Exploring the utility of indicators of uncomplicated malaria burden from routine health facility surveillance data in identifying and mapping high-risk areas for malaria in Uganda

SIMON PETER KIGOZI

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy

of the

University of London

SEPTEMBER 2020

Department of Disease Control

Faculty of Infectious & Tropical Diseases

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Funded by Fogarty International Center, US

Research group affiliation(s): LASER Group

Acknowledgement

Uttermost glory to my God – Mustard seed size is all you required!? I'm amazed!

To my lovely family: Ruth, Abi, Ami, Ann, & Asa – you cheered me on in this 'search', thank you.

To my supervisor, Dr. Rachel L. Pullan – With all the emails without number that I sent you, you did not give up on me, truly thankful. My associate supervisor Dr. Grant Dorsey - I could always count on your feedback, thank you.

To my advisory committee (Dr. Emanuele Giorgi, Dr. Katherine Halliday, Dr. Victor Alegana) and mentors (Dr. Joan Nankabirwa, Dr. Teun Bousema, & Dr. Chris Drakely) – you helped me make those rather foggy turns, many thanks.

To Dr. Isabel Rodriguez-Barraquer, it was great learning from you.

To my very dear friends Talemwa Naluggwa, Keron Ssebyala, Abrar Alasmari, Yovitha Sedekia, Wendy, Julian & Robin, Ian & Sue, Chris Musoke, Clive-Davis office mates, Room 449 office mates, FIC fellow-students. Thank you for all the encouragement!

To my brothers, sisters, nephews, and nieces – thanks for standing in for me & with my family.

To Jjaja SP and Erina Kigozi, thank you for the encouragement, prayers, and asking what it is I am doing every time – invaluable guideposts and friends you've always been.

To Dr. Sam Nsobya – You've groomed a brother, thank you.

To the friendliest yet most effective admins I'd ever encountered, Mary Marimoto and Helen White – you made my stay at LSHTM worth the while. You'll always be an important point of reference for me.

To the leadership and staff at IDRC, the ACT-PRIME study team (where it all began), PRISM study, and UMSP study teams. To Dr. Sarah Staedke, Dr. David Smith, Dr. Andy Tatem, and Mr. Geoff Lavoy for believing in me, thank you.

Finally, to the FIC leadership Dr. Phil Rosenthal, Dr. Arthur Mpimbaza, Ms. Deborah E. Ssebatta, Prof. Moses Kamya, and Ms. Rodah Namubiru for all the support, many thanks.

I, Simon Peter Kigozi, confirm that work presented within this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Table of Contents

List of Figures

List of Tables

Abstract

Background and aim: Routine surveillance is increasingly recognised as central to multi-dimensional malaria control efforts, especially for programme planning and impact assessment. Whilst it is global strategy to transform surveillance into a core programmatic component, essential in-depth interpretation of routine surveillance data remains limited, especially in higher transmission settings. I therefore aimed to explore utility of indicators of uncomplicated malaria burden from routine health facility surveillance data in identifying and mapping high-risk areas for malaria in Uganda.

Methods and data sources: To examine routine surveillance indicators of malaria burden, I first evaluated internal consistency between measures from three national reference health facilities, comparing incidence and test positivity rates over time and space. In addition, I examined impacts of control interventions on the age associated burden of malaria, stratified by endemicity and intervention. I then extended this to compare routine reporting data with concurrent community cohort incidence estimates across three sub-counties to evaluate potential sources of bias. Finally, using four years of national health management information system (HMIS)-reported confirmed malaria data in a Bayesian autoregressive analytical framework, I explored the space-time distribution of malaria, and estimated adjusted national and local HMIS-based incidence rates.

Primary findings: At the health facility level, HMIS-based incidence and test positivity rates showed similar trends and predicable relationships, with reduced transmission associated with increasing age of test confirmed malaria cases. Comparison of HMIS and cohort data suggested that HMIS data could provide a relatively unbiased proxy for true incidence - especially in lower-transmission, better performing surveillance systems settings. Lastly, space-time modelling of national HMIS data revealed high-burden and high-risk areas within health facility catchments, districts, and regions, highlighting the utility of routine surveillance data in identifying programmatically relevant heterogeneities in malaria burden in Uganda.

Conclusion: This thesis highlights the potential viability of routine data in evaluating endemic malaria risk with improved routine HMIS. This is shown by: similar trends of HMIS-based incidence with other measures; its unbiased relationship with community cohort incidence; and, its capacity to identify high case rate locations. To realize the potential of these data, coordinated efforts are needed towards high testing rates, complete and timely recording and reporting, and multilevel feedback within national malaria control programme systems. Further research opportunities include treatment or non-care seeking and non-reporting care alternatives impacts on surveillance-based indicators of malaria burden.

Statement of contribution

Work described in this thesis was conducted in an independent project supported by secondary data from multiple research and programmatic projects, with the primary project funded by the Fogarty international centre (FIC), Grant Numbers: D43TW7375 and D43TW010526. All data collection was conducted in Uganda under the Infectious diseases research collaboration (IDRC) with additional support from the USAID's Malaria Action Plan for Districts (MAPD). I led the primary project activities of data collection and cleaning under the supervision of Assoc. Prof. Rachel Pullan. Projects providing secondary data including ACT PRIME, PRISM, and UMSP were led by Professors Sarah Staedke (LSHTM), Grant Dorsey (UCSF), and Moses Kamya (MUK).

My role, as well as the contributions of other collaborators in the work that led to this thesis are briefly described here.

Chapter 1:

This chapter describes the background and introduction of this study, to provide the context around the key research problem to be addressed, given available evidence from work done through other studies. It also describes the study aim, objectives, and rationale for this study. I wrote this chapter.

Chapter 2:

This chapter summarizes the general overview of the data sets used in this study, outlining the preparation approaches to addressing the specific objectives of this study. I wrote this chapter.

Chapter 3:

This chapter explains the relationship between two key measures of burden from HMIS including malaria test positivity and incidence rates using enhanced routine surveillance data from three malaria reference centres. It is summarized into a published paper on malaria burden through routine reporting: relationship between incidence and test positivity rates. The Uganda Malaria Surveillance Project (UMSP) that provided the secondary data for this work was conceived by Moses Kamya, Grant Dorsey, and Sarah Staedke. Ruth Kigozi managed all the data in this project under the supervision of Geoff Lavoy. I designed the study under the guidance of Rachel Pullan. I conducted all the required data cleaning and preparation for this study and analysed the data. I obtained population raster data sets from which I extracted population estimates for the study sites and created the village shapefiles for the three sites. I wrote the manuscript with input from all the co-authors for publication.

