmans *et al.*, 2007, Guerra *et al.*, 2008, Guerra *et al.*, 2010).

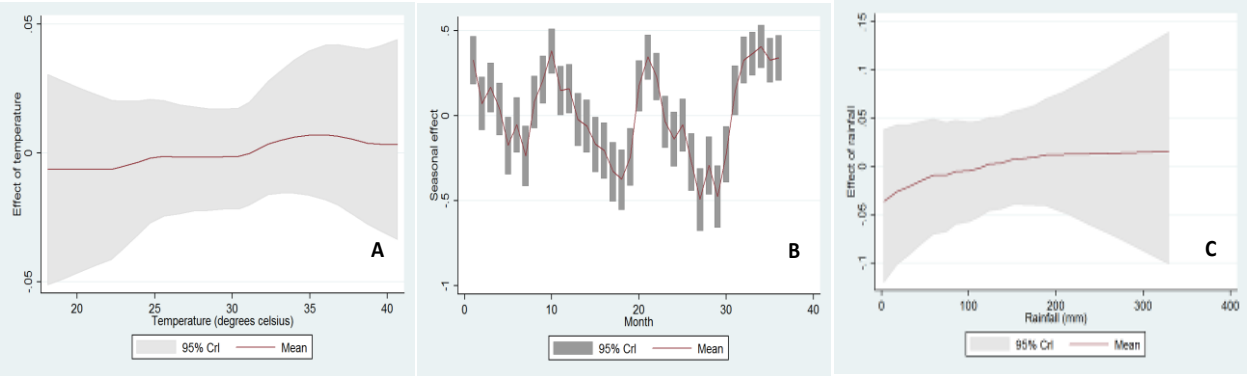

Figure 4.10: Non-linear effect of covariates on *P. falciparum*
The effect of temperature, seasonality and rainfall on P. falciparum incidence (solid center line) with 95% credible interval (shaded area and bars on the seasonal plot).

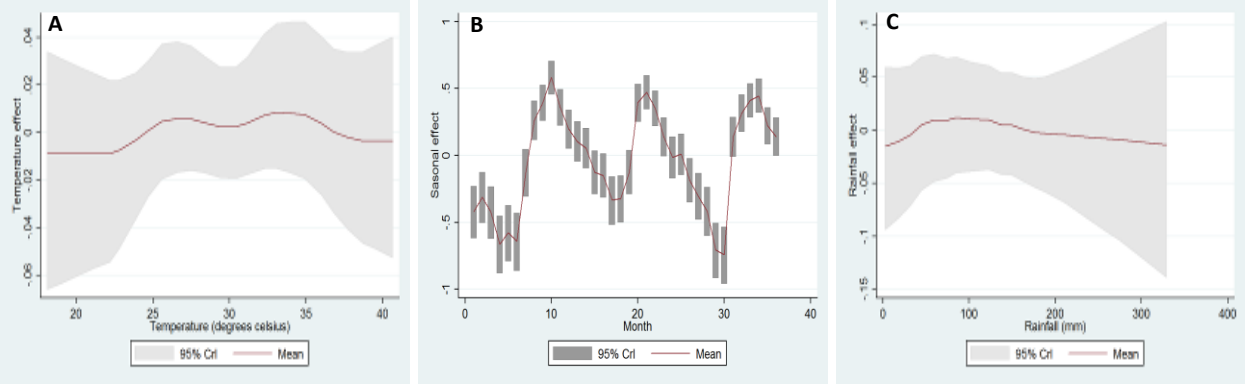

Figure 4.11: Non-linear effect of covariates on *P. vivax*
The effect of temperature, seasonality and rainfall on P. vivax (solid center line) with 95% credible interval (shaded area and bars on the seasonal plot).

4.4.5.2 The incidence of P. falciparum and P. vivax in Eritrea

Figure 4.12 (page 234) and Figure 4.13 (Page 235) shows the monthly maps of mean incidence of *P. falciparum* and *P. vivax* for the study period. *P. falciparum* constituted approximately 56.6% of the estimated malaria burden for the three years while 43.4% were *P. vivax*. There was a strong seasonal effect in the modelled incidence of *P. falciparum* and *P. vivax*. For both parasites, incidence peaked in September and October. Gash Barka region showed the most risk on average followed by Debub, while the Southern Red Sea region had the lowest estimated risk. The largest number of cases of *P. falciparum* was 6.1 cases per 1000 population in October 2012

compared to 5.6 cases per 1000 population for *P. vivax* in 2010. The correlation between the mean monthly incidence of *P. falciparum* and *P. vivax* was 0.69 (Pearson correlation).

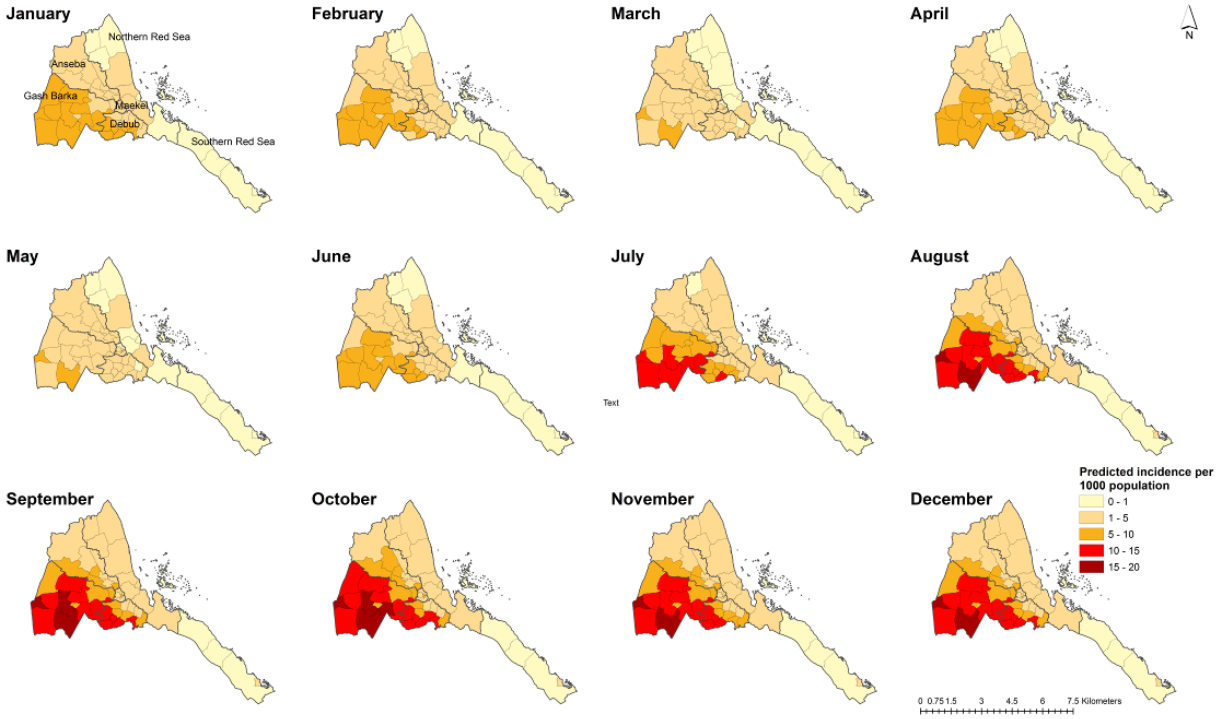

Figure 4.12: Spatio-temporal maps of incidence of *P. falciparum* per 1000 population in Eritrea
Mean monthly maps of monthly incidence of *P. falciparum* per 1000 population in Eritrea using a Bayesian spatio-temporal Poisson model. Districts with low risk are classified as < 5 cases per 1000 population) and moderate risk with > 5 cases per 1000 population.

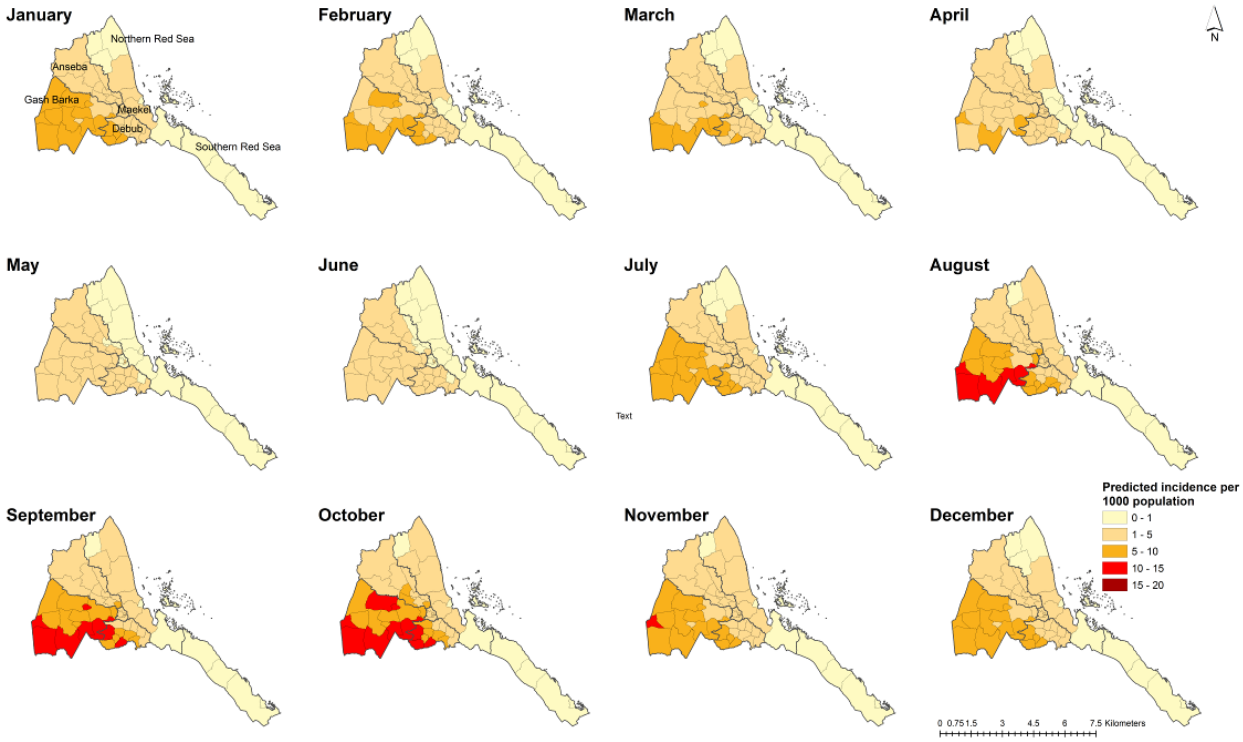

Figure 4.13: Spatio-temporal maps of incidence of *P. vivax* per 1000 population in Eritrea
Mean monthly maps of incidence of *P. vivax* per 1000 population in Eritrea using a Bayesian spatio-temporal Poisson model. *P. vivax* risk was lower compared to *P. falciparum* although peaked at a similar time to *P. falciparum*. Districts with low risk are classified as < 5 cases per 1000 population) and moderate risk with > 5 cases per 1000 population

Figure 4.14 shows the annualized maps of *P. falciparum* incidence per 1000 population while Figure 4.13 (b) represents the incidence of *P. vivax* over the three years (2010-2012). The mean annual estimates for *P. falciparum* were 3.60 (95% Crl 2.27 - 5.44) cases per 1000 population in 2010, 2.99 (1.90 - 4.49) cases per 1000 population in 2011 and 3.43 (2.17 - 5.16) cases per 1000 population in 2012. For *P.* vivax the annual mean incidence was 2.39 (1.44 - 3.72) cases per 1000 population in 2010, 2.77 (1.67 - 4.32) cases per 1000 population in 2011 and 2.53 (1.53 - 3.93) cases per 1000 population in 2012.

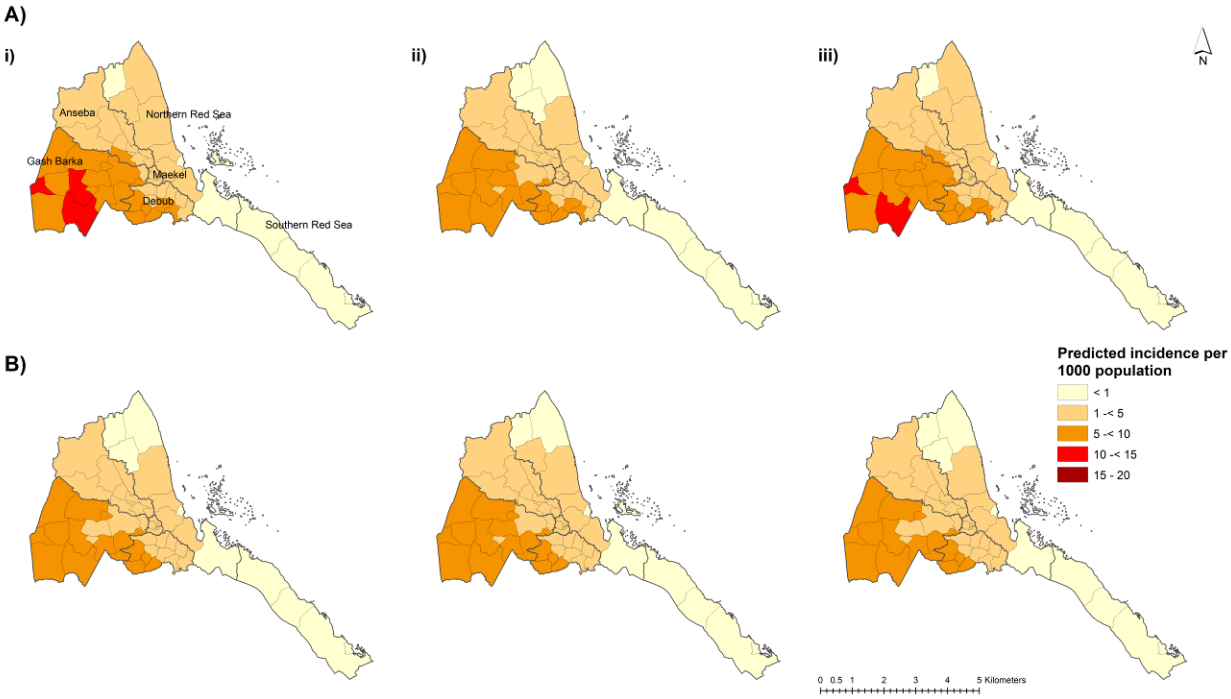

Figure 4.14: Annualised incidence of *P. falciparum* and *P. vivax* per 1000 population
Maps of incidence of by year (2012 – 2013) for (a) for *P. falciparum* and (b) for *P. vivax*. Overall, most risk was in Garsh Barka and Debub. The annual mean incidence of P. vivax did not exceed 5 cases per 1000 population in any Sub-Zoba.

Figure 4.15 (below, Page 236) show a comparison between the modelled incidence per 1000 population with the mean parasite prevalence modelled at community level (Noor *et al.*, 2014). There was a positive correlation between incidence and parasite rate (Pearson correlation, 0.7).

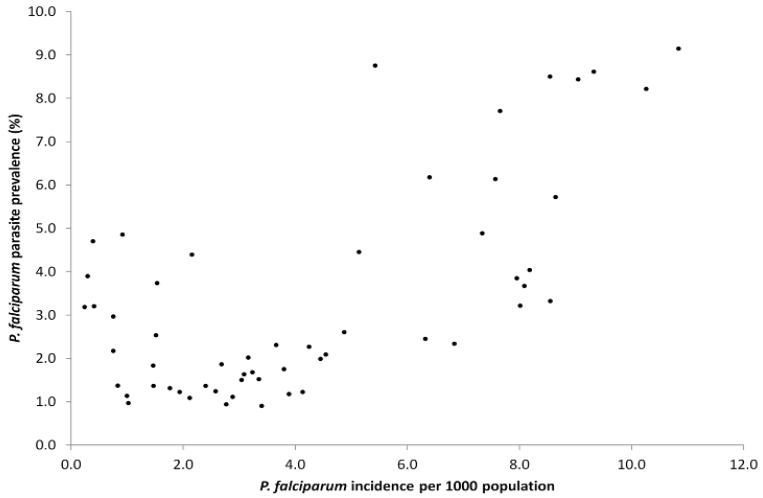

Figure 4.15: Comparison of incidence with community parasite prevalence

Scatter plot of *P. falciparum* parasite prevalence (rate) at community level with the modelled incidence per 1000 population (Pearson correlation =0.7)

4.4.5.3 The modelled slide positivity rates results

Table 4.9 lists the Mean Absolute Error (MAE) and the RMSE. The MAE for *P. falciparum* and for *P. vivax* was 0.59 and 0.50, respectively. Pearson correlation coefficient was greater than 0.6 in both modelling framework. The 95% nominal coverage for *P. falciparum* was 96.3% and 98.7% for *P. vivax* showing a tendency to slightly over predict at 95%. The spatial dependence as define by the variogram range parameter was also different for both parasites, approximately 10 km for *P. falciparum*. Only minimum spatial structure was left in the residuals after model fit (Figure 4.16).

Table 4.9: Model of slide positivity rate in Eritrea.
Bayesian model comparison based on separable covariance function (Product) (M1) and no-separable form (Product-sum) (M2) for *P. falciparum* (P*f*) and *P. vivax* (P*v*).

Model	DIC	P_D[1]	MAE	R.M.S.E	Probability of prediction interval (%)[2]	Model Range in km [95% Crl]	Correlation
P*f*	6892.50	122.56	0.59	0.95	96.33	9.95 [4.27 - 18.23]	0.62
P*v*	6564.02	92.47	0.50	0.79	98.68	8.42 [3.58 - 13.95]	0.64

1. P_D represent the effective number of parameters that represent model complexity
2. The nominal probability of prediction is 95%

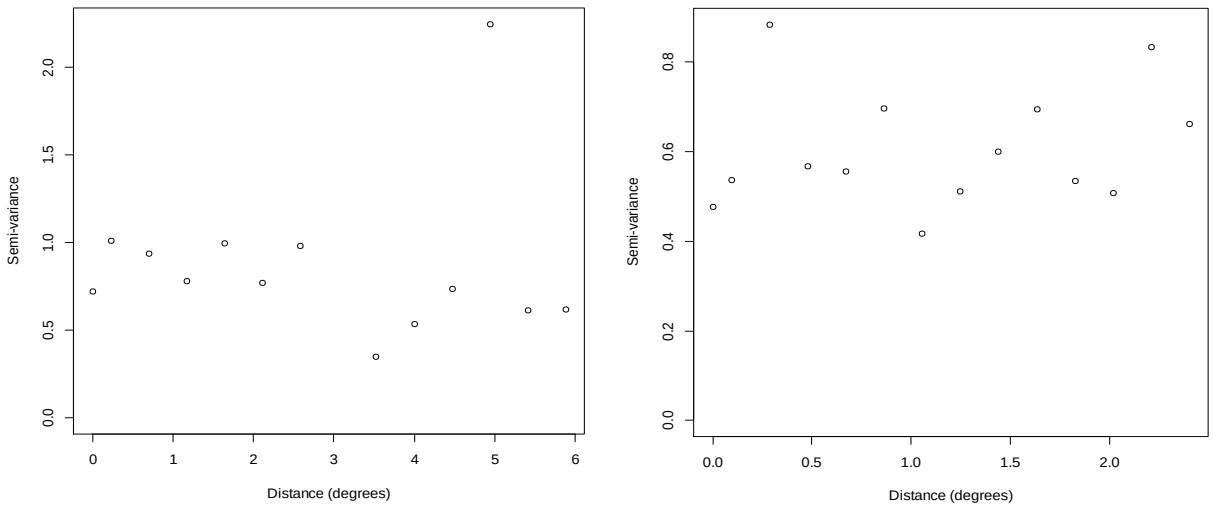

Figure 4.16: Semi-variogram of residuals of slide positivity rate

Semivariogram plots of the residuals in spatial domain for *P. falciparum* (a) and *P. vivax* (b) based on the predictions on the hold out set. The x-axis shows distance in degrees latitude and longitude (decimal degrees) whiles the y-axis shows semi-variance.

Table 4.10 shows the posterior distributions of the fixed effects along with Bayesian credible intervals and nominal range. None of the fixed effects covariates was a significant predictor of *P. falciparum* at 95%. For *P. vivax*, only EVI was significant marginally at 95% credible interval. All other random effects and the intercept were significant.

Table 4.10: Posterior distribution of parameters of slide positivity rate.
Distribution of posterior estimates (mean, standard deviation and quantiles) of the fixed components and random effects

Model	Parameter	Mean	SD[1]	5%	50%	95%
P. *falciparum*	Intercept	-1.13	0.3	-1.62	-1.12	-0.67
	Minimum temperature	-0.1	0.11	-0.28	-0.1	0.08
	Maximum temperature	-0.07	0.14	-0.31	-0.07	0.16
	Precipitation	0.03	0.03	-0.03	0.03	0.08
	Enhanced vegetation Index	0.02	0.04	-0.05	0.02	0.09
	Marginal variance (σ_w^2)	5.25	4.41	1.18	3.94	13.7
	Model Range (ϕ)	9.95	4.38	4.27	9.18	18.23
P. *vivax*	Intercept	-1.29	0.20	-1.60	-1.29	-0.99
	Minimum temperature	-0.18	0.14	-0.42	-0.18	0.05
	Maximum temperature	0.03	0.11	-0.15	0.03	0.22
	Precipitation	0.01	0.03	-0.04	0.01	0.07
	Enhanced vegetation Index	-0.06	0.04	-0.13	-0.06	-0.00
	Marginal variance (σ_w^2)	4.23	2.65	1.52	3.49	9.39
	Model Range (ϕ)	8.42	3.19	3.58	8.21	13.95

1. SD: Standard deviation

Figure 4.17 shows predicted map of slide positivity rate for 2012 at 1 x 1 km weighted by the predicted probability of health facility use for fever treatment. Panel A represents: the continuous; binned (six classes <1.0; 1 -< 5.0; 5.0 -< 10.0; 10.0 -< 15.0; 15 -< 20.0; and >20.0); and the standard deviation maps of *P. falciparum* while paned B is for *P. vivax*. The mean slide positivity by region is shown in Table 4.11. For *P. falciparum* mean slide positivity was 3.8% (range 12.8%) while for *P. vivax* the mean was 3.3% (range 25.5 %). Highest positivity rates were predicted in Maekel, Garsh Barka and Debub. Overall slide positivity rate were less than 5% in most parts of the country and less than 1% in some areas of Anseba, Northern Red Sea and

the Southern Red Sea. The highest uncertainties were in regions with no facilities and with low reporting rates.

Table 4.11: Slide positivity rate by region in Eritrea.
Summary (in percentage) of the mean predicted slide positivity rate (SPR) for *P. falciparum* and *P. vivax* by region

	P. falciparum					*P. vivax*				
Region	**Mean**	**SD**	**Minimum**	**Maximum**	**Range**	**Mean**	**SD**	**Minimum**	**Maximum**	**Range**
Anseba	2.2	2.0	0.1	10.4	10.3	1.6	1.7	0.0	10.8	10.7
Debub	4.9	2.3	0.5	16.1	15.6	5.0	3.2	0.7	29.8	29.1
Gash Barka	5.0	2.6	0.6	23.2	22.7	3.1	3.4	0.4	81.1	80.7
Maekel	6.0	3.1	2.8	14.0	11.3	6.8	1.9	2.0	14.4	12.4
Northern Red Sea	2.1	1.4	0.1	11.0	10.9	1.7	1.2	0.1	14.8	14.7
Southern Red Sea	2.5	1.4	0.3	6.3	6.0	1.7	1.0	0.2	5.6	5.4
Average	**3.8**	**2.1**	**0.7**	**13.5**	**12.8**	**3.3**	**2.1**	**0.6**	**26.1**	**25.5**

The comparison between the modelled slide positivity rates and community-level parasite prevalence is shown in Figure 4.18. In general there was a positive correlation similar to incidence comparison with community parasite prevalence (Pearson correlation =0.6).

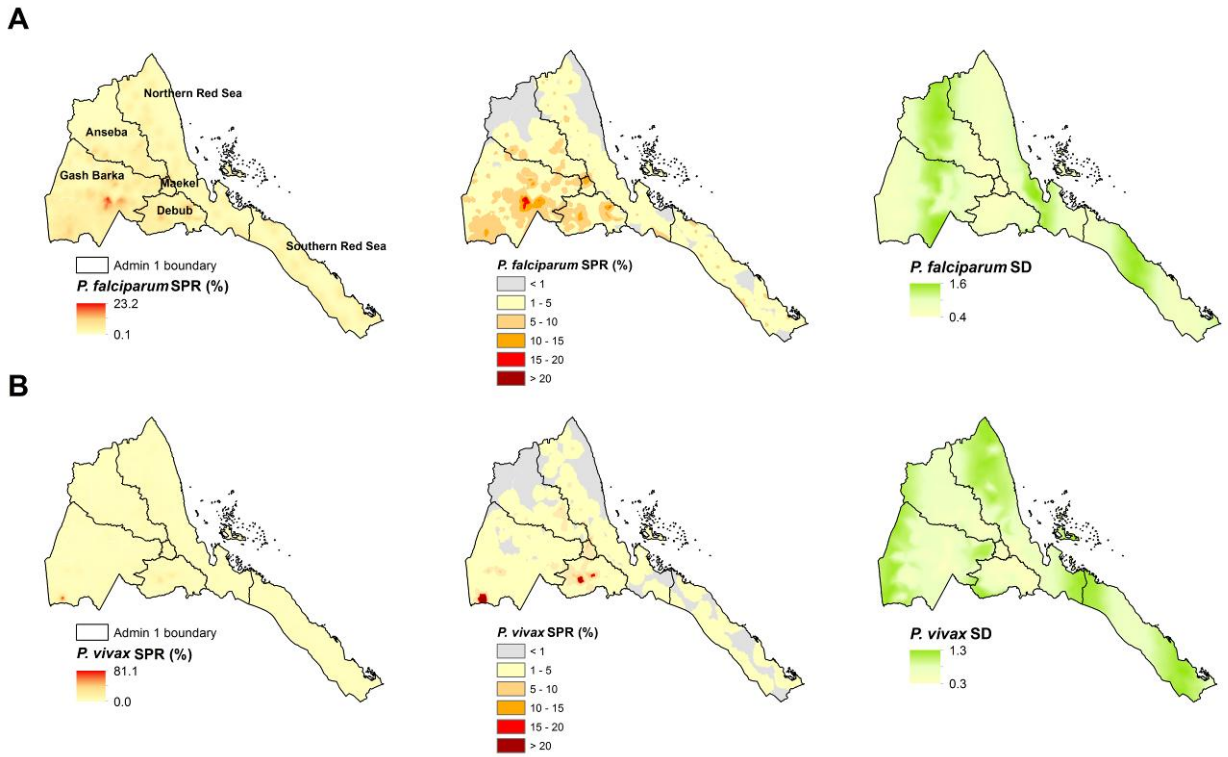

Figure 4.17: Posterior mean predictions of slide positivity rate in Eritrea

Bayesian Posterior mean predictions at 1km by 1 km of slide positivity rate (SPR) map of **(a)** *P. falciparum* and **(b)** *P. vivax*. The first map is continuous and second binned in five classes. The last map is the standard deviation showing some artefacts especially for areas with no or sparse data points.

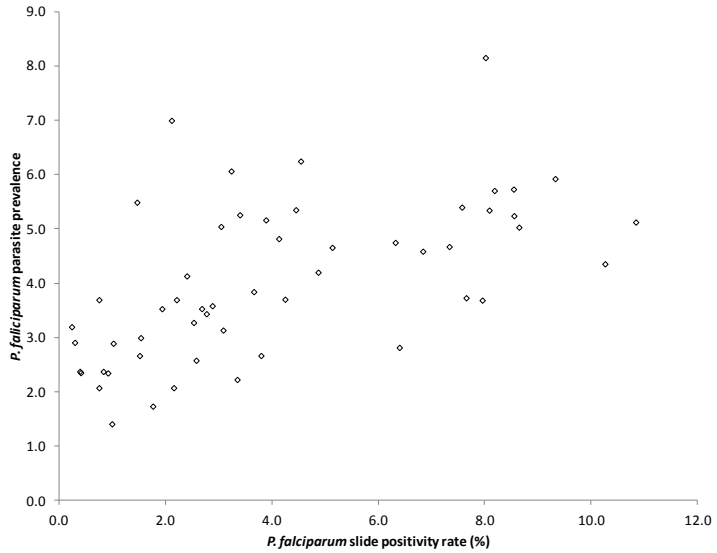

Figure 4.18: Comparison of predicted slide positivity rate and community-level parasite prevalence
Scatter plot of *P. falciparum* parasite prevalence (rate) at community level with the mean slide positivity rate at district level (Pearson correlation = 0.6).

4.4.6 Assessing the population at risk in Eritrea using incidence

The burden of both parasites in general population was assessed by comparing calculation

number of cases in a Zoba based on estimated incidence.

Table 4.12 and Table 4.13 shows the estimated clinical burden for *P. falciparum* and *P. vivax.* The most burdened region was Gash Barka where just over 10,000 cases in the overall population were estimated in 2010 and 2012 for *P. falciparum* compared to 6,866 and 7,764 estimated for *P. vivax*, respectively. The region with the second highest burden was Debub. The Southern Red Sea had the smallest number of estimated malaria cases (less than 100 throughout the three years).

Table 4.12: Estimated clinical burden of *P. falciparum* based on incidence.

Summary of clinical burden of *P. falciparum* in Eritrea by year based on the mean incidence per 1000 population.

Region	Estimated Population 2010	Crude P*f* incidence 2010	P*f* incidence 2010 Mean (95% Crl)	Estimated 2010 P*f* clinical burden	Estimated Population 2011	Crude Pf incidence 2011	P*f* incidence 2011 Mean (95% Crl)	Estimated 2011 P*f* clinical burden	Estimated Population 2012	Crude P*f* incidence 2012	Pf incidence 2012 Mean (95% Crl)	Estimated 2012 P*f* clinical burden
Anseba	689,385	9.1	3.7 (2.29 - 5.63)	2,551	711,802	6.9	3.11 (1.93 - 4.7)	2,214	734,948	11.5	3.54 (2.19 - 5.36)	2,602
Debub	1,258,464	20.1	5.19 (3.44 - 7.52)	6,531	1,299,386	7.7	4.32 (2.88 - 6.22)	5,613	1,341,639	12.6	4.94 (3.29 - 7.14)	6,628
Gash Barka	1,216,679	79.9	8.46 (5.63 - 12.39)	10,293	1,256,242	27.5	6.93 (4.66 - 10.04)	8,706	1,297,092	59.2	8.02 (5.36 - 11.7)	10,403
Maekel	467,642	2.0	2.45 (1.4 - 3.95)	1,146	482,849	2.0	2.08 (1.18 - 3.32)	1,004	498,550	2.2	2.35 (1.34 - 3.76)	1,172
Northern Red Sea	1,359,812	1.8	1.33 (0.7 - 2.23)	1,809	1,404,030	1.7	1.14 (0.6 - 1.9)	1,601	1,449,685	1.3	1.28 (0.67 - 2.14)	1,856
Southern Red Sea	205,609	0.3	0.45 (0.16 - 0.93)	93	212,295	0.8	0.38 (0.14 - 0.79)	81	219,198	0.3	0.43 (0.15 - 0.89)	94
Total	**5,197,591**	**18.9**	**3.60 (2.27 - 5.44)**	**18,711**	**5,366,603**	**7.8**	**2.99 (1.90 - 4.49)**	**16,046**	**5,541,112**	**14.5**	**3.43 (2.17 - 5.16)**	**19,006**

Crl: Bayesian credible interval, P*f: Plasmodium falciparum*

Table 4.13: Estimated clinical burden of *P. vivax* based on incidence.

Summary of clinical burden of *P. vivax* in Eritrea by year based on the mean incidence per 1000 population.

Region	Estimated Population 2010	Crude P*v* incidence 2010	*Pv* incidence 2010 Mean (95% Crl)	Estimated 2010 P*v* burden	Estimated Population 2011	Crude P*v* incidence 2011	P*v* incidence 2011 Mean (95% Crl)	Estimated 2011 P*v* burden	Estimated Population 2012	Crude P*v* incidence 2012	P*v* incidence 2012 Mean (95% Crl)	Estimated 2012 P*v* burden
Anseba	689,385	3.7	2.26 (1.32 - 3.59)	1,558	711,802	5.2	2.6 (1.51 - 4.13)	1,851	734,948	5.0	2.39 (1.39 - 3.78)	1,757
Debub	1,258,464	6.5	3.39 (2.17 - 5.01)	4,266	1,299,386	7.8	3.92 (2.51 - 5.81)	5,094	1,341,639	6.6	3.58 (2.3 - 5.29)	4,803
Gash Barka	1,216,679	8.7	5.64 (3.66 - 8.4)	6,862	1,256,242	16.0	6.6 (4.27 - 9.88)	8,291	1,297,092	17.7	5.99 (3.89 - 8.9)	7,770
Maekel	467,642	1.0	1.85 (0.98 - 3.11)	865	482,849	1.9	2.12 (1.13 - 3.58)	1,024	498,550	2.0	1.95 (1.04 - 3.27)	972
Northern Red Sea	1,359,812	0.8	0.89 (0.43 - 1.57)	1,210	1,404,030	0.9	1.01 (0.49 - 1.79)	1,418	1,449,685	1.3	0.93 (0.45 - 1.65)	1,348
Southern Red Sea	205,609	0.0	0.31 (0.11 - 0.66)	64	212,295	0.0	0.36 (0.12 - 0.75)	76	219,198	0.3	0.33 (0.11 - 0.7)	72
Total	**5,197,591**	**3.4**	**2.39 (1.44 - 3.72)**	**12,422**	**5,366,603**	**5.3**	**2.77 (1.67 - 4.32)**	**14,865**	**5,541,112**	**5.5**	**2.53 (1.53 - 3.93)**	**14,019**

Crl: Bayesian credible interval; P*v: Plasmodium vivax*

4.5 Discussion on incidence and slide positivity rate in Eritrea

The objective of this study was to estimate the burden of *P. falciparum* and *P. vivax* in Eritrea based on the HMIS data reported between 2010 and 2012. The results from this study support current efforts to move the control programme to pre-elimination given the overall low incidence for *P. falciparum* (3.0 cases per 1000 population) and 2.3 cases per 1000 population for *P. vivax*. *P. falciparum* constituted approximately 56.6% of the total estimated case burden on average for the three year period. Several other findings emerge from this study. First, there was a clear seasonal pattern of transmission with a peak in incidence in September and October. There was little evidence in support of a second malaria peak in January or for the eastern districts in Debub and Northern Red Sea as suggested by Ceccato *et al.* (2007). Although incidence was highest in Gash Barka and in Debub, the slide positivity rates were just higher in Maekel. Low incidence, similar to the slide positivity rate, was in the coastal regions bordering the Red Sea in western parts of the country.

4.5.1 Implication for malaria control and elimination in Eritrea

The spatial distribution of incidence and the slide positivity rate of *P. falciparum* reflects broadly the recent maps of prevalence from community datasets with higher rates in the Gash Barka and Debub regions while the North Red Sea and Southern Red Sea exhibited low prevalence (Noor *et al.*, 2014). *P. falciparum* remains the main parasite and constitutes approximately 56% of the malaria burden in Eritrea. In other low transmission countries outside Sub-Saharan Africa where the two parasites are endemic, such as Latin America and south and south east Asia, *P. vivax* dominates (Gething *et al.*, 2012) due to its biological characteristics that include a dormant liver stage (hypnozoite) that usually causes clinical relapses (Mueller *et al.*, 2009, White, 2011). The challenge for NMCP in Eritrea as the burden of malaria reduces towards the <1 case per

1000 population threshold is likely to be the management of *P. vivax* infections. It remains a challenge to detect vivax malaria at the liver stage which may form a reservoir of infections and there is a likelihood of missing some infections when routine diagnostics such as microscopy and RDTs are used (Mueller *et al.*, 2009). It is also not straightforward to prescribe primaquine, the recommended drug for clearing liver-stage infections, at higher doses due to a 4% glucose-6-phosphate dehydrogenase deficiency (G6PDd) prevalence in Eritrea (Howes *et al.*, 2012, Domingo *et al.*, 2013) with lack of routine individual testing for this blood disorder at health facility level. Additional challenges exist for vulnerable groups such as pregnant women. Appropriate control of *P. vivax* is likely to be prevention of mosquito inoculations in addition to resolving the relapsing infections. Low doses of primaquine have been recommended for *P. falciparum* in low transmission areas (White *et al.*, 2012) with trials starting to emerge on the performance of this drug (Steketee and ter Kuile, 2014).

Although there is a suggested lower impact of ITNs on *P. vivax* when compared to *P. falciparum* (Bockarie and Dagoro, 2006), ITNs in general had the greatest impact on reducing clinical malaria episodes in Eritrea for both parasites by reducing human – mosquito contact, in addition to other control measures such as larvae source management, change in antimalarial drug policy, and increasing awareness on malaria through public campaigns advocacy (East Africa Roll Back Malaria Network (EARN), 2013). ITNs are delivered generally through health facilities in Eritrea in addition to other mass campaigns. Previous findings suggest access to health facilities plays a role in ownership of ITNs (Macintyre *et al.*, 2006). There was one Long-Lasting Insecticidal Net (LLIN) for every 0.5 individuals in a household from the 2012 MIS (n=1 818) which is lower that the WHO recommendation of one LLIN per 1.8 people (World Health

Organization, 2012b). ITN use in children under the age of five years was high (60%), but use in all age-cohorts was somewhat lower at 55%. From the analysis of fever treatment patterns based on the MIS, utilisation declined significantly after 180 minutes, and at an increasing travel times from the health facility. The estimated case loads for the three years suggest a need to consolidate gains in the last 10 years and identify endemic districts to guide interventions to achieve pre-elimination. Interventions may include, but not limited to: raising awareness on the use of ITNs at appropriate time of the year lagged with seasonal trends observed here, conducting bed-net re-treatment, and selective targeting of IRS lagged with malaria seasonality. The GFTAM remains the main source of funding for malaria control and supports the universal coverage efforts. Increasing coverage in key *sub-zobas* and improving net use during the transmission season may be combined with mosquito net retreatment (for ITNs and LLINs older than 3 years). Figure 2(a) and Figure 2(b) suggest where and when this could be most appropriate.

Previous studies also suggest that the seasonal peaks identified in Eritrea resemble those observed in the neighbouring countries (Paulander *et al.*, 2009, van den Bogaart *et al.*, 2013). In addition, the recent malaria programme performance review suggested the need for Eritrea to implement cross-border collaboration with neighbouring countries (East Africa Roll Back Malaria Network (EARN), 2013). The clear identification of high malaria incidence in Debub and Garsh Barka supports these initiatives with Djibouti, Ethiopia and Sudan since imported infection may pose a threat to the long-term goal of elimination (Le Menach *et al.*, 2011). The low malaria incidence overall supports current strategies for pre-elimination starting potentially with the Southern Red Sea bordering Djibouti.

4.5.2 Slide positivity rates for 2012 and health facility utilisation

The mean slide positivity rate in 2012 was less than 5% for both *P. falciparum* and *P. vivax*. A comparison of the SPR at district level with parasite prevalence modelled from community level data showed a positive correlation. Less than 5% is a threshold for pre-elimination (World Health Organization, 2007a). Slide positivity reflects rates amongst population using health facilities and the correlation will be higher in areas where access and utilisation (of the public sector) is high. In Eritrea, approximately 80% of the population was estimated to be within a public health facility catchment and 61% of estimated febrile cases were likely to present at these facilities for treatment. Elsewhere, in China and Uganda, slide positivity rate was shown to be a strong covariate in estimation of incidence (Jensen *et al.*, 2009, Bi *et al.*, 2012) although the results do not distinguish between the indigenous infection within the district and the imported cases.

The relationship between slide positivity rate and incidence may not be linear as suggested by Jensen *et al.* (2009). This because the denominator, number examined, for slide positivity is more sensitive compared to incidence measured in the overall population. Changes in incidence may be more informative compared to relative changes in slide positivity rates. Nonetheless, the threshold levels estimated in this study provide thresholds for pre-elimination in Eritrea.

4.5.3 Modelling gains for malaria incidence and slide positivity analysis in Eritrea

The analysis here introduced nonlinear effects in specifying priors for the covariates rather than an assumption of fixed linear priors. Such a specification may have benefitted the extraction of

seasonal trends. In terms of crude incidence comparison, there was a wider difference between the crude estimates compared to smooth estimates by year for *P. falciparum* (Table 4.12). For *P. vivax* the difference were less (Table 4.13). First, the covariates were important in modelling incidence in data points with missing data (Table 4.8 which also provides the 95% confidence intervals). Thus, the spatial and temporal smoothing as discussed in Namibia and Afghanistan produce more reliable estimates compared to crude estimates.

There was a strong nonlinear effect in precipitation and temperature trends (Figure 4.10 and Figure 4.11) which suggest that the assumptions of linearity in Namibia or Afghanistan case studies may not be true. Previous research has already illustrated the important of these covariates in driving transmission (Noor *et al.*, 2014, Weiss *et al.*, 2015). However, the 'delayed or lagged effect' of these covariates on transmission is not well established and may not apply to all the settings. Some studies have shown a time-series decomposition of these climatic variables may provide useful insights into disease anomalies (Wardrop *et al.*, 2013) and the result here suggest more research is required to further understand the role of covariates as drivers of malaria transmission. Since the effective threshold for temperature and rainfall on malaria transmission is well established, (Paaijmans *et al.*, 2007, Guerra *et al.*, 2008, Guerra *et al.*, 2010) the non-linear trends maybe useful in prospective modelling of epidemics in low malaria transmission.

4.5.4 Limitations

The analysis presented here was limited to the three years of available time-series data and can potentially be extended by inclusion of longer space-time data sets. However, the data used here

were abundant spatially and within the temporal extent studied and allowed us to reveal a clear seasonal pattern that is likely to be useful in directing interventions. Secondly, there could be biases introduced due to the type of diagnosis at the facility level by using RDT or microscopy with varying sensitivities (WHO-FIND, 2009). This relates to a suggested higher slide positivity rate when using RDTs in low transmission settings, but also to concern surrounding the quality of microscopy in routine data (Hopkins *et al.*, 2013). However, health facility data used in this analysis represent parasite examination from febrile cases and, therefore, the likelihood of detecting infection is higher significantly. In addition, cases presenting at the health facility are drawn from multiple villages within the health facility catchment area. This increases geographic representativeness of estimated malaria burden compared to community-based cluster surveys where estimates are usually based on a single cluster sample. The MIS was used to adjust for fever burden seen in the public sector. In general, we assumed a single fever episode, and treatment-seeking at other times of year followed the pattern in the MIS implementation months of September to October.

4.6 Conclusion

In conclusion, we demonstrated the utility of routine HMIS data for malaria burden estimation in a low transmission setting and provide seasonal profile for *P. falciparum* and *P. vivax* in Eritrea. These maps are important as the NMCP aims for pre-elimination. HMIS has advantages of data being collected in an ongoing manner and can provide reliable assessment of monthly variability of symptomatic cases presenting at health facilities. The method used here demonstrates how this data can be used for estimation; quantify uncertainty around estimations while at same time adjusting for facility utilisation. The results from this analysis contribute to the characterization

and understanding of the epidemiology of *P. falciparum* and *P. vivax* malaria in Eritrea. The modelled distribution of incidence presented here suggests that a concerted effort is required in Debub and Garsh Barka in addition to implementing cross-border collaboration with Ethiopia and Sudan. The spatial distributions of *P. vivax* revealed here through Bayesian statistical modelling will present a challenge for pre-elimination and elimination especially when it comes to clearing the reservoir infections at the liver stage. Lastly, routine HMIS can be used to identify areas where active case detection can be targeted. However, the interaction of both species in co-distributed *sub-zobas* remains unexplored and future research should focus on this in addition to screening asymptomatic infections at community level.