SIMON P. KIGOZI, RUTH N. KIGOZI, ASADU SSERWANGA, JOANITER I. NANKABIRWA, SARAH G. STAEDKE, MOSES R. KAMYA, RACHEL L. PULLAN. 2019. Malaria burden through routine reporting: Relationship between incidence and test positivity rates. *Ajtmh*, 101, 137-147.

Chapter 4:

This chapter addresses the effect of malaria control interventions on the age distribution of malaria cases in four sites of Uganda with two sites receiving long lasting insecticidal nets (LLINs) and the other two receiving LLINs with indoor residual spraying (IRS) in addition. The Uganda Malaria Surveillance Project that provided the secondary data for this work was conceived by Moses Kamya, Grant Dorsey, and Sarah Staedke, in collaboration with the National Malaria Control Program. Ruth Kigozi managed all the data in this project under the supervision of Geoff Lavoy. Catherine Maiteki-Sebuguzi contributed to the collection of data on timing of malaria control interventions from multiple departments in the Uganda Ministry of health. I designed the study, cleaned, prepared, and analysed the data with support from Adrienne Epstein, Grant Dorsey, Bryan Greenhouse, and Isabel Rodriguez-Barraquer. I wrote the manuscript for publication with input from all co-authors.

SIMON P. KIGOZI, RUTH N. KIGOZI, ADRIENNE EPSTEIN, ARTHUR MPIMBAZA, ASADU SSERWANGA, ADOKE YEKA, JOANITER I. NANKABIRWA, KATHERINE HALLIDAY, RACHEL L. PULLAN, DAMIAN RUTAZAANA, CATHERINE M. SEBUGUZI, JIMMY OPIGO, MOSES R. KAMYA, SARAH G. STAEDKE, GRANT DORSEY, BRYAN GREENHOUSE & ISABEL RODRIGUEZ-BARRAQUER. Rapid shifts in the age-specific burden of malaria following successful control interventions in four regions of Uganda. *Malar J* **19,** 128 (2020). https://doi.org/10.1186/s12936-020-03196-7

Chapter 5:

This chapter addresses the relationship between incidence rates from health management information systems and cohorts. The three projects providing supplementary data including the UMSP, ACT-PRIME, and Program for resistance, immunology, surveillance & modelling of malaria (PRISM) projects were conceived by Moses Kamya, Sarah Staedke, and Grant Dorsey and I managed the GPS data from the ACT-PRIME and PRISM projects. I designed this study for this chapter, lead the fieldwork for data collection from the three sites included, managed, cleaned, and analysed the data that was collected. I wrote the manuscript for this study with input from all co-authors.

SIMON P. KIGOZI, EMANUELE GIORGI, ARTHUR MPIMBAZA, RUTH N. KIGOZI, TEUN BOUSEMA, EMMANUEL ARINAITWE, JOANITER I. NANKABIRWA, CATHERINE M. SEBUGUZI, MOSES R. KAMYA, SARAH G. STAEDKE, GRANT DORSEY, & RACHEL L. PULLAN. (2020). Practical Implications of a Relationship between Health Management Information System and Community Cohort–Based Malaria Incidence Rates. *The American Journal of Tropical Medicine and Hygiene*, tpmd190950

Chapter 6:

This chapter includes unpublished work on the identification of high-risk areas for malaria in Uganda using HMIS-based incidence rates. I designed the study with guidance from Rachel Pullan. Ruth Kigozi contributed to the collection of National HMIS data using DHIS-2, for the five years included. Jorge Cano contributed to the acquisition of environmental data sets from multiple sources. I cleaned and prepared the HMIS data as well as health facility geo-location data and extracted and prepared all the environmental and other covariates (freely available online), besides generating all the geographic datasets involved in this study. I analysed the data under

the guidance of Benn Sartorious for this work and wrote the manuscript for publication with input from all co-authors.

SIMON P. KIGOZI, RUTH N. KIGOZI, CATHERINE M. SEBUGUZI, JORGE CANO, DAMIAN RUTAZAANA, JIMMY OPIGO, TEUN BOUSEMA, ADOKE YEKA, ANNE GASASIRA, BENN SARTORIOUS, & RACHEL L. PULLAN (*Submitted*) Spatial-temporal patterns of malaria incidence in Uganda using HMIS data from 2015 to 2019

Chapter 7:

This chapter provides a summarized synthesis of findings from this work, general conclusions drawn from it, and recommendations as well as further research gaps identified

List of Acronyms

ACT – Artemisinin-based Combination Therapy

CDC – Centres for Disease Control

CRT – Cluster-randomised Control Trial

DHIS-2 – District Health Information System – version 2

DHO – District Health Officer

DSS – Demographic Surveillance Site

EANMAT – East Africa Network for Monitoring Antimalarial Treatment

GAMM – Generalised additive Mixed Model

GLM – Generalised Linear Model

HMIS – Health Management Information System

HRP2 – Histidine-rich Protein 2

iCCM – integrated Community Case Management

IDRC – Infectious Diseases Research Collaboration

IPTp – Intermittent Preventive Treatment during pregnancy

IRS – Indoor Residual Spraying with insecticide

ITN – Insecticide Treated Net

LAMP – Loop-mediated isothermal Amplification

LLIN – Long Lasting Insecticidal Net

LSHTM – London School of Hygiene & Tropical Medicine

MAP – Malaria Atlas Project

MAPD – USAID's Malaria Action Programme for Districts

MARA – Mapping Malaria Risk in Africa

MDA – Mass Drug Administration

MIS – Malaria Indicator Survey

MoH – Ministry of Health

MPF – Malaria Positive Fraction

mRDT – Rapid Diagnostic Test for Malaria

NMCP – National Malaria Control Programme

OPD – Outpatient Department

PCR – Polymerase Chain Reaction

PFP – Private for Profit

pLDH – Parasite Lactate Dehydrogenase

PMI – President's Malaria Initiative

PNFP – Private Not for Profit

RBC – Red Blood Cells

SMR – Standardised Morbidity Ration

TPR – Test Positivity Rate

UBOS – Uganda National Bureau of Statistics

UCSF – University of California San-Francisco

UMSP – Uganda Malaria Surveillance Project

UNCST – Uganda National Centre for Science and Technology

VHT – Village Health Team

WHO – World Health Organization

1 Background and Introduction

1.1 Background

Malaria remains a significant global public health challenge with sub-Saharan Africa and South East Asia as epi-centres of the burden [1]. Global malaria control efforts are multi-dimensional and include: vector control, effective malaria case management, vaccine development, preventive therapies, and above all, stakeholder commitments [2, 3].

The WHO Global Technical Strategy for Malaria 2016-2030 aims to reduce incidence of malaria by at least 90%, particularly by urging affected countries to make the most of available control tools and strategies [4]. Further, in view of United Nations' third sustainable development goal that seeks to ensure healthy lives and promote well-being for all at all ages, one key target is to end the malaria epidemic by 2030 [5]. These targets were heavily influenced by evidence of significant declines over the first 15 years of the 21st century and on this basis, milestones were set to reduce case incidence by 40% and 75% by 2020 and 2025, respectively [3, 4]. Unfortunately, however, malaria burden declines have stalled since 2016 due to global or context specific causes [6, 7]. Two of the identified possible causes that especially affect sub-Saharan Africa are: substandard performance of health systems and weak surveillance, monitoring and evaluation with which capacity to identify program coverage gaps or disease burden changes is diminished [4]. This thesis addresses the latter.

With strong evidence of the effectiveness of available control tools [8], to meet global targets, interventions need to be prioritised to target areas of greatest need, aided by strategic transformation of surveillance into a core intervention [4]. Routine health management information systems (HMIS) data is uniquely placed for this, given: its central place in surveillance, its spatial scalability, and longitudinal dimension. Notably, however, several studies have suggested these data to be imperfect and of limited utility [9-11]. This ongoing perception unwittingly hinders the ability of malaria control programmes to use routine health systems data for effective resource allocation or timely intervention impact evaluations. Whilst efforts have been undertaken to improve the most notable drawbacks, especially accessibility, timeliness, and completeness [12-14], estimates of burden from these data are not fully understood [15] and as such, neither have the prevailing perceptions been improved nor its likely utility been widely investigated.

This thesis, therefore, focuses on exploring the utility of indicators of uncomplicated malaria burden from routinely collected health facility data, using the high-burden example of Uganda. In this chapter, I provide an initial background literature review describing the epidemiological and public health situation of malaria in Uganda, as well as details on current diagnostics and control strategies for malaria. I then summarise the distribution of malaria in Uganda and provide a critique of contemporary mapping approaches applied at global, regional, and sub-national scales. Lastly, I provide an overview of how maps have historically been used in Uganda towards policy guidance and decision making for malaria control, in relation to other countries in the region.

1.2 Introduction to malaria

1.2.1 Global burden of malaria

Malaria is transmitted by female anopheline vectors carrying any of the four main *Plasmodium* parasite species known to infect humans - *P. falciparum, P. vivax, P. malariae,* and *P. ovale* [16]. Notably, however, two of these parasite species i.e. *P. falciparum* and *P. vivax* are responsible for the majority of global infections [3, 16]. While these main parasite species are largely territorial, with *P. falciparum* predominating Africa and *P. vivax* East Asia, mixed infections involving the two or one of these together with other less notable species are also common across all endemic settings [3]. In 2018, 228 million malaria cases were estimated globally, 93% of these from Africa alone [3]. Moreover, an estimated 405,000 fatalities from malaria were also reported globally, 94% of which were from Africa, and 67% of the global total being among children under 5 years of age [3].

1.2.2 Epidemiology of malaria

Malaria transmission involves four vital contexts including: the host, which is primarily humans; the parasite of which there are several species; the vector, which is the mosquito and there are many species of these; as well as the environment within which all the first three exist. Factors that influence any of the four contexts may impact the rate of transmission of malaria either independently or collectively, both favourably and otherwise. Successful transmission involves all four contexts as follows. As illustrated in Figure 1, once a healthy vector, female anopheline, takes a blood meal from a human and picks up gametocytes in that meal, gametocytes undergo transformation within the vector from micro to macrogametes which in turn are transformed to the zygote and then ookinete that penetrate the midgut of the vector [17]. Within the vector's midgut, the ookinete is transformed to oocysts which develop and burst into the salivary gland to produce sporozoites [18]. With a sporozoite ready vector, a blood meal from a host is potentially infective of the host, which marks the start of the parasite life cycle within the human host. Once sporozoites sufficiently circulate within the host's blood, they are transported to the liver where they develop into schizonts that later produce merozoites that are then introduced back into the blood from the liver [19]. Merozoites attack the host's red blood cells (RBCs) in order to reproduce and then attack more RBCs though some merozoites develop into gametocytes that are known as the sexual stage of the parasite. Once gametocytes are ingested by a viable vector, the cycle starts all over again within the vector and continues the process of malaria transmission.

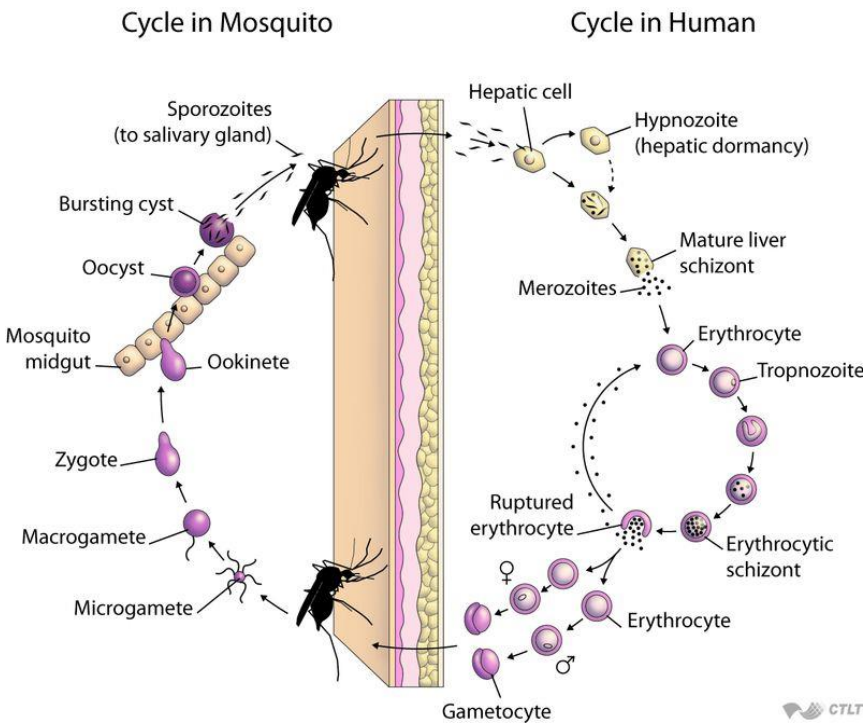


Figure 1. Malaria parasite life cycle

Image obtained from the Johns Hopkins School of Public Health at http://ocw.jhsph.edu. *Creative Commons BY-NC-SA.*