CHAPTER 5: Discussion and Conclusion

Discussion on advances in mapping malaria in low transmission using HMIS

5.1 Summary

In the last decade, there has been a substantial decline in malaria burden due to increased coverage of interventions. However, variation in the burden of disease exists nationally and at a sub-national level. A recent publication showed that 57% of the population in Africa live in areas with stable transmission (Noor *et al.*, 2014). In some countries, transmission has declined to a level requiring a re-orientation of the malaria programme from that focusing on universal insecticide-treated bednet coverage and control to geographic targeting of interventions. There is also renewed optimism for malaria elimination following the 2008 global call supported by the World Health Organization (World health Organizastion, 2008, Mendis *et al.*, 2009) and 36 countries are now aiming for elimination (Global Health Group, 2011). These include: Nine countries in Latin America and the Caribbean (Argentina, Belize, Costa Rica, Dominican Republic, El Salvador, Mexico, Nicaragua, Panama, Paraguay); 11 countries in the Asia-Pacific region (Bhutan, China, Democratic People's Republic of Korea, Malaysia, Philippines, Republic of Korea, Solomon Islands, Sri Lanka, Thailand, Vanuatu, Vietnam); northern Africa (Algeria), Europe (Azerbaijan, Georgia, Turkey), the Middle East (Iran, Iraq, Saudi Arabia), central Asia (Kyrgyzstan, Tajikistan, Uzbekistan); and six countries in SSA (Namibia, Botswana, South Africa and Swaziland, Cape Verde; and São Tomé and Príncipe).

This thesis has focused on modelling malaria incidence in low malaria transmission countries in Namibia, Afghanistan and Eritrea. These very low and seasonal transmission settings in general pose several challenges to malaria control and elimination. These include: the detection and treatment of symptomatic and asymptomatic infections using suitable tools; developing fine spatial resolution malaria distribution maps to guide malaria control; measuring transmission

patterns at very low parasite density; and dealing with the challenge posed by population movement. In addition, malaria tends to cluster in hotspots and it is harder to detect the low parasite density for both symptomatic and asymptomatic infections (Sturrock *et al.*, 2013b). Here, the focus was on symptomatic infections seen in the public sector. The ability to identify residual infections (asymptomatic) and quantify the population at risk is also critical for pre-elimination or elimination. Malaria elimination also requires tackling poverty, improving infrastructure and strengthening health systems (Smith *et al.*, 2013, World Health Organization, 2014a).

Until recently, HMIS are perceived as less reliable than household surveys (Gething *et al.*, 2006, RBM-MERG, 2008). Whilst symptomatic infections are captured in routine HMIS, in many countries, data are spatially and temporally incomplete due to sporadic reporting. Furthermore, only a proportion of febrile cases are usually seen in the formal health sector (i.e. the '*iceberg effect*') (Agyepong and Kangeya-Kayonda, 2004, Goodman *et al.*, 2007) raising questions about the burden not treated or using the informal sector. This thesis has demonstrated how incomplete HMIS can be used to estimate malaria burden in low transmission settings when coupled with accurate population denominators estimated by adjusting for geographical access and utilisation of healthcare facilities. Moreover, the thesis has contributed to development of novel Bayesian techniques that can be used to harness HMIS data efficiently (Alegana *et al.*, 2013, Alegana *et al.*, 2014). The Bayesian approach provides opportunities for dealing with the deficiencies in HMIS data by smoothing incidence rates in space and time; filling in gaps where no health reports have been assembled; and adjusting for the rate of facility utilisation since only a proportion of actual cases present in the public health sector (Cibulskis *et al.*, 2011). The method

also incorporates ecological or environmental drivers to estimate risk while at same time quantifying uncertainty associated with disease estimates (Banerjee and Fuentes, 2011).

Asymptomatic infections remain a challenge and many countries are yet to adopt active case detection (ACD). These reservoirs of infection can be responsible for sustaining transmission between seasons and cause resurgence if malaria control is withdrawn (Diallo *et al.*, 2012, Sturrock *et al.*, 2013b). Currently, population sample surveys such as the MIS incorporate parasitaemia modules. However, these surveys cannot reliably capture all the asymptomatic infections in their current form. Other drawback of these surveys is the relatively small sample sizes that do not meet the very large sample size requirement at very low transmission historically regarded as areas of <3% parasite prevalence (Yekutiel, 1960). In addition, under low transmission there is a need for a greater temporal frequency to adequately capture the variability of infection rates through time (either monthly or on an annual basis) (Yekutiel, 1960, Beier *et al.*, 1999). This is because malaria is highly seasonal in low transmission settings and driven by the changes in climate at geographically large and small scales. There is also an immediate challenge of current diagnostic tools. Microscopy and RDT remain the recommended tools for detecting infections in routine settings due to their availability and low cost (World Health Organization, 2014c). Although PCR is more sensitive compared to RDT or microscopy, its use in routine settings remains a challenge (Satoguina *et al.*, 2009). The alternative option of measuring EIR is not suitable due to difficulties in capturing sufficient number of infected mosquitoes (Stuckey *et al.*, 2013).

Estimating and relating the number of malaria cases to the population given time period (incidence) to estimate the clinical burden of disease is important to the national malaria control programmes. This is important for resource allocation and provides useful knowledge in carrying out targeted case detection. Targeted case detection has been demonstrated in Swaziland (Sturrock *et al.*, 2013a), Zambia (Davis *et al.*, 2011), Sri-Lanka (Rajakaruna *et al.*, 2010) and Peru (Branch *et al.*, 2005) where reactive case detection identified additional cases in the population where passive cases had been observed. It is, therefore, essential to outline methodological advances in mapping HMIS in low malaria transmission settings since HMIS remains the foundation for gathering evidence, tracking progress on malaria control and identifying areas for rapid response. The recent publication of malaria endemicity maps for Africa (Noor *et al.*, 2014) provides an additional opportunity to compare findings from mapping incidence using HMIS data to parasite prevalence. This discussion is generalized on the use of HMIS in low transmission settings in the three case studies.

5.2 Thesis contribution to malaria strategies in low malaria transmission

The thesis focussed on estimating incidence as an alternative measure to the widely used parasite prevalence. This is because parasite prevalence is less efficient when it comes to low transmission settings. There is a requirement for large sample sizes to detect the low parasite densities in low transmission settings and this are logistically challenging and costly to implement. There is also a need for a greater temporal frequency of parasite prevalence surveys to adequately capture the variability in infection rates due to seasonal variation (Yekutiel, 1960, Beier *et al.*, 1999). This is difficult to achieve with point prevalence surveys. The Force of Infection (FOI) defined as the rate of new infections in the population, such as the sero-

conversion rate (Bekessy *et al.*, 1976, Charlwood *et al.*, 1998), is an alternative measure in low malaria transmission settings. FOI, however, requires a follow up with a specific population group over a certain period of time (Yukich *et al.*, 2012). Further, the Entomological Inoculation Rate (EIR) is hampered by low numbers of positive-sporozoite mosquitoes in low transmission settings (Stuckey *et al.*, 2013). Estimating incidence from HMIS using hierarchical Bayesian spatio-temporal models is therefore an alternative to these approaches.

HMIS data has several advantages, for example, the disease burden (Figure 5.1 below) estimation at the population level using symptomatic cases (at a health facility) increases the likelihood of detecting infections. The temporal spread of data collected in an ongoing manner in HMIS makes it useful in identifying the seasonal patterns. In addition, the spatial representativeness of symptomatic cases seen in health facility data is improved by modelling the catchment populations which come from several clusters around the health facility compared to single cluster prevalence survey. HMIS is also increasingly being used in very low transmission settings to identify areas to target case detection and interventions (Rajakaruna *et al.*, 2010, Davis *et al.*, 2011, Sturrock *et al.*, 2013a). This suggests that HMIS may be instrumental in rolling out ACD and mapping symptomatic infections seen in the public sector is an important contribution.

The overall objective was to assess expectation (mean) incidence and translate this into cases at the population level. Looking at expectations made it possible to use Latent Gaussian Models in R-INLA. Given the nature and completeness of data in Namibia, model set up involved innovations only at the facility level with a single covariate (EVI) to minimise the statistical

problem of over-fitting. In a simplified form $y_1(s,t) = \alpha + \mu(s,t)\beta + e(s,t)$ where $\mu = X^T_{j=1}(s,t)\beta_j$ and $e(s,t) \sim N(0,\tau^2)$ as independent distributed parameter with a zero mean and precision parameter. The latter were assumed to be independently distributed. The $e(s,t)$ parameter was further specified in a hierarchical sense to include the spatial and temporal terms at constituency level via the Besag, York and Mollie model (Besag *et al.*, 1991). This modelling approach addressed several sources of uncertainty in comparison to the crude incidence estimates. First, the model was applied at facility level and, therefore, the method not only takes into account the nature of the facility, but also season and environmental factors (at facility level) in adjusting for under-reporting. Secondly, incidence was smoothed across the facility reports, thereby addressing the potential impact of model instability resulting from small numbers of reported cases, apparent in the facility data. Smoothing incidence also reduces the potential impacts of under-reporting of cases by facilities. Third, incorporating the environmental covariates explained spatial variation where data were absent in addition to providing information on the climatic suitability of malaria transmission in highly receptive districts. The approach 'elevated' incidence in areas where incidence would be below average and 'reduced' incidence in areas where there would be an overestimation. The innovation of the CAR model improved smoothing. For example, In Namibia the CAR estimates were closer to overall mean incidence (smoothing toward the global mean) compared to Afghanistan or Eritrea. This is likely to be a factor of using 0-1 weights as neighbour matrix during the model set up (Section 2.3.5). This means that neighbour areas are correlated and treated as independent if not a neighbour of the region of interest. Other studies have found differences in level of smoothing based on specification of the neighbourhood matrix (Earnest *et al.*, 2007). In the thesis, the problem was

mitigated by introducing higher level random effects at regional level in addition to the facility level effects.

The modelling framework in Afghanistan introduced additional random effects at the province level by modelling $e_2(s,t) = \tau_1 u_1(s,t) + \tau_2 u_2(s,t) + \tau_3 u_3(s,t)$ where $\tau_1 u_1(s,t)$ and $\tau_2 u_2(s,t)$ represented random effects at the district and province levels in addition to the facility random effect $\tau_3 u_3(s,t)$. Model complexity increased given the extra randomisation but improved the smoothing at the district level. In other words, smoothing was closer to the province mean rather than the global mean as seen in the Namibia study. In terms of covariates, both the Namibia and the Afghanistan model specifications treated this as fixed linear effects. The difference between the two model specifications to that for Eritrea was the use of nonlinear effect for covariates. Nonlinear smoothing functions (first order random walk priors) were used to smooth the seasonal effects and improve temporal estimates of incidence. Therefore, the final hierarchical model included complexities in terms of random effects (facility, district and province level) as well as the seasonal component introduced via the covariates. The thesis however, did not validate how well the optimal model could forecast future incidence estimates due to lack of data.

Lastly, the thesis quantifying the proportion of population attending the public sector for malaria treatment based on fever. Novel approaches were used in deriving spatially the public health facility attendance and subsequently deriving catchment population based on the observed patterns. It is important to understand extend of public sector use and infrastructure (in terms of availability, access, quality of services) in countries aiming for elimination.

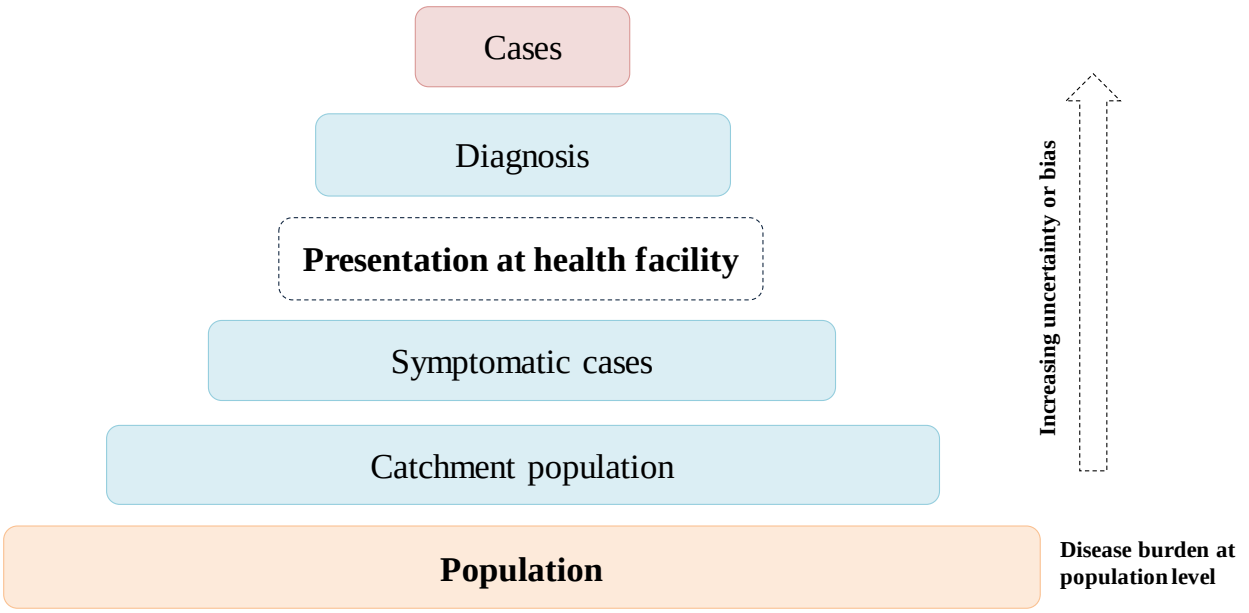

Figure 5.1: Burden estimation model.
While using HMIS data to estimate disease burden, it is vital to first define population catchments since only a proportion of case are observed in formal sector in a well-functioning HMIS. Some of the febrile events may not be due to malaria, after suspected or parasitological examination. Incidence estimated using the catchment population is translated to burden estimate at population level.

5.3 Common emerging themes across the three case study countries

5.3.1 Denominator estimation for modelling incidence by country

To measure incidence at a health facility, a reliable denominator of the catchment population from whom the cases are drawn was required. Secondly, it was also important to recognize that not all the febrile cases within the catchment area would seek treatment or use the public health sector. In practice, multiple sectors (formal or informal) are often used and most fevers are likely to be transient and resolve without treatment (Agyepong and Kangeya-Kayonda, 2004, Goodman *et al.*, 2007). An analysis of public health sector utilisation was used to derive catchments in the three case-study countries.

In the three case studies, there were similarities in healthcare attendance for fever treatment with declining utilisation as travel time increased. The probability of using the closest facility was low

for the population at a greater distance from a public health facility even when stratified by facility type (e.g in Afghanistan). Secondly, over 65% of the population was within three hours' travel time to the nearest health facility in the three settings. When three hours' travel time was used to assess population coverage by public healthcare facilities, these translated to coverage rates of 67% in Namibia, 79% in Eritrea and 85% in Afghanistan. These three hour threshold was used to delineate the boundaries of the catchment areas. In addition, the probability of fever treatment within the catchment areas was used to define the number of people by travel to the health facility and number of fever cases likely to be seen at the facility. Elsewhere in other health facility utilisation studies, the phenomenon of by-passing the nearest health facility has been observed (Akin and Hutchinson, 1999), but, this generally was attributed to seeking better quality of services or more specialised treatment. Some hospital-based studies have shown that longer travel times have also been associated with shorter time to death (Manongi *et al.*, 2014) or disease severity (Moisi *et al.*, 2011). Although the available data do not allow for investigation of the phenomenon of by-passing, the results from the three case studies in Namibia, Afghanistan and Eritrea generally suggested that the majority of cases travelled shorter distances to the nearest health facility based on the distance-decay pattern. It is also possible that long travel times in the three case studies could be associated with the phenomenon of by-passing for better quality healthcare or for cultural reasons.

5.3.2 Implications for malaria control and elimination based on estimated incidence in the three case study countries

In terms of modelling incidence, the thesis started with an analysis of the 2009 malaria case data from public health facilities in Namibia. In Eritrea data were available from the year 2010 to

2012. The focus in Afghanistan was to track progress towards the national target between 2006 and 2009 and provide estimates of clinical burden of *P. falciparum* and *P. vivax*. In Eritrea, a similar analysis to Afghanistan in terms of prevalence of both *P. falciparum* and *P. vivax* was conducted but also focused on seasonality. The spatio-temporal maps of malaria incidence for *P. falciparum* and *P. vivax* are important in quantifying populations at risk to target interventions.

Although Namibia is currently in pre-elimination, the mean *P. falciparum* incidence in 2009 was 12.5 (95% Crl 10.4-15.5) per 1000 population which was higher compared to the estimated mean incidence in Afghanistan (1.2, 95% Crl 0.4-2.9 per 1000 population) or in Eritrea (3.4, 95% Crl 2.2-5.2 per 1000 population). Only Afghanistan had *P. vivax* as the dominant malaria parasite with estimated incidence of 5.4 (95% Crl 3.2-9.2) per 1000 population. In Eritrea the incidence of *P. vivax* was 2.5 (95%Crl 1.5-3.9) per 1000 population and constituted approximately 43% of the estimated burden. Recent reports suggest that the three case studies countries are on path in reducing malaria cases by >75% (Smith Gueye *et al.*, 2014, World Health Organization, 2014d) and results here support orientation of national programmes to pre-elimination. Estimated malaria incidence, however, varied at a sub-national level. Elimination programmes should probably commence with low incidence districts, for example, the southern health districts in Namibia, the two red sea regions in Eritrea and the provinces bordering Tajikistan for the elimination of *P. falciparum* in Afghanistan. Tajikistan is already in elimination with only seven malaria cases reported in 2013 (World Health Organization, 2014d).

Malaria incidence in the three countries tended to cluster in marginalised populations, for example, in the border areas. This also suggested that malaria in these regions may be attributed

to cross border population movement. For Namibia, there was elevated incidence at the border with Angola and Zambia. In Eritrea incidence was higher in regions that bordered Ethiopia while for Afghanistan these were in districts close to Pakistan. In these countries, cross-border malaria may pose a threat to the elimination efforts. Namibia has subsequently started cross border initiatives with Angola and Zambia. These include the Trans-Kunene Malaria Initiative (TKMI) with Angola (Smith Gueye *et al.*, 2014) and the Trans-Zambezi Malaria Initiative (TZMI) (Ministry of Health and Social Services, 2010c, Trans-Zambezi Malaria Initiative (TZMI), 2012, Noor *et al.*, 2013b). For Afghanistan there is a cross border initiative with Tajikistan to support elimination of falciparum malaria (World Health Organization, 2007b). Eritrea is yet to start such coordination with neighbouring countries. The cross-border activities include screening at health facilities, treatment and distribution of LLINs to meet universal coverage targets.

The monthly trend in incidence was different across the sites. For Namibia, the temporal trend of *P. falciparum* was only for 2009 but showed that there was a peak in cases early in the year. In Afghanistan, the seasonal peak in incidence was different for the two malaria parasites. *P. vivax* incidence peaked in August while *P. falciparum* peaked later in the year, in November. In Eritrea, the incidence of *P. falciparum* and *P. vivax* peaked between September and October. The difference in the patterns observed between *P. falciparum* and *P. vivax* parasites and countries may be attributed to the transmission dynamics in different settings in addition to other factors such as infrastructure, poverty and health systems. A factor to consider is receptivity i.e. favourable conditions for transmission which may change the results from the case studies. For example in Namibia, epidemics have been observed in years when above average precipitation has been experienced (Noor *et al.*, 2013c). The temporal profile is useful in planning the timing

of interventions. Thus the difference in peak seasons should inform control strategies, for example, lagging the use of IRS with peak case incidence. Across the three case studies, IRS varies. In 2008 IRS use in Namibia was only 16% rising to approximately 41% by 2012 (Smith Gueye *et al.*, 2014), approximately 30% in Eritrea and was only launched in 2012 in Afghanistan (Ministry of Public Health, 2008b). Along with IRS, universal coverage should be targeted in areas with higher incidence and in line with recent WHO recommendation on managing insecticide resistance (World Health Organization, 2014b).

Relating the modelled incidence to current maps of parasite prevalence showed areas with higher incidence also exhibited high prevalence. For Eritrea and Afghanistan where both *P. falciparum* and *P. vivax* exist, there was a spatial co-distribution of the two parasites in malaria risk areas. Where there was elevated incidence, this was associated with high prevalence from the parasite prevalence surveys. The latter was modelled from community prevalence surveys age-standardised (2-10 years) (Noor *et al.*, 2014). The positive correlation between incidence and prevalence is expected and higher incidence is associated with higher parasite prevalence (Figure 5.2). In Eritrea and Afghanistan there was a spatial co-distribution of *P. vivax* and *P. falciparum*. This co-distribution poses a challenge in managing mixed infections, requiring careful case management strategies.

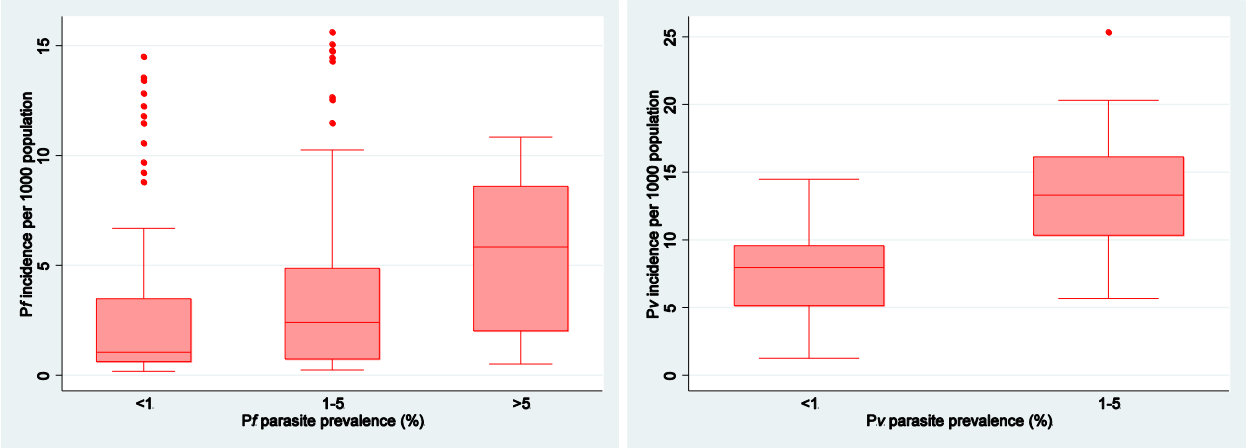

Figure 5.2: Comparisons of incidence and with community parasite prevalence
Box plots showing comparisons between modelled incidence per 1000 population and the mean parasite prevalence for *P. falciparum* and *P. vivax*. The mean incidence is higher in parasite prevalence greater than 5%. Parasite prevalence was based on community surveys for 2010. Incidence was estimated from health facility data (total cases as numerator) and catchment population as the denominator.

5.3.3 Reaction of the case study countries to findings

Overall, the methods used in this thesis aimed at producing maps at the district level relevant for decision making to various malaria control programmes. The national programmes were interested in the overall case burden for prospective planning of malaria interventions. The NMCPs were involved in each of the three case studies by providing data and reports used in this thesis. Both Namibia and Afghanistan studies have been jointly published with the NMCPs (journal articles in the appendices). For Namibia, some of study findings have been incorporated in nation policy documents (MoHSS, 2011) and cited in other studies (Lourenco *et al.*, 2013, Smith Gueye *et al.*, 2014). The Eritrea NMCP has used findings from this study (district level incidence) in prospective planning of malaria control activities. The Eritrea study, however, is yet to be published in a peer review journal.

5.4 Transferability, scalability and challenges for malaria elimination

5.4.1 Data availability

In terms of surveillance, health facilities remain the foundation for identifying and tracing the changing burden of disease. School sentinels remain an option to routine household surveys. Their utility has already been demonstrated in several studies (Ashton *et al.*, 2011, Gitonga *et al.*, 2012). However, there is still a gap in studies that have delineated and combined school catchments with disease incidence. This is in contrast to health facility-based data which are collected on a continuous basis, and provide information from symptomatic (febrile) infections. There is also an increasing availability of e-health frameworks in at least 46 countries globally (DHIS2, 2014) with improved systems of monitoring by district level health management information systems including the use of mobile phones. This system is complete in eight countries in Africa (Kenya, Rwanda, Uganda, Tanzania, Zanzibar, Zambia, Ghana, Gambia and Liberia) and in two countries in Asia (Bangladesh and India). Of these Rwanda and Zanzibar are amongst low malaria transmission countries in Africa (World Health Organization, 2014d). The rest are either at a pilot stage or have partially rolled out such a system nationally (Figure 5.3). Such open data initiatives will favour scaling up of methodologies developed in this thesis in updating and developing high space-time resolution disease maps in reduced transmission or support malaria elimination.

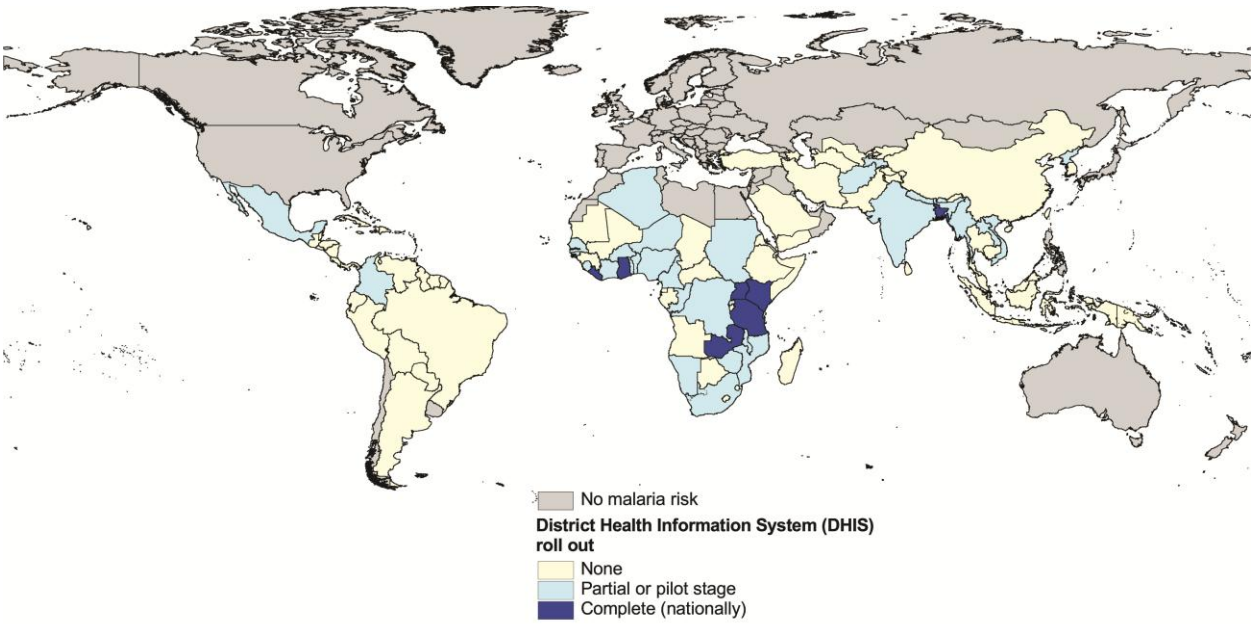

Figure 5.3: Global roll out of DHIS initiative
Countries where the e-health initiatives has been rolled out integrating national health facility data reporting systems with surveillance at a national level with (DHIS2, 2014). It is also important to note the availability of nationally representative household surveys in these countries in this report (Carolina *et al.*, 2013).

5.4.2 Development of Reactive Case Detection approaches in low transmission settings for elimination

Asymptomatic infections tend to cluster in hotspots in low transmission settings (Bousema *et al.*, 2012, Sturrock *et al.*, 2013b). These residuals infections in the population can be harmful to elimination programmes and may be responsible for maintaining transmission between seasons or cause resurgence of disease when control is withdrawn. The challenge in low transmission settings is that of identifying the asymptomatic cases usually via Active Case Detection (ACD). There two broad approaches: reactive case detection (RACD) and proactive case detection (PACD) (Sturrock *et al.*, 2013b). The reactive approach involves using the passive system (HMIS) to identify the origin of cases. It is most widely used in countries with better HMIS and where burden is reduced to just a few residual cases experiencing short epidemics. It is also suitable for low transmission where receptivity is high. The PACD does not involve the passive

system. The entire population is screened to identify residual infections, which can often be laborious and logistically challenging. The passive system should complement the PACD since incidence estimated at health facility catchment-level is translated to the wider population.

As demonstrated using HMIS in the three case studies, population catchments can be used to model incidence. The advantage of using health facilities is that symptomatic infections data are collected on a continuous basis. There is also an added advantage of reducing the cost of tracing cases within the wider population with limited resources. Studies in Zambia (Davis *et al.*, 2011) and Swaziland (Sturrock *et al.*, 2013a) provide examples of where passive and active case detection have been combined to identify asymptomatic infections. In Swaziland, 79 additional cases to those presenting in the formal health sector were identified among the wider population. In Zambia, a case-control study by Stressman *et al.* (2010) showed that there were more RDT-positive cases in the group targeted via RACD than in a selected random control group. These two studies illustrated the usefulness of targeted screening when identifying reservoirs of asymptomatic cases.

5.4.3 Implications for mapping *P. falciparum* and *P. vivax* in low transmission settings

Progress in identifying symptomatic cases within the population has important application for asymptomatic case detection. Both PACD and RACD will benefit from improved mapping of passive detected cases. The PACD usually targets peak transmission seasons in known geographic areas while RACD is triggered by cases seen in the formal health sector. However, the questions related to optimizing diagnostic tools and techniques remain. Current WHO recommendation is the use of RDT and microscopy for diagnosis even in low transmission

setting because of relative low cost and they are widely available (World Health Organization, 2014c). The use of more sensitive nucleic acid amplification (NAA) such as PCR should be applied in areas where the RDT and microscopy are already in use to support elimination.

There is however another challenge in areas where multiple parasites exist (Cotter *et al.*, 2013). Although most successful interventions have focused on *P. falciparum* more effective approaches need to be developed in detecting *P. vivax* infections since it is less responsive to current interventions. There has been some progress in mapping this parasite at the global level (Gething *et al.*, 2012). However, there is still a need to improve these maps at country level, especially outside sub-Saharan Africa. In countries where *P. vivax* exists, the challenges of elimination are considerably higher due to the biological characteristics of this parasite (Mueller *et al.*, 2009, Cotter *et al.*, 2013). For instance, *P. vivax* can exist at very low parasite densities that are difficult to detect using the recommended tools. It also exhibits a dormant liver stage responsible for most relapses weeks or months after an initial attack (White, 2011). This complicates the ability to detect asymptomatic vivax infections within the population, while treatment is compounded by possible adverse reactions to drugs in populations with Glucose-6-Phosphate Dehydrogenase (G6PD) deficiency (Howes *et al.*, 2012). Although there has been attempts to develop G6PD maps at a global level (Howes *et al.*, 2012), there is still a gap in fine resolution mapping at coutry level which may inform case magement of *P. vivax*.

5.5 Limitations

5.5.1 Thesis limitations and recommendations

The majority of the limitations in this thesis have been addressed in each case study fully. Here a summary of the limitations and recommendations in a general sense is provided.

The methodology and presentation of the thesis focused on three case studies (countries) with low malaria transmission. First, these three countries were selected because they met the experimental requirements (low transmission) and due to data availability. They were, thus, not selected randomly. However, the strength of the methodology presented and discussed in this thesis applies to similar transmission settings as discussed in the introduction section and in the respective case studies. Thus, the approaches are relevant mostly to countries pursuing pre-elimination or in the elimination stages at the margins of transmission (Feachem *et al.*, 2010).

Secondly, there was variation between the temporal spread in the data between countries. For example, in Namibia only one-year data was available compared to three years in Eritrea and four years for Afghanistan. This short temporal spread in each context was sufficient for developing the methods and studying disease incidence, but, was not sufficient in providing long-term assessment of change in incidence over time or evaluating the impact of interventions. This was also evident on the less significant effects of environmental covariates on disease modelling, for example, in Afghanistan. Although there an attempt was made to quantify change of incidence in Afghanistan, the short four year period was not enough to reveal significant impact of interventions on disease. Future studies should revisit this aspect of establishing changes in disease incidence over long periods of time (for example, 10 years) and quantify the

impact of interventions to benefit the goal of elimination. In addition, such long-term trends may also improve findings on the impact of covariates for disease modelling.

The thesis did not explore the potential uncertainty that may result from misclassification of cases especially those related to parasite classification. These were beyond the scope of the study given that such errors and other may occur at data entry stage at health facility level. While the model performance was satisfactory, these errors may contribute to uncertainty and unexplained variance.

Throughout the three case studies, the question of the limitations of current diagnostics approaches (RDTs and microscopy) remains and has also been outlined in this final chapter. It is not only the quality of routine microscopy or RDTs that remains a limitation but also their respective sensitivities when it comes to low parasite density (Satoguina *et al.*, 2009). A review of the literature showed that PCR and serology are preferred in most observational studies and in some longitudinal studies, but, over small areas. Routine use of these approaches at a national level remains a challenge. Nonetheless, microscopy and RDT provide some indication of reservoirs of infection in symptomatic cases seen at a health facility and are currently recommended by the WHO for low transmission. Thus, they provide a framework where a sophisticated diagnosis can be targeted (World Health Organization, 2014c).

Another limitation relates to the reliance on self-reported data from national household surveys on fever. There is the possibility of variation in perception and interpretation of fever episodes between individuals. In addition, the fever variable reported in these surveys is usually based on

the two-week recall period which could introduce a bias. Furthermore, the treatment seeking behaviour of self-reported two-week period fever may differ from that of suspected cases seen at the health facility (Cibulskis *et al.*, 2011). These limitations remain in the three case studies and could have an impact on 'true' fever burden. Modelling the fever burden was not a specific objective of this thesis. The interpretation of fever burden estimated in the thesis in each case study should therefore be undertaken with this in mind. In addition, fever episode could be multiple over the two week period resulting in a different cause of action depending on perceived need by the individual. The multiple fever episodes may also not necessarily be malaria-specific, since malarial fever declines in proportion to transmission intensity (Trape *et al.*, 2014). The multiple fever episodes often result in multiple treatments in different sectors (either formal or informal healthcare). Here, only public healthcare utilisation was of interest and future studies should endeavour to define healthcare use in other sectors.

Lastly, the aspect of population movement and impact on defining cases seen at the health facility has not been addressed by this study. Population movement plays an important role in transmission and is potentially responsible for most of the cases seen in the border areas in Namibia, Afghanistan and Eritrea. Several aspects warrant further investigation. For example, what proportion of cases in Kabul in Afghanistan is acquired locally or imported from neighbouring regions that have higher endemicity? This relationship between incidence in HMIS (passive surveillance) and internal (between regions) or external population movement was not addressed by this thesis due to lack of data. It, therefore, remains a gap to be filled by future studies.

5.6 Future research

The thesis focused mainly on developing Bayesian approaches that can use HMIS data to estimate incidence. A frequentist approach had been tried earlier using the Kenya HMIS data (Gething *et al.*, 2006). In the Kenya application, however, the denominator was a sum of total case burden seen at a health facility as a proxy to catchment population. Moreover, the setting was in a high endemic country. The Bayesian approach used in this thesis lays the foundation for modelling uncertainty in HMIS data using Gaussian latent models. The framework started by modelling the denominator (the catchment population) at the health facility regulated by the probability of seeking treatment when sick with fever followed by a Bayesian analysis of incidence. Although there was attempt to relate the estimated incidence to parasite prevalence in a frequentist approach, future studies could focus on modelling this using a model-based approach to incorporate uncertainty. There has already been attempt on quantifying such a relationship (Patil *et al.*, 2009) but not using estimates from the modelling framework used here.

Future research can also make improvement in relation to the assumptions laid out in the model set up and additional operational and research challenges on burden estimation in low transmission settings. The first relate to modelling the malaria fever burden at a fine resolution. Effectively, geostatistical approaches could be used in fever burden estimation at a fine resolution. A major challenge is the specificity of malaria-fever cases in relation to other co-morbidities. Thus, the resulting fever burden map will depend on the definition of fever cases attributed to malaria. A recent study conducted in Senegal suggested malaria-specific fever reduced with declining endemicity (Trape *et al.*, 2014). The improved fever burden maps could

potentially be useful in assessing population that is likely to seek treatment for malaria and eventual case loads seen at the peripheral health facility.

In addition, the fever-treatment seeking patterns defined in this study can be modelled differently given different data inputs. Firstly, a major recommendation is to explore models based on spatial interaction rather than the assumption based on distance. An easy option is to include a question of where treatment for malarial-fever was sought in the national representative surveys such as DHS or MIS. However, given the frequency of national surveys in areas with low malaria transmission, it is difficult to assess whether this would be representative for different time periods and with varying climatic and disease endemicities. A second alternative approach would be to explore the utility of mobile phones in defining the treatment seeking behaviour in febrile populations. The latter, however, requires a validation of the mobile phone data to test the observed patterns as actual febrile-case movements. The third alternative involves the use spatially varying parameters fitted to the treatment seeking pattern. This approach is different from using national estimated mean parameters applied to travel times or distance metrics. Treatment seeking behaviour is likely to vary by region, based on availability of health facilities and accessibility. Thus, the parameters driving the three parameter regression used in this study could vary substantially by region altering the treatment seeking patterns observed here.

There is, however, a larger policy relevant question on the effect of human mobility in low transmission settings. This is, however, dependent on the availability of long-term morbidity data as well as population movement information. Under morbidity, the main challenge would be to evaluate the impact of various interventions on incidence while at the same time accounting for

population movement. Regarding mobility, the challenge would involve combining both incidence and morbidity data to define risk within the population. The difficulty is in obtaining mobility data as these are not always available to the public domain regardless of whether they are census, national infrastructure surveys or from mobile phone records. Although some research has been undertaken using mobile phones (James and Versteeg, 2007, Le Menach *et al.*, 2011, Zurovac *et al.*, 2012, Tatem *et al.*, 2014), it would be interesting to assess how these mobile populations alter disease dynamics in different regions within and between the countries.

5.6.1 Improved Bayesian approach to downscaling incidence in low transmission settings

Bayesian approaches have attractive properties when it comes to modelling which has already been mentioned in the preceding chapters. This includes the ability to use hierarchical models with uncertainty handled at different levels, particularly the latent Gaussian models. This class of models is within the range of application of this thesis although different models can be used in practice. Overall, the methods used in this thesis aimed at producing maps at the district level relevant for decision making to various malaria control programmes. Moreover, these national programmes demanded figures translated into number of cases for planning malaria interventions. Thus, the advantages were two-fold in improving the methods in a Bayesian statistical sense as well as providing policy relevant tools for various national malaria control programmes.

The modelling approach can be improved in two different ways. Firstly, the modelling can aim at analysing the co-variation of the two parasites spatially and temporally in countries with two or more parasites, such as Eritrea and Afghanistan. In other words, this involves setting up a joint

modelling framework such that $\nu_1 = z_1 + z_2$ and $\nu_2 = z_1 - z_2$ where ν_1 and ν_2 assess the co-relation between vivax and falciparum in a hierarchical sense while z_1 , z_2 define the two expectations common for the former and difference for the latter. Thus, we have perfect correlation when z_2 is zero. Secondly, the incidence maps can be improved by downscaling incidence from polygon representation to pixel level. This modelling framework is different from that specified in this thesis where facility-catchment level incidence was aggregated to the district level. The latter proposal will involve disaggregating incidence observed at facility level to pixel level such that risk is defined at fine resolution in a linear model of co-regionalisation (LMC) approach (Barnerjee *et al.*, 2004). Examples of this model specification exist in the literature (Gelfand *et al.*, 2004). It is then worth looking at the distribution of risk within the catchment and population most at risk to improve burden estimation. Future modelling work will involve this risk disaggregation to further improve burden estimates.

5.7 Conclusion

The main aim of thesis was to model healthcare utilisation and estimate malaria incidence in three low transmission countries from incomplete HMIS data. This was achieved using case studies in Namibia, Eritrea and Afghanistan. A denominator, the heath facility catchment population, was estimated and a Bayesian approach used to model incidence with uncertainty. There was a need to adjust for health facility utilisation, incomplete reporting and the use of clinical diagnosis using a slide positivity rate where cases are not confirmed. The modelled malaria incidence was translated into overall clinical burden to identify areas where malaria control can be improved for each case study. The results have been made available to various national malaria control programmes to aid in resource allocation, planning and policy.

The methodology illustrated here is highly relevant to countries in similar settings (i.e. aiming for pre-elimination and elimination) and where changes in incidence are more relevant as opposed to point prevalence surveys. The approach demonstrated the usefulness of incomplete HMIS data for malaria burden estimation in low transmission settings. HMIS data are readily available via national reporting systems compared to community prevalence surveys which would require intensive sampling in low malaria transmission settings. The HMIS data are also relevant for tracking change in incidence over time - given a longer time-series of data. Tracking the change in incidence over time and seasonality is not only useful in identifying high risk regions in space and time, but could also be useful in evaluating programme performance and the impact of interventions.

An improvement in the quality of HMIS data and its availability through DHIS 2 initiatives may undoubtedly increase the use and modelling of HMIS for malaria burden estimation with approach demonstrated here. As malaria becomes highly marginalised in a few hotspots in low transmission settings, future research should consider combining approaches in this thesis combined with active case surveillance at community level for targeted control.

REFERENCES

Abdalla, S. H. and Pasvol, G. (Eds.) (2004) *Malaria: A hematological perspective,* London, Imperial college press.

Abouzahr, C. and Boerma, T. (2005) Health information systems: the foundations of public health. *Bull World Health Organ,* 83**,** 578-83.

Abu-Mourad, T., Alegakis, A., Shashaa, S., Koutis, A. and Philalithis, A. (2008) Individual determinants of primary healthcare utilisation in Gaza Strip, Palestine. *Journal of Epidemiology and Community Health (1979-),* 62**,** 701-707.

Acerra, J., Iskyan, K., Qureshi, Z. and Sharma, R. (2009) Rebuilding the health care system in Afghanistan: an overview of primary care and emergency services. *International Journal of Emergency Medicine,* 2**,** 77-82.

Aday, L. A. and Andersen, R. (1974a) A framework for the study of access to medical care. *Health Services Research,* 9**,** 208-220.

Aday, L. A. and Andersen, R. (1974b) A fremework for the study of access to medical care. *Health Serv Res,* 9**,** 208-220.

Adedoyin, M. A. and Fagbule, D. (1992) Splenomegaly, malarial parasitaemia and anaemia in two Nigerian villages. *Cent Afr J Med,* 38**,** 371-5.

Afghan Public Health Institute, Ministry of Public Health (APHI/MoPH) [Afghanistan], Central Statistics Organization (CSO) [Afghanistan], ICF Macro, Indian Institute of Health Management Research (IIHMR) [India] and [Egypt], W. H. O. R. O. f. t. E. M. W. E.(2011) Afghanistan Mortality Survey 2010, Calverton, Maryland, USA, APHI/MoPH, CSO, ICF Macro, IIHMR and WHO/EMRO.