Environmental factors play a significant role particularly in supporting vector abundance and capacity [20, 21]. For instance, rainfall in appropriate amounts and locations may enable the availability of vector breeding sites and thereby foster vector density. On the other hand, temperature when conducive facilitates vector development, adult survival and immunity, as well as parasite development within vector candidates (conducive within the range of 16 to 35^0C), thereby facilitating a competent vector for continued transmission [18]. Whilst rainfall, temperature and humidity tend to have a direct influence, other factors such as vegetation, urbanization, altitude, land use and cover may have secondary influence on vectors and vector capacity through their influence on direct factors and facilitating vector-host contact [22]. Given the variability of environmental factors across space and time, the unlimited interplay between multiple environmental factors facilitates and supports diversity in vector, vector habitat and behaviour, which may influence heterogeneity of malaria transmission and risk [23].

Together, these factors influence the distribution of malaria burden through a non-linear interaction between: environmental suitability for vector abundance and competence; host susceptibility to virulent parasites as well as host infectivity to vectors; population-level control activities and subsequent adherence to these control strategies; community population distribution; and, availability of, or accessibility to healthcare services and adequacy of case management commodities. These are augmented by the implementation of systems to collect timely, high quality and accessible routine data in synthesizable formats to support informed onward control decisions.

1.2.3 Detection and Diagnosis of malaria

Malaria diagnosis in endemic settings has undergone massive transformation over the years. For very long, diagnosis of malaria was performed presumptively especially among children [24]. However, this approach was increasingly associated with over-treatment of fever as malaria in many countries, including Uganda, due to the non-specific nature of malaria related symptoms (particularly fever) that are often caused by myriad other conditions [25]. Moreover, parasite resistance to antimalarials, particularly involving the fairly cheap and previously highly effective drug Chloroquine, globally [26] and in Uganda [27], showed that trends in over-treatment of fevers as malaria with newer antimalarials – ACTs, were a threat to the longevity of the high efficacy of these much more tolerable drugs [28]. To this effect, global recommendations were made for the use of diagnostic confirmation prior to treatment [29]. These facilitated the scale-up of research into diagnostic methods, aimed at overcoming shortcomings in the pre-existing testing method of microscopy.

Whilst detection of malaria parasites had been possible for hundreds of years using blood slide microscopy, the method is demanding, particularly for low resource settings. This gold-standard method requires a microscope, electric power supply, slides, reagents, and importantly a skilled technician. With several of these requirements being in short supply across the highest endemicity regions, diagnostic confirmation of malaria to scale using this method was unattainable. Moreover, other molecular methods in existence such as polymerase chain reaction (PCR), loop mediated isothermal amplification (LAMP), flow cytometry, and mass spectrometry, though highly sensitive are far more expensive and therefore, not among feasible alternatives within clinical practice in these settings [25]. Newer approaches involving rapid diagnostic methods of detecting malaria antigens were developed and introduced. The four major categories of the rapid diagnostic tests for malaria (mRDTs) developed included: *P. falciparum* specific histidine-rich protein 2 (HRP2); parasite lactate dehydrogenase (pLDH) that could be produced for each of the four main parasite species, given that each has a distinct isomer of this enzyme; Plasmodium aldolase, another that covers all the parasite species; and, another antigen specific to *P. vivax* that has been used in combination tests for *P. falciparum* and *P. vivax* [30]. The ease of use of mRDTs even among remote facilities and community health workers [31] has facilitated largescale implementation of the test and treat global approach [29], that was later revised to the test, treat, and track policy for improved surveillance and care or case management [32].

Until 2007 when mRDTs were introduced in Uganda, diagnostic testing for malaria depended on microscopy, particularly among adults in hospitals and high-level health facilities, where laboratory services were functional [33]. Among children under 5 years of age when febrile, presumptive diagnosis was highly encouraged and functional laboratory services availability among lower level facilities was estimated at only 30% by 2009 [34]. National policy adoption of parasitological diagnosis using either microscopy or mRDT, was instituted in 2011 [35]. Consistent with policy, the national 2010-2015 malaria strategic plan set a target of 90% parasitological diagnostic performance by 2015, and the country had attained 59% in a 2013 assessment [36]. Notably, however,

performance reached 85% later on in 2018/2019 [37], suggesting very slow adoption of the national 'test and treat' policy for management of suspected cases, regardless of age of patient or level and/or ownership of health facility [38]. Additionally, interrupted commodity (drugs and diagnostic testing materials) distribution, disregard of negative test results in the diagnosis of malaria, and insufficient support supervision still remain very concerning for progress [36, 37]. However, poor mRDT performance due to *Pf-HRP2*/*Pf-HRP3* gene deletions has also been reported in the region, particularly given that *Pf-HRP2*-based mRDTs are the recommended test kits in Uganda and these deletions lead to true cases turning out as false negatives [39-41]. The increasing use of mRDTs therefore, may be associated with large-scale reduced sensitivity of diagnostic confirmation of malaria cases.

The increased availability and accessibility of parasitological diagnostic testing has facilitated improved capacity to assess malaria burden from routine HMIS data with more reliable indicator accuracy. Whilst there are several derivate indicators of malaria burden in use, how they relate each with the other remains unclear. Moreover, very few studies have evaluated the effectiveness, utility, or relationships among HMIS-based indicators of malaria burden pairs or between these and indicators from other data sources. One study examined the relationship between current and lagged monthly HMIS-based incidence estimates to explore HMIS capacity for malaria burden forecasting in Burundi's regions with seasonal endemicity, using environmental covariates. Though it included seven years (1997-2003) of routine data and found a strong association between monthly incidence and maximum temperature in the previous month's estimates, the study could only define incidence using predominantly presumptive malaria cases, limiting the reliability of incidence rate estimates used [42]. Another study compared health centre and community survey metrics including *Plasmodium falciparum* (*P.f.*) parasite and gametocytes prevalence as well as seroprevalence among others, between wet and dry seasons in The Gambia. They reported stronger correlation between facility and community parasite prevalence estimates in the wet than dry seasons and noted versatility of and greater ease in collecting health facility than community survey data. Importantly, study sites were spread across the Gambia from coast to hinterland and paired on opposite sides of the national main river, providing good coverage of spatial diversity [43]. Yet another study described a weak link between relative changes in slide positivity and incidence rates over time, from a four-year cohort of children in Kampala - central Uganda. Though conducted at one site, the study straddled a duration of drastic changes in malaria burden having reported significant declines in incidence of malaria from 0.93 to 0.39 episodes per person per year from 2005 to 2009, respectively ($p<0.001$), therefore providing a good setting to understand temporal changes in the metrics compared. Besides not being HMIS-based, however, this study reported an indeterminate relationship between slide positivity and incidence rates - simply describing it as "neither linear nor proportional" [44]. However, another study conducted at one site in Western Uganda revealed a non-linear temporal relationship between test positivity rate (TPR) and HMIS-based incidence at a six-monthly temporal scale. Importantly, this was the first description of this non-linear relationship, best explained by an exponential function (compared to many other models fits) where correlation between the two indicators was stronger at higher

transmission levels [45]. These few studies available underscore the dearth of knowledge of the indicators of malaria burden derived from HMIS data though in wide use.

1.2.4 Malaria control strategies

Vector control has primarily involved the use of long-lasting insecticidal nets (LLINs), indoor residual spraying with insecticide (IRS), and larval stage management (including larvicide use or habitat modification) [46]. Owing to excessive parasite resistance to chloroquine that was widely used through the 1990's, global policy on malaria case management transitioned to other antimalarial monotherapies and then rapidly on to combination therapies, following quick failure of the monotherapies [29]. As regards chemoprevention, however, vaccine trials are in early stages in a few places like Ghana, Kenya and Malawi [46]; preventive therapies including mass drug administration (MDA) to reduce the parasite reservoir in the community [29, 47] and intermittent preventive treatment during pregnancy (IPTp) to address adverse birth outcomes due to malaria in both mother and new-borns [48] are in use. Importantly also, stakeholder commitments and global initiatives have been instrumental in achieving these multi-dimensional control efforts so far. These initiatives have included first, the global eradication of malaria initiative of the 1950's whose biggest success in Africa may have been the wide-scale availability of chloroquine, an effective antimalarial that was associated with reduced malaria mortality in Africa [49]. Others have included the Garki project, Roll back malaria, millennium - and later sustainable - development goals with health at the centre, and the WHO's "high burden to high impact" initiative [2, 50]. Each of these either have been or continue to be informed by available data, including surveillance data.

In Uganda, malaria is perennial and endemic in over 95% of the country, given prevalence of a diverse and versatile composition of competent vectors [51]. The main vector species in the country are *Anopheles gambiea* and *A. funestus* with some *A. arabiensis* [52-54] and predominant vector control methods have included LLINs in universal distribution campaigns and IRS in selected districts [55]. These have been consolidated by effective case management using artemisinin-based combination therapy (ACT) as first line treatment since 2004 [56, 57], on top of IPTp using Sulphadoxine pyrimethamine (SP) since 2001 [58]. While malaria risk remains high and widespread across the country, Uganda has reported considerable declines in malaria burden over time due to these interventions. For instance, national prevalence estimates declined from 42% during the Malaria Indictor Survey (MIS) of 2009 [59] to 9.1% from the most recent survey of 2018 (Figure 2), consistent with global and regional reported downward trends.