Agyepong, I. A. (1992) Malaria: ethnomedical perceptions and practice in an Adangbe farming community and implications for control. *Soc Sci Med,* 35**,** 131-7.

Agyepong, I. A. and Kangeya-Kayonda, J. (2004) Providing practical estimates of malaria burden for health planners in resource-poor countries. *Am J Trop Med Hyg,* 71**,** 162-7.

Agyepong, I. A. and Manderson, L. (1994) The diagnosis and management of fever at household level in the Greater Accra Region, Ghana. *Acta Tropica,* 58**,** 317-330.

Airey, T. (1989) The impact of road construction on hospital in-patient catchments in the Meru District of Kenya. *Soc Sci Med,* 29**,** 95-106.

Airey, T. (1992) The impact of road construction on the spatial characteristics of hospital utilization in the Meru district of Kenya. *Soc Sci Med,* 34**,** 1135-46.

Akin, J. S. and Hutchinson, P. (1999) Health-care facility choice and the phenomenon of bypassing. *Health Policy Plan,* 14**,** 135-51.

Akogun, O., Gundiri, M., Badaki, J., Njobdi, S., Adesina, A. and Ogundahunsi, O. (2012) Febrile illness experience among Nigerian nomads. *International Journal for Equity in Health,* 11**,** 5.

Akpede, G. O., Abiodun, P. O. and Sykes, R. M. (1992) Relative contribution of bacteraemia and malaria to acute fever without localizing signs of infection in under-five children. *J Trop Pediatr,* 38**,** 295-8.

Akpede, G. O., Abiodun, P. O. and Sykes, R. M. (1993) Acute fevers of unknown origin in young children in the tropics. *J Pediatr,* 122**,** 79-81.

Al-Taiar, A., Jaffar, S., Assabri, A., Al-Habori, M., Azazy, A., Al-Gabri, A., Al-Ganadi, M., Attal, B. and Whitty, C. J. (2008) Who develops severe malaria? Impact of access to healthcare, socio-economic and environmental factors on children in Yemen: a case-control study. *Trop Med Int Health,* 13**,** 762-70.

Alegana, V., Wright, J., Petrina, U., Noor, A., Snow, R. and Atkinson, P. (2012) Spatial modelling of healthcare utilisation for treatment of fever in Namibia. *International Journal of Health Geographics,* 11**,** 6.

Alegana, V. A., Atkinson, P. M., Wright, J. A., Kamwi, R., Uusiku, P., Katokele, S., Snow, R. W. and Noor, A. M. (2013) Estimation of malaria incidence in northern Namibia in 2009 using Bayesian conditional-autoregressive spatial-temporal models. *Spatial and Spatio-temporal Epidemiology,* 7**,** 25-36.

Alegana, V. A., Wright, J. A., Nahzat, S. M., Butt, W., Sediqi, A. W., Habib, N., Snow, R. W., Atkinson, P. M. and Noor, A. M. (2014) Modelling the Incidence of *Plasmodium vivax* and *Plasmodium falciparum* Malaria in Afghanistan 2006 - 2009. *PLoS One,* 9**,** e102304.

Allam, M. M., Alkadarou, T. A., Ahmed, B. G., Elkhair, I. S., Alansary, E. H., Ibrahim, M. E., Elhassan, A. M. and Elhassan, I. M. (2008) Hyper-reactive Malarial Splenomegaly (HMS) in malaria endemic area in Eastern Sudan. *Acta Trop,* 105**,** 196-9.

Amek, N., Bayoh, N., Hamel, M., Lindblade, K. A., Gimnig, J., Laserson, K. F., Slutsker, L., Smith, T. and Vounatsou, P. (2011) Spatio-temporal modeling of sparse geostatistical malaria sporozoite rate data using a zero-inflated binomial model. *Spatial and Spatio-temporal Epidemiology,* 2**,** 283-290.

Ameli, O. and Newbrander, W. (2008) Contracting for health services: effects of utilization and quality on the costs of the Basic Package of Health Services in Afghanistan. *Bull World Health Organ,* 86**,** 920-8.

Amin, A. A., Marsh, V., Noor, A. M., Ochola, S. A. and Snow, R. W. (2003) The use of formal and informal curative services in the management of paediatric fevers in four districts in Kenya. *Tropical Medicine & International Health,* 8**,** 1143-1152.

Amouzou, A., Habi, O. and Bensaid, K. (2012) Reduction in child mortality in Niger: a Countdown to 2015 country case study. *Lancet,* 380**,** 1169-78.

Andersen, R. M., McCutcheon, A., Aday, L. A., Chiu, G. Y. and Bell, R. (1983) Exploring dimensions of access to medical care. *Health serv Res.,* 18**,** 49-74.

Anderson, T. R. (1956) Potential models and the spatial distribution of population. *Papers in Regional Science,* 2**,** 175-182.

Angers, J.-F. o. and Biswas, A. (2003) A Bayesian analysis of zero-inflated generalized Poisson model. *Computational Statistics & Data Analysis,* 42**,** 37-46.

Apparicio, P., Abdelmajid, M., Riva, M. and Shearmur, R. (2008) Comparing alternative approaches to measuring the geographical accessibility of urban health services: Distance types and aggregation-error issues. *Int J Health Geogr,* 7**,** 7.

Apparicio, P. and Seguin, A.-M. (2006) Measuring the Accessibility of Services and Facilities for Residents of Public Housing in Montreal. *Urban Studies,* 43**,** 187-211.

Arino, O., Gross, D., Ranera, F., Bourg, L., Leroy, M., Bicheron, P., Latham, J., Di Gregorio, A., Brockman, C., Witt, R., Defourny, P., Vancutsem, C., Herold, M., Sambale, J., Achard, F., Durieux, L., Plummer, S. and Weber, J. L. (2007) GlobCover: ESA service for global land cover from MERIS. *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS) 2007.* Barcelona, IEEE International.

Ashton, R., Kefyalew, T., Tesfaye, G., Pullan, R., Yadeta, D., Reithinger, R., Kolaczinski, J. and 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.

Babyak, M. A. (2004) What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosom Med,* 66**,** 411-21.

Bailey, C. T. and Gatrell, C. A. (1995) *Interactive Spatial Data Analysis,* Essex, England, Longman Scientific & Technical.

Balk, D., Pozzi, F., Yetman, G., Deichmann, U. and Nelson, A.(2004) The distribution of people and the dimension of place: Methodologies to improve global population estimates in urban and rural areas. , New York, CIESIN, Columbia University.

Banerjee, S. and Carlin, B. P. (2003) Semiparametric spatio-temporal frailty modeling. *Environmetrics,* 14**,** 523-535.

Banerjee, S. and Fuentes, M. (2011) Bayesian modeling for large spatial datasets. *Wiley Interdisciplinary Reviews: Computational Statistics,* 4**,** 59-66.

Banerjee, S., Gelfand, A. E., Finley, A. O. and Sang, H. (2008) Gaussian predictive process models for large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 70**,** 825-848.

Barnerjee, S., Carling, P. B. and Gelfand, A. E. (2004) *Hierarchical modeling and analysis for spatial data,* London, Chapman & Hall/CRC.

Barnighausen, T. and Bloom, D. E. (2009) Financial incentives for return of service in underserved areas: a systematic review. *BMC Health Serv Res,* 9**,** 86.

Bastiaens, G. J. H., Bousema, T. and Leslie, T. (2014) Scale-up of Malaria Rapid Diagnostic Tests and Artemisinin-Based Combination Therapy: Challenges and Perspectives in Sub-Saharan Africa. *PLoS Med,* 11**,** e1001590.

Basu, S., Andrews, J., Kishore, S., Panjabi, R. and Stuckler, D. (2012) Comparative Performance of Private and Public Healthcare Systems in Low- and Middle-Income Countries: A Systematic Review. *PLoS Med,* 9.

Bauze, A. E., Tran, L. N., Nguyen, K. H., Firth, S., Jimenez-Soto, E., Dwyer-Lindgren, L., Hodge, A. and Lopez, A. D. (2012) Equity and geography: the case of child mortality in Papua New Guinea. *PLoS One,* 7**,** e37861.

Bayes, T. (1763) An Essay Towards Solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society of London,* 53**,** 370-418.

Beier, J. C., Killeen, G. F. and Githure, J. I. (1999) Short report: entomologic inoculation rates and Plasmodium falciparum malaria prevalence in Africa. *Am J Trop Med Hyg,* 61**,** 109-13.

Beiersmann, C., Sanou, A., Wladarsch, E., De Allegri, M., Kouyate, B. and Muller, O. (2007) Malaria in rural Burkina Faso: local illness concepts, patterns of traditional treatment and influence on health-seeking behaviour. *Malaria Journal,* 6**,** 106.

Bekessy, A., Molineaux, L. and Storey, J. (1976) Estimation of incidence and recovery rates of Plasmodium falciparum parasitaemia from longitudinal data. *Bull World Health Organ,* 54**,** 685-93.

Belay, T.(2010) Building on Early Gains in Afghanistan's Health, Nutrition and Population Sector: *Challenges and Options,* Washington (D.C.), World Bank.

Bell, D. R., Wilson, D. W. and Martin, L. B. (2005) False-positive results of a Plasmodium falciparum histidine-rich protein 2-detecting malaria rapid diagnostic test due to high sensitivity in a community with fluctuating low parasite density. *Am J Trop Med Hyg,* 73**,** 199-203.

Bennett, R. J. (1979) *Spatial time series: analysis-forecasting-control,* London, Pion Limited.

Bennett, S. (1992) Promoting the private sector: a review of developing country trends. *Health Policy and Planning,* 7**,** 97-110.

Berman, P., Nwuke, K., Hanson, K., Mbugua, K., Ngugi, T., Omurwa, J. and Ong'ayo, S.(1995) Kenya: Non-governmental healthcare provision. Data for decision making project Harvard School of Public Health., Washington, Havard University and USAID.

Bernardinelli, L., Clayton, D., Pascutto, C., Montomoli, C., Ghislandi, M. and Songini, M. (2007) Bayesian analysis of space-time variation in disease risk. *Stat Med,* 14**,** 2433-43.

Bernardinelli, L. and Montomoli, C. (1992) Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. *Stat Med,* 11**,** 983-1007.

Bernardinelli, L., Pascutto, C., Best, N. G. and Gilks, W. R. (1997) Disease mapping with errors in covariates. *Stat Med,* 16**,** 741-52.

Besag, J. (1974) Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society. Series B (Methodological),* 36**,** 192-236.

Besag, J., York, J. and Mollie, A. (1991) Bayesian image restoration, with two applications in spatial statistics. . *Annals of the Institute of Statistical Mathematics,* 43**,** 1-59.

Bhaduri, B., Bright, E., Coleman, P. and Urban, M. (2007) LandScan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics. *GeoJournal,* 69**,** 103-117.

Bi, Y., Hu, W., Liu, H., Xiao, Y., Guo, Y., Chen, S., Zhao, L. and Tong, S. (2012) Can slide positivity rates predict malaria transmission? *Malaria Journal,* 11**,** 117.

Black, J. A. (1981) *Urban Transport Planning: Theory and Practice,* Croom Helm.

Black, R. E., Morris, S. S. and Bryce, J. (2003) Where and why are 10 million children dying every year? *Lancet,* 361**,** 2226-34.

Blangiardo, M., Cameletti, M., Baio, G. and Rue, H. v. (2013) Spatial and spatio-temporal models with R-INLA. *Spatial and Spatio-temporal Epidemiology,* 4**,** 33-49.

Bockarie, M. and Dagoro, H. (2006) Are insecticide-treated bednets more protective against Plasmodium falciparum than Plasmodium vivax-infected mosquitoes? *Malaria Journal,* 5**,** 15.

Boerma, J. T. and Stansfield, S. K. (2007) Health statistics now: are we making the right investments? *Lancet,* 369**,** 779-86.

Bousema, T., Griffin, J. T., Sauerwein, R. W., Smith, D. L., Churcher, T. S., Takken, W., Ghani, A., Drakeley, C. and Gosling, R. (2012) Hitting Hotspots: Spatial Targeting of Malaria for Control and Elimination. *PLoS Med,* 9.

Branch, O., Casapia, W. M., Gamboa, D. V., Hernandez, J. N., Alava, F. F., Roncal, N., Alvarez, E., Perez, E. J. and Gotuzzo, E. (2005) Clustered local transmission and asymptomatic Plasmodium falciparum and Plasmodium vivax malaria infections in a recently emerged, hypoendemic Peruvian Amazon community. *Malaria Journal,* 4.

Breman, J. G. (2001) The ears of the hippopotamus: manifestations, determinants, and estimates of the malaria burden. *Am J Trop Med Hyg,* 64**,** 1-11.

Breman, J. G. and Holloway, C. N. (2007) Malaria surveillance counts. *Am J Trop Med Hyg,* 77**,** 36-47.

Briggs, D. J., Gulliver, J., Fecht, D. and Vienneau, D. M. (2007) Dasymetric modelling of small-area population distribution using land cover and light emissions data. *Remote Sensing of Environment,* 108**,** 451-466.

Brook, D. (1964) On the distinction between the conditional probability and the joint probability approaches in the specification of nearest-neighbour systems. *Biometrika,* 51**,** 481-483.

Brooker, S., Clarke, S., Njagi, J. K., Polack, S., Mugo, B., Estambale, B., Muchiri, E., Magnussen, P. and Cox, J. (2004) Spatial clustering of malaria and associated risk factors during an epidemic in a highland area of western Kenya. *Trop Med Int Health,* 9**,** 757-66.

Brooker, S., Hay, S. I., Issae, W., Hall, A., Kihamia, C. M., Lwambo, N. J., Wint, W., Rogers, D. J. and Bundy, D. A. (2001) Predicting the distribution of urinary schistosomiasis in Tanzania using satellite sensor data. *Trop Med Int Health,* 6**,** 998-1007.

Brooker, S., Hay, S. I., Tchuem Tchuente, L.-A. and Ratard, R. (2002) Using NOAA-AVHRR data to model human helminth distributions in planning disease control in Cameroon, West Africa. *Photogrammetric Engineering & Remote Sensing,* 68**,** 175-179.

Brugha, R. and Zwi, A. (1998) Improving the Quality of Private Sector Delivery of Public Health Services: Challenges and Strategies. *Health Policy and Planning,* 13**,** 107-120.

Bullen, N., Moon, G. and Jones, K. (1996) Defining localities for health planning: A GIS approach. *Social Science & Medicine,* 42**,** 801-816.

Buor, D. (2003) Analysing the primacy of distance in the utilization of health services in the Ahafo-Ano South district, Ghana. John Wiley & Sons, Ltd.

Burrough, P. A. and McDonnell, R. (1998) *Principles of Geographic Information Systems,* Oxford, Oxford University Press.

Cameletti, M., Lindgren, F., Simpson, D. and Rue, H. (2012) Spatio-temporal modeling of particulate matter concentration through the SPDE approach. *AStA Advances in Statistical Analysis***,** 1-23.

Campbell, J. B. and Wynne, R. H. (2011) *Introduction to Remote Sensing,* New York, London, The Guilford Press.

Cappellini, M. D. and Fiorelli, G. (2008) Glucose-6-phosphate dehydrogenase deficiency. *Lancet,* 371**,** 64-74.

Carapinha, J. L., Ross-Degnan, D., Desta, A. T. and Wagner, A. K. (2011) Health insurance systems in five Sub-Saharan African countries: medicine benefits and data for decision making. *Health Policy,* 99**,** 193-202.

Carolina, P.-H., Warren, J. L., Burgert, C. R. and Emch, M. E.(2013) Guidelines on the Use of DHS GPS Data, Calverton, Maryland, USA: , ICF International.

Ceesay, S. J., Casals-Pascual, C., Erskine, J., Anya, S. E., Duah, N. O., Fulford, A. J. C., Sesay, S. S. S., Abubakar, I., Dunyo, S., Sey, O., Palmer, A., Fofana, M., Corrah, T., Bojang, K. A., Whittle, H. C., Greenwood, B. M. and Conway, D. J. (2008) Changes in malaria indices between 1999 and 2007 in The Gambia: a retrospective analysis. *The Lancet,* 372**,** 1545-1554.

Center for International Earth Science Information Network (CIESIN).(2004) Global Rural-Urban Mapping Project (GRUMP): Urban Extents, New York, CIESIN, Columbia University.

Charlwood, J. D., Smith, T., Lyimo, E., Kitua, A. Y., Masanja, H., Booth, M., Alonso, P. L. and Tanner, M. (1998) Incidence of Plasmodium falciparum infection in infants in relation to exposure to sporozoite-infected anophelines. *The American Journal of Tropical Medicine and Hygiene,* 59**,** 243-51.

Chibwana, A. I., Mathanga, D. P., Chinkhumba, J. and Campbell, C. H., Jr. (2009) Socio-cultural predictors of health-seeking behaviour for febrile under-five children in Mwanza-Neno district, Malawi. *Malar J,* 8**,** 219.

Chowdhury, S., Banu, L. A., Chowdhury, T. A., Rubayet, S. and Khatoon, S. (2012) Achieving Millennium Development Goals 4 and 5 in Bangladesh. *BJOG: An International Journal of Obstetrics & Gynaecology,* 118**,** 36-46.

Christakos, G. (2000) *Modern Spatiotemporal Geostatistics,* New York, Oxford University Press.

Chuma, J., Abuya, T., Memusi, D., Juma, E., Akhwale, W., Ntwiga, J., Nyandigisi, A., Tetteh, G., Shretta, R. and Amin, A. (2009a) Reviewing the literature on access to prompt and effective malaria treatment in Kenya: implications for meeting the Abuja targets. *Malar J,* 8**,** 243.

Chuma, J., Musimbi, J., Okungu, V., Goodman, C. and Molyneux, C. (2009b) Reducing user fees for primary health care in Kenya: Policy on paper or policy in practice? *Int J Equity Health,* 8**,** 15.

Chuma, J. and Okungu, V. (2011) Viewing the Kenyan health system through an equity lens: implications for universal coverage. *International Journal for Equity in Health,* 10**,** 22.

Chuma, J., Okungu, V. and Molyneux, C. (2010) Barriers to prompt and effective malaria treatment among the poorest population in Kenya. *Malaria Journal,* 9**,** 144.

CIA. (2013). The world fact book [Online]. Place Published: UNDP. Available: https://www.cia.gov/library/publications/the-world-factbook/rankorder/2172rank.html [Accessed August 2 2013].

Cibulskis, R. E., Aregawi, M., Williams, R., Otten, M. and Dye, C. (2011) Worldwide Incidence of Malaria in 2009: Estimates, Time Trends, and a Critique of Methods. *PLoS Med,* 8**,** e1001142.

Cibulskis, R. E., Bell, D., Christophel, E. M., Hii, J., Delacollette, C., Bakyaita, N. and Aregawi, M. W. (2007) Estimating trends in the burden of malaria at country level. *Am J Trop Med Hyg,* 77**,** 133-7.

Ciocco, A., Altman, I. and Service, U. S. P. H. (1954) *Medical Service Areas and Distances Traveled for Physician Care in Western Pennsylvania*, U.S. Government Printing Office.

Clark, C., Wilson, F. and Bradley, J. (1969) Industrial location and economic potential in Western Europe. *Regional Studies,* 3**,** 197-212.

Clausen, F., Sandberg, E., Ingstad, B. and Hjortdahl, P. (2000) Morbidity and Health Care Utilisation among Elderly People in Mmankgodi Village, Botswana. *Journal of Epidemiology and Community Health (1979-),* 54**,** 58-63.

Clayton, D. and Kaldor, J. (1987) Empirical Bayes Estimates of Age-Standardized Relative Risks for Use in Disease Mapping. *Biometrics,* 43**,** 671-681.

Clements, A., Barnett, A., Cheng, Z., Snow, R. and Zhou, H. (2009) Space-time variation of malaria incidence in Yunnan province, China. *Malaria Journal,* 8**,** 180.

Cleveland, W. (1981) LOWESS: A Program for Smoothing Scatterplots by Robust Locally Weighted Regression. *The American Statistician,* 35.

Cleveland, W. S. (1979) Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association,* 74**,** 829-836.

Cliff, A. D., Haggett, P., Ord, J. K., Bassett, K. A. and Davies, R. B. (1975) *Elements of Spatial Structure: A Quantative Approach*, Cambridge University Press.

Coetzee, M., Craig, M. and le Sueur, D. (2000) Distribution of African malaria mosquitoes belonging to the Anopheles gambiae complex. *Parasitol Today,* 16**,** 74-7.

Cohen, J. M., Woolsey, A. M., Sabot, O. J., Gething, P. W., Tatem, A. J. and Moonen, B. (2012) Optimizing Investments in Malaria Treatment and Diagnosis. *Science,* 338**,** 612-614.

Comber, A., Brunsdon, C. and Radburn, R. (2011) A spatial analysis of variations in health access: linking geography, socio-economic status and access perceptions. *International Journal of Health Geographics,* 10**,** 44.

Connor, R. A., Kralewski, J. E. and Hillson, S. D. (1994) Measuring Geographic Access to Health Care in Rural Areas. *Medical Care Research and Review,* 51**,** 337-377.

Cook, J., Kleinschmidt, I., Schwabe, C., Nseng, G., Bousema, T., Corran, P. H., Riley, E. M. and Drakeley, C. J. (2011) Serological Markers Suggest Heterogeneity of Effectiveness of Malaria Control Interventions on Bioko Island, Equatorial Guinea. *PLoS One,* 6**,** e25137.

Corran, P., Coleman, P., Riley, E. and Drakeley, C. (2007) Serology: a robust indicator of malaria transmission intensity? *Trends Parasitol,* 23**,** 575-82.

Cosner, C., Beier, J. C., Cantrell, R. S., Impoinvil, D., Kapitanski, L., Potts, M. D., Troyo, A. and Ruan, S. (2009) The effects of human movement on the persistence of vector-borne diseases. *J Theor Biol,* 258**,** 550-60.

Cotter, C., Sturrock, H. J. W., Hsiang, M. S., Liu, J., Phillips, A. A., Hwang, J., Gueye, C. S., Fullman, N., Gosling, R. D. and Feachem, R. G. A. (2013) The changing epidemiology of malaria elimination: new strategies for new challenges. *The Lancet,* 382**,** 900-911.

Craig, M. H., Sharp, B. L., Mabaso, M. L. and Kleinschmidt, I. (2007) Developing a spatial-statistical model and map of historical malaria prevalence in Botswana using a staged variable selection procedure. *Int J Health Geogr,* 6**,** 44.

Craig, M. H., Snow, R. W. and le Sueur, D. (1999) A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitol Today,* 15**,** 105-11.

Crawley, J., English, M., Waruiru, C., Mwangi, I. and Marsh, K. (1998) Abnormal respiratory patterns in childhood cerebral malaria. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 92**,** 305-308.

Cressie, N. (1985a) Fitting variogram models by weighted least squares. *Mathematical Geology,* 17**,** 563-586.

Cressie, N. (1985b) When are relative variograms useful in geostatistics? *Mathematical Geology,* 17**,** 693-702.

Cressie, N. (1986) Kriging Nonstationary Data. *Journal of the American Statistical Association,* 81**,** 625-634.

Cressie, N. (1990) The origins of kriging. *Mathematical Geology,* 22**,** 239-252.

Cressie, N. (1993) *Statistics for spatial data,* New York, Wiley.

Cressie, N. and Huang, H.-C. (1999) Classes of Nonseparable, Spatio-Temporal Stationary Covariance Functions. *Journal of the American Statistical Association,* 94**,** 1330-1340.

Cressie, N. and Johannesson, G. (2008) Fixed rank kriging for very large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 70**,** 209-226.

Cressie, N. and Wikle, K. C. (2011) *Statistics for spatiotemporal data,* Hoboken, New Jersey, John Wiley & Sons, Inc.

Cressie, N. and Zimmerman, D. (1992) On the stability of the geostatistical method. *Mathematical Geology,* 24**,** 45-59.

Cromley, E. K. and McLafferty, S. L. (2002) *GIS and Public Health,* New York, Guilford Press.

Cromley, E. K. and McLafferty, S. L. (2010) *GIS and Public Health,* New York, Guilford Press.

Czado, C., Gneiting, T. and Held, L. (2009) Predictive model assessment for count data. *Biometrics,* 65**,** 1254-61.

Dao, N. V., Cuong, B. T., Ngoa, N. D., Thuy le, T. T., The, N. D., Duy, D. N., Dai, B., Thanh, N. X., Chavchich, M., Rieckmann, K. H. and Edstein, M. D. (2007) Vivax malaria: preliminary observations following a shorter course of treatment with artesunate plus primaquine. *Trans R Soc Trop Med Hyg,* 101**,** 534-9.

Das, A., Gupta, R. D., Friedman, J., Pradhan, M., Mohapatra, C. and Sandhibigraha, D. (2013) Community perceptions on malaria and care-seeking practices in endemic Indian settings: policy implications for the malaria control programme. *Malaria Journal,* 12**,** 39.

Davies, G. (1994). The Medical Culture of the Ovambo of Southern Angola and Northern Namibia. Doctor of Philosophy, University of Kent at Canterbury

Davis, R. G., Kamanga, A., Castillo-Salgado, C., Chime, N., Mharakurwa, S. and Shiff, C. (2011) Early detection of malaria foci for targeted interventions in endemic southern Zambia. *Malar J,* 10**,** 260.

De Cesare, L., Myers, D. E. and Posa, D. (2001) Product-sum covariance for space-time modeling: an environmental application. John Wiley & Sons, Ltd.

De Cesare, L., Myers, D. E. and Posa, D. (2002) FORTRAN programs for space-time modeling. *Computers & Geosciences,* 28**,** 205-212.

de Dios Ortzar, J. and Willumsen, L. G. *Modelling Transport,* John Wiley & Sons.

De Iaco, S., Maggio, S., Palma, M. and Posa, D. (2011) Towards an automatic procedure for modeling multivariate space - time data. *Computers & Geosciences*.

De Iaco, S., Myers, D. E., Palma, M. and Posa, D. (2010) FORTRAN programs for space - time multivariate modeling and prediction. *Computers & Geosciences,* 36**,** 636-646.

De Iaco, S., Myers, D. E. and Posa, D. (2001) Space-time analysis using a general product-sum model. *Statistics & Probability Letters,* 52**,** 21-28.

De Iaco, S., Palma, M. and Posa, D. (2005) Modeling and prediction of multivariate space - time random fields. *Computational Statistics & Data Analysis*, 48**,** 525-547.

De Iaco, S. and Posa, D. (2011) Predicting spatio-temporal random fields: Some computational aspects. *Computers & Geosciences*.

De Meillon, B. (1951) Malaria survey of South-West Africa. *Bull World Health Organ***,** 333-417.

De Vries, J. J., Nijkamp, P. and Rietveld, P. (2009) Exponential or power distance-decay for commuting? An alternative specification. *Environment and Planning A*, 41**,** 461-480.

Delawer, F. M., Isono, M., Ueki, H., Zhuben, M., Zafari, M., Seddiq, M. K., Habib, H. and Ayoubi, M. K. (2013) Management of paediatric tuberculosis in provincial and district hospitals in Afghanistan. *Eastern Mediterranean Health Journal*, 19.

Delfini, L. F. (1989) The first case of Plasmodium falciparum resistant to chloroquine treatment discovered in the Republic of Afghanistan. *Trans R Soc Trop Med Hyg*, 83**,** 316.

Densen, P. M. (1972) The measurement of utilization in health care programs. *Health Serv Rep*, 87**,** 958-68.

Department of Health (2009) World class commissioning for GP services: Improving GP access and responsiveness. IN Health, D. o. (Ed.). Leeds, UK.

Deressa, W. (2007) Treatment-seeking behaviour for febrile illness in an area of seasonal malaria transmission in rural Ethiopia. *Malaria Journal*, 6**,** 49.

Deutsch, C. V. and Journel, A. G. (1998) *Gslib: Geostatistical software library and user's guide*, New York., Oxford University Press.

DeVise, P. and Chicago, H. P. C. f. M. (1966) *Hospital Study Districts for Metropolitan Chicago: A Geographic Analysis and Methodology*, Hospital Planning Council for Metropolitan Chicago.

DHIS2. (2014). DHIS2 [Online]. Place Published. Available: http://www.dhis2.org/ [Accessed July, 20 2014].

Diallo, A., Ndam, N. T., Moussiliou, A., Dos Santos, S., Ndonky, A., Borderon, M., Oliveau, S., Lalou, R. and Le Hesran, J. Y. (2012) Asymptomatic carriage of plasmodium in urban Dakar: the risk of malaria should not be underestimated. *PLoS One*, 7**,** e31100.

Dietz, K. (1993) The estimation of the basic reproduction number for infectious diseases. *Stat Methods Med Res*, 2**,** 23-41.

Diggle, P., Moyeed, R., Rowlingson, B. and Thomson, M. (2002) Childhood Malaria in the Gambia: A Case-Study in Model-Based Geostatistics. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 51**,** 493-506.

Diggle, P. J., Tawn, J. A. and Moyeed, R. A. (1998) Model-based geostatistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 47**,** 299-350.

Dimitrakopoulos, R. and Lou, X. (Eds.) (1994) *Spatiotemporal modeling: covariances and ordinary kriging system*, Dordrecht, Kluwer Academic Publisher.

DIVA-GIS. (2011). Spatial Data Download Namibia [Online]. Place Published. Available: http://www.diva-gis.org/gData [Accessed January 17 2011].

DOMC.(2010) Ministry of Public Health and Sanitation: National Guidelines for Diagnosis, Treatment and Prevention of Malaria in Kenya, Nairobi, Ministry of Public Health and Sanitation.

Domingo, G., Satyagraha, A., Anvikar, A., Baird, K., Bancone, G., Bansil, P., Carter, N., Cheng, Q., Culpepper, J., Eziefula, C., Fukuda, M., Green, J., Hwang, J., Lacerda, M., McGray, S., Menard, D., Nosten, F., Nuchprayoon, I., Oo, N., Bualombai, P., Pumpradit, W., Qian, K., Recht, J., Roca, A., Satimai, W., Sovannaroth, S., Vestergaard, L. and Von Seidlein, L. (2013) G6PD

testing in support of treatment and elimination of malaria: recommendations for evaluation of G6PD tests. *Malaria Journal,* 12**,** 391.

Donaldson, L. J. and Gabriel, S. (2009) *Donaldsons' essential public health,* Oxford, Radcliffe Publishing Ltd.

Douglas, N. M., Anstey, N. M., Angus, B. J., Nosten, F. and Price, R. N. (2010) Artemisinin combination therapy for vivax malaria. *Lancet Infect Dis,* 10**,** 405-16.

Douglas, N. M., Nosten, F., Ashley, E. A., Phaiphun, L., van Vugt, M., Singhasivanon, P., White, N. J. and Price, R. N. (2011) Plasmodium vivax recurrence following falciparum and mixed species malaria: risk factors and effect of antimalarial kinetics. *Clin Infect Dis,* 52**,** 612-20.

Drakeley, C. J., Corran, P. H., Coleman, P. G., Tongren, J. E., McDonald, S. L., Carneiro, I., Malima, R., Lusingu, J., Manjurano, A., Nkya, W. M., Lemnge, M. M., Cox, J., Reyburn, H. and Riley, E. M. (2005) Estimating medium- and long-term trends in malaria transmission by using serological markers of malaria exposure. *Proc Natl Acad Sci U S A,* 102**,** 5108-13.

Duncan, L. (2011) A comparison of conditional autoregressive models used in Bayesian disease mapping. *Spatial and Spatio-temporal Epidemiology,* 2**,** 79-89.

Dutt, A. K., Dutta, H. M., Jaiswal, J. and Monroe, C. (1986) Assessment of service adequacy of primary health care physicians in a two county region of Ohio, USA. *GeoJournal,* 12**,** 443-455.

Dy, F. J. (1954) Present status of malaria control in Asia. *Bull World Health Organ,* 11**,** 725-63.

Dzator, J. and Asafu-Adjaye, J. (2004) A study of malaria care provider choice in Ghana. *Health Policy,* 69**,** 389-401.

Earnest, A., Morgan, G., Mengersen, K., Ryan, L., Summerhayes, R. and Beard, J. (2007) Evaluating the effect of neighbourhood weight matrices on smoothing properties of Conditional Autoregressive (CAR) models. *International Journal of Health Geographics,* 6**,** 54.

East Africa Roll Back Malaria Network (EARN).(2013) Eritrea malaria Programme performance review, Asmara. http://www.rollbackmalaria.org/countryaction/aideMemoire/Eritrea-The-malaria-program-performance-review-2013.pdf.

Edward, A., Kumar, B., Kakar, F., Salehi, A. S., Burnham, G. and Peters, D. H. (2011) Configuring Balanced Scorecards for Measuring Health System Performance: Evidence from 5 Years' Evaluation in Afghanistan. *PLoS Med,* 8**,** e1001066.

Egbert, G. D. and Lettenmaier, D. P. (1986) Stochastic Modeling of the Space-Time Structure of Atmospheric Chemical Deposition. *Water Resour. Res.,* 22**,** 165-179.

Eilers, P. H. C. and Marx, B. D. (1996) Flexible smoothing with B-Splines and penalties. *Statistical Science,* 11**,** 89-121.

Einterz, E. M. and Bates, M. E. (1997) Fever in Africa: do patients know when they are hot? *The Lancet,* 350**,** 781-781.

Eisele, T. P., Macintyre, K., Yukich, J. and Ghebremeskel, T. (2006) Interpreting household survey data intended to measure insecticide-treated bednet coverage: results from two surveys in Eritrea. *Malar J,* 5**,** 36.

El Obeid, S., Mendelsohn, J. M., Lejars, M., Forster, N. and Brulé, G. 2001. Health in Namibia: Progress and challenges. [Online]. Available: http://www.raison.com.na/page_htm_raison/downloads.html [Accessed 2011, May 17].

Elmardi, K. A., Noor, A. M., Githinji, S., Abdelgadir, T. M., Malik, E. M. and Snow, R. W. (2011) Self-reported fever, treatment actions and malaria infection prevalence in the northern states of Sudan. *Malar J,* 10**,** 128.

English, M., Waruiru, C., Amukoye, E., Murphy, S., Crawley, J., Mwangi, I., Peshu, N. and Marsh, K. (1996) Deep Breathing in Children with Severe Malaria: Indicator of Metabolic Acidosis and Poor Outcome. *The American Journal of Tropical Medicine and Hygiene,* 55**,** 521-524.

Ensor, T. and Cooper, S. (2002) Overcoming barriers to health service access: influencing the demand side. *Health Policy and Planning,* 19**,** 69-79.

ESA. (2010). GlobCover 2009 (Global Land Cover Map) [Online]. Place Published. Available: http://ionia1.esrin.esa.int/ [Accessed January 15 2011].

Eshghy, N. and Nushin, M. (1978) Evaluation of malathion as a residual spray for the control of Anopheles culicifacies in the province of Helmand, southeast Afghanistan. *Mosquito News,* 38**,** 268-274.

Fahrmeir, L. and Kneib, T. (2009) Propriety of posteriors in structured additive regression models: Theory and empirical evidence. *Journal of Statistical Planning and Inference,* 139**,** 843-859.

Fahrmeir, L. and Lang, S. (2001) Bayesian inference for generalized additive mixed models based on Markov random field priors *Journal of the Royal Statistical Society, Series C,* 2**,** 201-220.

Falk, N., Maire, N., Sama, W., Owusu-Agyei, S., Smith, T., Beck, H. P. and Felger, I. (2006) Comparison of PCR-RFLP and Genescan-based genotyping for analyzing infection dynamics of Plasmodium falciparum. *Am J Trop Med Hyg,* 74**,** 944-50.

FAO. (2000). Land Cover Classification System (LCCS): Classification Concepts and User Manual [Online]. Place Published: Natural Resources Management and Environment Department. Available at: http://www.fao.org/docrep/003/x0596e/x0596e00.htm [Accessed 17 May 2011]. Available: http://www.fao.org/docrep/003/x0596e/x0596e00.htm [Accessed 17 May 2011].

Feachem, R. G. A., Phillips, A. A., Hwang, J., Cotter, C., Wielgosz, B., Greenwood, B. M., Sabot, O., Rodriguez, M. H., Abeyasinghe, R. R., Ghebreyesus, T. A. and Snow, R. W. (2010) Shrinking the malaria map: progress and prospects. *The Lancet,* 376**,** 1566-1578.

Fiedler, J. L. and Wight, J. B. (2003) Privatization and the allure of franchising: a Zambian feasibility study. *Int J Health Plann Manage,* 18**,** 179-204.

Flessa, S., Moeller, M., Ensor, T. and Hornetz, K. (2011) Basing care reforms on evidence: The Kenya health sector costing model. *BMC Health Serv Res,* 11**,** 128.

Fortney, J., Thill, J.-C., Zhang, M., Duan, N. and Rost, K. (2001) Provider Choice and Utility Loss due to Selective Contracting in Rural and Urban Areas. *Medical Care Research and Review,* 58**,** 60-75.

Free Software Foundation, I. (2007). GNU General Public License [Online]. Place Published. Available: http://www.gnu.org/copyleft/gpl.html [Accessed 1 December 2012].

Fuentes, M. (2002) Spectral Methods for Nonstationary Spatial Processes. *Biometrika,* 89**,** 197-210.

Fuentes, M. (2005) A formal test for nonstationarity of spatial stochastic processes. *Journal of Multivariate Analysis,* 96**,** 30-54.

Furrer, Reinhard, Genton, Marc, G., Nychka and Douglas (2006) Covariance Tapering for Interpolation of Large Spatial Datasets. *Journal of Computational & Graphical Statistics,* 15**,** 502-523.

Gaughan, A. E., Stevens, F. R., Linard, C., Jia, P. and Tatem, A. J. (2013) High Resolution Population Distribution Maps for Southeast Asia in 2010 and 2015. *PLoS One,* 8**,** e55882.

Gelfand, A., Schmidt, A., Banerjee, S. and Sirmans, C. F. (2004) Nonstationary multivariate process modeling through spatially varying coregionalization. *Test,* 13**,** 263-312.

Gelfand, A. E. and Ghosh, S. K. (1998) Model choice: A minimum posterior predictive loss approach. *Biometrika,* 85**,** 1-11.

Gelfand, A. E., Kim, H.-J., Sirmans, C. F. and Banerjee, S. (2003) Spatial Modeling with Spatially Varying Coefficient Processes. *Journal of the American Statistical Association,* 98**,** 387-396.

Gelfand, A. E. and Vounatsou, P. (2003) Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics,* 4**,** 11-25.

Gelman, A. and Rubin, D. B. (1992) Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science,* 7**,** 457-472.

Gelman, A. and Rubin, D. B. (1996) Markov chain Monte Carlo methods in biostatistics. *Stat Methods Med Res,* 5**,** 339-55.

Gemperli, A. (2003). Development of spatial statistical methods for modelling point-referenced spatial data in malaria epidemiology. PhD Doctoral Dissertation, University of Basel.

Gemperli, A., Vounatsou, P., Sogoba, N. and Smith, T. (2006) Malaria Mapping Using Transmission Models: Application to Survey Data from Mali. *American Journal of Epidemiology,* 163**,** 289-297.

Geonames. (2010). The GeoNames geographical database [Online]. Place Published. Available: http://www.geonames.org/search.html?q=&country=NA [Accessed December 12 2010].

Gething, P., Atkinson, P., Noor, A., Gikandi, P., Hay, S. and Nixon, M. (2007) A local space-time kriging approach applied to a national outpatient malaria dataset. *Comput Geosci,* 33**,** 1337-1350.

Gething, P., Van Boeckel, T., Smith, D., Guerra, C., Patil, A., Snow, R. and Hay, S. (2011a) Modelling the global constraints of temperature on transmission of Plasmodium falciparum and P. vivax. *Parasites & Vectors,* 4**,** 92.

Gething, P. W., Elyazar, I. R. F., Moyes, C. L., Smith, D. L., Battle, K. E., Guerra, C. A., Patil, A. P., Tatem, A. J., Howes, R. E., Myers, M. F., George, D. B., Horby, P., Wertheim, H. F. L., Price, R. N., MÃ¼eller, I., Baird, J. K. and Hay, S. I. (2012) A Long Neglected World Malaria Map: *Plasmodium vivax* Endemicity in 2010. *PLoS Negl Trop Dis,* 6**,** e1814.

Gething, P. W., Noor, A. M., Gikandi, P. W., Hay, S. I., Nixon, M. S., Snow, R. W. and Atkinson, P. M. (2008) Developing geostatistical space-time models to predict outpatient treatment burdens from incomplete national data. *Geogr Anal,* 40**,** 167-188.

Gething, P. W., Noor, A. M., Gikandi, P. W., Ogara, E. A., Hay, S. I., Nixon, M. S., Snow, R. W. and Atkinson, P. M. (2006) Improving imperfect data from health management information systems in Africa using space-time geostatistics. *PLoS Med,* 3**,** e271.

Gething, P. W., Noor, A. M., Zurovac, D., Atkinson, P. M., Hay, S. I., Nixon, M. S. and Snow, R. W. (2004) Empirical modelling of government health service use by children with fevers in Kenya. *Acta Trop,* 91**,** 227-37.

Gething, P. W., Patil, A. P., Smith, D. L., Guerra, C. A., Elyazar, I. R., Johnston, G. L., Tatem, A. J. and Hay, S. I. (2011b) A new world malaria map: Plasmodium falciparum endemicity in 2010. *Malar J,* 10**,** 378.

Ghosh, S. K., Mukhopadhyay, P. and Lu, J.-C. (2006) Bayesian analysis of zero-inflated regression models. *Journal of Statistical Planning and Inference,* 136**,** 1360-1375.

Gilles, H. M. (Ed.) (2002) *Historical outline,* London, Arnold.

Gilson, L. (1997) The lessons of user fee experience in Africa. *Health Policy Plan,* 12**,** 273-85.

Girt, J. L. (1973) Distance to general medical practice and its effect on revealed ill-health in a rural environment. *Canadian Geographer / Le Géographe canadien,* 17**,** 154-166.

Githeko, A. K., Mbogo, C. N. M., Curtis, C. F., Lines, J. and Lengeler, C. (1996) Entomological monitoring of large-scale vector-control interventions. *Parasitology Today,* 12**,** 127-128.

Gitonga, C. W., Kihara, J. H., Njenga, S. M., Awuondo, K., Noor, A. M., Snow, R. W. and Brooker, S. J. (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.

Global Health Group. (2007). The Malaria Elimination Group (MEG) [Online]. Place Published. Available: http://www.malariaeliminationgroup.org/ [Accessed.

Global Health Group.(2011) Atlas of malaria eliminating countries, 2011, San Francisco, California. http://www.malariaeliminationgroup.org/sites/default/files/atlas/global/Global-Atlas-2011.pdf.

Gneiting, T. (2002) Nonseparable, stationary covariance functions for space time data. *Journal of the American Statistical Association,* 97**,** 590-600.

Gneiting, T. and Raftery, A. E. (2007) Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association,* 102**,** 359-378.

Goodman, C., Brieger, W., Unwin, A., Mills, A., Meek, S. and Greer, G. (2007) Medicine sellers and malaria treatment in sub-Saharan Africa: what do they do and how can their practice be improved? *Am J Trop Med Hyg,* 77**,** 203-18.

Google. (2013). Google Maps [Online]. Place Published: Google Inc. Available: http://www.google.co.uk/intl/en_uk/earth/ [Accessed December 15 2010].

Goovaerts, P. (1997) *Geostatistics for Natural Resources Evaluation,* New York; Oxford, Oxford University Press.