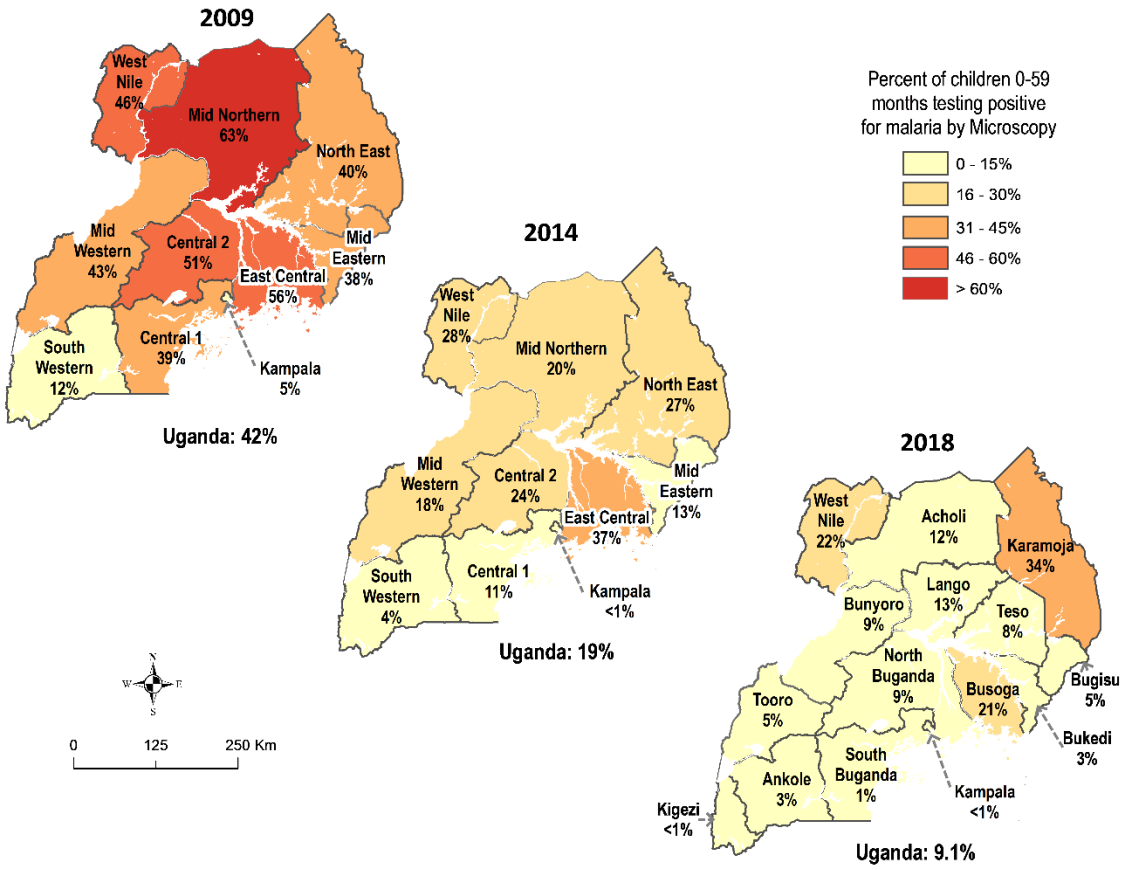


Figure 2. Parasite prevalence by microscopy among children 0-59 months of age based on the 2009, 2014 and 2018 Malaria indicator surveys.

The 2009 MIS was the first national malaria indicator survey conducted in Uganda covering ten defined regions of the country. Results showed that prevalence of malaria parasitaemia by microscopy among children under five years of age ranged from 5 to 63% in Kampala and Mid-Northern regions, respectively [59]. The 2014 MIS suggested a reduction in the prevalence of malaria parasitaemia in the same age group ranging from <1 to 37% in Kampala and East Central regions, respectively [60].

The 2018 MIS (third and most recent survey) covered 15 regions and recorded further declines in the prevalence of malaria parasitaemia, ranging from <1 in Kampala and Kigezi to 34% in Karamoja [61]. There was a marked decline in national parasite prevalence by microscopy from 42 to 19% for 2009 to 2014-15, respectively and then down to 9% during 2018-19.

Overall, whilst regional boundaries changed over time, reduction was still evident across all regions. For instance, prevalence of malaria parasitaemia reduced in Kampala from 5 to <1% and in the mid-northern region from 63% to a regional average of 13% between 2009 and 2018, respectively.

By WHO reports, Uganda ranked 3rd largest contributor of cases and 7th of malaria related deaths by 2018 [3], down from 4th in terms of number of malaria cases and 11th in terms of number of malaria related deaths by 2015 and 2016 [62, 63]. Nevertheless, national HMIS-based reports have documented declines in incidence of

confirmed cases down to 14 cases per 1000 population per year, in 2018/19 compared to 478 cases per 1000 in 2015/16 [37].

1.3 Understanding the distribution of malaria in Uganda

Geographical representation of the distribution of disease burden and/or risk is critical in understanding and designing plans of action to minimise public health disease impact. Our understanding of the geographical distribution of malaria in Uganda has been informed by various sources of data. Historically, these included data from small available studies across the country coupled with expert opinion, which served a purpose in the absence of robust national datasets to generate more representative maps [64]. These could only provide a general overview of the distribution of malaria with very limited capacity to inform targeted control and therefore, hardly put to known extensive use. More recently, data from large malaria indicator surveys (such as the 2009, 2014-15, and 2018-19 rounds) have been utilised for mapping the distribution of malaria, forming the primary basis for geographical burden reference. These, however, may only reliably inform the coverage of previously implemented interventions, treatment seeking practices among one high-risk group of children under five years of age, and provide some indication of general malaria endemicity strata by region [61]. This limitation is determined by the cluster-level sampling design (based on 10 to 15 regions of the country) of these infrequent surveys, implying that results are principally limited to regional summaries, less helpful for local onward planning.

For on-going control activities within the Ministry of Health (MoH), HMIS was instituted with the objectives of supporting evidence-based decision making, setting performance targets, and assessing health sector performance [65, 66]. Data summaries in the form of trend plots and other dashboard summary outputs are assessed within the district health information system (DHIS-2) framework, that provides the necessary data [14]. These are supplemented by reports and information from development partners and stakeholders such as: the World Health Organization, the United States' Centres for Disease Control/President's Malaria Initiative (CDC/PMI) [67], the Uganda malaria surveillance project (UMSP) conducting sentinel surveillance and providing regular reports [68] and the USAID's malaria action program for districts (MAPD) operational across a network of districts through the convergence of a variety of expertise in Uganda to support MoH efforts in control and diagnosis of malaria [69], among others.

The extensive focus on regional or district level assessments, coupled with reported disconnect between survey-based and on-going HMIS reports [70, 71] indicates that presently, malaria control managers are without a reliable source of fine-scale information. Consequently, the potential for important timely assessment of the spatial distribution of malaria burden, using HMIS data, remains unappreciated, and opportunities for improved decision making are missed.

1.3.1 The use of maps in policy and decision-making

Historical use of malaria risk maps in Uganda is limited. Figure 3 below, for instance, was used for nearly a decade in official malaria policy reports in Uganda, including multiple national malaria strategy documents [55, 72]. The map (Figure 3) would have been generated in the early 2000's from available data at the time. It was first used in the 2005-2009 Malaria Strategic Plan by Ministry of Health referring to it as "most recent one based on available data" [57, 72].

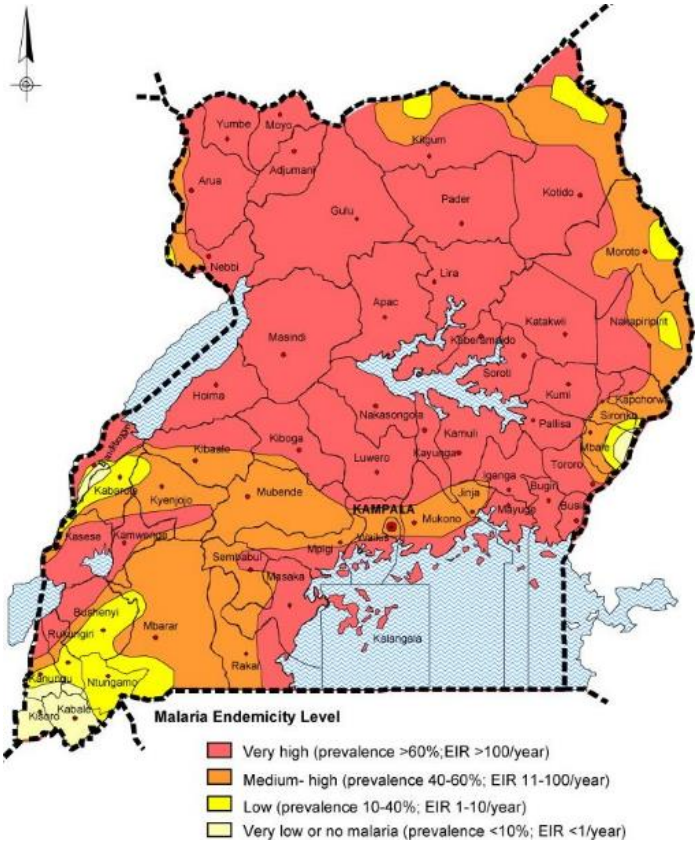


Figure 3. Risk map used between 2005 and 2014 – adapted from Talisuna et al. [64]