Goovaerts, P. (1998) Ordinary Cokriging Revisited. *Mathematical Geology,* 30**,** 21-42.

Goovaerts, P. (2006) Geostatistical analysis of disease data: accounting for spatial support and population density in the isopleth mapping of cancer mortality risk using area-to-point Poisson kriging. *Int J Health Geogr,* 5**,** 52.

Goovaerts, P. (2010) Combining Areal and Point Data in Geostatistical Interpolation: Applications to Soil Science and Medical Geography. *Math Geosci,* 42**,** 535-554.

Goovaerts, P. and Chiang, C. N. (1993) Temporal Persistence of Spatial Patterns for Mineralizable Nitrogen and Selected Soil Properties. *Soil Sci. Soc. Am. J.,* 57**,** 372-381.

Gosoniu, L., Vounatsou, P., Sogoba, N. and Smith, T. (2006) Bayesian modelling of geostatistical malaria risk data. *Geospat Health,* 1**,** 127-39.

Government of Bangladesh.(1998) Fifth Five Year Plan 1997-2002, Dhaka.

Greenwood, B. M., Groenendaal, F., Bradley, A. K., Greenwood, A. M., Shenton, F., Tulloch, S. and Hayes, R. (1987) Ethnic differences in the prevalence of splenomegaly and malaria in The Gambia. *Ann Trop Med Parasitol,* 81**,** 345-54.

Guagliardo, M. F. (2004) Spatial accessibility of primary care: concepts, methods and challenges. *Int J Health Geogr,* 3**,** 3.

Guerra, C., Hay, S., Lucioparedes, L., Gikandi, P., Tatem, A., Noor, A. and Snow, R. (2007) Assembling a global database of malaria parasite prevalence for the Malaria Atlas Project. *Malaria Journal,* 6**,** 17.

Guerra, C. A., Gikandi, P. W., Tatem, A. J., Noor, A. M., Smith, D. L., Hay, S. I. and Snow, R. W. (2008) The limits and intensity of Plasmodium falciparum transmission: implications for malaria control and elimination worldwide. *PLoS Med,* 5**,** e38.

Guerra, C. A., Howes, R. E., Patil, A. P., Gething, P. W., Van Boeckel, T. P., Temperley, W. H., Kabaria, C. W., Tatem, A. J., Manh, B. H., Elyazar, I. R. F., Baird, J. K., Snow, R. W. and Hay, S. I. (2010) The International Limits and Population at Risk of *Plasmodium vivax* Transmission in 2009. *PLoS Negl Trop Dis,* 4**,** e774.

Guttorp, P., Meiring, W. and Sampson, P. D. (1994) A space-time analysis of ground-level ozone data. *Environmetrics,* 5**,** 241-254.

Guyatt, H. L. and Snow, R. W. (2004) The management of fevers in Kenyan children and adults in an area of seasonal malaria transmission. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 98**,** 111-115.

Haas, J. S., Phillips, K. A., Sonneborn, D., McCulloch, C. E., Baker, L. C., Kaplan, C. P., Perez-Stable, E. J. and Liang, S. Y. (2004) Variation in access to health care for different racial/ethnic groups by the racial/ethnic composition of an individual's county of residence. *Med Care,* 42**,** 707-14.

Hadley, M. (2011) Does increase in utilisation rates alone indicate the success of a user fee removal policy? A qualitative case study from Zambia. *Health Policy,* 103**,** 244-254.

Hamedi, Y., Safa, O., Zare, S., Tan-ariya, P., Kojima, S. and Looareesuwan, S. (2004) Therapeutic efficacy of artesunate in Plasmodium vivax malaria in Thailand. *Southeast Asian J Trop Med Public Health,* 35**,** 570-4.

Hansford, C. F. (1972) Recent trends in the control and treatment of malaria. *S Afr Med J,* 46**,** 635-7.

Harrison, G. (1978) *Mosquitoes, malaria and man: A history of the hostilities since 1880,* London, John Murray.

Hastie, T., Tibshirani, R. and Friedman, J. (2009) *The Elements of Statistical Learning–Data Mining, Inference and Prediction,* New York:, Springer.

Hay, S. I., Guerra, C. A., Gething, P. W., Patil, A. P., Tatem, A. J., Noor, A. M., Kabaria, C. W., Manh, B. H., Elyazar, I. R., Brooker, S., Smith, D. L., Moyeed, R. A. and Snow, R. W. (2009a) A world malaria map: Plasmodium falciparum endemicity in 2007. *PLoS Med,* 6**,** e1000048.

Hay, S. I., Guerra, C. A., Gething, P. W., Patil, A. P., Tatem, A. J., Noor, A. M., Kabaria, C. W., Manh, B. H., Elyazar, I. R. F., Brooker, S., Smith, D. L., Moyeed, R. A. and Snow, R. W. (2009b) A World Malaria Map: Plasmodium falciparum Endemicity in 2007. *Plos Medicine,* 6.

Hay, S. I., Omumbo, J. A., Craig, M. H. and Snow, R. W. (2000a) Earth observation, geographic information systems and Plasmodium falciparum malaria in sub-Saharan Africa. *Adv Parasitol,* 47**,** 173-215.

Hay, S. I., Packer, M. J. and Rogers, D. J. (1997) The impact of remote sensing on the study and control of invertebrate intermediate hosts and vectors for disease. *Int. j. remote sensing,* 18**,** 2899 ± 2930.

Hay, S. I., Rogers, D. J., Toomer, J. F. and Snow, R. W. (2000b) Annual Plasmodium falciparum entomological inoculation rates (EIR) across Africa: literature survey, Internet access and review. *Trans R Soc Trop Med Hyg,* 94**,** 113-27.

Hay, S. I., Smith, D. L. and Snow, R. W. (2008) Measuring malaria endemicity from intense to interrupted transmission. *Lancet Infect Dis,* 8**,** 369-78.

Hay, S. I. and Snow, R. W. (2006) The malaria Atlas Project: developing global maps of malaria risk. *PLoS Med,* 3**,** e473.

Hay, S. I., Snow, R. W. and Rogers, D. J. (1998) Predicting malaria seasons in Kenya using multitemporal meteorological satellite sensor data. *Trans R Soc Trop Med Hyg,* 92**,** 12-20.

Hay, S. I., Tatem, A. J., Graham, A. J., Goetz, S. J., Rogers, D. J., Simon I. Hay, A. G. and David, J. R. (2006) Global Environmental Data for Mapping Infectious Disease Distribution. *Advances in Parasitology.* Academic Press.

Haynes, R., Bentham, C., Lovett, A. and Gale, S. (1999) Effects of distances to hospital and GP surgery on hospital inpatient episodes, controlling for needs and provision. *Soc Sci Med,* 49**,** 425 - 433.

Hengl, T., Minasny, B. and Gould, M. (2009) A geostatistical analysis of geostatistics. *Scientometrics,* 80**,** 491-514.

Higgs, G. (2004) A Literature Review of the Use of GIS-Based Measures of Access to Health Care Services. *Health Services and Outcomes Research Methodology,* 5**,** 119-139.

Hijmans, R., Cameron, S., Parra, J., Jones, P. and Jarvis, A. (2005) WORLDCLIM - a set of global climate layers (climate grids).

Hogmander, H. and Moller, J. (1995) Estimating Distribution Maps from Atlas Data Using Methods of Statistical Image Analysis. *Biometrics,* 51**,** 393-404.

Hopkins, H., GonzÃ¡lez, I. J., Polley, S. D., Angutoko, P., Ategeka, J., Asiimwe, C., Agaba, B., Kyabayinze, D. J., Sutherland, C. J., Perkins, M. D. and Bell, D. (2013) Highly sensitive detection of malaria parasitemia in an endemic setting: Performance of a new LAMP kit in a remote clinic in Uganda. *Journal of Infectious Diseases*.

Howes, R. E., Piel, F. d. r. B., Patil, A. P., Nyangiri, O. A., Gething, P. W., Dewi, M., Hogg, M. M., Battle, K. E., Padilla, C. D., Baird, J. K. and Hay, S. I. (2012) G6PD Deficiency Prevalence and Estimates of Affected Populations in Malaria Endemic Countries: A Geostatistical Model-Based Map. *PLoS Med,* 9**,** e1001339.

Hudson, G. and Wackernagel, H. (1994) Mapping temperature using kriging with external drift: Theory and an example from scotland. *International Journal of Climatology,* 14**,** 77-91.

Huerta Munoz, U. and Kallestal, C. (2012) Geographical accessibility and spatial coverage modeling of the primary health care network in the Western Province of Rwanda. *International Journal of Health Geographics,* 11**,** 40.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X. and Ferreira, L. G. (2002) Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment,* 83**,** 195-213.

Huffman, G. J. (1997) Estimates of root-mean-square random error contained in finite sets of estimated precipitation. . *J. Appl. Meteor.,* 36**,** 191-1201.

Huffman, G. J. and Bolvin, D. T.(2011) TRMM and Other Data Precipitation Data Set Documentation, Goddard, Maryland, United States, Laboratory for Atmospheres, NASA Goddard Space Flight Center and Science Systems and Applications, Inc.

Huggel, C., Schneider, D., Miranda, P. J., Delgado Granados, H. and Kääb, A. (2008) Evaluation of ASTER and SRTM DEM data for lahar modeling: A case study on lahars from Popocatépetl Volcano, Mexico. *Journal of Volcanology and Geothermal Research,* 170**,** 99-110.

Husk, G. and Waxman, D. A. (2004) Using data from hospital information systems to improve emergency department care. *Acad Emerg Med,* 11**,** 1237-44.

Illian, J. B., Sørbye, S. H. and Rue, H. (2012) A toolbox for fitting complex spatial point process models using Integrated Nested Laplace Approxiamation (INLA). *The Annals of Applied Statistics,* 6**,** 1499-1530.

Imwong, M., Nakeesathit, S., Day, N. and White, N. (2012) A review of mixed malaria species infections in anopheline mosquitoes. *Malaria Journal,* 10**,** 253.

INDEPTH. (2012). The INDEPTH Network [Online]. Place Published. Available: http://www.indepth-network.org/ [Accessed March 12 2012].

Ingram, D. R., Clarke, D. R. and Murdie, R. A. (1978) Distance and the decision to visit an emergency department. *Soc Sci Med,* 12**,** 55-62.

Isaacks, H. E. and Srivastava, R. M. (1989) *An introduction to applied geostatistics,* Oxford, Oxford University Press.

Jacobs, B., Ir, P., Bigdeli, M., Annear, P. L. and Van Damme, W. (2012) Addressing access barriers to health services: an analytical framework for selecting appropriate interventions in low-income Asian countries. *Health Policy and Planning,* 27**,** 288-300.

James, J. and Versteeg, M. (2007) Mobile phones in Africa: how much do we really know? *Soc Indic Res,* 84**,** 117-126.

Jayawardene, R. (1993) Illness perception: Social cost and coping-strategies of malaria cases. *Social Science & Medicine,* 37**,** 1169-1176.

Jehlik, P. J. and McNamara, R. L. (1952) The relation of distance to the differential use of certain health personnel and facilities to the extent of bed illness. *Rural Sociology,* 17**,** 261-265.

Jenkins, C. and Campbell, J. (1996) Catchment areas in general practice and their relation to size and quality of practice and deprivation: a descriptive study in one London borough. *BMJ,* 313**,** 1189-92.

Jensen, T., Bukirwa, H., Njama-Meya, D., Francis, D., Kamya, M., Rosenthal, P. and Dorsey, G. (2009) Use of the slide positivity rate to estimate changes in malaria incidence in a cohort of Ugandan children. *Malaria Journal,* 8**,** 213.

Jie, S., Shaoqing, Z., Bozhao, X., Feng, C., Sujiang, P., Jianjun, Y., Ming, J., Shuiqing, L. and Webber, R. H. (1998) Surveillance for low-level malaria. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 92**,** 3-6.

Jima, D., Wondabeku, M., Alemu, A., Teferra, A., Awel, N., Deressa, W., Adissie, A., Tadesse, Z., Gebre, T., Mosher, A., Richards, F. and Graves, P. (2012) Analysis of malaria surveillance data in Ethiopia: what can be learned from the Integrated Disease Surveillance and Response System? *Malaria Journal,* 11**,** 330.

Jones, C., Abeku, T. A., Rapuoda, B., Okia, M. and Cox, J. (2008) District-based malaria epidemic early warning systems in East Africa: Perceptions of acceptability and usefulness among key staff at health facility, district and central levels. *Social Science & Medicine,* 67**,** 292-300.

Jordan, H., Roderick, P., Martin, D. and Barnett, S. (2004) Distance, rurality and the need for care: access to health services in South West England. *International Journal of Health Geographics,* 3**,** 21.

Joseph, A. E. and Phillips, D. R. (1984) *Accessibility and utilization: Geographical perspectives on health care delivery,* London, UK, Harper & Row.

Kallander, K., Nsungwa-Sabiiti, J. and Peterson, S. (2004) Symptom overlap for malaria and pneumonia--policy implications for home management strategies. *Acta Trop,* 90**,** 211-4.

Kaufman, C. G., Schervish, M. J. and Nychka, D. W. (2008) Covariance Tapering for Likelihood-Based Estimation in Large Spatial Data Sets. *Journal of the American Statistical Association,* 103**,** 1545-1555.

Kazembe, L. N. (2007) Spatial modelling and risk factors of malaria incidence in northern Malawi. *Acta Tropica,* 102**,** 126-137.

Kazembe, L. N., Appleton, C. C. and Kleinschmidt, I. (2007) Choice of treatment for fever at household level in Malawi: examining spatial patterns. *Malar J,* 6**,** 40.

Khan, A. A. (1992) An integrated approach to measuring potential spatial access to health care services. *Socioecon Plann Sci,* 26**,** 275-87.

Khan, A. A. and Bhardwaj, S. M. (1994) Access to health care. A conceptual framework and its relevance to health care planning. *Eval Health Prof,* 17**,** 60-76.

Kirby, M. J. and Lindsay, S. W. (2009) Effect of temperature and inter-specific competition on the development and survival of Anopheles gambiae sensu stricto and An. arabiensis larvae. *Acta Trop,* 109**,** 118-23.

Kiwanuka, S. N., Ekirapa, E. K., Peterson, S., Okui, O., Rahman, M. H., Peters, D. and Pariyo, G. W. (2008) Access to and utilisation of health services for the poor in Uganda: a systematic review of available evidence. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 102**,** 1067-1074.

Kizito, J., Kayendeke, M., Nabirye, C., Staedke, S. and Chandler, C. (2012) Improving access to health care for malaria in Africa: a review of literature on what attracts patients. *Malaria Journal,* 11**,** 55.

Kleinschmidt, I., Bagayoko, M., Clarke, G. P., Craig, M. and Le Sueur, D. (2000) A spatial statistical approach to malaria mapping. *Int J Epidemiol,* 29**,** 355-61.

Kleinschmidt, I., Schwabe, C., Benavente, L., Torrez, M., Ridl, F. C., Segura, J. L., Ehmer, P. and Nchama, G. N. (2009) Marked increase in child survival after four years of intensive malaria control. *Am J Trop Med Hyg,* 80**,** 882-8.

Kleinschmidt, I., Sharp, B., Mueller, I. and Vounatsou, P. (2002) Rise in Malaria Incidence Rates in South Africa: A Small-Area Spatial Analysis of Variation in Time Trends. *American Journal of Epidemiology,* 155**,** 257-264.

Kloek, T. and Dijk, H. K. v. (1978) Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo. *Econometrica,* 46**,** 1-19.

Kloos, H. (1990) Utilization of selected hospitals, health centres and health stations in central, southern and western Ethiopia. *Soc Sci Med,* 31**,** 101-14.

Knorr-Held, L. (2000) Bayesian modelling of inseparable space-time variation in disease risk. *Statistics in Medicine,* 19**,** 2555-2567.

Kolaczinski, J., Graham, K., Fahim, A., Brooker, S. and Rowland, M. (2005) Malaria control in Afghanistan: progress and challenges. *The Lancet,* 365**,** 1506-1512.

Konde-Lule, J., Gitta, S. N., Lindfors, A., Okuonzi, S., Onama, V. O. and Forsberg, B. C. (2010) Private and public health care in rural areas of Uganda. *BMC Int Health Hum Rights,* 10**,** 29.

Kumaranayake, L., Mujinja, P., Hongoro, C. and Mpembeni, R. (2000) How do countries regulate the health sector? Evidence from Tanzania and Zimbabwe. *Health Policy Plan,* 15**,** 357-67.

Kwan, M.-P. (1998) Space-Time and Integral Measures of Individual Accessibility: A Comparative Analysis Using a Point-based Framework. *Geographical Analysis,* 30**,** 191-216.

Kyriakidis, P. and Journel, A. (1999) Geostatistical Space–Time Models: A Review. *Mathematical Geology,* 31**,** 651-684.

Kyriakidis, P. C. and Journel, A. G. (2001) Stochastic modeling of atmospheric pollution: a spatial time-series framework. Part I: methodology. *Atmospheric Environment,* 35**,** 2331-2337.

Lambert, D. (1992) Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics,* 34**,** 1-14.

Laudicella, M., Siciliani, L. and Cookson, R. (2012) Waiting times and socioeconomic status: evidence from England. *Soc Sci Med,* 74**,** 1331-41.

Le Goff, G., Carnevale, P., Fondjo, E. and Robert, V. (1997) Comparison of three sampling methods of man-biting anophelines in order to estimate the malaria transmission in a village of south Cameroon. *Parasite,* 4**,** 75-80.

Le Menach, A., Tatem, A. J., Cohen, J. M., Hay, S. I., Randell, H., Patil, A. P. and Smith, D. L. (2011) Travel risk, malaria importation and malaria transmission in Zanzibar. *Sci. Rep.,* 1.

Lee, D. (2011) A comparison of conditional autoregressive models used in Bayesian disease mapping. *Spat Spatiotemporal Epidemiol,* 2**,** 79-89.

Lee, P.-W., Liu, C.-T., do Rosario, V., de Sousa, B., Rampao, H. and Shaio, M.-F. (2010) Potential threat of malaria epidemics in a low transmission area, as exemplified by São Tomé and Príncipe. *Malaria Journal,* 9**,** 1-12.

Leonard, K., Mliga, G. R. and Mariam, D. H.(2002) Bypassing health centers in Tanzania: Revealed preferences for observable and unobservable quality, Columbia University, Department of Economics. http://econpapers.repec.org/RePEc:clu:wpaper:0102-43.

Leslie, T., Mayan, I., Mohammed, N., Erasmus, P., Kolaczinski, J., Whitty, C. J. and Rowland, M. (2008) A randomised trial of an eight-week, once weekly primaquine regimen to prevent relapse of plasmodium vivax in Northwest Frontier Province, Pakistan. *PLoS One,* 3**,** e2861.

Linard, C., Gilbert, M., Snow, R. W., Noor, A. M. and Tatem, A. J. (2012) Population Distribution, Settlement Patterns and Accessibility across Africa in 2010. *PLoS One,* 7**,** e31743.

Linard, C., Gilbert, M. and Tatem, A. (2010) Assessing the use of global land cover data for guiding large area population distribution modelling. *GeoJournal***,** 1-14.

Lindelow, M. and Wagstaff, A.(2003) Health facility surveys: An introduction, Washington, D.C., The World Bank development research group, public services.

Lindgren, F. (2013) Continuous domain spatial models in R-INLA. *ISBA Bulletin.*

Lindgren, F. and Rue, H. (2008) On the Second-Order Random Walk Model for Irregular Locations. *Scandinavian Journal of Statistics,* 35**,** 691-700.

Lindgren, F. and Rue, H. 2013. Bayesian Spatial and Spatio-temporal Modelling with R-INLA. [Online]. Available: http://people.bath.ac.uk/fl353/tutorials/oulu/tut_spde_lindgren.pdf [Accessed.

Lindgren, F., Rue, H. and Lindström, J. (2011) An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 73**,** 423-498.

Littrell, M., Gatakaa, H., Evance, I., Poyer, S., Njogu, J., Solomon, T., Munroe, E., Chapman, S., Goodman, C., Hanson, K., Zinsou, C., Akulayi, L., Raharinjatovo, J., Arogundade, E., Buyungo, P., Mpasela, F., Adjibabi, C. B., Agbango, J. A., Ramarosandratana, B. F., Coker, B., Rubahika, D., Hamainza, B., Shewchuk, T., Chavasse, D. and O'Connell, K. A. (2011) Monitoring fever treatment behaviour and equitable access to effective medicines in the context of initiatives to improve ACT access: baseline results and implications for programming in six African countries. *Malar J,* 10**,** 327.

Liu, X., Hotchkiss, D. R. and Bose, S. (2008) The effectiveness of contracting-out primary health care services in developing countries: a review of the evidence. *Health Policy Plan,* 23**,** 1-13.

Loader, C. and Switzer, P. (Eds.) (1992) *Spatial covariance estimation for monitoring data,* London, Edward Arnold.

Loevinsohn, B. and Harding, A. (2005) Buying results? Contracting for health service delivery in developing countries. *Lancet,* 366**,** 676-81.

Loevinsohn B, S. G. (2008) Lessons from the health sector in afghanistan: How progress can be made in challenging circumstances. *JAMA,* 300**,** 724-726.

Loha, E. and Lindtjorn, B. (2010) Model variations in predicting incidence of Plasmodium falciparum malaria using 1998-2007 morbidity and meteorological data from south Ethiopia. *Malaria Journal,* 9**,** 166.

Looareesuwan, S., White, N. J., Chittamas, S., Bunnag, D. and Harinasuta, T. (1987) High rate of Plasmodium vivax relapse following treatment of falciparum malaria in Thailand. *Lancet,* 2**,** 1052-5.

Lourenco, C., Uusiku, P., Haidula, L., Kandula, D., Ward, A. and Cohen, J. (2013) Improving Malaria Diagnosis and Appropriate Treatment Through Adaptable Training Interventions in Namibia *American Society of Tropical Medicine and Hygiene*.

Lowenthal, M. N., Hutt, M. S., Jones, I. G., Mohelsky, V. and O'Riordan, E. C. (1980) Massive splenomegaly in Northern Zambia. I. Analysis of 344 cases. *Trans R Soc Trop Med Hyg,* 74**,** 91-8.

Lozano, R., Wang, H., Foreman, K. J., Rajaratnam, J. K., Naghavi, M., Marcus, J. R., Dwyer-Lindgren, L., Lofgren, K. T., Phillips, D., Atkinson, C., Lopez, A. D. and Murray, C. J. (2011) Progress towards Millennium Development Goals 4 and 5 on maternal and child mortality: an updated systematic analysis. *Lancet,* 378**,** 1139-65.

Lu, C., Tsai, S., Ruhumuriza, J., Umugiraneza, G., Kandamutsa, S., Salvatore, P. P., Zhang, Z., Binagwaho, A. and Ngabo, F. (2014) Tracking Rural Health Facility Financial Data in Resource-Limited Settings: A Case Study from Rwanda. *PLoS Med,* 11**,** e1001763.

Luo, W. (2004) Using a GIS-based floating catchment method to assess areas with shortage of physicians. *Health Place,* 10**,** 1-11.

Luo, W. and Qi, Y. (2009) An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians. *Health Place,* 15**,** 1100-7.

Luo, W. and Wang, F. (2003) Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. *Environment and Planning B: Planning and Design,* 30**,** 865-884.

Lysenko, A. J. and Semashko, I. N. (Eds.) (1968) *Geography of malaria. A medico-geographic profile of an ancient disease [in Russian],* Moscow, Academy of Sciences, USSR.

MacDonald, G. (1950) The analysis of malaria parasite rates in infants. *Trop Dis Bull* 47**,** 915-938.

Macdonald, G. (1956) Epidemiological basis of malaria control. *Bull World Health Organ,* 15**,** 613-26.

Macdonald, G. (1957) *The epidemiology and control of malaria,* London, Oxford University Press.

Macintyre, K., Keating, J., Okbaldt, Y. B., Zerom, M., Sosler, S., Ghebremeskel, T. and Eisele, T. P. (2006) Rolling out insecticide treated nets in Eritrea: examining the determinants of possession and use in malarious zones during the rainy season. *Trop Med Int Health,* 11**,** 824-33.

Macro International Inc.(1996) Sampling manual: DHS III basic documentation, Carlverton, Maryland. 6

Malaria Elimination 8 Ministerial Meeting. 2009. Resolution. [Online]. Available: http://www.sadc.int/files/3313/1823/5293/E8_Ministerial_Resolution_20090303.pdf [Accessed.

Manh, B. H., Clements, A. C. A., Thieu, N. Q., Hung, N. M., Hung, L. X., Hay, S. I., Hien, T. T., Wertheim, H. F. L., Snow, R. W. and Horby, P. (2011) Social and environmental determinants of malaria in space and time in Viet Nam. *International Journal for Parasitology,* 41**,** 109-116.

Manongi, R., Mtei, F., Mtove, G., Nadjm, B., Muro, F., Alegana, V., Noor, A. M., Todd, J. and Reyburn, H. (2014) Inpatient child mortality by travel time to hospital in a rural area of Tanzania. *Trop Med Int Health*.

MapCruzin. (2013). Free Afghanistan ArcGIS shapefile map layers [Online]. Place Published. Available: http://www.mapcruzin.com/free-afghanistan-arcgis-maps-shapefiles.htm [Accessed February, 14 2013].

MARA/ARMA. (2004). Mapping Malaria Risk in Africa [Online]. Place Published. Available: http://www.mara.org.za/ [Accessed 17 December 2013].

Marsh, K. and Snow, R. W. (1999) Malaria transmission and morbidity. *Parassitologia,* 41**,** 241-6.

Martin, D. and Atkinson, P. M. (2001) Investigating the Spatial Linkage of Primary School Performance and Catchment Characteristics. *Geographical and Environmental Modelling,* 5**,** 67-83.

Martin, D., Jordan, H. and Roderick, P. (2008) Taking the bus: incorporating public transport timetable data into health care accessibility modelling. *Environment and Planning A,* 40**,** 2510-2525.

Martin, D., Roderick, P., Diamond, I., Clements, S. and Stone, N. (1998) Geographical aspects of the uptake of renal replacement therapy in England. *Int J Population Geography,* 4**,** 227 - 242.

Martin, D., Wrigley, H., Barnett, S. and Roderick, P. (2002) Increasing the sophistication of access measurement in a rural healthcare study. *Health & Place,* 8**,** 3 - 13.

Martinez, B. and Gilabert, M. A. (2009) Vegetation dynamics from NDVI time series analysis using the wavelet transform. *Remote Sensing of Environment,* 113**,** 1823-1842.

Martins, T., Simpson, D., lindgren, F. and Rue, H.(2013) Bayesian computing with INLA: new features, Trondheim, Norway, Department of Mathematical Sciences, Norwegian University of Science and Technology. 29

Mashal, T., Takano, T., Nakamura, K., Kizuki, M., Hemat, S., Watanabe, M. and Seino, K. (2008) Factors associated with the health and nutritional status of children under 5 years of age in Afghanistan: family behaviour related to women and past experience of war-related hardships. *BMC Public Health,* 8**,** 301.

Mauch, V., Weil, D., Munim, A., Boillot, F., Coninx, R., Huseynova, S., Powell, C., Seita, A., Wembanyama, H. and van den Hof, S. (2010) Structure and management of tuberculosis control programs in fragile states--Afghanistan, DR Congo, Haiti, Somalia. *Health Policy,* 96**,** 118-27.

Mayxay, M., Pukrittayakamee, S., Newton, P. N. and White, N. J. (2004) Mixed-species malaria infections in humans. *Trends Parasitol,* 20**,** 233-40.

Mbeeli, T., Samahiya, M., Ravishankar, N., Zere, E. and Kirigia, J. (2011) Resource flows for health care: Namibia reproductive health sub-accounts. *International Archives of Medicine,* 4**,** 41.

McCombie, S. C. (1996) Treatment seeking for malaria: a review of recent research. *Soc Sci Med,* 43**,** 933-45.

McCombie, S. C. (2002) Self-treatment for malaria: the evidence and methodological issues. *Health Policy Plan,* 17**,** 333-44.

McLeod, A. I. and Xu, C. (2008) bestglm: Best Subset GLM.

MDA. (2013). EarthSat GeoCover LC Overview [Online]. Place Published: MD: MacDonald, Dettwiler and Associates Ltd. Available: http://www.mdafederal.com/geocover/geocoverlc/gclcoverview [Accessed 27 February 2013].

MEASURE DHS. (2011). Demographic and Health Surveys [Online]. Place Published: Measure DHS. Available: http://www.measuredhs.com [Accessed December 15 2010].

Mehlotra, R. K., Lorry, K., Kastens, W., Miller, S. M., Alpers, M. P., Bockarie, M., Kazura, J. W. and Zimmerman, P. A. (2000) Random distribution of mixed species malaria infections in Papua New Guinea. *Am J Trop Med Hyg,* 62**,** 225-31.

Mendelsohn, J. M. and El Obeid, S.(2001) A preliminary profile of Kavango region, Windhoek. , Namibia Nature Foundation.

Mendelsohn, J. M., El Obeid, S. and Roberts, C. S.(2000) A profile of north-central Namibia, Windhoek, Gamsberg MacMillan.

Mendelsohn, J. M. and Roberts, C. S.(1997) An environmental atlas and profile of Caprivi, Windhoek, Ministry of Environment and Tourism, Windhoek.

Mendis, K., Rietveld, A., Warsame, M., Bosman, A., Greenwood, B. and Wernsdorfer, W. H. (2009) From malaria control to eradication: The WHO perspective. *Tropical Medicine & International Health,* 14**,** 802-809.

Ministry of Finance.(2010) Afghanistan Millennium Development Goals report 2010, Kabul, Afghanistan National Development Strategy Directorate. 59

Ministry of Health.(2010) Annual report: July 2009 to June 2010., Kigali, Government of Rwanda.

Ministry of Health and Social Services.(1995) Integrated Health Care Delivery: The Challenge of Implementation. A Situation Analysis and Practical Implementation Guide, Windhoek, Ministry of Health and Social Services.

Ministry of Health and Social Services.(1998) Towards Achieving Health and Social Well-being for All Namibians: A Policy Framework, Windhoek, Ministry of Health and Social Services.

Ministry of Health and Social Services.(2007) Community based health care, Windhoek, Ministry of Health and Social Services.

Ministry of Health and Social Services.(2008a) Health and Social services system review, Windhoek, Ministry of Health and Social Services, Republic of Namibia.

Ministry of Health and Social Services (2008b) Namibia Demographic and Health Survey, 2006-07. Windhoek, Namibia, Ministry of Health and Social Services.

Ministry of Health and Social Services.(2010a) Namibia Health Facility Census, 2009, Windhoek, Namibia, Ministry of Health and Social Services and ICF Macro. 585

Ministry of Health and Social Services.(2010b) Namibia malaria indicator survey 2009, Windhoek. 86

Ministry of Health and Social Services.(2010c) Namibia malaria strategic plan 2010-2016, Windhoek. 34

Ministry of Health and Social Services.(2010d) National health policy framework 2010-2020, Windhoek, Ministry of Health and Social Services.

Ministry of Health and Social Services and World Health Organization (1996) *Malaria in Namibia. An information pack prepared for the national malaria awareness week,* Windhoek, Namibia, WHO,
Ministry of Health and Social Services.

Ministry of Health and Social Services (MoHSS) and ICF Macro.(2010) Namibia Health Facility Census (HFC) 2009, Windhoek, Namibia, MoHSS and ICF Macro. http://www.measuredhs.com/publications/publication-spa16-spa-final-reports.cfm.

Ministry of Health, E.(2012) Malaria indicator and prevalence surveys, 2012, Asmara, National malaria control programme.

Ministry of Health, Z. and HMN.(2007) Assessment of the Health Information System in Zambia, Lusaka, Department of Planning & development Monitoring & Evaluation Unit.

Ministry of public Health.(2002) Afghanistan National Health Resources Assessment (ANHRA), Kabul. http://www.msh.org/afghanistan/ANHRA_2002_LITE.pdf.

Ministry of Public Health.(2003) A basic package of health services for Afghanistan. Kabul: Islamic Transitional Government of Afghanistan, Kabul. 1382

Ministry of Public Health.(2006) National Health Management Information Systems (HMIS) : Procedures manual, Kabul.

Ministry of Public Health.(2008a) Afghanistan Health Survey 2006: Estimates of Priority Indicators for Rural Afghanistan, Kabul, Johns Hopkins University/Indian Institute of Health Management Research and Islamic Republic of Afghanistan, Ministry of Public Health.

Ministry of Public Health.(2008b) National Malaria Strategic Plan 2008 – 2013, General Directorate of Preventive Medicine and Primary Health Care, National Malaria and Leishmaniasis Control Programme.

Ministry of Public Health.(2009a) Malaria Indicator Survey, Ministry of Public Health, Islamic Republic of Afghanistan.

Ministry of Public Health.(2009b) National strategy on health care financing and sustainability 2009-2013, Kabul.

Ministry of Public Health. 2010a. A Basic Package of Health Services for Afghanistan. [Online]. Available: [Accessed.

Ministry of Public Health. 2010b. National malaria treatment guidelines. [Online]. Available: [Accessed.

Ministry of Public Health.(2011) National Health Accounts Afghanistan 2008-2009: Preliminary findings, Kabul. 38 http://www.who.int/nha/country/afg/afg-nha_2008-2009.pdf.

Ministry of Public Health (2012) The second Malaria Indicator Survey in the Islamic Republic of Afghanistan, 2011. World Health Organization, Health Net TPO, BRAC, HPRO, KEMRI/Wellcome Trust Research Programme.

Mirghani, S., Nour, B., Bushra, S., Elhassan, I., Snow, R. and Noor, A. (2010) The spatial-temporal clustering of Plasmodium falciparum infection over eleven years in Gezira State, The Sudan. *Malaria Journal,* 9**,** 172.

MOH.(2008) Health Information System Strategic Plan (HISSP) 2009-2013. http://www.who.int/healthmetrics/library/countries/HMN_ERI_StrPlan_Final_2008_12_en.pdf.

MoH.(2010) National health policy. http://www.who.int/medical_devices/survey_resources/health_technology_national_policy_eritrea.pdf.

MoH.(2013) Annual health service activity report of year 2012, Asmara. http://www.scribd.com/doc/164234199/Eritrea-Annual-Health-Service-Activity-Report-of-Year-2012.

MoH and RBM.(2005) Five year strategic plan 2005-2009, Asmara. http://www.rbm.who.int/countryaction/nsp/eritrea1.pdf.

MoHSS.(2011) Namibia malaria programme perfomance review. http://www.rbm.who.int/countryaction/aideMemoire/Namibia-The-malaria-program-performance-review-2011.pdf.

Moisi, J. C., Nokes, D. J., Gatakaa, H., Williams, T. N., Bauni, E., Levine, O. S. and Scott, J. A. (2011) Sensitivity of hospital-based surveillance for severe disease: a geographic information system analysis of access to care in Kilifi district, Kenya. *Bull World Health Organ,* 89**,** 102-11.

Molineaux, L. (1988) The epidemiology of human malaria as an explanation of its distribution, including some implications for its control. IN Wernsdorfer, W. and McGregor, I. L. (Eds.) *Malaria: Principles and Practice of Malariology.* London, Churchill Livingstone.

Monmonier, M. S. and Schnell, G. A. (1984) Land-use and land-cover data and the mapping of population density. *The International Yearbook of Cartography,* 24**,** 115-21.

Mueller, I., Galinski, M. R., Baird, J. K., Carlton, J. M., Kochar, D. K., Alonso, P. L. and del Portillo, H. A. (2009) Key gaps in the knowledge of Plasmodium vivax, a neglected human malaria parasite. *Lancet Infect Dis,* 9**,** 555-66.

Mueller, I., Schoepflin, S., Smith, T. A., Benton, K. L., Bretscher, M. T., Lin, E., Kiniboro, B., Zimmerman, P. A., Speed, T. P., Siba, P. and Felger, I. (2012) Force of infection is key to understanding the epidemiology of Plasmodium falciparum malaria in Papua New Guinean children. *Proceedings of the National Academy of Sciences*.

Mueller, I., Slutsker, L. and Tanner, M. (2011) Estimating the Burden of Malaria: The Need for Improved Surveillance. *PLoS Med,* 8**,** e1001144.

Mufunda, J., Nyarango, P., Usman, A., Gebremeskel, T., Mebrahtu, G., Ogbamariam, A., Kosia, A., Ghebrat, Y., Gebresillosie, S., Goitom, S., Araya, E., Andemichael, G. and Gebremichael, A. (2007) Roll back malaria--an African success story in Eritrea. *S Afr Med J,* 97**,** 46-50.

Müller, I., Smith, T., Mellor, S., Rare, L. and Genton, B. (1998) The effect of distance from home on attendance at a small rural health centre in Papua New Guinea. *International Journal of Epidemiology,* 27**,** 878-884.

Mundry, R. and Nunn, C. L. (2009) Stepwise model fitting and statistical inference: turning noise into signal pollution. *Am Nat,* 173**,** 119-23.

Murray, C. J., Lopez, A. D. and Wibulpolprasert, S. (2004) Monitoring global health: time for new solutions. *BMJ,* 329**,** 1096-100.

Murtaugh, P. A. (2009) Performance of several variable-selection methods applied to real ecological data. *Ecology Letters,* 12**,** 1061-1068.

Muthaka, D. I., Kimani, D. N., Mwaura, S. and Manda, D. K.(2004) A Review of the regulatory framework for private healthcare services in Kenya, Nairobi, Kenya Institute for Public Policy Research and Analysis.

Nabyonga-Orem, J., Karamagi, H., Atuyambe, L., Bagenda, F., Okuonzi, S. A. and Walker, O. (2008) Maintaining quality of health services after abolition of user fees: A Uganda case study. *BMC Health Serv Res,* 8**,** 102.

Nájera, J. A., González-Silva , M. and Alonso, P. L. (2011) Some Lessons for the Future from the Global Malaria Eradication Programme (1955â€“1969). *PLoS Med,* 8**,** e1000412.

Nakamura, H., Ikeda, N., Stickley, A., Mori, R. and Shibuya, K. (2011) Achieving MDG 4 in sub-Saharan Africa: what has contributed to the accelerated child mortality decline in Ghana? *PLoS One,* 6**,** e17774.

NASA. (2011). TRMM [Online]. Place Published: NASA Goddard Space Flight Center. Available: http://trmm.gsfc.nasa.gov/ [Accessed November 8 2011].

NASA. (2012). ASTER [Online]. Place Published. Available: http://asterweb.jpl.nasa.gov/gdem.asp [Accessed November 15, 2013.

Nateghpour, M., Akbarzadeh, K., Farivar, L. and Amiri, A. (2011) Detection of asymptomatic malaria infection among the Afghani immigrant population in Iranshahr district of southeastern Iran. *Bulletin de la SociÃ©tÃ© de pathologie exotique,* 104**,** 321-323.

National planning Commission.(2004) Namibia vision 2030: Policy framework for long term development, Windhoek, Office of the president.

National Planning Commission.(2008) The 2nd Millenium development goals report: Namibia, Windhoek.

National Planning Commission.(2012) Namibia 2011 population and housing census preliminary results, Windhoek, Namibia. www.npc.gov.na.

National Planning Commission and UNDP.(2010) Namibia Millennium Development Goals: Third report Windhoek.

National Planning Commission Secretariat. (2013). Namibia vision 2030. [Online]. Place Published. Available: www.npc.gov.na/vision/vision_2030bgd.htm. [Accessed August 2 2013].

National Statistical Office (NSO) and ICF Macro.(2011) Malawi Demographic and Health Survey 2010, Zomba, Malawi and Calverton, Maryland, USA, NSO and ICF Macro.

National Statistics and Evaluation Office (NSEO) [Eritrea] and Macro, O.(2003) Eritrea Demographic and Health Survey 2002, Calverton, Maryland, USA, National Statistics and Evaluation Office and ORC Macro.

Nemet, G. F. and Bailey, A. J. (2000) Distance and health care utilization among the rural elderly. *Soc Sci Med,* 50**,** 1197-208.

Neuman, M. D., David, G., Silber, J. H., Sanford Schwartz, J. and Fleisher, L. A. (2011) Changing Access to Emergency Care for Patients Undergoing Outpatient Procedures at Ambulatory Surgery Centers: Evidence From Florida. *Medical Care Research and Review,* 68**,** 247-258.

Newbrander, W., Yoder, R. and Debevoise, A. B. (2007) Rebuilding health systems in post-conflict countries: estimating the costs of basic services. *The International Journal of Health Planning and Management,* 22**,** 319-336.

Ngo, A. D., Alden, D. L., Pham, V. and Phan, H. (2010) The impact of social franchising on the use of reproductive health and family planning services at public commune health stations in Vietnam. *BMC Health Serv Res,* 10**,** 54.

Ngui, A. and Apparicio, P. (2011) Optimizing the two-step floating catchment area method for measuring spatial accessibility to medical clinics in Montreal. *BMC Health Services Research,* 11**,** 166.

Nguyen, D. S., Dao, B. H., Nguyen, P. D., Nguyen, V. H., Le, N. B., Mai, V. S. and Meshnick, S. R. (1993) Treatment of malaria in Vietnam with oral artemisinin. *Am J Trop Med Hyg,* 48**,** 398-402.

Nobre, A. A., Schmidt, A. M. and Lopes, H. F. (2005) Spatio-temporal models for mapping the incidence of malaria in Pará. *Environmetrics,* 16**,** 291-304.

Noor, A., Alegana, V., Kamwi, R., Hansford, C., Ntomwa, B., Katokele, S. and Snow, R. (2013a) Malaria control and the intensity of Plasmodium falciparum transmission in Namibia 1969 -1992. *PLoS One,* 8.

Noor, A., Clements, A., Gething, P., Moloney, G., Borle, M., Shewchuk, T., Hay, S. and Snow, R. (2008) Spatial prediction of Plasmodium falciparum prevalence in Somalia. *Malaria Journal,* 7**,** 159.

Noor, A., Mohamed, M., Mugyenyi, C., Osman, M., Guessod, H., Kabaria, C., Ahmed, I., Nyonda, M., Cook, J., Drakeley, C., Mackinnon, M. and Snow, R. (2011) Establishing the extent of malaria transmission and challenges facing pre-elimination in the Republic of Djibouti. *BMC Infectious Diseases,* 11**,** 121.

Noor, A., Mutheu, J., Tatem, A., Hay, S. and Snow, R. (2009a) Insecticide-treated net coverage in Africa: mapping progress in 2000-07. *Lancet,* 373**,** 58 - 67.

Noor, A., Uusiku, P., Kamwi, R., Katokele, S., Ntomwa, B., Alegana, V. and Snow, R. (2013b) The receptive versus current risks of Plasmodium falciparum transmission in Northern Namibia: implications for elimination. *BMC Infectious Diseases,* 13**,** 184.

Noor, A. M., Alegana, V. A., Gething, P. W. and Snow, R. W. (2009b) A spatial national health facility database for public health sector planning in Kenya in 2008. *Int J Health Geogr,* 8**,** 13.

Noor, A. M., Alegana, V. A., Kamwi, R. N., Hansford, C. F., Ntomwa, B., Katokele, S. and Snow, R. W. (2013c) Malaria Control and the Intensity of *Plasmodium falciparum* Transmission in Namibia 1969 - 1992. *PLoS One,* 8**,** e63350.

Noor, A. M., Amin, A. A., Gething, P. W., Atkinson, P. M., Hay, S. I. and Snow, R. W. (2006) Modelling distances travelled to government health services in Kenya. *Trop Med Int Health,* 11**,** 188-96.