This malaria risk map was generated using data availed from small studies, two of which were: (1) A drug efficacy study under the East African Network for Monitoring Antimalarial Treatment (EANMAT) conducted in seven locations including: Arua, Apac, Tororo, Mubende, Kabarole, Rukungiri and Jinja [73]. This study involved surveys conducted between September and December 1999. (2) An entomological study that included the same EANMAT sites, where 11 entomological surveys involving mosquito collections by human landing collection method, was conducted between June 2001 and May 2002 [53]. Together with these data, other historical data from the 1960's were also used to inform the final output [64].

Subsequent risk maps used in MoH documents (Figure 4), however, were generated via geo-statistical models with mean population adjusted *Plasmodium falciparum* parasite rates, among children aged 2 to 10 years old from

surveys conducted between 2000 and 2010 across the country, together with a selection of climatic metrics as explanatory variables [72]. The role of age in malaria transmission is highlighted here as being strongly associated with parasite rates, attributable to acquired immunity [74, 75] due to manifold exposure. This approach of using an age standardizing algorithm to control the effect of varied age ranges on detectable infection rates in a particular age-range, is classical with risk mapping across the endemic world [74, 76], and is applied to both *P. falciparum* and *P. vivax* wherever they predominate [77, 78].

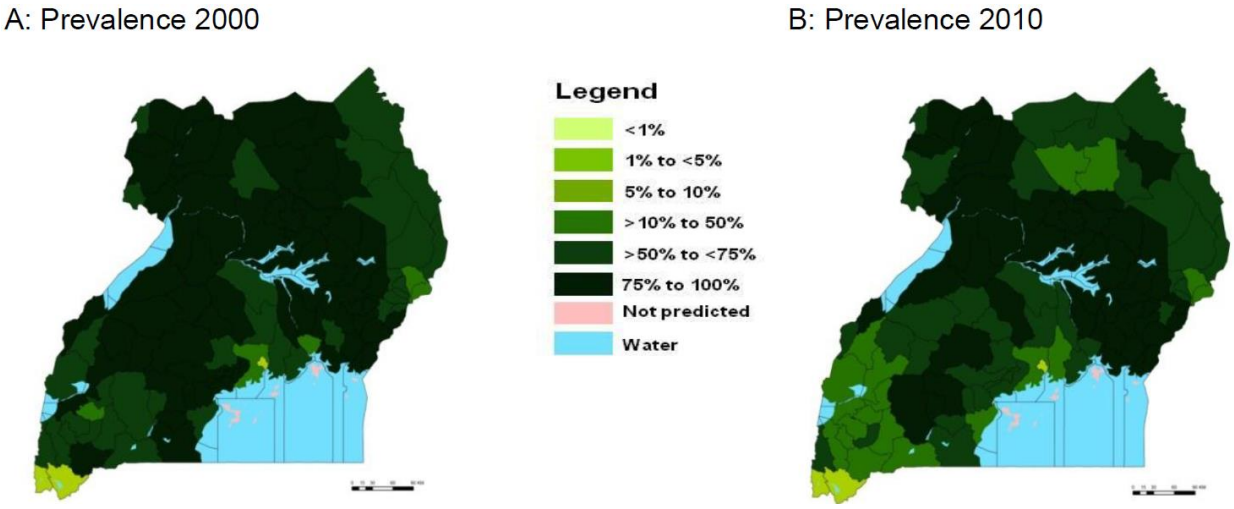


Figure 4. Malaria map in use by the Ministry of Health between 2014 and 2017, in multiple policy reports of malaria risk representation in Uganda

While use and inclusion of malaria risk maps in official MoH documentation is increasing, maps in previous use were seldom updated with a single risk map used across multiple years [72]. Moreover, these recent malaria risk maps at these district spatial scales [79, 80] have been recognised as difficult to use for intervention implementation, potentially due to masking of important fine-scale heterogeneity and thus undermining effective response action [64].

However, progress in using routine data for risk maps is evident in the MoH's national annual report of 2017/18 (Figure 5), which included HMIS-based incidence figures presented for comparative year-to-year progress [37]. Furthermore, the soon to be launched national Malaria strategic plan 2021 – 2025 for Uganda has proposed a shift of focus from universal to targeted implementation of control interventions under the 'High Burden to High Impact' initiative. Importantly, the included new map of district-level malaria incidence from 2019 routine reported data was cited as a key input in this decision process. Here, districts were stratified by specific combinations of control tools for intervention, in response to WHO advice in the national bid for malaria funding, "to use strategic information to drive impact" (Figure 6) [81]. This provides an indication of recent utility of malaria burden maps for decision support in Uganda.