Noor, A. M., Gething, P. W., Alegana, V. A., Patil, A. P., Hay, S. I., Muchiri, E., Juma, E. and Snow, R. W. (2009c) The risks of malaria infection in Kenya in 2009. *BMC Infect Dis,* 9**,** 180.

Noor, A. M., Kinyoki, D. K., Mundia, C. W., Kabaria, C. W., Mutua, J. W., Alegana, V. A., Fall, I. S. and Snow, R. W. (2014) The changing risk of Plasmodium falciparum malaria infection in Africa: 2000-10: a spatial and temporal analysis of transmission intensity. *The Lancet,* 383**,** 1739-1747.

Noor, A. M., Rage, I. A., Moonen, B. and Snow, R. W. (2009d) Health service providers in Somalia: their readiness to provide malaria case-management. *Malar J,* 8**,** 100.

Noor, A. M., Zurovac, D., Hay, S. I., Ochola, S. A. and Snow, R. W. (2003) Defining equity in physical access to clinical services using geographical information systems as part of malaria planning and monitoring in Kenya. *Trop Med Int Health,* 8**,** 917-26.

NoorAli, R., Luby, S. and Rahbar, M. H. (1999) Does use of a government service depend on distance from the health facility? *Health Policy Plan,* 14**,** 191-7.

Nyarango, P. M., Gebremeskel, T., Mebrahtu, G., Mufunda, J., Abdulmumini, U., Ogbamariam, A., Kosia, A., Gebremichael, A., Gunawardena, D., Ghebrat, Y. and Okbaldet, Y. (2006) A steep decline of malaria morbidity and mortality trends in Eritrea between 2000 and 2004: the effect of combination of control methods. *Malar J,* 5**,** 33.

O'Connell, K., Hom, M., Aung, T., Theuss, M. and Huntington, D. (2011) Using and joining a franchised private sector provider network in Myanmar. *PLoS One,* 6**,** e28364.

O'Meara, W. P., Bejon, P., Mwangi, T. W., Okiro, E. A., Peshu, N., Snow, R. W., Newton, C. R. J. C. and Marsh, K. (2008) Effect of a fall in malaria transmission on morbidity and mortality in Kilifi, Kenya. *The Lancet,* 372**,** 1555-1562.

O'Meara, W. P., Mangeni, J. N., Steketee, R. and Greenwood, B. (2010) Changes in the burden of malaria in sub-Saharan Africa. *The Lancet Infectious Diseases,* 10**,** 545-555.

O'Meara, W. P., Noor, A., Gatakaa, H., Tsofa, B., McKenzie, F. E. and Marsh, K. (2009) The impact of primary health care on malaria morbidity - defining access by disease burden. *Trop Med Int Health,* 14**,** 29-35.

OECD, D. C.-o. D. D.-D. (2012). Partnership in statistics for development in the 21st century (PARIS21) [Online]. Place Published. Available: http://www.paris21.org/ [Accessed 13 March 2012].

Okiro, E., Alegana, V., Noor, A., Mutheu, J., Juma, E. and Snow, R. (2009) Malaria paediatric hospitalization between 1999 and 2008 across Kenya. *BMC Medicine,* 7**,** 75.

Okwaraji, Y. B., Cousens, S., Berhane, Y., Mulholland, K. and Edmond, K. (2012) Effect of Geographical Access to Health Facilities on Child Mortality in Rural Ethiopia: A Community Based Cross Sectional Study. *PLoS One,* 7**,** e33564.

Ollivier, L., Nevin, R. L., Darar, H. Y., Bougere, J., Saleh, M., Gidenne, S., Maslin, J., Anders, D., Decam, C., Todesco, A., Khaireh, B. A. and Ahmed, A. A. (2011) Malaria in the Republic of Djibouti, 1998-2009. *Am J Trop Med Hyg,* 85**,** 554-9.

Omer, K., Cockcroft, A. and Andersson, N. (2011) Impact of a hospital improvement initiative in Bangladesh on patient experiences and satisfaction with services: two cross-sectional studies. *BMC Health Services Research,* 11**,** S10.

Omumbo, J., Ouma, J., Rapuoda, B., Craig, M. H., le Sueur, D. and Snow, R. W. (1998) Mapping malaria transmission intensity using geographical information systems (GIS): an example from Kenya. *Ann Trop Med Parasitol,* 92**,** 7-21.

Omumbo, J. A., Hay, S. I., Goetz, S. J., Snow, R. W. and Rogers, D. J. (2000) Updating historical maps of malaria transmission intensity in East Africa using Remote Sensing. *Photogrammetric Engineering & Remote Sensing,* 68**,** 161-166.

Omumbo, J. A., Hay, S. I., Guerra, C. A. and Snow, R. W. (2004) The relationship between the Plasmodium falciparum parasite ratio in childhood and climate estimates of malaria transmission in Kenya. *Malar J,* 3**,** 17.

Omumbo, J. A., Hay, S. I., Snow, R. W., Tatem, A. J. and Rogers, D. J. (2005) Modelling malaria risk in East Africa at high-spatial resolution. *Tropical Medicine & International Health,* 10**,** 557-566.

Onwujekwe, O., Hanson, K. and Uzochukwu, B. (2011) Do poor people use poor quality providers? Evidence from the treatment of presumptive malaria in Nigeria. *Trop Med Int Health,* 16**,** 1087-98.

Organizastion, W. H.(1969) Re-examination of the Global Strategy of Malaria Eradication, Provisional Agenda Item 2.4, Geneva.

Otten, M., Aregawi, M., Were, W., Karema, C., Medin, A., Bekele, W., Jima, D., Gausi, K., Komatsu, R., Korenromp, E., Low-Beer, D. and Grabowsky, M. (2009) Initial evidence of reduction of malaria cases and deaths in Rwanda and Ethiopia due to rapid scale-up of malaria prevention and treatment. *Malar J,* 8**,** 14.

Owen, K. K., ObregÃ³n, E. J. and Jacobsen, K. H. (2011) A geographic analysis of access to health services in rural Guatemala. *International Health,* 2**,** 143-149.

Owusu-Agyei, S., Koram, K. A., Baird, J. K., Utz, G. C., Binka, F. N., Nkrumah, F. K., Fryauff, D. J. and Hoffman, S. L. (2001) Incidence of symptomatic and asymptomatic Plasmodium falciparum infection following curative therapy in adult residents of northern Ghana. *The American Journal of Tropical Medicine and Hygiene,* 65**,** 197-203.

Oyakhirome, S., Profanter, K. and Kremsner, P. G. (2010) Assessment of Fever in African Children: Implication for Malaria Trials. *The American Journal of Tropical Medicine and Hygiene,* 82**,** 215-218.

Paaijmans, K. P., Wandago, M. O., Githeko, A. K. and Takken, W. (2007) Unexpected High Losses of *Anopheles gambiae* Larvae Due to Rainfall. *PLoS One,* 2**,** e1146.

Palmer, N., Strong, L., Wali, A. and Sondorp, E. (2006) Contracting out health services in fragile states. *BMJ,* 332**,** 718-721.

Parenteau, M.-P. and Sawada, M. (2011) The modifiable areal unit problem (MAUP) in the relationship between exposure to NO2 and respiratory health. *International Journal of Health Geographics,* 10**,** 58.

Parsons, E., Chalkley, B. and Jones, A. (2000) School Catchments and Pupil Movements: A case study in parental choice. *Educational Studies,* 26**,** 33-48.

Patil, A. P., Okiro, E. A., Gething, P. W., Guerra, C. A., Sharma, S. K., Snow, R. W. and Hay, S. I. (2009) Defining the relationship between Plasmodium falciparum parasite rate and clinical disease: statistical models for disease burden estimation. *Malar J,* 8**,** 186.

Paulander, J., Olsson, H., Lemma, H., Getachew, A. and San Sebastian, M. (2009) Knowledge, attitudes and practice about malaria in rural Tigray, Ethiopia. *Global Health Action.*

Penchansky, R. and Thomas, J. W. (1981) The concept of access: definition and relationship to consumer satisfaction. *Med Care,* 19**,** 127-40.

Perry, H., McDonnell, S., Alemu, W., Nsubuga, P., Chungong, S., Otten, M., Lusamba-dikassa, P. and Thacker, S. (2007) Planning an integrated disease surveillance and response system: a matrix of skills and activities. *BMC Medicine,* 5**,** 24.

Peters, D. H., Noor, A. A., Singh, L. P., Kakar, F. K., Hansen, P. M. and Burnham, G. (2007) A balanced scorecard for health services in Afghanistan. *Bull World Health Organ,* 85**,** 146-51.

Pinheiro, J. and Bates, D. (2002) *Mixed Effects Models in S and S-Plus*, Springer.

Prata, N., Montagu, D. and Jefferys, E. (2005) Private sector, human resources and health franchising in Africa. *Bull World Health Organ,* 83**,** 274-9.

Price, R. N., Tjitra, E., Guerra, C. A., Yeung, S., White, N. J. and Anstey, N. M. (2007) Vivax malaria: neglected and not benign. *Am J Trop Med Hyg,* 77**,** 79-87.

Pull, J. H. (1972) Malaria surveillance methods, their uses and limitations. *Am J Trop Med Hyg,* 21**,** 651-7.

Quaye, R. (2010) *Balancing public and private health care systems: The Sub-Saharan African experience,* Maryland, University Press of America.

R Development Core Team (2010) R : A language and environment for statistical computing. Vienna, Austria, R foundation for statistical computing.

Rahman, S. U. and Smith, D. K. (1999) Deployment of Rural Health Facilities in a Developing Country. *The Journal of the Operational Research Society,* 50**,** 892-902.

Rajakaruna, R. S., Alifrangis, M., Amerasinghe, P. H. and Konradsen, F. (2010) Pre-elimination stage of malaria in Sri Lanka: assessing the level of hidden parasites in the population. *Malaria Journal,* 9.

Ramachandra, T. (1951) Malaria control using indoor residual sprays in the Eastern Province of Afghanistan. *Bull World Health Organ,* 3**,** 639-61.

Ramiro Ruiz-Cárdenas, R., Krainski, E. T. and Rue, H. (2012) Direct fitting of dynamic models using integrated nested Laplace approximations â€" INLA. *Computational Statistics & Data Analysis,* 56**,** 1808-1828.

Raso, G., Schur, N., Utzinger, J., Koudou, B., Tchicaya, E., Rohner, F., N'Goran, E., Silue, K., Matthys, B., Assi, S., Tanner, M. and Vounatsou, P. (2012) Mapping malaria risk among children in Cote d'Ivoire using Bayesian geo-statistical models. *Malaria Journal,* 11**,** 160.

Ray, A. P. and Beljaev, A. E. (1984) Epidemiological surveillance: a tool for assessment of malaria and its control. *J Commun Dis,* 16**,** 197-207.

Ray, N. and Ebener, S. (2008) AccessMod 3.0: computing geographic coverage and accessibility to health care services using anisotropic movement of patients. *International Journal of Health Geographics,* 7**,** 63.

RBM-MERG.(2008) A guide to Malaria Indicator Surveys (MIS). http://malariasurveys.org/toolkit.cfm.

Regional Office for Eastern and Mediteranean Region, W. (2013). Malaria control and elimination [Online]. Place Published. Available: http://www.emro.who.int/entity/malaria-control-and-elimination/ [Accessed January 20 2013].

Reid, H. L., Haque, U., Roy, S., Islam, N. and Clements, A. C. (2012) Characterizing the spatial and temporal variation of malaria incidence in Bangladesh, 2007. *Malar J,* 11**,** 170.

Reilley, B., Frank, T., Prochnow, T., Puertas, G. and Van Der Meer, J. (2004) Provision of health care in rural Afghanistan: needs and challenges. *Am J Public Health***,** 1686-1688.

Rinaldi, L., Cascone, C., Sibilio, G., Musella, V., Taddei, R. and Cringoli, G. (2004) Geographical Information Systems and remote sensing technologies in parasitological epidemiology. *Parassitologia,* 46**,** 71-4.

Roberts, L. and Enserink, M. (2007) Did They Really Say ... Eradication? *Science,* 318**,** 1544-1545.

Robinson, W. S. (1950) Ecological Correlations and the Behavior of Individuals. *American Sociological Review,* 15**,** 351-357.

Rodrigue, J.-P., Comtois, C. and Slack, B. (2009) *The Geography of Transport Systems,* New York, Routledge.

Roghmann, K. J. and Zastowny, T. R. (1979) Proximity as a factor in the selection of health care providers: Emergency room visits compared to obstetric admissions and abortions. *Social Science & Medicine. Part D: Medical Geography,* 13**,** 61-69.

Rohuani, S. and Hall, T. J. (Eds.) (1989) *Space–time kriging of groundwater data,* Dordrecht, Kluwer Academic Publishers.

Roll Back Malaria Monitoring and Evaluation Reference Group, World Health Organization, United Nations Children's Fund, MEASURE DHS, MEASURE Evaluation and Prevention, U. S. C. f. D. C. a. 2005. Malaria Indicator Survey: Basic documentation for survey design and implementation. [Online]. Available: [Accessed.

Roll Back Malaria partnership.(2011) A decade of partnership and results, Geneva. http://www.rbm.who.int/ProgressImpactSeries/docs/report8-en.pdf.

Roos, N. P. (1993) Linking patients to hospitals. Defining urban hospital service populations. *Med Care,* 31**,** YS6-15.

Rooth, I. and Bjorkman, A. (1992) Fever episodes in a holoendemic malaria area of Tanzania: parasitological and clinical findings and diagnostic aspects related to malaria. *Trans R Soc Trop Med Hyg,* 86**,** 479-82.

Rosenberg, M. W. and Hanlon, N. T. (1996) Access and utilization: a continuum of health service environments. *Soc Sci Med,* 43**,** 975-83.

Ross, R. (1910) *The prevention of malaria,* London, John Murray.

Rouhani, S. and Myers, D. (1990) Problems in space-time kriging of geohydrological data. *Mathematical Geology,* 22**,** 611-623.

Rowland, M., Hewitt, S., Durrani, N., Bano, N. and Wirtz, R. (1997) Transmission and control of vivax malaria in Afghan refugee settlements in Pakistan. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 91**,** 252-255.

Rowland, M., Mohammed, N., Rehman, H., Hewitt, S., Mendis, C., Ahmad, M., Kamal, M. and Wirtz, R. (2002) Anopheline vectors and malaria transmission in eastern Afghanistan. *Trans R Soc Trop Med Hyg,* 96**,** 620-6.

Rue, H. and Held, L. (2005) *Gaussian Markov Random Fields: Theory and Applications (Chapman & Hall/CRC Monographs on Statistics & Applied Probability),* Chapman and Hall/CRC.

Rue, H., Martino, S. and Chopin, N. (2009) Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 71**,** 319-392.

Rue, H. v. and Martino, S. (2007) Approximate Bayesian inference for hierarchical Gaussian Markov random field models. *Journal of Statistical Planning and Inference,* 137**,** 3177-3192.

Rutebemberwa, E., Kallander, K., Tomson, G., Peterson, S. and Pariyo, G. (2009) Determinants of delay in care-seeking for febrile children in eastern Uganda. *Trop Med Int Health,* 14**,** 472-9.

Sabri, B., Siddiqi, S., Ahmed, A. M., Kakar, F. K. and Perrot, J. (2007) Towards sustainable delivery of health services in Afghanistan: options for the future. *Bull World Health Organ,* 85**,** 712-8.

Safi, N., Hameed, H., Sediqi, W. and Himmat, E. (2009a) NMLEP Annual Report 2008. *Afghanistan Annual Malaria Journal***,** 8-14.

Safi, N., Leslie, T. and Rowland, M. (2009b) Progress and Challenges to Malaria Control in Afghanistan. *Afghanistan Annual Malaria Journal***,** 15-29.

Sahu, S. K., Baffour, B., Harper, P. R., Minty, J. H. and Sarran, C. (2013) A hierarchical Bayesian model for improving short-term forecasting of hospital demand by including meteorological information. *Journal of the Royal Statistical Society: Series A (Statistics in Society),* 177**,** 39–61.

Sahu, S. K. and Bakar, K. S. (2012) Hierarchical Bayesian autoregressive models for large space–time data with applications to ozone concentration modelling. *Applied Stochastic Models in Business and Industry,* 28**,** 395-415.

Saksena, P., Antunes, A. F., Xu, K., Musango, L. and Carrin, G. (2011) Mutual health insurance in Rwanda: Evidence on access to care and financial risk protection. *Health Policy,* 99**,** 203-209.

Satoguina, J., Walther, B., Drakeley, C., Nwakanma, D., Oriero, E. C., Correa, S., Corran, P., Conway, D. J. and 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.

Scharlemann, J. r. P. W., Benz, D., Hay, S. I., Purse, B. V., Tatem, A. J., Wint, G. R. W. and Rogers, D. J. (2008) Global Data for Ecology and Epidemiology: A Novel Algorithm for Temporal Fourier Processing MODIS Data. *PLoS One,* 3**,** e1408.

Schonfeld, H. K., Heston, J. F. and Falk, I. S. (1972) Numbers of Physicians Required for Primary Medical Care. *New England Journal of Medicine,* 286**,** 571-576.

Schrödle, B. and Held, L. (2010) A primer on disease mapping and ecological regression using INLA. *Computational Statistics,* 26**,** 241-258.

Schuurman, N., Fiedler, R. S., Grzybowski, S. C. W. and Grund, D. (2006) Defining rational hospital catchments for non-urban areas based on travel-time. *Int J Health Geogr,* 5**,** 43.

Shah, I., Rowland, M., Mehmood, P., Mujahid, C., Razique, F., Hewitt, S. and Durrani, N. (1997) Chloroquine resistance in Pakistan and the upsurge of falciparum malaria in Pakistani and Afghan refugee populations. *Annals of Tropical Medicine and Parasitology,* 91**,** 591-602.

Shannon, G. W., Bashshur, R. L. and Metzner, C. A. (1969) The concept of distance as a factor in accessibility and utilisation of health care. . *Med. Care Rev. ,* 26**,** 143-161.

Shannon, G. W., Spurlock, C. W., Gladin, S. T. and Skinner, J. L. (1975) A method for evaluating the geographic accessibility of health services. *The Professional Geographer,* 27**,** 30-36.

Sharan, M., Ahmed, S., Ghebrehiwet, M. and Rogo, K. (2011) The quality of the maternal health system in Eritrea. *Int J Gynaecol Obstet,* 115**,** 244-50.

Shepard, D. (1968) A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 23rd ACM national conference.* ACM.

Short Fabic, M., Choi, Y. and Bird, S. (2012) A systematic review of Demographic and Health Surveys: data availability and utilization for research. *Bull World Health Organ,* 90**,** 604-12.

Shukla, M., Singh, N. and Singh, M. (2011) Spleen rates and infant parasite rates as surveillance tool for malaria control in remote hard to reach areas of central India. *Malaria Journal,* 10**,** 381.

Silverman, B. W. (1981) Using Kernel Density Estimates to Investigate Multimodality. *Journal of the Royal Statistical Society. Series B (Methodological),* 43**,** 97-99.

Silverman, B. W. (1986) *Density estimation for statistics and data analysis,* London, Chapman & Hall.

Sinclair, D., Gogtay, N., Brand, F. and Olliaro, P. (2011) Artemisinin-based combination therapy for treating uncomplicated Plasmodium vivax malaria. *Cochrane Database Syst Rev***,** CD008492.

Singh, P., Singh, L., Kumar, C. and Rai, R. (2013) Correlates of maternal healthcare service utilisation among adolescent women in Mali: analysis of a nationally representative cross-sectional survey, 2006. *Journal of Public Health,* 21**,** 15-27.

Smith, D. L., Cohen, J. M., Chiyaka, C., Johnston, G., Gething, P. W., Gosling, R., Buckee, C. O., Laxminarayan, R., Hay, S. I. and Tatem, A. J. (2013) A sticky situation: the unexpected stability of malaria elimination. *Philos Trans R Soc Lond B Biol Sci,* 368**,** 20120145.

Smith, D. L., Dushoff, J., Snow, R. W. and Hay, S. I. (2005) The entomological inoculation rate and Plasmodium falciparum infection in African children. *Nature,* 438**,** 492-5.

Smith, D. L., Hay, S. I., Noor, A. M. and Snow, R. W. (2009) Predicting changing malaria risk after expanded insecticide-treated net coverage in Africa. *Trends Parasitol,* 25**,** 511-6.

Smith, D. L., McKenzie, F. E., Snow, R. W. and Hay, S. I. (2007) Revisiting the Basic Reproductive Number for Malaria and Its Implications for Malaria Control. *PLoS Biol,* 5**,** e42.

Smith Gueye, C., Gerigk, M., Newby, G., Lourenco, C., Uusiku, P. and Liu, J. (2014) Namibia's path toward malaria elimination: a case study of malaria strategies and costs along the northern border. *BMC Public Health,* 14**,** 1190.

Smith, T., Killen, G., Lengeler, C. and Tanner, M. (2004) Relationships between the outcome of plasmodium falciparum and the intensity of transmission in Africa. *The American Journal of Tropical Medicine and Hygiene,* 71**,** 80-86.

Snow, R. W., Alegana, V. A., Makomva, K., Reich, A., Uusiku, P., Katokele, S., Gething, P. W., Linard, C., Tatem, A. J., Moonen, B. and Noor, A. M.(2010a) Estimating the distribution of malaria in Namibia in 2009: assembling the evidence and modeling risk, Ministry of Health & Social Services, Republic of Namibia & Malaria Atlas Project.

Snow, R. W., Amratia, P., Kabaria, C. W., Noor, A. M. and Marsh, K. (2012) The changing limits and incidence of malaria in Africa: 1939-2009. *Adv Parasitol,* 78**,** 169-262.

Snow, R. W., Eckert, E. and Teklehaimanot, A. (2003) Estimating the needs for artesunate-based combination therapy for malaria case-management in Africa. *Trends Parasitol,* 19**,** 363-9.

Snow, R. W., Guerra, C. A., Noor, A. M., Myint, H. Y. and Hay, S. I. (2005) The global distribution of clinical episodes of Plasmodium falciparum malaria. *Nature,* 434**,** 214-7.

Snow, R. W. and Marsh, K. (2002) The consequences of reducing transmission of Plasmodium falciparum in Africa. *Adv Parasitol,* 52**,** 235-64.

Snow, R. W. and Marsh, K. (2010) Malaria in Africa: progress and prospects in the decade since the Abuja Declaration. *Lancet,* 376**,** 137-9.

Snow, R. W., Okiro, E. A., Gething, P. W., Atun, R. and Hay, S. I. (2010b) Equity and adequacy of international donor assistance for global malaria control: an analysis of populations at risk and external funding commitments. *Lancet,* 376**,** 1409-16.

Snow, R. W., Omumbo, J. A., Lowe, B., Molyneux, C. S., Obiero, J. O., Palmer, A., Weber, M. W., Pinder, M., Nahlen, B., Obonyo, C., Newbold, C., Gupta, S. and Marsh, K. (1997) Relation between severe malaria morbidity in children and level of Plasmodium falciparum transmission in Africa. *Lancet,* 349**,** 1650-4.

Southern Africa Roll Back Malaria Network (SARN).(2010) Report of the elimination - 8: Malaria elimination technical meeting, Maputo, SARN. 2012 http://www.malariaeliminationgroup.org/sites/default/files/E8_MAPUTO_REPORT_SEND.pdf.

Southern African Development Community. (2012). Malaria Elimination Eight [Online]. Place Published. Available: http://www.sadc.int/english/regional-integration/shdsp/sarn/elimination-eight-e8/ [Accessed.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Van Der Linde, A. (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 64**,** 583-639.

Stansfield, S. (2005) Structuring information and incentives to improve health. *Bull World Health Organ,* 83**,** 562.

Stein, M. L. (2005) Space: Time Covariance Functions. *Journal of the American Statistical Association,* 100**,** 310-321.

Stein, M. L., Chi, Z. and Welty, L. J. (2004) Approximating likelihoods for large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 66**,** 275-296.

Steinhardt, L. C., Aman, I., Pakzad, I., Kumar, B., Singh, L. P. and Peters, D. H. (2011) Removing user fees for basic health services: a pilot study and national roll-out in Afghanistan. *Health Policy and Planning,* 26**,** ii92-ii103.

Stekelenburg, J., Jager, B. E., Kolk, P. R., Westen, E. H. M. N., Kwaak, A. v. d. and Wolffers, I. N. (2005) Health care seeking behaviour and utilisation of traditional healers in Kalabo, Zambia. *Health Policy,* 71**,** 67-81.

Steketee, R. W. and ter Kuile, F. (2014) Single low-dose primaquine to reduce malaria transmission. *The Lancet Infectious Diseases,* 14**,** 91-92.

Stern, H. S. and Cressie, N. (Eds.) (1999) *Inference for extremes in disease mapping,* Chichester, Wiley.

Stock, R. (1983) Distance and the utilization of health facilities in rural Nigeria. *Soc Sci Med,* 17**,** 563-70.

Stouffer, S. A. (1940) Intervening opportunities: a theory relating mobility to distance. . *Am. Sot. Rec.,* 5**,** 845.

Stuckey, E. M., Smith, T. A. and Chitnis, N. (2013) Estimating malaria transmission through mathematical models. *Trends in parasitology,* 29**,** 477-482.

Sturrock, H. J., Novotny, J. M., Kunene, S., Dlamini, S., Zulu, Z., Cohen, J. M., Hsiang, M. S., Greenhouse, B. and Gosling, R. D. (2013a) Reactive case detection for malaria elimination: real-life experience from an ongoing program in Swaziland. *PLoS One,* 8**,** e63830.

Sturrock, H. J. W., Hsiang, M. S., Cohen, J. M., Smith, D. L., Greenhouse, B., Bousema, T. and Gosling, R. D. (2013b) Targeting Asymptomatic Malaria Infections: Active Surveillance in Control and Elimination. *PLoS Med,* 10**,** e1001467.

Tabios, G. Q. and Salas, J. D. (1985) A comparative analysis of techniques for spatial interpolation of precipitation. *JAWRA Journal of the American Water Resources Association,* 21**,** 365-380.

Tanser, F., Gijsbertsen, B. and Herbst, K. (2006) Modelling and understanding primary health care accessibility and utilization in rural South Africa: an exploration using a geographical information system. *Soc Sci Med,* 63**,** 691-705.

Tanser, F., Hosegood, V., Benzler, J. and Solarsh, G. (2001) New approaches to spatially analyse primary health care usage patterns in rural South Africa. *Trop Med Int Health,* 6**,** 826-38.

Tanzania Commission for AIDS (TACAIDS), Zanzibar AIDS Commission (ZAC), Office of Chief Government Statistician (OCGS) and Macro International Inc.(2008) Tanzania HIV/AIDS and Malaria Indicator Survey 2007-08, Dar es Salaam, TACAIDS, ZAC, NBS, OCGS and Macro International Inc.

Tatem, A., Adamo, S., Bharti, N., Burgert, C., Castro, M., Dorelien, A., Fink, G., Linard, C., Mendelsohn, J., Montana, L., Montgomery, M., Nelson, A., Noor, A., Pindolia, D., Yetman, G. and Balk, D. (2012) Mapping populations at risk: improving spatial demographic data for infectious disease modeling and metric derivation. *Population Health Metrics,* 10**,** 8.

Tatem, A., Huang, Z., Narib, C., Kumar, U., Kandula, D., Pindolia, D., Smith, D., Cohen, J., Graupe, B., Uusiku, P. and Lourenco, C. (2014) Integrating rapid risk mapping and mobile phone call record data for strategic malaria elimination planning. *Malaria Journal,* 13**,** 52.

Tatem, A. J., Noor, A. M., von Hagen, C., Di Gregorio, A. and Hay, S. I. (2007) High Resolution Population Maps for Low Income Nations: Combining Land Cover and Census in East Africa. *PLoS One,* 2**,** e1298.

Tatem, A. J., Smith, D. L., Gething, P. W., Kabaria, C. W., Snow, R. W. and Hay, S. I. (2010) Ranking of elimination feasibility between malaria-endemic countries. *The Lancet,* 376**,** 1579-1591.

Taylor, T. E. and Molyneux, M. E. (2002) Clinical features of malaria in children. IN Warrell, D. and Gilles, H. M. (Eds.) *Essential Marialology.* London, Arnold.

Teich, J. M. (1998) Clinical information systems for integrated healthcare networks. *Proc AMIA Symp***,** 19-28.

The Global Fund. (2013a). The global fund to fight AIDS, tuberculosis and malaria. Afghanistan: Country Grant Portfolio [Online]. Place Published. Available: http://portfolio.theglobalfund.org/en/Country/Index/AFG [Accessed August 2 2013].

The Global Fund. (2013b). The global fund to fight AIDS, tuberculosis and malaria. Eritrea: Country Grant Portfolio [Online]. Place Published. Available: http://portfolio.theglobalfund.org/en/Country/Index/ERI [Accessed February 04 2014].

The Global Fund. (2013c). The global fund to fight AIDS, tuberculosis and malaria. Namibia: Country Grant Portfolio [Online]. Place Published. Available: http://portfolio.theglobalfund.org/en/Contacts/CountryCoordinatingMechanisms/NAM [Accessed August 2 2013].

The Malaria Elimination Group (Ed.) (2009) *Shrinking the Malaria Map: A Prospectus on Malaria Elimination,* San Francisco, The Global Health Group, UCSF Global Health Sciences.

The malERA, C. G. o. D. and Diagnostics (2011) A Research Agenda for Malaria Eradication: Diagnoses and Diagnostics. *PLoS Med,* 8**,** e1000396.

The PLoS Medicine Editors (2011) Health Care Systems and Conflict: A Fragile State of Affairs. *PLoS Med,* 8**,** e1001065.

The World Bank (2012) Harmonized list of fragile situations FY13. *Fragile, conflict and violence.*

The World Bank. (2013). Namibia [Online]. Place Published. Available: http://www.worldbank.org/en/country/namibia [Accessed June 20 2013].

Thill, J.-C. and Kim, M. (2005) Trip making, induced travel demand, and accessibility. *Journal of Geographical Systems,* 7**,** 229-248.

Thurmond, M. C., Branscum, A. J., Johnson, W. O., Bedrick, E. J. and Hanson, T. E. (2005) Predicting the probability of abortion in dairy cows: a hierarchical Bayesian logistic-survival model using sequential pregnancy data. *Preventive Veterinary Medicine,* 68**,** 223-239.

Tierney, L. and Kadane, J. B. (1986) Accurate Approximations for Posterior Moments and Marginal Densities. *Journal of the American Statistical Association,* 81**,** 82-86.

Tobler, W.(1993) Three presentations on geographical analysis and modeling: National Center for Geographic Information and Analysis, Santa Barbara, CA93106-4060, University of California, Santa Barbara. Technical report 93-1

Trani, J.-F., Bakhshi, P., Noor, A. A., Lopez, D. and Mashkoor, A. (2010) Poverty, vulnerability, and provision of healthcare in Afghanistan. *Social Science & Medicine,* 70**,** 1745-1755.

Trani, J. F. and Barbou-des-Courieres, C. (2012) Measuring equity in disability and healthcare utilization in Afghanistan. *Med Confl Surviv,* 28**,** 219-46.

Trans-Zambezi Malaria Initiative (TZMI). (2012). Implementation Plan for the Trans-Zambezi Malaria Initiative (TZMI),a multi-national cross-border control and elimination initiative involving Angola, Botswana, Namibia, Zambia and Zimbabwe [Online]. Place Published. Available: https://www.sadc.int/files/4013/1824/0952/TZMI_CONCEPT_-_SEND.pdf [Accessed April 25 2013].

Trape, J.-F., Tall, A., Sokhna, C., Ly, A. B., Diagne, N., Ndiath, O., Mazenot, C., Richard, V., Badiane, A., Dieye-Ba, F., Faye, J., Ndiaye, G., Diene Sarr, F., Roucher, C., Bouganali, C., Bassène, H., Touré-Baldé, A., Roussilhon, C., Perraut, R., Spiegel, A., Sarthou, J.-L., Da Silva, L. P., Mercereau-Puijalon, O., Druilhe, P. and Rogier, C. (2014) The rise and fall of malaria in a west African rural community, Dielmo, Senegal, from 1990 to 2012: a 22 year longitudinal study. *The Lancet Infectious Diseases,* 14**,** 476-488

UN.(2012a) Report of the United Nations conference on sustainable development Rio de Janeiro, Brazil http://www.uncsd2012.org/content/documents/814UNCSD%20REPORT%20final%20revs.pdf.

UN.(2012b) The United Nations development strategy beyond 2015, New York. http://www.un.org/en/development/desa/policy/cdp/cdp_publications/2012cdppolicynote.pdf.

UN.(2014) The Millenium Development Goals report, 2014, New York. http://mdgs.un.org/unsd/mdg/Default.aspx.

UN General Assembly.(2000) Resolution adopted by the General Assembly: United Nations Millenium Declaration, New York, General Assembly. 55/2 http://www.un.org/millennium/declaration/ares552e.pdf.

UN Population Division. (2013). United Nations Population Division (2007) World Urbanization Prospects: The 2011 Revision [Online]. Place Published: Available at : http://esa.un.org/unpp/ [Accessed 17 February, 2013]. Available: http://data.un.org/Data.aspx?d=PopDiv&f=variableID%3A47 [Accessed.

UNDP.(2003) The millennium development goals: a compact among nations to end human poverty, New York, Oxford, United Nations Development Programme. http://hdr.undp.org/en/media/hdr03_complete.pdf.

UNDP.(2013a) Human development report 2013, Washington DC. 22 http://hdr.undp.org/en/media/HDR2013_EN_Summary.pdf.

UNDP. (2013b). International human development indicators [Online]. Place Published. Available: http://hdrstats.undp.org/en/countries/profiles/NAM.html [Accessed August 8 2013].

UNEP. (2002). GRID Africa GeoPortal [Online]. Place Published. Available: http://gridnairobi.unep.org/Portal/ [Accessed November 15, 2013.

United Nations.(2010) The Millenium Development Goals report, 2010, New York. http://www.un.org/millenniumgoals/pdf/MDG%20Report%202010%20En%20r15%20-low%20res%2020100615%20-.pdf.

United Nations.(2012) The Millenium Development Goals report, 2012, New York. http://www.un.org/millenniumgoals/11_MDG%20Report_EN.pdf.

United Nations Development Programme.(2003) The millennium development goals: a compact among nations to end human poverty, New York, Oxford, United Nations Development Programme. http://hdr.undp.org/en/media/hdr03_complete.pdf.

United Nations, G. A.(2000) Resolution adopted by the General Assembly: United Nations Millenium Declaration, New York, General Assembly. 55/2 http://www.un.org/millennium/declaration/ares552e.pdf.

van den Bogaart, E., Berkhout, M., Nour, A., Mens, P., Talha, A.-B., Adams, E., Ahmed, H., Abdelrahman, S., Ritmeijer, K., Nour, B. and Schallig, H. (2013) Concomitant malaria among visceral leishmaniasis in-patients from Gedarif and Sennar States, Sudan: a retrospective case-control study. *BMC Public Health,* 13**,** 332.

Vecchia, A. V. (1988) Estimation and Model Identification for Continuous Spatial Processes. *Journal of the Royal Statistical Society. Series B (Methodological),* 50**,** 297-312.

Vounatsou, P., Raso, G., Tanner, M., N'Goran E, K. and Utzinger, J. (2009) Bayesian geostatistical modelling for mapping schistosomiasis transmission. *Parasitology,* 136**,** 1695-705.

Wagner, H. and Tüchler, R. (2010) Bayesian estimation of random effects models for multivariate responses of mixed data. *Computational Statistics & Data Analysis,* 54**,** 1206-1218.

Wakefield, J. (2003) Sensitivity analyses for ecological regression. *Biometrics,* 59**,** 9-17.

Waldman, R., Strong, L. and Wali, A.(2006) Afghanistan's health system sinse 2001: condition improvement, prognosis cautiously optimistic., Afghanistan Research and Evaluation Unit briefing paper series. http://dspace.cigilibrary.org/jspui/bitstream/123456789/15408/1/Afghanistans%20Health%20System%20Since%202001%20Condition%20Improved%20Prognosis%20Cautiously%20Optimistic%202006.pdf?1.

Walter Carr, J. (1892) Enlargement of the spleen in young children. *The Lancet,* 139**,** 907-909.

Walter, Z. (2008). Bicycle speed and power calculation [Online]. Place Published. Available: http://www.kreuzotter.de/english/espeed.htm [Accessed 25 February 2013].

Wang, L. (2007) Immigration, ethnicity, and accessibility to culturally diverse family physicians. *Health Place,* 13**,** 656-71.

Wardlow, B. D. and Egbert, S. L. (2010) A comparison of MODIS 250-m EVI and NDVI data for crop mapping: a case study for southwest Kansas. *International Journal of Remote Sensing,* 31**,** 805-830.

Wardlow, B. D., Egbert, S. L. and Kastens, J. H. (2007) Analysis of time-series MODIS 250Â m vegetation index data for crop classification in the U.S. Central Great Plains. *Remote Sensing of Environment,* 108**,** 290-310.

Wardrop, N., Barnett, A., Atkinson, J.-A. and Clements, A. (2013) Plasmodium vivax malaria incidence over time and its association with temperature and rainfall in four counties of Yunnan Province, China. *Malaria Journal,* 12**,** 452.

Water Resource Programme.(2007) Analysis of Present and Future Water demand in Namibia: Namibia Water Resources Management Review, Windhoek, Windhoek Consulting Engineers.

Weiss, D., Mappin, B., Dalrymple, U., Bhatt, S., Cameron, E., Hay, S. and Gething, P. (2015) Re-examining environmental correlates of Plasmodium falciparum malaria endemicity: a data-intensive variable selection approach. *Malaria Journal,* 14**,** 68.

White, N. (2011) Determinants of relapse periodicity in Plasmodium vivax malaria. *Malaria Journal,* 10**,** 297.

White, N., Qiao, L., Qi, G. and Luzzatto, L. (2012) Rationale for recommending a lower dose of primaquine as a Plasmodium falciparum gametocytocide in populations where G6PD deficiency is common. *Malaria Journal,* 11**,** 418.

Whittingham, M. J., Stephens, P. A., Bradbury, R. B. and Freckleton, R. P. (2006) Why do we still use stepwise modelling in ecology and behaviour? *Journal of Animal Ecology,* 75**,** 1182-1189.

Whittle, P. (1954) On Stationary Processes in the Plane. *Biometrika,* 41**,** 434-449.

Whittle, P. (1963) Stochastic processes in several dimensions. *Bulletin of the International Statistical Institute,* 40**,** 974-994.

WHO-FIND.(2009) Malaria rapid diagnostic test performance: results of WHO product testing of malaria RDTs: round 1 (2008), Geneva, World Health Organization Special Programme for Tropical Diseases.

WHO/AFRO.(2001) Integrated disease surveillance in the African region: A regional strategy for communicable diseases 1999-2003, Harare, , World Health Organization, Regional Office for Africa.

WHO/AFRO.(2007) Implementation of Indoor Residual Spraying of insecticides for malaria control in the WHO African Region, Vector Biology and Control Unit.

WHO/EMRO.(2003) Use of fish for mosquito control, Cairo. http://applications.emro.who.int/dsaf/dsa205.pdf.

WHO/Regional Office for Africa.(2001) Integrated disease surveillance in the African region: A regional strategy for communicable diseases 1999-2003, Harare, , World Health Organization, Regional Office for Africa.

WHO/Regional Office for South-East Asia.(2003) Regional strategy for integrated disease surveillance, New Delhi, World Health Organization, Regional Office for South-East Asia. ICP CSR 001 43

Williams, H. A. and Jones, C. O. (2004) A critical review of behavioral issues related to malaria control in sub-Saharan Africa: what contributions have social scientists made? *Soc Sci Med,* 59**,** 501-23.

Wilson, I. C. and Metzler, W. H. (1938) *Sickness and Medical Care in an Ozark Area in Arkansas*, University of Arkansas, College of Agriculture, Agricultural Experiment Station.

Winch, P. J., Makemba, A. M., Kamazima, S. R., Lurie, M., Lwihula, G. K., Premji, Z., Minjas, J. N. and Shiff, C. J. (1996) Local terminology for febrile illnesses in Bagamoyo District, Tanzania and its impact on the design of a community-based malaria control programme. *Social Science & Medicine,* 42**,** 1057-1067.

World Bank. (2014). The World Bank [Online]. Place Published. Available: http://www.worldbank.org/ [Accessed 10 December 2012].

World Health Organisation (2007) The quest to be malaria free. *Bulletting of World Health Organisation,* 85**,** 507-508.

World health Organisation.(2012a) Disease surveillance for malaria control: an operational manual, Geneva.

World Health Organisation.(2012b) WHO global health expenditure atlas, Geneva. http://www.who.int/nha/atlas.pdf.

World health Organizastion (1957) Expert Committee on Malaria: sixth report. *World Health Organ Tech Rep Ser,* 38**,** 3-84.

World Health Organizastion.(2006) Bridging the "Know–Do" gap meeting on knowledge translation in Global Health, Geneva, Switzerland, The World Health Organizastion. WHO/EIP/KMS/2006.2 http://www.who.int/kms/WHO_EIP_KMS_2006_2.pdf.

World Health Organizastion.(2007) Malaria elimination: A field manual for low and moderate endemic countries, Geneva.

World health Organizastion. 2008. Global malaria control and elimination: report of a technical review. [Online]. Available: http://whqlibdoc.who.int/publications/2008/9789241596756_eng.pdf [Accessed.

World Health Organizastion.(2009) International Travel and Health, Geneva. http://www.who.int/ith/ITH_2009.pdf.

World Health Organizastion.(2011) Universal access to malaria diagonistic testing: An operational manual, Geneva. 160

World Health Organization.(2000a) Health systems: Improving performance, Geneva, World Health Organization.

World Health Organization.(2000b) Twentieth Report. Expert Committee on Malaria, Geneva, WHO.

World Health Organization.(2007a) Malaria elimination: a field manual for low and moderate endemic countries, Geneva, World Health organization. http://whqlibdoc.who.int/publications/2007/9789241596084_eng.pdf.

World Health Organization.(2007b) Regional workshop on malaria elimination and malaria-free initiatives, Dubai, UAE, World Health organization. http://applications.emro.who.int/docs/who_em_mal_350_e_en.pdf?ua=1.

World Health Organization.(2008) Framework and standards for country Health Information Systems, Switerland, World health organizations. 63 http://www.who.int/healthmetrics/en/.

World Health Organization (2010a) *World Health Organization: Guidelines for the treatment of malaria*.

World Health Organization.(2010b) World Malaria Report: 2010, Geneva, World Health organization. http://www.who.int/malaria/publications/atoz/9789241564106/en/index.html.

World Health Organization.(2011) World Malaria Report: 2011, Geneva, World Health organization. http://www.who.int/malaria/publications/atoz/9789241564106/en/index.html.

World health Organization.(2012a) Disease surveillance for malaria control: an operational manual, Geneva.

World Health Organization.(2012b) World Malaria Report: 2012, Geneva, World Health organization. http://www.who.int/malaria/publications/world_malaria_report_2012/report/en/index.html.

World Health Organization. (2013a). Country profile, Afghanistan [Online]. Place Published: World Health Organization. Available: http://www.who.int/countries/afg/en/ [Accessed 25 February 2013].

World Health Organization.(2013b) World Malaria Report 2013, Geneva, World Health organization. http://www.who.int/malaria/publications/world_malaria_report_2012/report/en/index.html.

World Health Organization.(2014a) From malaria control to malaria elimination: A manual for elimination scenario planning, Geneva, World Health organization. http://apps.who.int/iris/bitstream/10665/112485/1/9789241507028_eng.pdf.

World Health Organization.(2014b) WHO guidance for countries on combining indoor residual spraying and long-lasting insecticidal nets, Geneva. WHO/HTM/GMP/MPAC/2014.2 http://www.who.int/malaria/publications/atoz/who-guidance-combining-irs_llins-mar2014.pdf?ua=1.

World Health Organization.(2014c) WHO policy recommendation on malaria diagnostics in low transmission settings, Geneva. http://www.who.int/malaria/publications/atoz/who-recommendation-diagnostics-low-transmission-settings-mar2014.pdf?ua=1.

World Health Organization.(2014d) World Malaria Report 2014, Geneva, World Health organization. http://www.who.int/malaria/publications/world_malaria_report_2014/report/en/.

World Health Organization and UNICEF.(2012) Countdown to 2015 Maternal Newborn & Child Survival: Building a future for Women and Children, Geneva, UNICEF. http://www.countdown2015mnch.org/reports-and-articles/2012-report.

Worldpop. (2010). The WorldPop project [Online]. Place Published. Available: http://www.worldpop.org.uk/ [Accessed December 01 2013].

Xu, C. (2010). Model Selection with Information Criteria. Doctor of Philosophy, University of Western Ontario.

Yekutiel, P. (1960) Problems of epidemiology in malaria eradication. *Bull World Health Organ,* 22**,** 669-83.

Young, J. T. (2004) Illness behaviour: a selective review and synthesis. *Sociology of Health & Illness,* 26**,** 1-31.

Youssef, R. M., Alegana, V. A., Amran, J., Noor, A. M. and Snow, R. W. (2010) Fever prevalence and management among three rural communities in the North West Zone, Somalia. *East Mediterr Health J,* 16**,** 595-601.

Yukich, J., BriÃ«t, O., Bretscher, M. T., Bennett, A., Lemma, S., Berhane, Y., Eisele, T. P., Keating, J. and 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.

Zakeri, S., Kakar, Q., Ghasemi, F., Raeisi, A., Butt, W., Safi, N., Afsharpad, M., Memon, M. S., Gholizadeh, S., Salehi, M., Atta, H., Zamani, G. and Djadid, N. D. (2010) Detection of mixed Plasmodium falciparum & P. vivax infections by nested-PCR in Pakistan, Iran & Afghanistan. *Indian J Med Res,* 132**,** 31-5.

Zere, E., Kirigia, J. M., Duale, S. and Akazili, J. (2012) Inequities in maternal and child health outcomes and interventions in Ghana. *BMC Public Health,* 12**,** 252.

Zere, E., Mbeeli, T., Shangula, K., Mandlhate, C., Mutirua, K., Tjivambi, B. and Kapenambili, W. (2006) Technical efficiency of district hospitals: evidence from Namibia using data envelopment analysis. *Cost Eff Resour Alloc,* 4**,** 5.

Zhang, H. (2004) Inconsistent Estimation and Asymptotically Equal Interpolations in Model-Based Geostatistics. *Journal of the American Statistical Association,* 99**,** 250-261.

Zhou, F., Guo, H.-C., Ho, Y.-S. and Wu, C.-Z. (2007) Scientometric analysis of geostatistics using multivariate methods. *Scientometrics,* 73**,** 265-279.

Zhou, G., Sirichaisinthop, J., Sattabongkot, J., Jones, J., Bjornstad, O. N., Yan, G. and Cui, L. (2005) Spatio-temporal distribution of Plasmodium falciparum and p. Vivax malaria in Thailand. *Am J Trop Med Hyg,* 72**,** 256-62.

Zikusooka, C. M., Kyomuhang, R., Orem, J. N. and Tumwine, M. (2009) Is health care financing in Uganda equitable? *Afr Health Sci,* 9 Suppl 2**,** S52-8.

Zimmerman, D. (1993) Another look at anisotropy in geostatistics. *Mathematical Geology,* 25**,** 453-470.

Zinszer, K., Jauvin, C., Verma, A., Bedard, L., Allard, R., Schwartzman, K., de Montigny, L., Charland, K. and Buckeridge, D. L. (2010) Residential address errors in public health surveillance data: A description and analysis of the impact on geocoding. *Spatial and Spatio-temporal Epidemiology,* 1**,** 163-168.

Zurovac, D., Talisuna, A. O. and Snow, R. W. (2012) Mobile phone text messaging: tool for malaria control in Africa. *PLoS Med*, 9**,** e1001176.

APPENDICES

APPENDIX: 1 Publications

Author's personal copy

Spatial and Spatio-temporal Epidemiology 7 (2013) 25–36

Contents lists available at ScienceDirect

Spatial and Spatio-temporal Epidemiology

journal homepage: www.elsevier.com/locate/sste

Estimation of malaria incidence in northern Namibia in 2009 using Bayesian conditional-autoregressive spatial–temporal models ☆

Victor A. Alegana [a,b,*], Peter M. Atkinson [b], Jim A. Wright [b], Richard Kamwi [c], Petrina Uusiku [c], Stark Katokele [c], Robert W. Snow [a,d], Abdisalan M. Noor [a,d,*]

[a] Malaria Public Health Department, KEMRI-Wellcome Trust-University of Oxford Collaborative Programme, P.O. Box 43640, 00100 GPO Nairobi, Kenya
[b] Centre for Geographical Health Research, Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK
[c] Directorate of Special Programmes, National Vector-borne Diseases Control Programme, Ministry of health and Social Services, Windhoek, Namibia
[d] Centre for Tropical Medicine, Nuffield Department of Clinical Medicine, University of Oxford, CCVTM, Oxford OX3 7LJ, UK

ARTICLE INFO

Article history:
Received 23 January 2013
Revised 5 August 2013
Accepted 5 September 2013
Available online 17 September 2013

Keywords:
Namibia
Malaria
Spatio-temporal
Conditional-autoregressive

ABSTRACT

As malaria transmission declines, it becomes increasingly important to monitor changes in malaria incidence rather than prevalence. Here, a spatio-temporal model was used to identify constituencies with high malaria incidence to guide malaria control. Malaria cases were assembled across all age groups along with several environmental covariates. A Bayesian conditional-autoregressive model was used to model the spatial and temporal variation of incidence after adjusting for test positivity rates and health facility utilisation. Of the 144,744 malaria cases recorded in Namibia in 2009, 134,851 were suspected and 9893 were parasitologically confirmed. The mean annual incidence based on the Bayesian model predictions was 13 cases per 1000 population with the highest incidence predicted for constituencies bordering Angola and Zambia. The smoothed maps of incidence highlight trends in disease incidence. For Namibia, the 2009 maps provide a baseline for monitoring the targets of pre-elimination.

© 2013 The Authors. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Maps of malaria transmission intensity are increasingly being used for planning, monitoring and evaluation, and resource allocation (Hay et al., 2009; Noor et al., 2010; Omumbo et al., 2013). In countries where malaria elimination is feasible, the World Health Organisation (WHO) proposes a transition from measuring risk by malaria

Abbreviations: ACD, active case detection; CAR, conditional auto-regressive; CPO, conditional predictive ordinate; DIC, deviance information criterion; ESRI, Environmental System Research Institute; EVI, enhanced vegetation index; GF, Gaussian field; GIS, geographic information system; GMRF, Gaussian markov random field; GPS, global positioning system; GRUMP, Global Rural and Urban Mapping Project; HMIS, Health Management Information System; INLA, Integrated Nested Laplace Approximation; JAXA, Japan Aerospace Exploration Agency; MAUP, Modifiable Areal Unit Problem; MCMC, Markov Chain Monte Carlo; MODIS, MODerate-resolution Imaging Spectro-radiometer; MoHSS, Ministry of Health and Social Services; NASA, National Aeronautics and Space Administration; NVBDCP, National Vector-Borne and Disease Control Programme; PCD, passive case detection; PHS, public health sector; RDT, Rapid Diagnostic Test; SPA, Service Provision Assessments; TRMM, Tropical Rainfall Measuring Mission; TSI, temperature suitability index; WHO, World Health Organisation; ZIP, Zero-Inflated Poisson.

☆ This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

* Corresponding author. Address: Malaria Public Health Department, KEMRI-Wellcome Trust-University of Oxford Collaborative Programme, P.O. Box 43640, 00100 GPO Nairobi, Kenya. Tel.: +254 727326587.
E-mail addresses: vaa1g10@soton.ac.uk, valegana@kemri-wellcome.org (V.A. Alegana).

1877-5845/$ - see front matter © 2013 The Authors. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.sste.2013.09.001

Author's personal copy

prevalence surveys to surveillance through a combination of routine health management information systems (HMIS) and active case detection (World Health Organization, 2007). The year 2009 has a special significance for the fight against malaria in Namibia. This is when the Elimination Eight (E8) initiative was launched, under which eight southern African countries decided to collaborate to eliminate malaria in Namibia, Botswana, South Africa and Swaziland. Under this initiative, Namibia formally declared the ambition to eliminate malaria by 2020 (Noor et al., 2013a,b; Southern Africa Roll Back Malaria Network (Sarn), 2010). These ambitions were motivated by reported substantial declines in malaria burden in the four eliminating countries and by the 2008 global call for malaria elimination (World health Organizastion, 2008). A Namibian malaria indicator survey (MIS) conducted in 2009 showed a mean community *Plasmodium falciparum* prevalence of approximately 3% nationally (Ministry of Health and Social Services, 2010b). This is a threshold at which countries are advised to use case incidence data for measuring malaria risk (Hay et al., 2008; Yekutiel, 1960). In 2010, Namibia launched a national malaria strategy for the period 2010–2016 (Ministry of Health and Social Services, 2010c). The aim was to reduce malaria case incidence to 10 persons per 1000 population by 2013 and to move the country to pre-elimination status by 2016 where case incidence will be less than 1 person per 1000 population (Ministry of Health and Social Services, 2010c, d).

Most malaria eliminating countries in Africa, including Namibia, are yet to adopt active case-detection (ACD) systems (World Health Organization, 2012) and the main source of data for measuring disease incidence is from passive case detection (PCD), assembled through the public health sector (PHS). Such data, however, have deficiencies that limit their use for estimating overall case incidence accurately. A substantial proportion of malaria cases are treated outside of the PHS (Cibulskis et al., 2011; Cibulskis et al., 2007), while only a proportion of health facilities in the PHS submit returns and even fewer report every month of the year, making the data incomplete spatially and temporally (Gething et al., 2008; Gething et al., 2006; Murray et al., 2004; Stansfield, 2005). Third, only a subset of reported cases is diagnosed parasitologically and most of these cases are fevers that have been diagnosed presumptively as malaria (Cibulskis et al., 2011; Cibulskis et al., 2007). The use of such data therefore requires approaches that adjust for the non-utilisation of the PHS, incomplete data reporting which underestimate burden and the presumptive diagnosis which inflate incidence (Alegana et al., 2012; Cibulskis et al., 2011). In addition, these approaches must harness the spatial and temporal autocorrelation of the available data to predict at locations and periods where data are missing as well as estimate robustly the uncertainties of these predictions (Loha and Lindtjorn, 2010; Reid et al., 2012).

Bayesian hierarchical conditional auto-regressive (CAR) models can improve the quality of HMIS data at a national level, where routine surveillance is inefficient, by representing risk via a set of environmental or ecological factors and random effects using CAR priors (Barnerjee et al., 2004; Gelfand and Vounatsou, 2003; Gething et al., 2006). Examples of such approaches have been used previously in modelling spatial–temporal variation of disease risk in Yunnan province in China (Clements et al., 2009) and in identifying social and ecological factors driving malaria risk in Vietnam (Manh et al., 2011). These methods handle uncertainty in a coherent manner, are able to predict risk in areas where data are not recorded while at the same time smoothing variability where the denominator (population) is small (Gelfand and Vounatsou, 2003; Reid et al., 2012). These approaches are used in this study with the primary aim of predicting malaria incidence at second administrative unit level (constituencies) in northern Namibia where malaria is considered endemic (Ministry of Health and Social Services, 2010c). In addition, a novel approach is used to adjust PHS utilisation rates to estimate catchment population. Secondary aims of this study were to calculate populations at risk to determine areas where interventions can be targeted to provide universal coverage and to evaluate the use of environmental factors such as rainfall and vegetation indices in predicting incidence.

2. Methods

2.1. Study area

Namibia is divided into 13 regions (administrative level 1) and 108 constituencies (Ministry of Health and Social Services, 2010c; Zere et al., 2006) (Fig. 1). The country is largely dry and sparsely populated with an estimated 2.1 million people in 2009 living in an area of approximately 0.83 million km^2 (National Planning Commission, 2012). The risk of malaria is constrained by aridity (Ministry of Health and Social Services, 2010c; Snow et al., 2010) with the larger and sparsely populated south made up of four regions, Karas, Hardap, Khomas and Erongo, considered either malaria-free or supporting high focal very low transmission intensity (Ministry of Health and Social Services, 1995, 2010c). The majority of the population resides in the other nine northern regions of the country that are also considered to contribute almost the entire malaria burden in Namibia (Ministry of Health and Social Services, 2010b, c, e). In this study, analysis of malaria incidence was restricted to the 78 constituencies in the nine northern regions (Fig. 1).

2.2. Assembly of malaria case data

Monthly data (January to December) for 2009 on confirmed and suspected (clinically diagnosed) cases of malaria among patients of all ages were obtained from the Ministry of Health and Social Services (MoHSS) after a national Service Provision Assessment (SPA) survey was conducted (Ministry of Health and Social Services (MoHSS) and Icf Macro, 2010). The health facility survey covered 273 facilities in the north comprising of hospitals, health centres, clinics and sick bays that are managed by the Ministry of Health and Social Services (MoHSS), missions, Non-Governmental Organisations (NGOs), the private sector and Ministry of Defence (MoD) and police. Of these, only

Author's personal copy

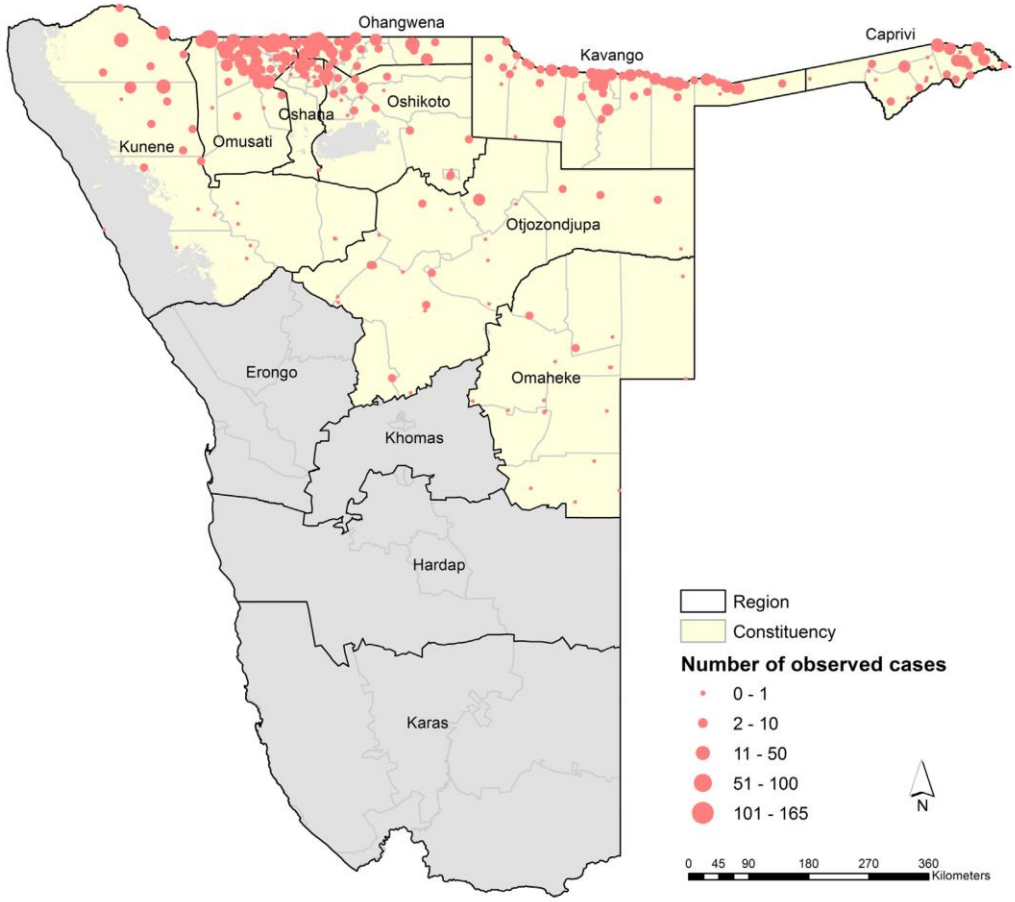

Fig. 1. Map showing the number of cases observed at a public health facility superimposed on the 78 constituency boundaries (Administrative level 2) in the northern regions (Administrative level 1) of Namibia in 2009. The four southern regions namely Erongo, Khomas, Hadarp and Karas are considered as '*malaria free*' while the grey areas in the north correspond to desert arid areas where the MODIS-derived enhanced vegetation index (EVI) was <0.1 and were, thus, considered unsuitable for transmission and masked out (Scharlemann et al., 2008). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

13 were private health facilities, all located in urban centers. Three constituencies had no facility data and were treated as missing data. During the survey, a health system questionnaire was used to collect data on suspected and confirmed malaria cases for a 12-month period from patient registers. Each facility was also geo-located using a handheld global positioning system (GPS) device. Rapid Diagnostic Tests (RDTs) were used to examine blood samples from most patients at primary health facilities although a few, mostly at tertiary facilities, were examined using microscopy (Ministry of Health and Social Services, 2010a).

2.3. Assembling data on environmental predictors of incidence

The incidence of malaria is usually a function of its underlying transmission intensity (Patil et al., 2009) which in turn is driven by factors such as rainfall, temperature and human habitation that influence the development and survival of the malaria parasite and vector (Molineaux, 1988). The annual mean enhanced vegetation index (EVI) for 2009 derived from MODerate-resolution Imaging Spectroradiometer (MODIS) sensor imagery was used as a measure of vegetation cover (Scharlemann et al., 2008). Monthly 2009 precipitation data were obtained from the Tropical Rainfall Measuring Mission (TRMM 3B43) [http://trmm.gsfc.nasa.gov/], a joint collaboration between NASA and the Japan Aerospace Exploration Agency (JAXA) (Huffman and Bolvin, 2011; NASA, 2011). TRMM 3B43 [http://trmm.gsfc.nasa.gov/] is a gridded mean monthly precipitation product in mm h^{-1} at 0.25° × 0.25° spatial resolution (Huffman, 1997). It is produced after TRMM multi-satellite precipitation analysis (TMPA) (Huffman and Bolvin, 2011) that combines both satellite sensor data

Author's personal copy

and observations from at least 6700 rain gauges from global reports and country-specific reports. A 1 km × 1 km surface depicting a temperature suitability index (TSI) for malaria transmission (Gething et al., 2011) ranging from 0 (not suitable) to 1 (most suitable) was also obtained from the Malaria Atlas Project [http://www.map.ox.ac.uk]. The annual mean values of EVI, precipitation and TSI were computed for each constituency. Finally, the proportion of urban population within each constituency was extracted based on urban extent from the Global Rural Urban Mapping Project (GRUMP) (Balk et al., 2004, Center for International Earth Science Information Network (CIESIN), 2004) overlaid on a 100 m × 100 m resolution population surface developed by Afripop (Balk et al., 2004, Center for International Earth Science Information Network (CIESIN), 2004) and available at [http://www.afripop.org/]. The assembled covariates were re-sampled to 1 × 1 km spatial resolution and a value extracted for each facility in ArcGIS 10 (ESRI, Redlands, CA, USA).

2.4. Analysis

2.4.1. Adjusting observed malaria cases based on test positivity rates and PHS utilisation

The calculation of malaria incidence requires accurate estimates of both the number of parasitologically confirmed positive cases and the size of the population from which the cases originate. The malaria cases were computed as the sum of the parasitologically diagnosed cases presented at public sector health facilities and the suspected (clinical) cases adjusted using the *P. falciparum* positivity (microscopy or RDT) rate per facility. The focus was on the public sector, which constituted majority of surveyed health facilities (96%) and is mainly sponsored by the government and public resources. To define the catchment population two factors were considered: (a) only a subset of the population was likely to use the public health sector and; (b) these would vary geographically within a catchment area and by constituency. In Namibia, the MIS of 2009 recorded treatment seeking behaviour for fevers and showed that only 52% of all individuals who had a fever in the last 2 weeks sought treatment in the public health sector and the utilisation rate varied by region (Ministry of Health and Social Services, 2010b). To define public health facility catchment populations empirically, the treatment seeking data from the MIS were used subsequently to develop a utilisation model that defined, at every 1 km × 1 km grid cell, the probability that a febrile individual will use a public health facility using a three-parameter logistic regression model (Alegana et al., 2012). These probabilities were applied to a population surface of similar resolution (Afripop, 2010) to estimate the 2009 population seeking treatment for fever at public health sector facilities. The adjusted population counts were then used in modelling incidence.

2.4.2. Preliminary analysis of environmental covariates

A non-spatial Poisson regression model was used to test the univariate and multivariate associations of assembled environmental covariates and crude incidence in R version 2.15.2 [http://www.r-project.org/]. The environmental covariates were used in the continuous form in a generalized linear regression model with the response variable being the observed crude incidence rates assuming that the expected cases have a Poisson distribution; $Y_{ij} \sim$ Poisson μ_{ij} for the ith observation in facility j. Wald's P-values and goodness-of-fit statistics with associated confidence intervals were assessed. Variables significant at a P-value of <0.05 were selected for inclusion into the predictive model.

2.4.3. Bayesian space–time zero-inflated CAR model for malaria incidence

Environmental covariates selected via the preliminary analysis, the reported cases and catchment population per public sector health facility were used in a Bayesian spatio-temporal zero-inflated conditional autoregressive (CAR) model using Integrated Nested Laplace Approximation (INLA) (Martins et al., 2013; Rue et al., 2009) to predict incidence at the constituency level. A Zero-Inflated Poisson (ZIP) model was used following the example of studies in low transmission settings, to handle count data with a lot of structural or excess zeros (Lambert, 1992; Manh et al., 2011). In Namibia, no malaria cases were reported in 65.3% of the facility level monthly returns, with 43% of facilities reporting no cases in March and over 60% from May to December. The ZIP models have also been applied previously in mapping the malaria vector sporozoite rate (Amek et al., 2011; Nobre et al., 2005) as well as in schistosomiasis (Vounatsou et al., 2009), but with inference made using the Markov Chain Monte Carlo (MCMC) approach. In this study, however, inference was made using INLA via the Gaussian Markov Random Field (GMRF) (Rue and Martino, 2007) that reduces computation time significantly (Kneib et al., 2010; Rue et al., 2009; Rue and Martino, 2007). In addition, a facility random effect model was fitted to allow for variation between two or more facilities in the same constituency. For the ZIP model, the probability of observing zero (Böhning, 1998; Ghosh et al., 2004; Neelon et al., 2010) is;

$$P_{ij}(y_i = 0) = (1 - P_{ij}) + P_{ij}e^{-\mu_{ij}} \quad 0 \le p \le 1$$

$$P_{ij}(y_i = k) = P_{ij}\frac{\mu_{ij}^k e^{-\mu_{ij}}}{k!} \quad k = 1, \ldots, \infty$$

The term $(1-P_{ij})$ in the first part represents the probability of observing a true zero and, therefore, when $P = 1$ the equation reduces to a general Poisson model and zero is inflated when $P < 1$. Covariates were introduced via a log linear model for μ_{ij} while maps of predicted monthly and annual incidence were produced at constituency level. In the model, the observed variables y_i, $i = 1, \ldots, n$ and the linear predictor η_i were modelled with additive effects as (Rue et al., 2009; Rue and Martino, 2007; Schrödle and Held, 2010):

$$\eta_i = \alpha + \sum_{j=1}^{nf} f^{(j)}(u_{ji}) + \sum_{k=1}^{n\beta} \beta_k z_{ki} + \varepsilon_i$$

where, $f^{(j)}$ is a linear function on some variables u, β_k are the coefficients for the covariates Z and ε represents the

Author's personal copy

unstructured effects. Rue and Martino (2007) show that the posterior marginal can be estimated as:

$$\tilde{\pi}(x_i|y) = \sum_k \tilde{\pi}(x_i|\theta_k, y)\tilde{\pi}(\theta_k|y)\Delta_k$$

with the sum evaluated using appropriate weights Δ_k solved at suitable reference points θ_k (Rue and Martino, 2007; Schrödle and Held, 2010). The posterior marginal $\pi(\theta|y)$ of the hyper-parameters are evaluated as:

$$\tilde{\pi}(\theta|y)\alpha\frac{\pi(x,\theta,y)}{\tilde{\pi}G(x|\theta,y)}|x = x^*(\theta)$$

with the denominator as a Gaussian approximation of $\pi(x|\theta,y)$ and $x^*(\theta)$ being the mode of the full conditional $\pi(x|\theta,y)$ (Schrödle and Held, 2010). The final log-relative risk model was represented as:

$$\eta_i = \log(E_i) + \mu + Z_{ij}^T\beta_{ij} + f(s_u) + \psi_i + f(t)$$

with the E_i being the expected number of confirmed and suspected cases adjusted for slide positivity at each facility i, the term μ represents the intercept, with the $f(s_u)$ and $f(t)$ terms representing the spatially unstructured effects and seasonal effects, respectively. The conditional autoregressive prior ψ_i was included to account for the assumption that neighboring polygons have similar incidence (Barnerjee et al., 2004). This specification ensures a smoothed map of risk with geographically reliable estimates (Bernardinelli et al., 1997; Kleinschmidt et al., 2002). Full Bayesian specifications were completed by specifying priors for the fixed effects and random components. The conditional prior for neighbouring regions $(\phi_j, j\neq i)$ was specified following Bernardinelli et al. (1997) as $(\phi_i \sim N(\mu_{\phi i}, \sigma^2_{\phi i})$ where $\mu_{\phi i} = \Sigma_{j\neq i}W_{ij}\phi_j/\Sigma_{j\neq i}W_{ij})$; $\sigma^2_{\phi i} = 1/\gamma_\phi\Sigma_{j\neq i}W_{ij})$. The W_{ij} is the adjacency matrix of weights assigned as $W_{ij} = 1$ for two neighbouring regions or $W_{ij} = 0$ otherwise. A full treatment on CAR modelling theory can be found elsewhere (Barnerjee et al., 2004; Gelfand and Vounatsou, 2003). The random effects component was specified as a set of vague normal priors.

Two CAR models were fitted: Model 1 included a spatio-temporal component but excluded environmental covariates such as vegetation indices, whilst Model 2 included these environmental covariates in addition to spatio-temporal structure.

2.4.4. Computing the cross-validation statistics and proper scoring rules

The performance of both CAR models was compared using the deviance information criterion (DIC) (Spiegelhalter et al., 2002). Predictive model assessment was conducted using the probability integral transform (PIT) and the conditional predictive ordinate (CPO), a leave-one-out cross-validation approach in which a prediction is validated based on the fitted model and the remaining data only (Czado et al., 2009; Spiegelhalter et al., 2002). The CPO, defined as the probability of observing a value given all other data, was examined for all observations in a full Laplace model (Martins et al., 2013). Both these measures assess the calibration (statistical consistency) and sharpness (concentration) of the predictive model. The predictive measures of fit have been shown to fail if the approximation of the latent Gaussian Field (GF) is not accurate (Czado et al., 2009). Model scoring rules such as the square error score (SES) and the ranked probability score (Gneiting and Raftery, 2007) as well as Pearson correlation of observed and predicted incidence were computed. The latter was based on 26 health facilities selected randomly as validation set. Proper Bayesian scoring rules are discussed by Gneiting and Raftery (2007) and implemented using a predictive distribution (Supplementary information). For example, the RPS generalizes the absolute error and is minimum for true predictions (Czado et al., 2009).

2.5. Population at risk

To estimate the population at risk at varying levels of malaria incidence, the total population resident in constituencies living in the six predicted endemicity classes of: less than 1; 1–5; 5–10; 10–15; 15–20 and greater than 20 cases per 1000 population was calculated. The population surface was obtained from Afripop (Afripop, 2010) which had been developed from a combination of census, population settlements and land cover by disaggregation of census data to improve their spatial resolution (Linard et al., 2010; Linard et al., 2012). The population surface had also been used in mapping health facility catchment population in Namibia (Alegana et al., 2012). The original population surface, produced for 2010 from Afripop, was back-projected to 2009 using the United Nations' inter-censual growth rates (http://esa.un.org/unup/) and categorized according to estimated risk in northern Namibia.

3. Results

3.1. Malaria incidence and facility attendance characteristics

A summary of the assembled malaria incidence data and modeled estimates of public health sector utilization are shown by health district in Table 1. Overall, only 17 PHS facilities had no malaria reports in 2009 and the remaining health outlets returned complete reports every month. Most health facilities were located in Caprivi, Kavango, Ohangwena, Oshana and Omusati regions where population density is greatest (Table 1 and Fig. 1). The spatial distribution of reported cases, including suspected cases adjusted for test positivity rates, is shown in Fig. 1 and indicates higher caseloads in the northern regions. In total, 134,851 cases were clinically diagnosed while 90,835 individuals were examined for malaria parasites of which, 9893 were positive. The mean test positivity rate was 11.2 [95% CI 6.7–15.7] (Table 1). Crude annual incidence based on the parasitological and clinically diagnosed cases, the latter corrected for slide positivity rate, was 16 cases per 1000 population. This was highest in the first 4 months of the year and peaked in March (Fig. 4). The highest crude incidence was in constituencies in Caprivi, Ohangwena and Kunene that border Angola where test positivity rates were also highest (Table 1).

Author's personal copy

Table 1
Summary of malaria incidence and modeled facility attendance by administrative region and health district in northern Namibia.

Region	Health district	Number of health facilities (number with missing data)	Number of constituencies	Confirmed malaria cases	Suspected malaria cases	Mean slide positivity rate (95% CI)	Population 2009	Percent of population attending a PHF[a] modelled
Caprivi	Katima	27(2)	6	954	10,605	21.1 (17.9–24.3)	87088	68
Kavango	Andara	10(0)	1	309	4293	9.2 (7.0–11.3)	26,677	71.1
	Nankudu	11(1)	2	244	7662	8.4 (6.0–10.8)	48,715	64.2
	Nyangana	8(0)	1	665	3063	25 (20.1–29.9)	19,815	71.9
	Rundu	23(1)	5	1176	34,608	16.4 (13.4–19.4)	119,855	71.1
Kunene	Khorixas	8(0)	1	1	89	2.7 (-0.5–6.1)	12,469	61.4
	Opuwo	14(0)	3	539	856	47.3 (40.7–53.8)	52,485	52.5
	Outjo	4(0)	2	1	53	1.1 (-0.2–2.5)	20,395	53.4
Ohangwena	Eenhana	10(1)	4	379	3956	7.1 (4.8–9.4)	80,419	68.2
	Engela	16(0)	6	916	13,774	9.8 (7.8–11.9)	131,744	74.2
	Kongo	4(1)	1	529	1788	24.3 (15.8–32.8)	24,744	61.5
Omaheke	Gobabis	14(2)	7	11	96	13.8 (9.5–18.1)	68,433	62.1
Omusati	Okahao	9(1)	2	384	9066	4.1 (2.1–6.1)	29,964	73.6
	Oshikuku	19(0)	5	436	10,315	3.6 (2.6–4.7)	101,587	75.2
	Outapi	10(0)	2	1,970	9846	8.4 (6.6–10.2)	48,812	70.8
	Tsandi	10(1)	3	617	5339	8.4 (6.2–10.6)	54,418	70.1
Oshana[b]	Oshakati[b]	19(4)	10	353	9133	3.1 (1.9–4.3)	169,053	75.4
Oshikoto	Onandjokwe	16(0)	8	266	8516	2.3 (1.6–3.1)	146,436	69.8
	Tsumeb	5(1)	2	28	628	5.1 (1.8–8.4)	29,094	67.4
Otjozondjupa	Grootfontein	6(0)	2	59	547	13.7 (7.1–20.3)	33,347	61.3
	Okahandja	2(1)	2	3	110	5.2 (0.2–10.1)	40,209	64.2
	Okakarara	5(0)	1	17	189	11.6 (5.0–18.2)	21,748	56.6
	Otjiwarongo	10(1)	2	36	319	5.4 (2.6–8.1)	42,336	67.3
Total		260(17)	78	9,893	134,851	11.2 (6.7–15.7)	1,409,841	65.3[c]

[a] PHF is an abbreviation for 'Public Health Facility', which in this case does not include private facilities or privates for profit.
[b] Two constituencies in Oshana region (Okatyali and Ompundja) did not have any health facilities, thus, the polygons where treated as missing data.
[c] Public health facility attendance for treatment of fever based on probability of attendance and the distance decay effect. Description outlined in Alegana et al. (2012).

Author's personal copy

3.2. Preliminary model involving environmental covariates

Of the selected environmental variables, univariate non-spatial regression analysis showed that the EVI (coefficient of regression, 95% CI: 6.55, 4.25–8.87, $p < 0.001$), TSI (7.57, 5.34–9.96, $p < 0.001$) and precipitation (0.02, 0.01–0.03, $p = 0.002$) were significant predictors of crude incidence. In addition, the percentage of urban resident population produced a negative and significant association with incidence (−0.01, −0.01 to −0.00, $p < 0.001$). In the multivariate model, that included all four covariates, only EVI (14.29, 9.24–19.42, $p < 0.001$) was positively associated with crude incidence and was included in the final model. The number of environmental covariates was minimized in the final model to achieve a parsimonious space–time model and due to the observed large correlation between some covariates, for example altitude and temperature or vegetation indices and rainfall (Craig et al., 2007; Pascutto et al., 2000).

3.3. CAR model predictions of monthly and annual incidence for 2009

Two spatio-temporal models of incidence were implemented. Model 2 included EVI while Model 1 excluded the covariate information. Table 2 lists Bayesian model parameters for the two CAR models with and without environmental covariate. Overall, Bayesian model parameters for seasonal random effects (2.02 with Crl 0.16–5.79), facility random effects (6.95, Crl 2.65–13.22) and unstructured random effects (0.20, Crl 0.02–0.57) were all significant at 95% Crl (Bayesian credible interval). There were also marginal differences in the overall mean: −1.80 Crl (−1.98 to −1.64) and −1.76 Crl (−1.93 to −1.58) for model with and without covariate information respectively.

Table 3 compares these two models based on the DIC, which represents a trade-off between model complexity and goodness-of-fit, and SES. The EVI improved the model fit marginally, as indicated by the lower DIC for Model 2 in Table 3. The SES for M2 (1.61) was lower than that for M1 (1.70) suggesting a better predictive performance for M2 although only marginally. The conditional predictive ordinate (CPO), a cross-validation logarithmic score, was also calculated for each prediction. For both models the CPO score was 0.22 (Table 3) and since a smaller CPO value usually indicates greater predictive accuracy (Schrödle and Held, 2010), this also suggests a small difference between the two fitted models. However, in view of its lower DIC, Model 2 (with EVI) is used as the basis for presenting subsequent model outputs. The Pearson correlation coefficient for this model based on a hold out set was 0.56.

Overall malaria incidence peaked in the months of March and April and was highest in Kunene, Kavango, Caprivi and in a few constituencies in Ohangwena region as shown in Fig. 2, based on Model 2. Fig. 3 shows a map of mean annual incidence based on this model choice. The predicted mean annual incidence of the Bayesian CAR model was 13 cases per 1000 population in the 78 constituencies in northern Namibia. The highest predictions were between 15 and 20 cases per 1000 population (Fig. 4).

3.4. Population at risk

Based on Model 2, 383,632 people (27.2% of the population) lived in areas where case incidence was greater than 15 cases per 1000 population; slightly more than half 745,903 (52.9%) lived in areas where case incidence was between 10 and 15 cases per 1000 population; approximately 216,512 (15.4%) resided in regions with an average of 5–10 cases per 1000 population; 49,005 (3.5%) in areas with greater than 1 case, but less than 5 cases per 1000 population and 1% of population lived in regions with less than 1 case per 1000 population. Population density was highest in the northern border constituencies.

Table 2
Parameters for two Bayesian zero-inflated CAR models of malaria incidence in northern Namibia on a log scale.

Parameter	Model 1 Without covariates: posterior mean, median, (95% Crl[1])	Model 2 With environmental covariate: posterior mean, median, (95% Crl[1])
μ (Intercept)	−1.763, −1.760 (−1.932 to −1.581)	−1.803,−1.800 (−1.980 to −1.639)
Enhanced vegetation index (EVI)	–	0.093, 0.093 (−0.028–0.211)
ϑ (parameter for Zero-inflation)	0.843, 0.843 (0.833–0.856)	0.843, 0.843 (0.833–0.854)
τ_m (seasonal random effect)	1.546, 1.023 (0.137–4.692)	2.015, 1.427 (0.161–5.789)
τ_f (facility random effect)	6.912, 5.836 (2.605–14.830)	6.952, 6.388 (2.641–13.220)
Υ (unstructural random effect)	0.190, 0.136 (0.020–0.542)	0.200, 0.144 (0.019–0.568)
ϕ (structural random effect)	0.081, 0.045 (0.003–0.278)	0.080, 0.004 (0.030–0.276)

[1] Crl is abbreviation for Bayesian credible interval.

Table 3
Posterior mean deviance, the number of effective parameters, the DIC and CPO score for each implemented model.

Model	Mean deviance	Number of effective parameters	DIC	CPO	SES
Model 1 (without covariate)	3113.22	9.79	3123.89	0.229	1.704
Model 2 (with covariate)	3112.08	10.68	3123.75	0.229	1.609

Author's personal copy

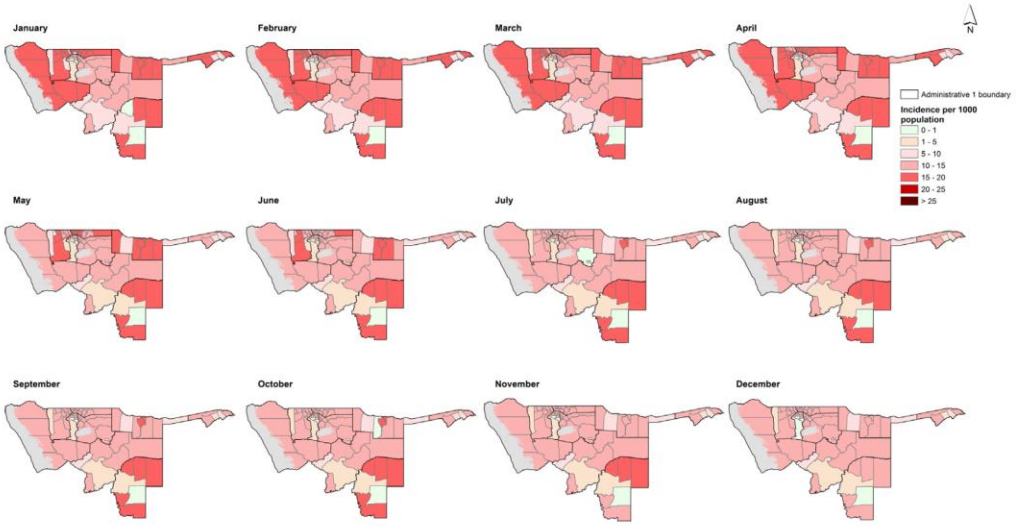

Fig. 2. Map showing the predicted monthly malaria incidence per 1000 population at constituency level for regions in the north of Namibia in 2009 using Bayesian CAR with environmental covariates (Model 2).

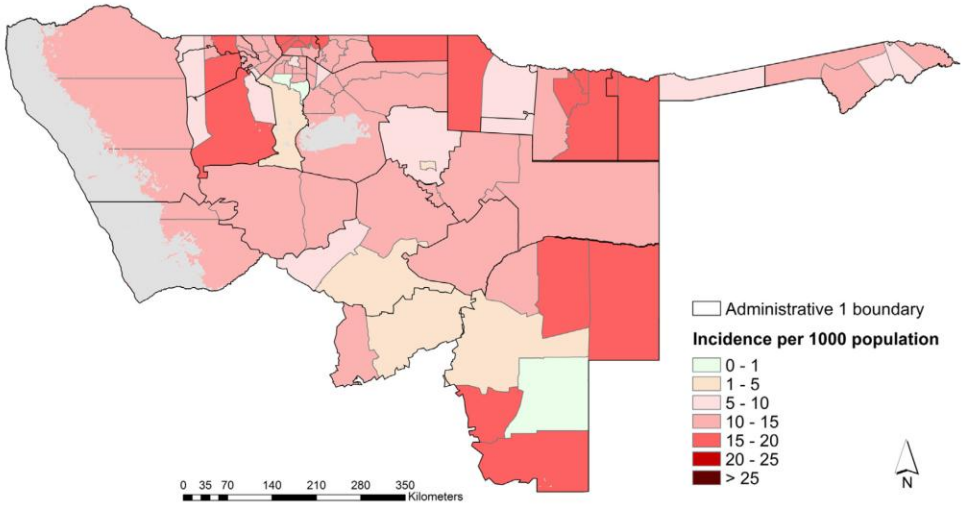

Fig. 3. Map showing the mean annual incidence prediction based on Bayesian CAR with environmental covariates (Model 2).

4. Discussion

The evaluation of pre-elimination status requires a detailed description of local epidemiology of malaria transmission patterns. From the predicted monthly maps of Namibia (Fig. 2 and Fig. 4), a higher incidence of malaria was observed between January and April in the constituencies bordering Angola and Zambia, while, lower values were observed for the July and December period. The overall mean incidence was 13 cases per 1000 population for 2009 (Fig. 3). The model included the unstructured random component to explain unobserved effects and the inclusion of the structural effects via the GMRF introduced dependence resulting in spatial and temporal smoothing of seasonal variation (Banerjee and Carlin, 2003; Rue and Held, 2005). The Bayesian CAR approach has the advantage of addressing several sources of uncertainty. The model was applied at facility level and, therefore, the method not only

Author's personal copy

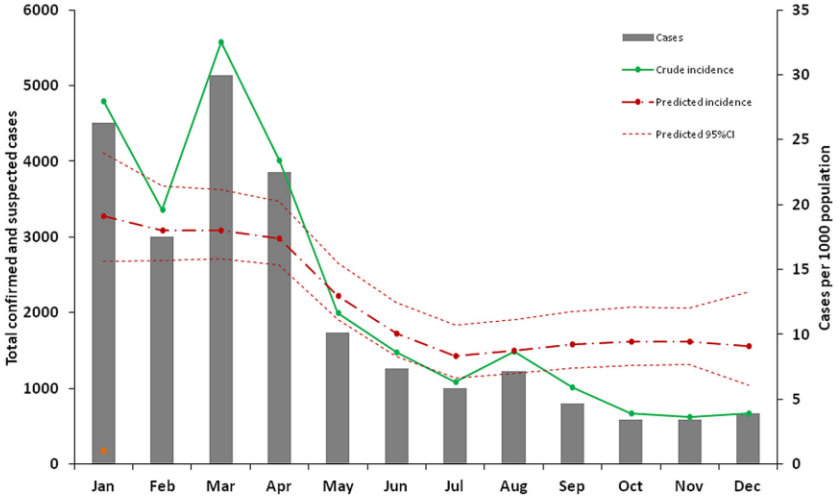

Fig. 4. Plot of the reported cases by month in northern Namibia in 2009 (vertical dark grey bar), the calculated crude incidence (green line) derived from combined confirmed and suspected cases and the predicted incidence per 1000 population (dashed-dotted red line) with 95% CrI upper and lower limits. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

takes into account the nature of the facility, but also season and environmental factors in adjusting for under-reporting. In addition, the CAR model smoothed incidence, thereby addressing the potential impact of model instability resulting from small numbers of reported cases, apparent in the facility data presented in Fig. 1. Smoothing incidence also reduces the potential impacts of under-reporting of cases by facilities. Secondly, incorporating the environmental covariate explained spatial variation where data were absent in addition to providing information on the climatic suitability of malaria transmission, for example, in Omaheke region (Craig et al., 1999; Guerra et al., 2008). This suggested that the inclusion of environmental covariates improved the model estimates for a few constituencies (in Kunene and Omaheke), but only marginally.

The mean incidence observed for 2009 was highest in constituencies in Omusati, Kavango and Omaheke region bordering Angola and Botswana. Historical *P. falciparum* data for Namibia between 1969 and 1992 (Noor et al., 2013a,b) suggest a parasite prevalence of greater than 5% in Kavango and other northern regions along the border with Angola. In addition, Craig and others showed that in Botswana, the area along the north-western border areas with Namibia had relatively high prevalence (Craig et al., 2007). For these border constituencies concerted efforts with neighbouring countries have to be put in place to realize the pre-elimination targets (Noor et al., 2013b,c). Incidence in these regions could well be driven by cross border population movement (Cosner et al., 2009). Similar suggestions were made for two districts in South Africa close to the Mozambique border (Kleinschmidt et al., 2002) and in Yunnan province in China that borders Myanmar, Laos and Vietnam (Clements et al., 2009).

The approach presented here drew upon a comparatively data rich setting and the facility census used may not be available in many countries. The recent improvements in case management in Namibia in which all suspected malaria fevers are diagnosed parasitologically before treatment will reduce the need for adjustment for test positivity rate. In addition, planned improvements in HMIS reporting and quality and transition to active case detection mean smaller adjustments for treatment seeking and reporting will be required in future. This is may also be useful for external validation, with additional resources, of approaches used in this study. These factors will, therefore, contribute to the precision of routine malaria case data in estimating disease burden in the future. More precise incidence estimates should provide a basis for targeting active case detection efforts at specific locations and in specific months, potentially making such resource-intensive efforts more cost-effective. A comparison of our predictions with the standard WHO approach shows that the latter estimates a higher annual malaria incidence of 23 per 1000 population in 2009 in Namibia and generally followed a pattern close to that of the crude incidence (Supplementary information). The WHO approach is described in detail elsewhere (Cibulskis et al., 2011). The main difference in our approach is the use of health facility as a random effect in the model and the utilization of the spatial and temporal autocorrelation in the data resulting in smoothing of the predictions.

Bayesian hierarchical models are often implemented using numerical statistical methods such as Markov Chain Monte Carlo (MCMC) and Laplace transformation amongst others (Cressie and Wikle, 2011) (p. 238). When large data set are involved, MCMC computation can be demanding and the Gaussian Markov Random Field (GMRF) (Rue and

Author's personal copy

Martino, 2007) offer an alternative approach due to the sparseness of resulting covariance matrixes. Thus, they are computationally faster and with desirable Markov properties (Kneib et al., 2010; Rue et al., 2009; Rue and Martino, 2007). GMRF are implementable in INLA (Martins et al., 2013), although, the result are more accurate if the number of hyperparameters in model implemented is small typically less than 12 (Kneib et al., 2010; Rue et al., 2009; Rue and Martino, 2007).

One drawback of many studies analyzing areal data, and one common to the Bayesian approach used in this study, is the modified areal unit problem (MAUP), a well-known analytical problem in geography that could affect the observed statistical results with a change in shape or size of spatial polygons used in the analysis (Barnerjee et al., 2004; Robinson, 1950; Wakefield, 2003). In this study constituencies were selected as the basis for presenting predictions, with the aim of providing information at this level to health authorities, though the model was fitted at facility level. There is therefore a potential impact of MAUP both in terms of the shape of constituencies and in predicting at constituency level from facility level data. Secondly, the data used for this study were obtained from the Namibia HMIS which covers the majority of public health facilities in the north. This means that the findings are relevant only for the 12-month time-series in 2009. The results could be improved by inclusion of more data and at different time points to draw more stable long-term spatio-temporal patterns (Zhou et al., 2005). In addition, the modelling approach excluded the effects of population movements between regions, especially across borders, while the relations between the environmental variables could change across space and at shorter time periods than those considered (Hay et al., 2008). Finally, some sources of uncertainty remain. In particular, the underlying the care-seeking behaviour data used to adjust denominator populations relate to children under 5 years, not the whole population. Utilisation rates were estimated from cross-sectional surveys and therefore may not capture temporal changes in care-seeking behaviour. The underlying utilisation data also relate to fever rather than malaria *per se*.

5. Conclusion

Although Namibia faces a significant malaria case incidence in the border regions, the results of this analysis suggest that the country may be within the pre-elimination targets in most parts of the northern region. The NVDCP has initiated a process of creating a malaria-free buffer extending approximately 25 km across the border with Angola as well as with Zambia and Botswana (Ministry of Health and Social Services, 2010c; Noor et al., 2013b, Trans-Zambezi Malaria Initiative (Tzmi), 2012). This study provides additional information to identify the highest malaria risk areas in Namibia and when used together with evidence from modelled community parasite prevalence surveys on receptive and contemporary malaria risk, should support malaria control and elimination initiatives in the country.

Competing interests

Authors declare no competing interests.

Authors' contributions

VAA was responsible for study design, data cleaning, analysis, interpretation, drafting and production of the final manuscript. RK and BN contributed to the data assembly, cleaning and contributed to the final manuscript. PMA and JW were responsible for analysis, interpretation and production of the final manuscript. AMN and RWS were responsible for overall scientific management, analysis, interpretation and preparation of the final manuscript. All authors read and approved the final manuscript.

Funding

VAA is supported by a Commonwealth fellowship (KECS-2012-601). AMN is supported by the Wellcome Trust as an Intermediate Research Fellow (#095127). RWS is supported by the Wellcome Trust as Principal Research Fellow (#079080). This work was partly funded by a grant from the Namibia Ministry of Health and Social Services-Global Fund Programme and a Wellcome Trust Major Overseas Programme grant to the KEMRI/Wellcome Trust Research Programme (#092654). The funders played no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

The authors are grateful to the National Vector-Borne Disease Control Programme, who provided the main health facility data used in this study. We particularly thank John Mendelsohn of RAISON for his help with most of the spatial data, Kudzai Makomva and Andreas Reich for help with data assembly and Caroline Kabaria, Betsy Makena and Viola Kirui for help with geo-coding and data extraction. We are also grateful to Jacob Ouko for help with the assembly of ancillary data sets and the useful inputs from the anonymous reviewers.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.sste.2013.09.001.

References

Afripop. The AfriPop project. In: Tatem A, Linard C., editors; 2010.
Alegana V, Wright J, Petrina U, Noor A, Snow R, Atkinson P. Spatial modelling of healthcare utilisation for treatment of fever in Namibia. Int J Health Geogr 2012;11:6.
Amek N, Bayoh N, Hamel M, Lindblade KA, Gimnig J, Laserson KF, Slutsker L, Smith T, Vounatsou P. Spatio-temporal modeling of sparse geostatistical malaria sporozoite rate data using a zero-inflated binomial model. Spat Spatio-Temp Epidemiol 2011;2:283–90.
Balk D, Pozzi F, Yetman G, Deichmann U, Nelson A. The distribution of people and the dimension of place: methodologies to improve global population estimates in urban and rural areas. New York: CIESIN, Columbia University; 2004.

Author's personal copy

Banerjee S, Carlin BP. Semiparametric spatio-temporal frailty modeling. Environmetrics 2003;14:523–35.

Barnerjee S, Carling PB, Gelfand AE. Hierarchical modeling and analysis for spatial data. London: Chapman & Hall/CRC; 2004.

Bernardinelli L, Pascutto C, Best NG, Gilks WR. Disease mapping with errors in covariates. Stat Med 1997;16:741–52.

Böhning D. Zero-inflated Poisson models and C.A. MAN: a tutorial collection of evidence. Biom J 1998;40:833–43.

Center for International Earth Science Information Network (CIESIN). Global rural-urban mapping project (GRUMP): urban extents. New York: CIESIN, Columbia University; 2004.

Cibulskis RE, Aregawi M, Williams R, Otten M, Dye C. Worldwide incidence of malaria in 2009: estimates, time trends, and a critique of methods. PLoS Med 2011;8:e1001142.

Cibulskis RE, Bell D, Christophel EM, Hii J, Delacollette C, Bakyaita N, Aregawi MW. Estimating trends in the burden of malaria at country level. Am J Trop Med Hyg 2007;77:133–7.

Clements A, Barnett A, Cheng Z, Snow R, Zhou H. Space-time variation of malaria incidence in Yunnan province, China. Malar J 2009;8:180.

Cosner C, Beier JC, Cantrell RS, Impoinvil D, Kapitanski L, Potts MD, Troyo A, Ruan S. The effects of human movement on the persistence of vector-borne diseases. J Theor Biol 2009;258:550–60.

Craig MH, Sharp BL, Mabaso ML, Kleinschmidt I. Developing a spatial-statistical model and map of historical malaria prevalence in Botswana using a staged variable selection procedure. Int J Health Geogr 2007;6:44.

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

Cressie N, Wikle KC. Statistics for spatiotemporal data. Hoboken, New Jersey: John Wiley & Sons, Inc.; 2011.

Czado C, Gneiting T, Held L. Predictive model assessment for count data. Biometrics 2009;65:1254–61.

Gelfand AE, Vounatsou P. Proper multivariate conditional autoregressive models for spatial data analysis. Biostatistics 2003;4:11–25.

Gething P, Van Boeckel T, Smith D, Guerra C, Patil A, Snow R, Hay S. Modelling the global constraints of temperature on transmission of *Plasmodium falciparum* and *P. vivax*. Parasites Vectors 2011;4:92.

Gething PW, Noor AM, Gikandi PW, Hay SI, Nixon MS, Snow RW, Atkinson PM. Developing geostatistical space-time models to predict outpatient treatment burdens from incomplete national data. Geogr Anal 2008;40:167–88.

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

Ghosh SK, Mukhopadhyay P, Lu J-C. Bayesian analysis of zero-inflated regression models. J Stat Plan Infer 2004;136:1360–75.

Gneiting T, Raftery AE. Strictly proper scoring rules, prediction, and estimation. J Am Stat Assoc 2007;102:359–78.

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

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

Hay SI, Smith DL, Snow RW. Measuring malaria endemicity from intense to interrupted transmission. Lancet Infect Dis 2008;8:369–78.

Huffman GJ. Estimates of root-mean-square random error contained in finite sets of estimated precipitation. J Appl Meteor 1997;36:191–1201.

Huffman GJ, Bolvin DT. TRMM and other data precipitation data set documentation, Maryland, United States. Goddard: Laboratory for Atmospheres, NASA Goddard Space Flight Center and Science Systems and Applications, Inc; 2011.

Kleinschmidt I, Sharp B, Mueller I, Vounatsou P. Rise in malaria incidence rates in South Africa: a small-area spatial analysis of variation in time trends. Am J Epidemiol 2002;155:257–64.

Kneib, T., Tutz, G., Held, L., Schrödle, B. and Rue, H. Posterior and Cross-validatory Predictive Checks: A Comparison of MCMC and INLA. Statistical Modelling and Regression Structures. Physica-Verlag HD, 2010.

Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics 1992;34:1–14.

Linard C, Alegana V, Noor A, Snow R, Tatem A. A high resolution spatial population database of Somalia for disease risk mapping. Int J Health Geogr 2010;9:45.

Linard C, Gilbert M, Snow RW, Noor AM, Tatem AJ. Population distribution, settlement patterns and accessibility across Africa in 2010. PLoS ONE 2012;7:e31743.

Loha E, Lindtjorn B. Model variations in predicting incidence of *Plasmodium falciparum* malaria using 1998–2007 morbidity and meteorological data from south Ethiopia. Malar J 2010;9:166.

Manh BH, Clements ACA, Thieu NQ, Hung NM, Hung LX, Hay SI, Hien TT, Wertheim HFL, Snow RW, Horby P. Social and environmental determinants of malaria in space and time in Viet Nam. Int J Parasitol 2011;41:109–16.

Martins T, Simpson D, Lindgren F, Rue H. Bayesian computing with INLA: new features, Norway. Trondheim: Department of Mathematical Sciences, Norwegian University of Science and Technology; 2013.

Ministry of Health and Social Services. National policy and strategy for malaria control. Windhoek: Ministry of Health and Social Services and World Health Organization; 1995.

Ministry of Health and Social Services. Namibia health facility census, 2009. Windhoek, Namibia: Ministry of Health and Social Services and ICF Macro; 2010a.

Ministry of Health and Social Services. Namibia malaria indicator survey 2009. Windhoek: Ministry of Health and Social Services; 2010b.

Ministry of Health and Social Services. Namibia malaria strategic plan 2010–2016. In National Vector-Borne Diseases Control Programme (Ed.). Windhoek; 2010c.

Ministry of Health and Social Services. National health policy framework 2010–2020. Towards quality health and social welfare services. Windhoek: Ministry of Health and Social Services; 2010d.

Ministry of Health and Social Services. National malaria monitoring and evaluation plan (2010–2016). In: Programme NV-B DC., editor. Windhoek: National Vector-borne Diseases Control Programme; 2010e.

Ministry of Health and Social Services (Mohss) and Icf Macro. Namibia health facility census (HFC) 2009. Windhoek: Namibia, MoHSS and ICF Macro; 2010.

Molineaux L. The epidemiology of human malaria as an explanation of its distribution, including some implications for its control. In: Wernsdorfer, W. and Mcgregor, I. L. (Eds.) Malaria: principles and practice of malariology; 1988, London, Churchill Livingstone.

Murray CJ, Lopez AD, Wibulpolprasert S. Monitoring global health: time for new solutions. BMJ 2004;329:1096–100.

Nasa. TRMM. Goddard, Maryland, USA: NASA Goddard Space Flight Center; 2011.

National Planning Commission. Namibia 2011 population and housing census preliminary results. Windhoek, Namibia; 2012.

Neelon BH, O'malley AJ, Normand SL. A Bayesian model for repeated measures zero-inflated count data with application to outpatient psychiatric service use. Stat Model 2010;10:421–39.

Nobre AA, Schmidt AM, Lopes HF. Spatio-temporal models for mapping the incidence of malaria in Pará. Environmetrics 2005;16:291–304.

Noor A, Alegana V, Kamwi R, Hansford C, Ntomwa B, Katokele S, Snow R. Malaria control and the intensity of *Plasmodium falciparum* transmission in Namibia 1969–1992. PLoS ONE 2013a;8.

Noor A, Uusiku P, Kamwi R, Katokele S, Ntomwa B, Alegana V, Snow R. The receptive versus current risks of *Plasmodium falciparum* transmission in Northern Namibia: implications for elimination. BMC Infect Dis 2013b;13:184.

Noor AM, Alegana VA, Kamwi RN, Hansford CF, Ntomwa B, Katokele S, Snow RW. Malaria control and the intensity of *Plasmodium falciparum* transmission in Namibia 1969–1992. PLoS ONE 2013c;8:e63350.

Noor AM, Alegana VA, Patil AP, Snow RW. Predicting the unmet need for biologically targeted coverage of insecticide-treated nets in Kenya. Am J Trop Med Hyg 2010;83:854–60.

Omumbo JA, Noor AM, Fall IS, Snow RW. How well are malaria maps used to design and finance malaria control in Africa? PLoS ONE 2013;8:e53198.

Pascutto C, Wakefield JC, Best NG, Richardson S, Bernardinelli L, Staines A, Elliott P. Statistical issues in the analysis of disease mapping data. Stat Med 2000;19:2493–519.

Patil AP, Okiro EA, Gething PW, Guerra CA, Sharma SK, Snow RW, Hay SI. Defining the relationship between *Plasmodium falciparum* parasite rate and clinical disease: statistical models for disease burden estimation. Malar J 2009;8:186.

Reid HL, Haque U, Roy S, Islam N, Clements AC. Characterizing the spatial and temporal variation of malaria incidence in Bangladesh, 2007. Malar J 2012;11:170.

Robinson WS. Ecological correlations and the behavior of individuals. Am Sociol Rev 1950;15:351–7.

Author's personal copy

Rue H, Held L. Gaussian markov random fields: theory and applications (Chapman & Hall/CRC Monographs on Statistics & Applied Probability). Chapman and Hall/CRC; 2005.

Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J R Stat Soc B 2009;71:319–92.

Rue HV, Martino S. Approximate Bayesian inference for hierarchical Gaussian Markov random field models. J Stat Plan Infer 2007;137:3177–92.

Scharlemann JRPW, Benz D, Hay SI, Purse BV, Tatem AJ, Wint GRW, Rogers DJ. Global data for ecology and epidemiology: a novel algorithm for temporal Fourier processing MODIS data. PLoS ONE 2008;3:e1408.

Schrödle B, Held L. A primer on disease mapping and ecological regression using INLA. Comput Stat 2010;26:241–58.

Snow RW, Alegana VA, Makomva K, Reich A, Uusiku P, Katokele S, et al. Estimating the distribution of malaria in Namibia in 2009: assembling the evidence and modeling risk. Ministry of Health & Social Services, Republic of Namibia & Malaria Atlas Project; 2010.

Southern Africa Roll Back Malaria Network (Sarn). 2010. Report of the elimination - 8: Malaria elimination technical meeting. IN Mudambo, K. and Lesaso, B. (Eds.). Maputo, SARN. Available: http://www.malariaeliminationgroup.org/sites/default/files/E8_MAPUTO_REPORT_SEND.pdf

Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. J R Stat Soc B 2002;64:583–639.

Stansfield S. Structuring information and incentives to improve health. Bull World Health Organ 2005;83:562.

Trans-Zambezi Malaria Initiative (Tzmi). Implementation plan for the trans-Zambezi malaria initiative (TZMI), a multi-national cross-border control and elimination initiative involving Angola, Botswana, Namibia, Zambia and Zimbabwe, 2012.

Vounatsou P, Raso G, Tanner M, N'goran E, Utzinger J. K. Bayesian geostatistical modelling for mapping schistosomiasis transmission. Parasitology 2009;136:1695–705.

Wakefield J. Sensitivity analyses for ecological regression. Biometrics 2003;59:9–17.

World Health Organizastion. Malaria elimination: a field manual for low and moderate endemic countries. Geneva: World Health Organizastion; 2007.

World Health Organizastion. Global malaria control and elimination: report of a technical review. Geneva: World Health Organizastion; .

World Health Organization. World malaria report: 2012. Geneva: World Health Organization; 2012.

Yekutiel P. Problems of epidemiology in malaria eradication. Bull World Health Organ 1960;22:669–83.

Zere E, Mbeeli T, Shangula K, Mandlhate C, Mutirua K, Tjivambi B, Kapenambili W. Technical efficiency of district hospitals: evidence from Namibia using data envelopment analysis. Cost Eff Resour Alloc 2006;4:5.

Zhou G, Sirichaisinthop J, Sattabongkot J, Jones J, Bjornstad ON, Yan G, Cui L. Spatio-temporal distribution of *Plasmodium falciparum* and *P. vivax* malaria in Thailand. Am J Trop Med Hyg 2005;72:256–62.

OPEN ACCESS Freely available online

PLOS ONE

Modelling the Incidence of *Plasmodium vivax* and *Plasmodium falciparum* Malaria in Afghanistan 2006–2009

Victor A. Alegana[1,2]*, Jim A. Wright[2], Sami M. Nahzat[3], Waqar Butt[4], Amad W. Sediqi[3], Naeem Habib[4], Robert W. Snow[1,5], Peter M. Atkinson[2], Abdisalan M. Noor[1,5]*

1 Spatial Health Metrics Group, Department of Public Health, KEMRI-Wellcome Trust, Nairobi, Kenya, **2** Centre for Geographical Health Research, Geography and Environment, University of Southampton, Highfield Southampton, United Kingdom, **3** National Malaria and Leishmaniasis Control Programme, Ministry of Public Health, Kabul, Afghanistan, **4** Malaria and Leishmaniasis, WHO Office, Kabul, Afghanistan, **5** Centre for Tropical Medicine, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, United Kingdom

Abstract

Background: Identifying areas that support high malaria risks and where populations lack access to health care is central to reducing the burden in Afghanistan. This study investigated the incidence of *Plasmodium vivax* and *Plasmodium falciparum* using routine data to help focus malaria interventions.

Methods: To estimate incidence, the study modelled utilisation of the public health sector using fever treatment data from the 2012 national Malaria Indicator Survey. A probabilistic measure of attendance was applied to population density metrics to define the proportion of the population within catchment of a public health facility. Malaria data were used in a Bayesian spatio-temporal conditional-autoregressive model with ecological or environmental covariates, to examine the spatial and temporal variation of incidence.

Findings: From the analysis of healthcare utilisation, over 80% of the population was within 2 hours' travel of the nearest public health facility, while 64.4% were within 30 minutes' travel. The mean incidence of *P. vivax* in 2009 was 5.4 (95% CrI 3.2–9.2) cases per 1000 population compared to 1.2 (95% CrI 0.4–2.9) cases per 1000 population for *P. falciparum*. *P. vivax* peaked in August while *P. falciparum* peaked in November. 32% of the estimated 30.5 million people lived in regions where annual incidence was at least 1 case per 1,000 population of *P. vivax;* 23.7% of the population lived in areas where annual *P. falciparum* case incidence was at least 1 per 1000.

Conclusion: This study showed how routine data can be combined with household survey data to model malaria incidence. The incidence of both *P. vivax* and *P. falciparum* in Afghanistan remain low but the co-distribution of both parasites and the lag in their peak season provides challenges to malaria control in Afghanistan. Future improved case definition to determine levels of imported risks may be useful for the elimination ambitions in Afghanistan.

Citation: Alegana VA, Wright JA, Nahzat SM, Butt W, Sediqi AW, et al. (2014) Modelling the Incidence of *Plasmodium vivax* and *Plasmodium falciparum* Malaria in Afghanistan 2006–2009. PLoS ONE 9(7): e102304. doi:10.1371/journal.pone.0102304

Editor: Luzia H. Carvalho, Centro de Pesquisa Rene Rachou/Fundação Oswaldo Cruz (Fiocruz-Minas), Brazil

Received March 31, 2014; **Accepted** June 16, 2014; **Published** July 17, 2014

Copyright: © 2014 Alegana et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that, for approved reasons, some access restrictions apply to the data underlying the findings. The malaria case data used in this study were obtained from Afghanistan Health Management Information System (HMIS) through the National Malaria and Leishmaniasis Control Programme (NMLCP). The database is summarised in the supplementary information. The summary includes a time series plot of cases in relation to environmental covariate. However, raw data can be made available through a request to the Afghanistan NMCP via Dr. Mohamed Sami Nahzat (address provided on the manuscript). Ancilliary data (e.g. Precipitation) can be obtained from online web sources (links provided in the manuscript and supporting information).

Funding: VAA is supported by a Commonwealth fellowship (KECS-2012-601). AMN is supported by the Wellcome Trust as an Intermediate Research Fellow (#095127). RWS is supported by the Wellcome Trust as Principal Research Fellow (#079080). This work was partly funded by a Wellcome Trust Major Overseas Programme grant to the KEMRI/Wellcome Trust Research Programme (#092654). The funders played no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors would like to confirm that one of the co-authors, Dr. Abdisalan M. Noor, is a PLOS ONE Editorial Board member. This does not alter the authors' adherence to PLOS ONE Editorial policies and criteria.

* Email: valegana@kemri-wellcome.org (VAA); anoor@kemri-wellcome.org (AMN)

Background

Since the Soviet invasion in 1979, Afghanistan has experienced prolonged periods of insecurity and political instability. Consequently it has some of the poorest socio-economic and health status indicators globally. The country is ranked the thirteen lowest on the human development index [1] and has a child mortality rate of 97 deaths before the age of five years for every 1000 children born [2,3]. In Afghanistan, malaria is an important disease with approximately half the population at risk [4,5,6].

Malaria transmission in the country is constrained by altitude, the rugged topography, patchy rainfall and extreme aridity [7]. There is no active malaria transmission in areas greater than 2000 metres above mean sea level [8], while transmission is unstable in

areas with limited annual rainfall. There are at least six malaria vectors in Afghanistan namely: the *Anopheles superpictus, An. culicifacies, An. hycranus, An. Pulcherimus, An. fluviatilis, and An. stephensi*. The latter two are mainly found in the eastern provinces [4,9]. Malaria infections are predominantly due to the *Plasmodium vivax* parasite although *Plasmodium falciparum* infections exist [8].

Afghanistan has a long history of malaria control dating back to the formation of the Directorate General of Preventive Medicine and Primary Health Care in 1948 [7]. Earlier vector control efforts focused on spraying using dichlorodiphenyltrichloroethane (DDT) and by 1970 the *An. superpictus* was almost eradicated [7,10,11]. After the Soviet invasion, the national program gradually weakened and had almost ceased to function [7,12]. Chloroquine resistance and population movement, mainly from returning refugees, contributed to an increase in malaria burden in Afghanistan [12,13,14]. Since 2000, however, substantial resources have been invested in malaria control in Afghanistan with support from the Global Fund to fight AIDS, Tuberculosis and Malaria, the United States Agency for International Development (USAID) as well as other agencies [15]. Despite the insecurity and infrastructure challenges, progress has been made in reducing the burden [5]. A recent malaria indicator survey (MIS) conducted in 2011 showed an average prevalence of less than 1% for both *P. vivax* and *P. falciparum* nationally while 76% of household clusters had no residents infected.

In the national malaria strategy of 2008 2013, Afghanistan aimed to reduce, by 60%, the malaria morbidity by 2013 and reduce *P. falciparum* cases to near zero with the aim of eventually interrupting its transmission [15]. The main interventions were coverage with vector control, parasitological diagnosis and treatment with effective antimalarials. In addition, a cross-border initiative was launched with Tajikistan to reduce the risk of imported infections to Tajikistan and to eliminate *P. falciparum* malaria in three border districts.

To track progress towards the national targets, the National Malaria and Leishmaniasis Control Programme (NMLCP) and partners established a routine information system to report monthly malaria cases by health facility [16]. The system, however, captured passively detected case data from only the public health system and contained both clinically diagnosed and parasitologically confirmed *P. vivax* and *P. falciparum* cases. Passive case detection, usually from HMIS, is hindered by the challenges of the low parasite confirmation rates which inflate reported malaria caseloads. In addition, low reporting rates tend to underestimate disease burdens because of the spatially and temporally incomplete data [17]. To provide more reliable estimates of disease burden, techniques are required that can adjust for these deficiencies by smoothing crude incidence rates; filling in gaps where no health reports have been assembled; and adjusting for the rate of facility utilisation since only a proportion of actual cases present at a facility [18].

In this study, a formal spatial and temporal approach, that incorporates a variety of data sources to estimate malaria incidence by district from 2006 2009 in Afghanistan, was developed. First, nationally representative household survey data from the 2011 MIS were used to characterize the utilisation of public health facilities and subsequently develop the denominator (catchment population) weighted by probability of health facility use for fever treatment. Secondly, malaria cases reported at the health facilities were used to model incidence of *P. vivax* and *P. falciparum* spatially and temporally using a Bayesian approach [19]. The clinically reported cases were adjusted using species-specific slide positivity rates observed at the facility and combined with parasite species confirmed cases to calculate the numerator. Slide positivity is the ratio of the number of positive malaria cases to the total number of people examined usually expressed as a percentage (rate). The combination of the adjusted cases and catchment populations were then used to compute the incidence of both *P. vivax* and *P. falciparum*.

Methods

Health management information structure in Afghanistan

Afghanistan is divided into 34 administrative provinces. Healthcare is delivered mainly through the Basic Package for Health Services (BPHS) and the Essential Package for Hospital Services (EPHS) constituted in 2002 by the Ministry of Public Health (MoPH) [20,21,22]. In a bid to increase coverage, the BPHS was expanded through the contracting out of services to NGOs and MoPH partners [21,23]. The BHC constitutes clinics, health posts and Maternal Child Health (MCH) centres and Comprehensive Health Centres (CHC). This is linked to EPHS made up of the District Hospitals (DH) (first referral level) and regional or provincial (tertiary) hospitals. At village level community health workers manage the health posts and treat mild conditions and, in some cases, Mobile Health Teams (MHTs) are used [20,24]. In terms of data reports, tally sheets are filled at these lower-tier facilities and aggregated at the next tier facilities (CHC) which are then forwarded to regional directorates [16]. Thus, the health posts serve as a support network for the health centres and sometimes malaria cases are reported at the health centre rather than the individual health unit. The basic health centres link the basic service providers at the community level with the next service tier (the CHC) that are, in turn, linked to district hospitals and regional referral hospitals. Thus, where no regional or tertiary facility exists, district hospitals are the main referral centres. HMIS reports are also compiled the regional level and distributed to the national management level. Inpatient facilities are provided mainly at the tertiary level [20]. Parasitological diagnosis is conducted at higher tier facilities (Hospitals) where laboratory facilities exist while clinical diagnosis is predominantly used at health posts. The 2010 national malaria treatment guidelines outline the scale up of diagnostics at all health facilities to ensure diagnosis prior to treatment.

Data

The malaria case data were obtained from HMIS through the Afghanistan National Malaria and Leishmaniasis Control Programme (NMLCP). This consisted of records from 1,629 public health facilities for a 48-month period from 2006 to 2009. Data represented aggregate monthly cases of *P. falciparum* and *P. vivax*. Of the 1,629 health facilities, 1,587 had reported malaria cases based on both clinical and parasitology examination. Parasitological diagnosis (microscopy or RDTs) was conducted at higher-tier facilities (hospitals and health centres) where laboratory facilities exist while clinical diagnosis was predominantly used at lower-level facilities such as health posts (File S1). No cases were examined or reported for 228 facilities which were treated as missing data while data for mobile units ($n = 93$) were omitted from the final analysis since they serve as outreach centres from major facilities. The missing spatial and temporal structures of data were imputed as '*NAs*' and predictions made at missing locations. The spatial coordinates of health facilities were obtained from the Afghan Management Information Systems (AMIS) (http://www.aims.org.af/), which was formerly managed by the United Nations Office for the Coordination of Humanitarian

Affairs (UNOCHA) and the United Nations Development Programme (UNDP) in the early 2000s, but became a national independent Non-Governmental Organisation (NGO) in 2008. These facilities were either mapped using non-differential hand-held global positioning systems (GPS) receivers during the assessment surveys or in some cases the longitude and latitude were established using a village or settlement database. For analysis, the facilities were classified into three broad categories that combined: basic facilities made up of health posts (HPs), clinics and maternal health centres (MCH); health centres; and hospitals.

Data for modelling health care utilisation for treatment of fever was obtained from the national MIS carried out between September and October 2011 ($n = 15{,}442$ individuals)[25]. The MIS was conducted in 21 provinces, across the diverse malaria strata (medium to high risks; low risk; and very low or potentially malaria free areas) in Afghanistan, but excluded the southern regions for security reasons. A multi-stage probability sampling design was adopted in line with other MIS surveys conducted in sub-Saharan countries [26]. At the first stage clusters or villages were selected randomly in a district via probability sampling while at the second stage, households within the selected clusters were sampled randomly [25]. Self-reported treatment seeking behaviour, disaggregated by healthcare sector, was recorded for all household members that reported an episode of fever two weeks prior to the survey. A gridded population surface for Afghanistan was obtained from Asiapop at 100 m x 100 m spatial resolution (http://www.worldpop.org.uk/)[27].

Analysis

Analysis of public sector utilisation and defining the denominator for modelling incidence

A combination of land cover, elevation, road and river data layers was used to generate a gridded cost surface of travel time between patient origins (households) and destinations (public health facilities) as described elsewhere [28] and in File S1. Travel times were extracted for each MIS cluster and used to predict the probability of health facility attendance based on reported fever treatment. A probability of attendance was modelled spatially at 1 km by 1 km resolution and combined with population density to generate a population-weighted surface for fever treatment. The population-weighted counts, used in modelling incidence, were extracted based on a 2-hour cut off based on the modelled distance decay curves (SI). The catchment population was adjusted for reporting rates at the facilities calculated as a ratio of received reports to the expected number over the four-year period.

Modelling incidence of *P. falciparum* and *P. vivax* in Afghanistan

To model the incidence of malaria, HMIS data were compiled from cases aggregated at each facility for each month. A number of environmental covariates such temperature suitability index (TSI), precipitation and enhanced vegetation index (EVI) that are known to affect malaria transmission were assembled (SI). The selection of covariates was based on previous studies [19] as well as aiming for a minimum set to achieve parsimony based on *bestglm* package in R [29]. These covariates were extracted and matched to each data point in space and time. Environmental covariates were used in a Bayesian zero-inflated conditional autoregressive (CAR) model to predict incidence at the district level. Since 60% of data were zeros, a zero inflated Poisson distribution was used, generalized as [30,31];

$$f(y,\varphi,\mu) \sim \begin{cases} \varphi_i + (1-\varphi_i)\ e^{-\lambda} & y_i = 0 \\ (1-\varphi_i)\ Po(y,\ \mu) & y_i > 0 \end{cases} \quad (1)$$

for the i^{th} space-time observation and $0 < \varphi_i < 1$. The probability is defined via a two-component mixed model such that the probability is φ_i with 'structural' zero or defaulting to a general Poisson model ($\Pr(X=k) = \lambda^k e^{-\lambda}/k!$). In general μ_i can depend on a set of covariates such that:

$$f(y,p,m) = (1-p_{ij})Po(y,0) + p_{ij}Po\left(y, \exp\left[a + b^T X_{ij}\right]\right) \quad (2)$$

with α (equation 2) forming the intercept modified by a $q \times 1$ vector of X_{ij}covariates with unknown coefficients β. Further, $\log(\mu_i) = \sum_{j}^{k} x_{ij}\beta_j$ while $\log it(p_{ij}) = \log[p_{ij}/(1-p_{ij})]$. Thus, the zero-inflated probability increases the chance of predicting 'structural' zero [32,33]. Random effects were introduced at three levels of the health facility, district and province in the Bayesian framework. A spatial effect prior $\varphi_{st(i)} \sim N(0, \tau^2_{str} Q^{-1})$ was introduced at the district level to account for spatial heterogeneity. Model specification was completed by assigning priors to the remaining hyper-parameters (the unstructured random effects). Inverse Gamma priors $IG(a,b)$ were assigned to precision hyperparameters for these unstructured effects components $\theta_{unstr} \sim N(0,\tau^2)$. The time interaction was modelled as a first-order auto-regressive process, $\rho Y(s_i, t_{i-1})$ with the first term coming from a stationary distribution $N(0, \sigma^2_w \sum)$ that depends on past values for $0 > \rho < 1$ [34]. Full details of implementation can be found in File S1.

Posterior predictions were made at the district level along with associated standard errors. Four spatio-temporal models were compared to assess the effect of the introduced random effects at province and facility level as well as the inclusion of the covariates. The first two models (referred to as M1 and M2) did not include any covariates with random effects excluded for the first model (M1). The other two models (M3 and M4) included environmental covariates, with M3 excluding random effects at facility and province level. Comparisons were made using the deviance information criterion (DIC) [35]. This approach simplifies model selection to a single value, which can be easily tabulated for comparison with proper Bayesian interpretation. A subset comprising 10% of the data selected randomly was used independently to compare posterior prediction against the crude incidence. Additionally, model checking was implemented by assessing the variance and the standard error of the predictive distribution [36].

Results

Data characteristics and public sector utilisation in Afghanistan

In modelling healthcare utilisation, a list of the universe of public health facilities was used ($n = 1{,}581$) from the 34 provinces. There were more health posts ($n = 754$) compared to health centres ($n = 698$) and hospital ($n = 129$). The majority were run by NGOs that work in partnership with the Ministry of Public Health (MOPH). The malaria case reporting rate was low for basic health facilities (an average of 33% for the four years) compared to hospitals and health centres where the reporting rate was >70%. Of the estimated population (32.3 million) in 2011, 27.8 million (85.8%) were estimated to be within 2 hours' of travel of a public

health facility; 17.9 million (64.4%) were within 30 minutes. Approximately 13.1 million (47.4%) were within distances where the probability of attendance was ≥60% (SI).

Posterior predictions of incidence of *P. falciparum* and *P. vivax*

Table 1 lists the four Bayesian spatio-temporal models implemented along with associated model parameters for both *P. vivax* and *P. falciparum*. According to the DIC, the fourth model (M4) provided the best trade-off between model fit and parsimony compared to the other three models, although with more effective parameters P_D. For both *P. vivax* and *P. falciparum*, the standard error in M4 of the predictive distribution was also lower. This model was subsequently selected for analysis of incidence of *P. vivax* and *P. falciparum*. Overall mean error of the crude incidence and the predicted incidence per 1000 population per year, based on a 10% validation set was −0.30 and −0.44 for *P. vivax* and *P. falciparum* respectively showing an overall tendency to under-predict by less than 0.5 incident cases per 1000 population. The Pearson correlation based on the validation set was 0.63 for *P. vivax* and 0.62 for *P. falciparum*. Table 2 lists the posterior summaries of the parameters representing the fixed effects, the unstructured components, and the temporal and spatial parameters for both the *P. vivax* and *P. falciparum* models. None of the covariate parameters were significant at 95% Bayesian credible interval (Crl) based on the *P. falciparum* model but temperature suitability (0.123, 95% Crl 0.046 0.202) was significant based on the *P. vivax* model. All other model parameters were significant at 95% Crl.

Figure 1 shows the monthly variation of incidence for *P. vivax* and *P. falciparum* for the four-year period. The incidence of *P. vivax* peaked in August (7.611 95% Crl 4.849 11.721) compared to *P. falciparum* which peaked in November (mean incidence per 1,000 population was 2.403 95% Crl 0.929 5.276). *P. falciparum* was lowest in May (0.830 95% Crl 0.303 1.783). Figure 2 and Figure 3 shows maps of monthly incidence of *P. vivax* and *P. falciparum*, respectively, at district level. The incidence of *P. falciparum* was generally very low compared to *P. vivax*. Nangahar, Kabul and Kunar had highest estimated clinical burden of *P. vivax* and *P. falciparum* while lowest estimated burden was in districts bordering Iran in Nimroz and Farah and in northern Afghanistan. The predicted mean incidence in the most recent data year (2009) for *P. vivax* was 5.4 (95% Crl 3.2 9.2) cases per 1,000 population and 1.2 (95% Crl 0.4 2.9) cases per 1,000 population for *P. falciparum*. Comparison between the baseline in 2006 and in 2009 showed small change in incidence (4.9, 95% Crl 3.0 7.8 and 5.1, 95% Crl 3.2 8.1 respectively for *P. vivax*; 1.1, 95% Crl 0.3 2.4 and 1.1, 95% Crl 0.3 2.5 respectively for *P. falciparum*) (Figure 4). However, there was a slight increase in malaria incidence in 2008 for both *P. vivax* and *P. falciparum* as predicted by the model, but, subsequently dropped to the 2006 level in 2009. The mean percentage change in incidence in the 34 provinces between the baseline year and 2009 for *P. vivax* was 3.0 and 5.9 for *P. falciparum* (Table 3). *P. vivax* reduced in 17 of the 34 provinces in Afghanistan while *P. falciparum* reduced in 13 provinces.

Table 3 provides summaries of population at risk by region. Of the 30.6 million people in 2009, the estimated burden of *P. vivax* in 2009 was 165,712 compared to 36,077 for *P. falciparum*. Approximately 32% of the population lived in regions where *P. vivax* was greater than 1 case per 1000 population compared to 23.7% for *P. falciparum*. About 1.3% of the population were estimated to live in districts with <1 case per 1,000 population and the majority (66.7%) in districts of 1 to<5 *P. vivax* cases per 1,000

Table 1. Comparison of four Bayesian models with and without random effects and covariates (M1 with no random effects or environmental covariates; M2: with random effects but no environmental covariates but no random effects; M3: with environmental covariates and environmental covariates (Rainfall, TSI and EVI).

	Model	DIC	P_D	Mlik (Integration)	Variance of predictive distribution	Std error of predictive distribution	Mean Error	R^2
P. falciparum	M1	3670.00	86.80	−1824.57	0.002	1.026	-	-
	M2	3596.90	95.60	−1824.94	0.005	1.042	-	-
	M3	3599.48	90.64	−1821.78	0.002	1.026	-	-
	M4	3570.76	96.85	−1804.94	0.002	1.022	−0.442	0.619
P. vivax	M1	20933.49	203.48	−10571.74	0.001	1.054	-	-
	M2	20781.31	301.97	−10538.10	0.001	1.049	-	-
	M3	20935.46	206.49	−10593.93	0.001	1.052	-	-
	M4	20780.64	301.46	−10554.87	0.001	1.047	−0.308	0.629

DIC: Deviance information Criterion, **P_D**: Effective number of parameters, **Mlik**: Maximum likelihood estimate.

doi:10.1371/journal.pone.0102304.t001

Table 2. Parameters of the selected Bayesian models (M4) for both *P. falciparum* and *P. vivax* (sequentially as intercept β_0, EVI, TSI, Precipitation, random effects at (facility, district and province), temporal parameter and spatial CAR prior effect φ, SD is the Standard Deviation).

	Parameter	Mean	SD	5%	50%	95%
P. falciparum	Intercept (β_0)	−3.630	0.387	−4.244	−3.633	−3.008
	EVI (β_1)	−0.031	0.079	−0.162	−0.031	0.099
	TSI (β_2)	0.164	0.127	−0.042	0.163	0.334
	Precipitation (β_3)	0.008	0.051	−0.077	0.008	0.091
	Facility random effect (τ_1)	1.940	1.903	0.192	1.380	5.534
	District random effect (τ_2)	2.484	0.829	1.355	2.369	4.010
	Province random effect (τ_3)	3.668	1.164	2.040	3.521	5.838
	Rho for the month (ρ)	0.849	0.117	0.617	0.881	0.969
	Spatial effect (φ)	5.492	4.535	0.698	2.376	20.970
P. vivax	Intercept (β_0)	−2.065	0.240	−2.451	−2.069	−1.662
	EVI (β_1)	−0.026	0.019	−0.058	−0.026	0.005
	TSI (β_2)	0.124	0.048	0.046	0.124	0.202
	Precipitation (β_3)	0.013	0.011	−0.005	0.013	0.031
	Facility random effect (τ_1)	8.383	1.778	6.095	8.057	11.750
	District random effect (τ_2)	2.081	1.976	0.181	1.500	5.888
	Province random effect (τ_3)	7.972	3.953	3.897	6.922	15.530
	Rho for the month (ρ)	0.728	0.098	0.551	0.737	0.872
	Spatial effect (φ)	3.141	0.983	1.759	3.024	4.933

The betas represent the fixed effects of the covariates.
doi:10.1371/journal.pone.0102304.t002

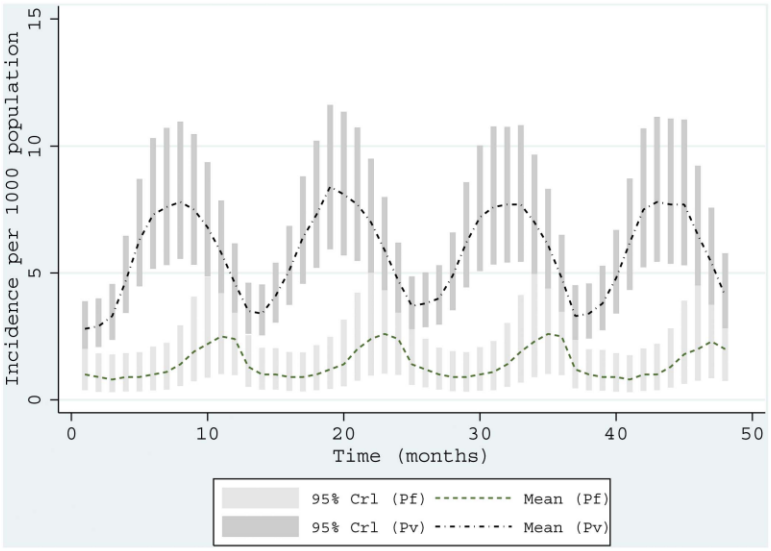

Figure 1. Time series of two malaria parasites. Plots showing the predicted monthly ($n = 48$ months) incidence for (2006–2009) for *P. vivax* (mean as top dash-dot line) and for *P. falciparum* (mean as green dash line) with error bars for each moth showing 95% Bayesian credible interval (Crl). *P. vivax* formed the most burden in Afghanistan and its incidence peaked in July and August compared to *P. falciparum* that peaked later in the year in November.
doi:10.1371/journal.pone.0102304.g001

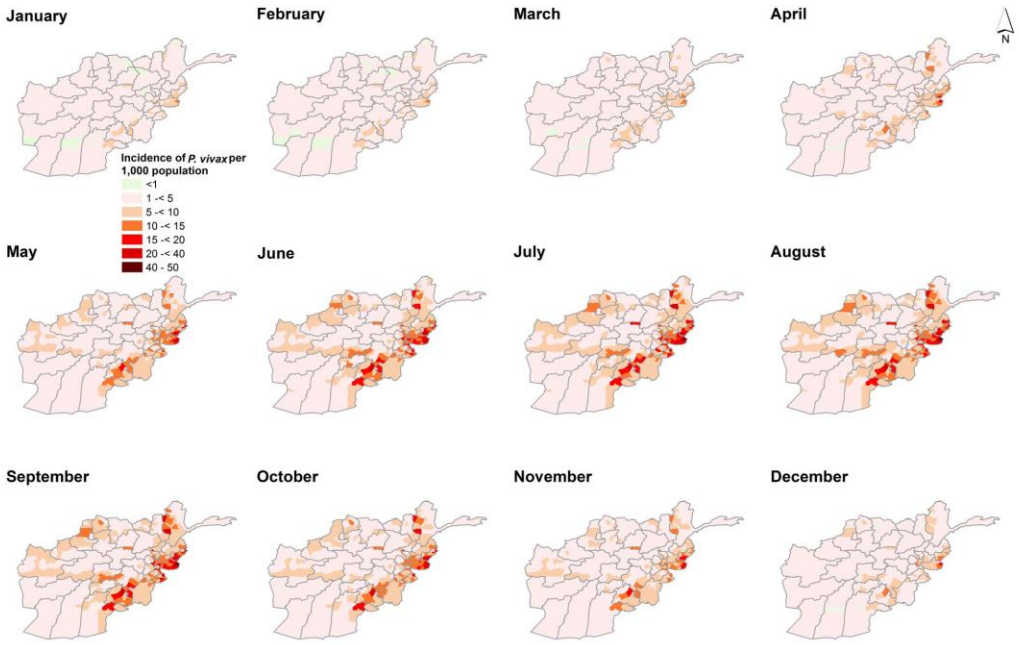

Figure 2. Monthly maps of *P. vivax*. Maps showing the predicted posterior mean monthly incidence of *P. vivax* per 1000 population for Afghanistan in 2009 using a Bayesian CAR model with environmental covariates (rainfall, TSI and EVI). Cases comprised of parasitologically confirmed and clinical cases corrected for slide positivity rates at the facility for a four-year period (2006–2009). Random unstructured effects were included at the facility level to account for regional heterogeneity. The highest burden of *P. vivax* (exceeding 15 cases per 1000 population) was in southeastern and the eastern regions bordering Pakistan.
doi:10.1371/journal.pone.0102304.g002

population. Of the remaining population, only 23.3% were in districts with 5 to<10 cases per 1,000 population, 8.4% in 10 to< 20 cases per 1,000 population and 0.3% of the population. The latter comprised of populations in eastern Afghanistan in Kunar and Nangarhar provinces. For *P. falciparum* case incidence, 76.3% of the population lived in districts where annual incidence was <1 per 1,000 population, while 20.9% lived in areas were incidence of *P. falciparum* was 1 to<5 cases per 1,000 population. A minority (2.8%) of the population lived in districts with an estimated annual incidence of 5 to<10 *P. falciparum* cases per 1,000 population.

Discussion

In this study we have developed a modelling approach that combines household and routine HMIS data within a Bayesian hierarchical spatial-temporal model, to compute the annual incidence of *P. vivax* and *P. falciparum* malaria across 398 districts in Afghanistan. The findings demonstrate a strong geographic co-distribution of *P. vivax* and *P. falciparum* malaria morbidity in Afghanistan (Figure 2 and Figure 3). There was no significant change in the mean annual incidence between 2006 and 2009. The incidence of *P. vivax* and *P. falciparum* in 2009 were estimated to be 5.4 and 1.2 per 1000 population respectively. The incidence (for both parasites) was higher in the south-eastern and eastern parts of Afghanistan bordering Pakistan and lowest in northern districts. In addition, the analysis showed that malaria in Afghanistan exhibits a strong seasonal peak between July and November. *P. vivax* tended to peak in August (mean incidence of 7.611 95% CrI 4.849–11.721) compared to *P. falciparum* which peaked in November (mean incidence 2.403 95% CrI 0.929–5.276). However, incidence was low in the winter months between January and May for both parasites. Slightly more than 76% of districts in Afghanistan had predicted incidence of<1 per 1000 population for *P. falciparum* which is a threshold for pre-elimination.

Using the 2006 data as baseline estimates, 17 and 13 provinces had already reduced *P. vivax* and *P. falciparum* incidence respectively by 2009. No reduction in incidence was predicted for Nangahar, Balkh, Sari Pul, Khost and Hirat Nangahar and Khost provinces in south-eastern regions of Afghanistan were amongst those with highest predicted incidence for both parasites. A range of malaria control strategies are implemented at a national level in Afghanistan. LLIN, for example, is targeted in the high to medium risk districts in Badakhshan, Badghes, Baghlan, Balkh, Faryab, Herat, Helmand, Kandahar, Khost, Kunar, Kunduz, Laghman, Nangarhar and Takhar. From the MIS undertaken in 2008, Nangahar had an estimated long lasting insecticidal nets (LLINs) coverage of 19% while no LLINs use was observed in Hirat [37]. Sari Pul region, for example, had some of lowest rates of long lasting insecticidal nets (LLINs) coverage and access to treatment of care. In the districts where indoor residual spraying (IRS) is used as the main vector control approach or to complement LLINs, the targeting of this intervention should be informed by the lag in the

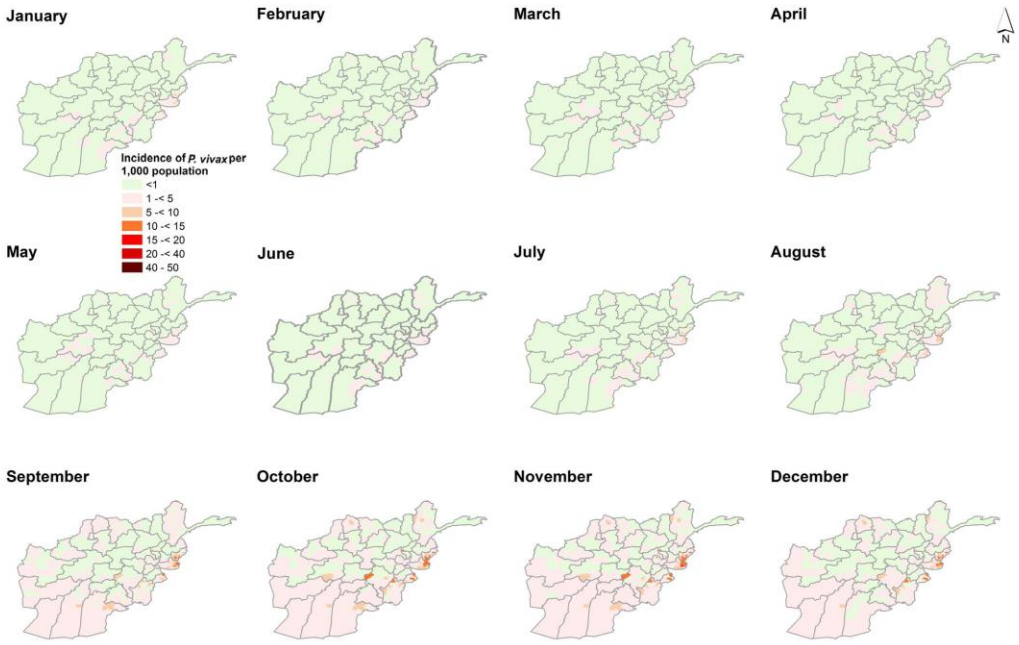

Figure 3. Monthly maps of *P. falciparum*. Maps showing the predicted monthly incidence of *P. falciparum* per 1000 population for Afghanistan in 2009 using a Bayesian CAR model with environmental covariates (rainfall, TSI and EVI). Malaria cases comprised of parasitologically confirmed and clinical cases corrected for slide positivity rates at the facility. Random unstructured effects were included at the facility level to account for regional heterogeneity. *P. falciparum* constitutes less than 10% of the malaria burden in Afghanistan and experienced a late peak in the year (November).
doi:10.1371/journal.pone.0102304.g003

peak season of the two main malaria parasites. *P. vivax* peaks in August while *P. falciparum* peaks in November. IRS campaigns should therefore be planned in such away the insecticide are efficacious through the two peak seasons.

Of the 34 provinces of Afghanistan, five (Bamyan, Ghazni, Ghor, Panjsheer and Nuristan) were considered to be malaria free based on altitude thresholds [38]. These provinces, however, accounted for 9.7% of all estimated cases in 2009 indicating a

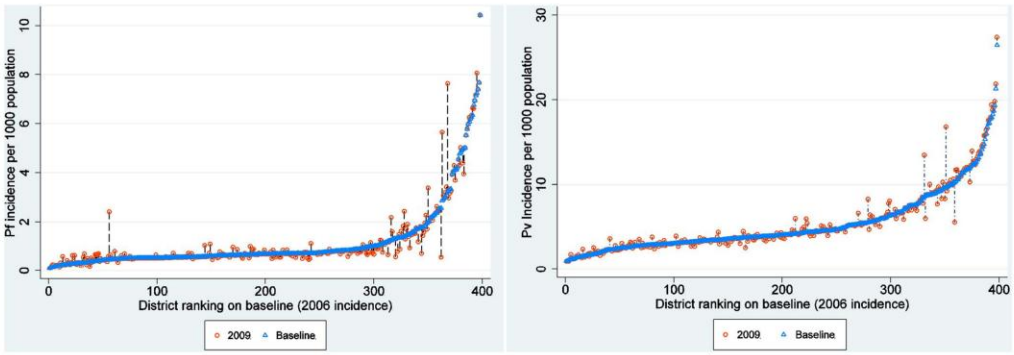

Figure 4. Incidence change plot at district level. Plot showing the differences in malaria incidence per 1000 population (*y*-axis) between the baseline year (2006) plotted as blue triangles and incidence for 2009 (hollow red circles). The *x*-axis represents districts ($n = 398$). The positive change denoting increase in plotted vertically upwards from the baseline year while negative denoting a reduction in incidence is vertically downwards from the baseline with the length indicating magnitude of change. Overall percentage change for *P. vivax* was 3.0 and 5.9 for *P. falciparum*.
doi:10.1371/journal.pone.0102304.g004

Table 3. Predicted mean malaria incidence per 1000 population for 2009 and estimated population at risk by Province for *P. vivax and P. falciparum*.

	Plasmodium vivax incidence per 1,000 population								*Plasmodium falciparum* incidence per 1,000 population							
Province	Estimated Pv Clinical burden	Mean incidence 2009	sd	Percentage change Baseline (2006 and 2009	<1 (%)	1 -<5 (%)	5 -<10 (%)	≥10 (%)	Estimated Pf Clinical burden	Mean incidence 2009	sd	Percentage change Baseline (2006 and 2009	<1 (%)	1 -<5 (%)	5 -<10 (%)	Total Population
Kabul	19,788	4.3	1.1	0.59	0	4,513,039 (98.1)	72,584 (1.6)	16222 (0.4)	5,016	1.1	0.7	-0.52	4,132,170 (89.8)	425,758 (9.3)	43,917 (1)	4,601,845
Kapisa	2,497	5.4	0.8	0.59	0	215,825 (46.8)	245,789 (53.2)	0	268	0.6	0.5	-0.21	461,613 (100)	0	0	461,613
Logar	1,988	4.4	1.1	-2.01	0	308,882 (68.2)	143,913 (31.8)	0	317	0.7	0.6	-5.99	425,918 (94.1)	26,877 (5.9)	0	452,795
Panjshir	777	5.5	1.5	1.24	0	0	123,418 (86.7)	18879 (13.3)	203	1.4	0.8	1.74	99,628 (70.0)	23,790 (16.7)	18,879 (13.3)	142,298
Parwan	2,806	3.4	0.9	-4.04	0	778,082 (93.2)	57,015 (6.8)	0	576	0.7	0.6	-3.74	778,082 (93.2)	57,015 (6.8)	0	835,096
Wardak	2,439	3.8	1.1	-1.13	0	515,633 (79.9)	129,611 (20.1)	0	400	0.6	0.5	-3.68	573,533 (88.9)	71,711 (11.1)	0	645,244
Bamyan	2,172	4.3	1.3	-4.79	0	465,596 (92.2)	0	39633 (7.8)	273	0.5	0.5	-8.76	505,228 (100)	0	0	505,228
Day Kundi	2,280	4.2	1.5	0.19	0	143,333 (26.2)	403,394 (73.8)	0	306	0.6	0.6	-0.02	546,727 (100)	0	0	546,727
Kunar	6,897	13.3	2.0	2.61	0	0	0	517,421 (100)	2,013	3.9	1.3	-0.03	175,647 (33.9)	182,013 (35.2)	159,762 (30.9)	517,421
Laghman	4,114	8.1	0.6	2.91	0	48,935 (9.7)	456,482 (90.3)	0	510	1.0	0.5	1.68	386,431 (76.5)	118,986 (23.5)	0	505,417
Nangarhar	23,043	13.3	0.9	5.48	0	0	733,865 (42.5)	993,460 (57.5)	6,391	3.7	1.2	3.94	0	1,453,897 (84.2)	273,428 (15.8)	1,727,324
Nuristan	1,043	5.9	1.1	3.20	0	35,830 (20.3)	112,523 (63.6)	28511 (16.1)	108	0.6	0.5	2.67	176,863 (100)	0	0	176,863
Badakhshan	6,895	5.4	1.4	0.59	0	247,984 (19.4)	785,904 (61.5)	243,005 (19)	1,302	1.0	0.8	0.93	618,838 (48.5)	658,054 (51.5)	0	1,276,892
Baghlan	3,039	2.9	1.0	2.05	0	1,040,766 (100)	0	0	385	0.4	0.4	5.79	1,040,766 (100)	0	0	1,040,766
Kunduz	4,639	4.1	0.7	-0.82	0	755,118 (66.4)	381,799 (33.6)	0	432	0.4	0.4	2.59	1,136,917 (100)	0	0	1,136,917
Takhar	3,452	3.1	0.7	1.73	0	1,128,142 (100)	0	0	463	0.4	0.3	3.65	1,128,142 (100)	0	0	1,128,142
Balkh	4,091	2.9	1.1	10.56	408,202 (28.9)	1,002,618 (71.1)	0	0	818	0.6	0.5	15.39	1,410,820 (100)	0	0	1,410,820
Faryab	5,209	4.8	1.3	-0.05	0	498,575 (45.8)	591,181 (54.2)	0	708	0.7	0.5	0.25	971,677 (89.2)	118,079 (10.8)	0	1,089,756
Jawzjan	2,583	4.1	1.3	-2.39	0	521,918 (83)	106,563 (17)	0	760	1.2	0.9	8.69	255,757 (40.7)	372,724 (59.3)	0	628,480
Samangan	1,219	2.8	1.0	-1.86	0	441,833 (100)	0	0	239	0.5	0.5	-4.16	441,833 (100)	0	0	441,833
Sari Pul	2,548	3.7	1.0	3.69	0	604,485 (88.5)	78,646 (11.5)	0	266	0.4	0.4	20.25	683,132 (100)	0	0	683,132
Khost	5,613	8.5	1.3	10.31	0	57,439 (8.7)	343,942 (52.2)	257,366 (39.1)	1,719	2.6	1.1	17.70	13,623 (2.1)	426,758 (64.8)	218,366 (33.1)	658,747
Paktika	3,037	6.0	1.9	-4.60	0	25,624 (5.0)	453,431 (89.1)	29,674 (5.8)	829	1.6	1.0	-13.29	171,714 (33.8)	216,884 (42.6)	120,131 (23.6)	508,729
Paktya	4,458	6.9	1.1	-0.44	0	115,091 (17.8)	432,258 (67)	97,766 (15.2)	387	0.6	0.5	1.12	645,114 (100)	0	0	645,114
Ghazni	9,033	6.3	1.2	2.46	0	530,263 (36.7)	739,121 (51.2)	173,664 (12)	2,165	1.5	1.0	16.31	321,390 (22.3)	1,084,947 (75.2)	36,711 (2.5)	1,443,048
Hilmand	2,859	2.8	1.1	-0.26	0	1,039,697 (100)	0	0	1,071	1.0	0.8	1.90	930,158 (89.5)	109,539 (10.5)	0	1,039,697
Kandahar	5,644	4.1	1.2	-0.08	0	1,161,551 (84.2)	178,308 (12.9)	40,003 (2.9)	1,421	1.0	0.8	-4.06	1,130,774 (81.9)	249,088 (18.1)	0	1,379,862
Nimroz	577	3.1	1.2	-2.38	0	187,872 (100)	0	0	178	1.0	0.8	-7.58	86,898 (46.3)	100,974 (53.7)	0	187,872
Uruzgan	1,601	4.0	1.1	-0.88	0	306,799 (75.9)	97,595 (24.1)	0	311	0.8	0.7	0.87	332,451 (82.2)	71,944 (17.8)	0	404,395
Zabul	2,621	7.5	1.6	-6.88	0	20,586 (5.9)	144,845 (41.3)	185,415 (52.8)	379	1.1	0.8	-9.01	149,910 (42.7)	200,936 (57.3)	0	350,846

Table 3. Cont.

	Plasmodium vivax incidence per 1,000 population								*Plasmodium falciparum* incidence per 1,000 population							
Province	Estimated Pv Clinical burden	Mean incidence 2009	sd	Percentage change Baseline (2006 and 2009	<1 (%)	1 -<5 (%)	5 -<10 (%)	≥10 (%)	Estimated Pf Clinical burden	Mean incidence 2009	sd	Percentage change Baseline (2006 and 2009	<1 (%)	1 -<5 (%)	5 -<10 (%)	Total Population
Badghis	2,257	4.1	0.9	1.92	0	534,183 (96.8)	17,642 (3.2)	0	353	0.6	0.5	7.60	551,825 (100)	0	0	551,825
Farah	1,459	2.5	1.0	-3.19	0	581,449 (100)	0	0	494	0.9	0.8	0.59	484,949 (83.4)	96,500 (16.6)	0	581,449
Ghor	3,141	4.1	1.4	-1.07	0	468,809 (60.9)	300,924 (39.1)	0	708	0.9	0.9	2.08	459,819 (59.7)	309,915 (40.3)	0	769,733
Hirat	7,532	3.6	1.2	16.37	0	2,098,175 (100)	0	0	1,133	0.5	0.5	23.14	209,8175 (100)	0	0	2,098,175
	165,712	**5.4**	**1.2**	**1.62**	**408,202 (1.3)**	**20,394,129 (66.7)**	**7,130,753 (23.3)**	**2,641,017 (8.6)**	**36,077**	**1.2**	**0.7**	**2.26**	**23,326,522 (76.3)**	**6,376,386 (20.9)**	**871,194 (2.8)**	**30,574,102**

SD: Standard Deviation.

doi:10.1371/journal.pone.0102304.t003

potential problem of importation of malaria cases due to human population movement in Afghanistan or foci transmission in valleys where climatic conditions are favourable. The available data, however, do not provide malaria case definitions and it is impossible to distinguish between imported and local cases. In the malaria free provinces, suspected imported infections should be documented and algorithms, based mainly on travel history, could be used as the basis for case definitions [39]. In addition, health advice and chemoprophylaxis for travellers from the malaria free to endemic provinces should be initiated as an additional package for malaria prevention. An incidence of less than 1 *P. falciparum* case per 1000 individuals is considered to be the threshold for pre-elimination by the WHO [40]. By 2009, 21 provinces in Afghanistan had already achieved such a threshold. However, the biggest challenge is likely to be operational and a comprehensive analysis of overall feasibility of *P. falciparum* elimination [41].

The results of our analysis also have important applications to the design and allocation of resources for malaria case management. Mixed infections especially with *P. vivax* and *P. falciparum* present a challenge for treatment [42]. *P. vivax* infections relapse from dormant liver-stage hypnozoites months after primary infections and are often difficult to diagnose and treat and define as true incident infections [43]. Chloroquine is used as first line treatment of *P. vivax* in Afghanistan as recommended for countries where it remains efficacious and where parasites can be isolated [44], while Artesunate + Sulfadoxine-Pyrimethamine (AS+SP) is used for *P. falciparum* [15]. The incidence maps developed in our study can be used to quantify the need for the number of treatment dosages required for both parasites in Afghanistan. The prevalence of Glucose-6-Phosphate Dehydrogenase (G6PD) deficiency is estimated to be 8% in Afghanistan [45] which complicates the use of the reccomended 14 day regiment of primaquine (PQ) for *P.vivax* [46]. The use of PQ in patients with G6PD deficiency can cause severe haemolysis [47]. To improve the existing information on the prevalence of G6PD deficiency these maps can be used to provide *P. vivax* incidence information that would be useful for the design of G6PD surveys. Where both species are endemic, the use of ACTs has been proposed [48,49] and other clinical studies have shown a faster parasite clearance rate when ACTs were used [50]. Figure 2 and Figure 3 indicates Kunar, Khost, Laghman, Nangarhar, Takhar and some districts in Kandahar could benefit from such a case management approach.

Although HMIS data used in this analysis represent individuals presenting with fever (symptomatic cases) and with a greater chance of detecting an infection, diagnosis was mainly based on either microscopy or Rapid Diagnostic Test (RDT), both of which have varying sensitivities [51]. In low transmission setting, the rate of false positivity rates, when using RDTs, may be higher and the quality of microscopy in routine HMIS data with varying laboratory conditions may also vary [52]. In addition, it was not possible to identify which diagnosis was used at each facility to adjust the sensitivities. A combination of factors might have limited the effect of environmental covariates especially for *P. falciparum*. This relates to assumptions of linearity during modelling even though rainfall was lagged by four months (SI) and given the short time-series of the data (four years). We used an autoregressive time-varying factor in the model assuming that present state evolves from previous values, but modified by the set of spatial and spatio-temporal covariates[34]. Future studies should relax the linearity assumptions of the fixed effects. Another limitation of the data presented here is that the effects of migration or travel between various regions were not incorporated into the modelling

framework. A study in south-eastern Afghanistan showed higher asymptomatic infections in the migrant population [53]. Modelling migration patterns at national level was beyond the scope of this study. It was assumed that individuals would seek treatment at the nearest facility or at least within a district or one of its neighbours. In this study we have modelled malaria morbidity at the facility level and explicitly modelled healthcare attendance at individual facilities as part of our methodology. Security or conflict remains an important factor affecting utilisation of healthcare facilities. This was not modelled in the study due to lack of data, but future studies should consider accounting for this effect on reported health events. Given the additional spatial precision resulting from the facility-level analysis representing febrile individuals within facility catchments, maps of both malaria species are useful for concerted planning.

Conclusion

This study demonstrates how HMIS and household survey data can be assembled, integrated and interpolated to identify districts with high malaria burden spatially and temporally. Maps were produced at the level of decision-making units, which are potentially useful to the malaria control programme in assessing the changing burden of disease in Afghanistan, targeting malaria interventions at the population most at risk, and planning health resources. The districts identified with high burden should be the focus for targeting vector control. Districts with both *P. vivax* and *P. falciparum* and high rates of mixed infections should be investigated and careful case management strategies adopted. Improved case definition to determine levels of imported risks in malaria free areas is necessary for the elimination ambitions in Afghanistan.

Supporting Information

File S1 The analysis of public healthcare utilisation for treatment of fever and Bayesian model specification for modelling incidence is provided in the supplementary information. For healthcare coverage and utilisation, this includes the modelling of probability of attendance for fever treatment and delineation of public health facility catchments to estimate population using public health facilities. For incidence modelling, the supplementary information includes details on model specification, parameter estimation and validation.
(DOC)

Acknowledgments

The authors are grateful to the National Malaria and Leishmaniasis Control Programme (NMLCP) under the Ministry of Public Health in Afghanistan, who provided the health facility data used in this study. We also thank Caroline Kabaria, Viola Otieno, Jacob Ouko for help with geocoding and the assembly of ancillary data sets. We would like to thank Dr Hoda Atta and Dr Ghasem Zamani of the WHO-EMRO for their comments.

Author Contributions

Conceived and designed the experiments: VAA AMN RWS. Performed the experiments: VAA. Analyzed the data: VAA. Contributed reagents/materials/analysis tools: VAA WB SMN AWS NH. Contributed to the writing of the manuscript: VAA PMA JW AMN RWS.

References

1. UNDP (2013) Human development report 2013. Washington DC. 22 p.
2. Afghan Public Health Institute, Ministry of Public Health (APHI/MoPH) [Afghanistan], Central Statistics Organization (CSO) [Afghanistan], ICF Macro, Indian Institute of Health Management Research (IIHMR) [India], et al. (2011) Afghanistan Mortality Survey 2010. Calverton, Maryland, USA: APHI/MoPH, CSO, ICF Macro, IIHMR and WHO/EMRO.
3. Rasooly MH, Govindasamy P, Aqil A, Rutstein S, Arnold F, et al. (2013) Success in reducing maternal and child mortality in Afghanistan. Glob Public Health.
4. Safi N, Leslie T, Rowland M (2009) Progress and Challenges to Malaria Control in Afghanistan. Afghanistan Annual Malaria Journal: 15 29.
5. World Health Organization (2013) World Malaria Report 2013. Geneva: World Health organization.
6. World Health Organization (2013) Country profile, Afghanistan. World Health Organization.
7. Kolaczinski J, Graham K, Fahim A, Brooker S, Rowland M (2005) Malaria control in Afghanistan: progress and challenges. The Lancet 365: 1506 1512.
8. Brooker S, Leslie T, Kolaczinski K, Mohsen E, Mehboob N, et al. (2006) Spatial epidemiology of Plasmodium vivax, Afghanistan. Emerg Infect Dis 12: 1600 1602.
9. Rowland M, Mohammed N, Rehman H, Hewitt S, Mendis C, et al. (2002) Anopheline vectors and malaria transmission in eastern Afghanistan. Trans R Soc Trop Med Hyg 96: 620 626.
10. Dy FJ (1954) Present status of malaria control in Asia. Bull World Health Organ 11: 725 763.
11. Ramachandra T (1951) Malaria control using indoor residual sprays in the Eastern Province of Afghanistan. Bull World Health Organ 3: 639 661.
12. Rowland M, Hewitt S, Durrani N, Bano N, Wirtz R (1997) Transmission and control of vivax malaria in Afghan refugee settlements in Pakistan. Transactions of the Royal Society of Tropical Medicine and Hygiene 91: 252 255.
13. Delfini LF (1989) The first case of Plasmodium falciparum resistant to chloroquine treatment discovered in the Republic of Afghanistan. Trans R Soc Trop Med Hyg 83: 316.
14. Shah I, Rowland M, Mehmood P, Mujahid C, Razique F, et al. (1997) Chloroquine resistance in Pakistan and the upsurge of falciparum malaria in Pakistani and Afghan refugee populations. Annals of Tropical Medicine and Parasitology 91: 591 602.
15. Ministry of Public Health (2008) National Malaria Strategic Plan 2008 2013. General Directorate of Preventive Medicine and Primary Health Care, National Malaria and Leishmaniasis Control Programme.
16. Ministry of Public Health (2006) National Health Management Information Systems (HMIS): Procedures manual. Kabul.
17. Gething PW, Noor AM, Gikandi PW, Ogara EA, Hay SI, et al. (2006) Improving imperfect data from health management information systems in Africa using space-time geostatistics. PLoS Med 3: e271.
18. Cibulskis RE, Aregawi M, Williams R, Otten M, Dye C (2011) Worldwide Incidence of Malaria in 2009: Estimates, Time Trends, and a Critique of Methods. PLoS Med 8: e1001142.
19. Alegana VA, Atkinson PM, Wright JA, Kamwi R, Uusiku P, et al. (2013) Estimation of malaria incidence in northern Namibia in 2009 using Bayesian conditional-autoregressive spatial-temporal models. Spatial and Spatio-temporal Epidemiology 7: 25 36.
20. Ministry of Public Health (2010) A Basic Package of Health Services for Afghanistan. Kabul: Islamic Republic of Afghanistan, Ministry of Public Health.
21. Waldman R, Strong L, Wali A (2006) Afghanistan's health system sinse 2001: condition improvement, prognosis cautiously optimistic.: Afghanistan Research and Evaluation Unit briefing paper series.
22. Ministry of Public Health (2003) A basic package of health services for Afghanistan. Kabul: Islamic Transitional Government of Afghanistan. Kabul.
23. Steinhardt LC, Aman I, Pakzad I, Kumar B, Singh LP, et al. (2011) Removing user fees for basic health services: a pilot study and national roll-out in Afghanistan. Health Policy and Planning 26: ii92 ii103.
24. Belay T (2010) Building on Early Gains in Afghanistan's Health, Nutrition and Population Sector: *Challenges and Options*. Washington (D.C.): World Bank.
25. Ministry of Public Health (2012) The second Malaria Indicator Survey in the Islamic Republic of Afghanistan, 2011. World Health Organization, Health Net TPO, BRAC, HPRO, KEMRI/Wellcome Trust Research Programme.
26. RBM-MERG (2008) A guide to Malaria Indicator Surveys (MIS). 11.
27. Gaughan AE, Stevens FR, Linard C, Jia P, Tatem AJ (2013) High Resolution Population Distribution Maps for Southeast Asia in 2010 and 2015. PLoS One 8: e55882.
28. Alegana V, Wright J, Petrina U, Noor A, Snow R, et al. (2012) Spatial modelling of healthcare utilisation for treatment of fever in Namibia. International Journal of Health Geographics 11: 6.
29. McLeod AI, Xu C (2008) bestglm: Best Subset GLM.
30. Ghosh SK, Mukhopadhyay P, Lu J-C (2006) Bayesian analysis of zero-inflated regression models. Journal of Statistical Planning and Inference 136: 1360 1375.
31. Lambert D (1992) Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics 34: 1 14.
32. Ghosh SK, Mukhopadhyay P, Lu J-C (2004) Bayesian analysis of zero-inflated regression models. Journal of Statistical Planning and Inference 136: 1360 1375.
33. Manh BH, Clements ACA, Thieu NQ, Hung NM, Hung LX, et al. (2011) Social and environmental determinants of malaria in space and time in Viet Nam. International Journal for Parasitology 41: 109 116.

34. Sahu SK, Bakar KS (2012) Hierarchical Bayesian autoregressive models for large space time data with applications to ozone concentration modelling. Applied Stochastic Models in Business and Industry 28: 395 415.
35. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A (2002) Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64: 583 639.
36. Neelon BH, O'Malley AJ, Normand SL (2010) A Bayesian model for repeated measures zero-inflated count data with application to outpatient psychiatric service use. Stat Modelling 10: 421 439.
37. Ministry of Public Health (2009) Malaria Indicator Survey. Ministry of Public Health, Islamic Republic of Afghanistan.
38. Ministry of Public Health (2010) National malaria treatment guidelines. Kabul: Islamic Republic of Afghanistan, Ministry of Public Health.
39. Tatem A, Huang Z, Narib C, Kumar U, Kandula D, et al. (2014) Integrating rapid risk mapping and mobile phone call record data for strategic malaria elimination planning. Malaria Journal 13: 52.
40. World Health Organization (2007) Malaria elimination: a field manual for low and moderate endemic countries. Geneva: World Health organization.
41. Feachem RGA, Phillips AA, Hwang J, Cotter C, Wielgosz B, et al. (2010) Shrinking the malaria map: progress and prospects. The Lancet 376: 1566 1578.
42. Douglas NM, Nosten F, Ashley EA, Phaiphun L, van Vugt M, et al. (2011) Plasmodium vivax recurrence following falciparum and mixed species malaria: risk factors and effect of antimalarial kinetics. Clin Infect Dis 52: 612 620.
43. White N (2011) Determinants of relapse periodicity in Plasmodium vivax malaria. Malaria Journal 10: 297.
44. World Health Organization (2012) World Malaria Report: 2012. Geneva: World Health organization.
45. Howes RE, Piel FdrB, Patil AP, Nyangiri OA, Gething PW, et al. (2012) G6PD Deficiency Prevalence and Estimates of Affected Populations in Malaria Endemic Countries: A Geostatistical Model-Based Map. PLoS Med 9: e1001339.
46. World Health Organization (2010) World Health Organization: Guidelines for the treatment of malaria; Organization WH, editor.
47. Cappellini MD, Fiorelli G (2008) Glucose-6-phosphate dehydrogenase deficiency. Lancet 371: 64 74.
48. Sinclair D, Gogtay N, Brand F, Olliaro P (2011) Artemisinin-based combination therapy for treating uncomplicated Plasmodium vivax malaria. Cochrane Database Syst Rev: CD008492.
49. Douglas NM, Anstey NM, Angus BJ, Nosten F, Price RN (2010) Artemisinin combination therapy for vivax malaria. Lancet Infect Dis 10: 405 416.
50. Dao NV, Cuong BT, Ngoa ND, Thuy le TT, The ND, et al. (2007) Vivax malaria: preliminary observations following a shorter course of treatment with artesunate plus primaquine. Trans R Soc Trop Med Hyg 101: 534 539.
51. WHO-FIND (2009) Malaria rapid diagnostic test performance: results of WHO product testing of malaria RDTs: round 1 (2008). Geneva: World Health Organization Special Programme for Tropical Diseases.
52. Hopkins H, GonzÃ¡lez IJ, Polley SD, Angutoko P, Ategeka J, et al. (2013) Highly sensitive detection of malaria parasitemia in an endemic setting: Performance of a new LAMP kit in a remote clinic in Uganda. Journal of Infectious Diseases.
53. Nateghpour M, Akbarzadeh K, Farivar L, Amiri A (2011) Detection of asymptomatic malaria infection among the Afghani immigrant population in Iranshahr district of southeastern Iran. Bulletin de la SociÃ©tÃ© de pathologie exotique 104: 321 323.

APPENDIX 2: Useful mathematical relations

A2.1: Conjugate prior for Poisson distribution

The Poisson distribution has a conjugate gamma prior. Remember for Bayesian inference, the posterior distribution is given as a product of the likelihood and the prior distribution. In practice there will be more than one prior for the posterior. The likelihood of $i.i.d$ samples has a Poisson distribution. Therefore starting with the Poisson likelihood, we have y_i, $i = 1,.....N$ as observations,

$$L(\lambda, y_i) = \prod_{i=1}^{N} \frac{\lambda^{yi} e^{-\lambda}}{y_i!}$$

$$= \frac{\lambda^{\Sigma y_i} e^{-n\lambda}}{(y_1!.....y_n!)}$$

With log likelihood given as;

$$= \left\{ \sum_{1}^{N} y_i \log \lambda - n\lambda \right\}$$

For the Gamma prior ;

$$Gamma(a,b) = \frac{b^a}{\Gamma a} \lambda^{a-1} \cdot e^{-b\lambda}$$

$$posterior = p(\lambda \mid y) \alpha P(y \mid \lambda) p(\lambda)$$

$$= \{\lambda^{\Sigma y_i} e^{-n\lambda}\} \{e^{-b\lambda} \lambda^{a-1}\}$$

$$= \lambda^{\Sigma y_i + a - 1} e^{-\lambda(n+b)}$$

$$posterior \Rightarrow Ga(\Sigma y_i + a, n + b)$$

A2.2: Jeffrey prior for Poisson

The Jeffrey prior which does not change and is non-informative with respect to parametrisation,

$$\sqrt{I(\theta)} = \sqrt{-E\frac{\partial^2}{\partial\theta^2}\log f(y\,|\,\theta)|\,\theta}$$

Writing inform of the posterior $f(y\,|\,\mu)$ and linearising by taking the logarithm we have

$$\log f(y_n\,|\,\mu) = \Sigma y_i \log\mu - n\mu - \log\sum_i^n y_i!$$

$$\frac{\partial}{\partial\mu}\log f(y_n\,|\,\mu) = \Sigma y_i\frac{1}{\mu} - n$$

$$\frac{\partial^2}{\partial\mu^2}\log f(y_n\,|\,\mu) = \frac{-\Sigma y_i}{\mu^2}$$

With expectations for the log of the likelihood as:

$$\sqrt{E\left(\frac{\partial}{\partial\mu^2}\log f(y_n\,|\,\mu)\right)} = -\frac{n}{\mu}$$

$$\alpha\frac{1}{\mu}$$

Where $\Sigma y_i = n$ a constant

A2.3: The inverse Gamma Prior on Poisson

Drawing from the gamma distribution,

$$P(\theta) = \frac{\beta^\alpha\theta^{\alpha-1}}{\Gamma(\alpha)}e^{-\beta\theta}$$

The inverse gamma prior can be written in terms of θ as

$$\phi = \frac{1}{\theta}$$

$$\frac{\partial\theta}{\partial\phi} = \frac{\partial}{\partial\phi}\left(\frac{1}{\phi}\right) = -\frac{1}{\phi^2}$$

The density as,

$$P(\phi) = \frac{\beta^{\alpha}\left(\frac{1}{\phi}\right)^{\alpha-1} e^{\frac{-\beta}{\phi}}}{\Gamma(\alpha)} \cdot \frac{1}{\phi^2}$$

$$= \frac{\beta^{\alpha}\phi^{-(\alpha+1)} e^{\frac{-\beta}{\phi}}}{\Gamma(\alpha)}$$

$$\phi \sim inverse - Gamma(\alpha, \beta)$$

A2.4: Conjugate prior for Binomial distribution

The binomial distribution has the form, with y as observations,

$$f(y \mid n, \pi) = \frac{n!}{y!(n-y)!}\pi^{y}(1-\pi)^{n-y};$$

$$\pi \in [0,1]$$

With the quantity on right hand side, $\frac{n!}{y!(n-y)!}$, sometimes simply written as $\binom{n}{y}$ and

$\pi^{y}(1-\pi)^{n-y}$ dictating the shape of the distribution.

The conjugate prior for binomial is the beta distribution valid in same interval, thus,

$$Beta(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\alpha^{a-1}(1-\alpha)^{b-1}$$

The prior for $\pi \in [0,1]$ via substitution

$$f(\pi) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\pi^{a-1}(1-\pi)^{b-1}$$

With posterior evaluated as product of prior and likelihood

$$\alpha \quad \pi^{a-1}(1-\pi)^{b-1} \times \pi^{y}(1-\pi)^{n-y}$$

$$\alpha \quad \pi^{a+y-1}(1-\pi)^{b+n-y-1}$$

This evaluates to

$$Beta(a+y, b+n-y)$$

$$\frac{\Gamma(a+b+n)}{\Gamma(a+y)\Gamma(b+n-y)}\alpha^{a+y-1}(1-\alpha)^{b+n-y-1}$$

APPENDIX 3: Data tabulation: sources, spatial and temporal resolutions and application

	Source (URL)	Type or resolution (Degree)	Description	Year downloaded or Assembled
Population	"WorldPop Project" (http://www.worldpop.org.uk/data/data_sources/)	0.000833	Population counts per 100m pixel	2010
Elevation	Advanced Spaceborne Thermal Emission and Reflection Radiometer-Global Digital Elevation Model (ASTER-GDEM) (http://asterweb.jpl.nasa.gov/data.asp)	0.008332	GTOPO30 Digital Elevation Model [meter]	2010
Vegetation	Enhanced Vegetaion Index (EVI) (monthly product)	0.008333	MODIS Mean Enhanced Vegetation Index	Monthly
Vegetation	Long term average (WoldClim, http://www.worldclim.org/)	30 seconds	Vegetation index	1950-2000
Vegetation	Normalised Difference vegetation Index (NDVI) (Long term mean)	0.008333	Vegetation index	Monthly
Precipitation	Tropical Rainfall Measuring Mission (TRMM 3B43)(http://mirador.gsfc.nasa.gov/collections/TRMM_3B43__006.shtml)	0.25x0.25	The rate of precipitation	Monthly
Precipitation	Long term average (WoldClim, http://www.worldclim.org/)	30 seconds	Precipitation	1950-2000
Temperature suitability index (TSI)	Malaria Atlas Project (http://www.map.ox.ac.uk/browse-resources/)	0.008333	Index (0 not suitable for malaria transmission, 1 suitable)	2011
Temperature	Long term average (WoldClim, http://www.worldclim.org/)	30 seconds	Tempearture in degrees celsius	1950-2000
Urban/Rural	WorldPop Project (http://www.worldpop.org.uk)	Vector	urban and rural areas - Africa extent	2010
Urban/Rural	CIESIN Global Rural Urban Mapping Project (GRUMP) (http://sedac.ciesin.columbia.edu/data/sets/browse)	0.008333	Urban areas [N/A]	2000
Ancilliary GIS Data	DIVA-GIS (http://www.diva-gis.org/)	Vector	Country boundaries, roads, rives, Gazetters	2011
Ancilliary GIS Data	Mapcruzin (http://www.mapcruzin.com/)	Vector	Country boundaries, roads, rives, Gazetters	2011
Gazetters	Alexandria (http://www.alexandria.ucsb.edu/)	Database	Alexandria Gazetteer Server Client	2011
	Falling Rain Genomics: (http://www.fallingrain.com/world/)	Database	Falling Rain Genomics: World Gazetter	2011
	GeoNames (http://www.geonames.org/)	Database	GeoNames Gazetter	2011
	Getty Thesaurus of Geographic Names (TGN) (http://www.getty.edu/research/tools/vocabularies/index.html)	Database	A structured vocabulary of geographic names for indexing art and architecture.	2011
	Google maps or earth search engine (http://www.google.co.uk/intl/en_uk/earth/)	Database	Google	2011
	NIMA GNS Public Page (http://earth-info.nga.mil/gns/html/)	Database	NGA GEOnet Names	2011
Statistics	R (http://cran.r-project.org/)	Software	Statistical software	2009
Land cover	ESA – "GlobCover Project" (http://due.esrin.esa.int/globcover/)	0.002778	MERIS global land cover classification [N/A]	2009