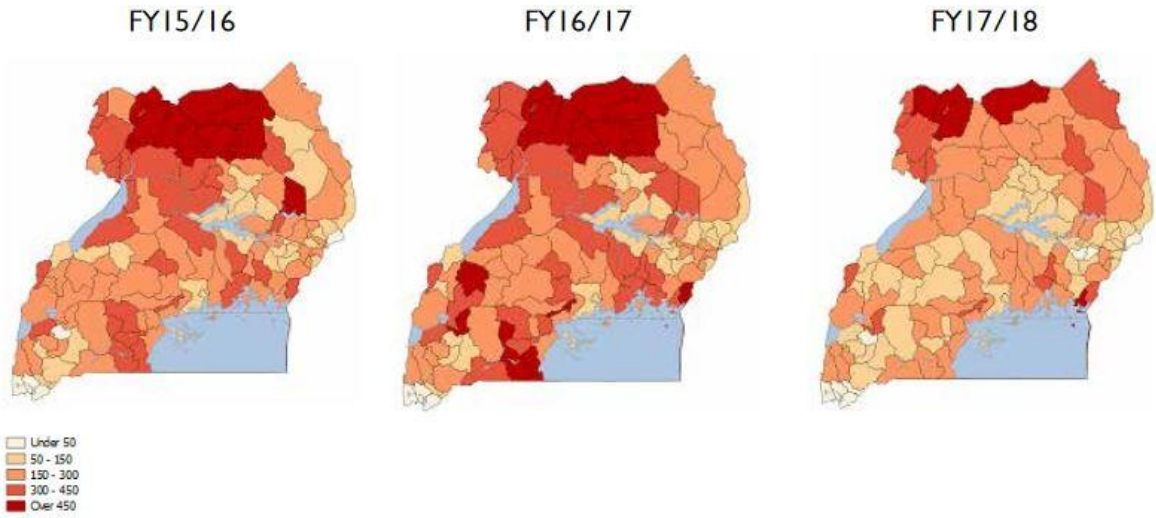


Figure 5. Map of malaria incidence rates in Uganda in use by 2017/18 from the first national annual malaria report.

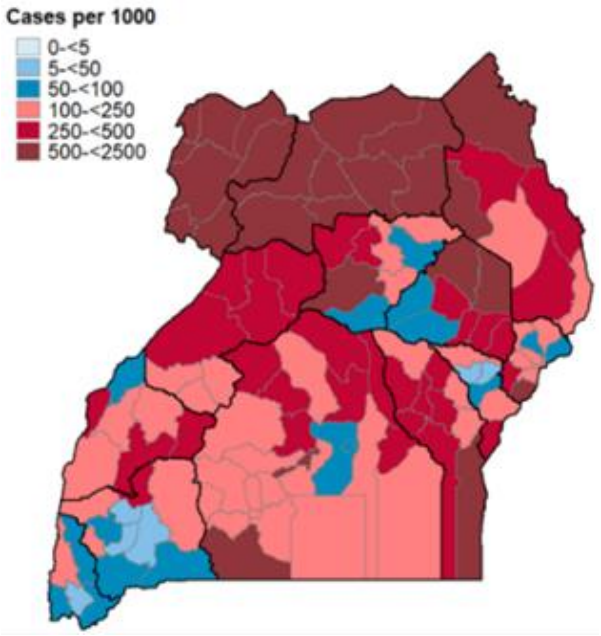


Figure 6. Malaria incidence rates by district as estimated from 2019 routine reported data.

This map provided some evidence of the distribution of malaria burden by district across the country, which was reported as vital to the determination of district strata for targeted intervention approaches. These interventions are intended for implementation during the 2021-2025 national malaria control strategies for Uganda supported by Global Fund, among others.

Though challenging to evaluate fully, particularly for day-to-day activities, the use of risk maps for decision support in Uganda may be otherwise demonstrated by the inclusion of these maps in national health reports and may also suggest an increasing appreciation of geo-spatial output for malaria control in Uganda. However, for their viability as an important tool for surveillance support, risk maps remain heavily underutilised.

1.4 Contemporary mapping approaches

Population-based prevalence surveys: Maps of malaria risk support decision-making for control and intervention, especially concerning geographical scope and feasibility. Typically, these maps are developed using population-based surveys due to their simple and rapid representation of disease prevalence [82, 83]. Whilst these survey-based estimates of burden are only generalizable to regional scales, often among very large regions [84], geo-spatial modelling approaches have been developed and used to improve inference at finer spatial scales. In this process, parasite rates are utilised together with environmental predictors (explanatory variables) in statistical models that predict disease burden estimates associated with geographical variability, known as geo-statistical models. From these models, parasite rates and other associated indicators are interpolated at un-sampled locations, and often output as map surfaces or images. Explanatory variables can include rainfall, vapour pressure or humidity, temperature, vegetation amounts, land use or land cover, land surface moisture, elevation, and their derivatives [85, 86].

Using a comprehensive collection of survey data spanning decades, through formal and grey literature databases and contacts with research scientists and officials globally, global malaria burden maps have been generated using multiple derivative indicators within the malaria atlas project (MAP) [83]. These maps have provided valuable information especially for global endemicity stratification overview and distribution of parasite specific burden, which have aided large-scale intervention planning. A notable milestone of this work, for instance, was the identification of regions where liver-stage infection clearing anti-malarial drugs like primaquine would be beneficial or harmful due to prevalence of the Duffy negative blood group phenotype [87]. Whilst this blood group variant largely confers protection against *P. vivax* infections where prevalence of the phenotype is high, individuals are not totally immune to vivax infections that are characterised by relapses of malaria due to uncleared infections in the liver [88, 89]. Ill-advised treatment of these infections with this effective drug for liver stage parasite clearance poses a risk among individuals with this blood group variant. The analyses showing spatial distribution of this blood group variant, therefore, have been important in the design and implementation of region-appropriate policies. These approaches have also been adopted in the Mapping Malaria Risk in Africa (MARA) project, which implemented geostatistical models to generate point prevalence-based risk maps for the sub-Saharan African region and provided survey data from across the region for similar studies [86, 90]. However, limitations of geostatistical outputs, such as these, include: infrequency and sparsity of surveys – for instance only eight countries provided 100 or more survey sites and a large majority of countries far fewer than 50; large differences in timing and seasonality of the surveys; varied age of participants; design, size and generalizability of surveys included; and, potential underrepresentation of specific parasite species surveys by region – for instance, very few *P. vivax*-specific surveys in Africa or *P. falciparum*-specific surveys in South East Asia were included [82].

Whilst geostatistical approaches were historically computationally intensive for high precision of modelled estimates, particularly with Markov Chain Monte Carlo simulations for Bayesian inference, increasingly, a more

summarized and computationally efficient approach in integrated nested Laplace approximation (INLA) for Bayesian inference has been adopted [91]. Besides lingering computational demands, however, the capacity to incorporate maximum likelihood, prior information [92], and the neighbourhood structure through conditional autoregression [93-95] for model estimates has not only facilitated identification of important environmental factors for malaria risk assessment such as rainfall, temperature, and vegetation, but also the credible presentation of geographical patterns of malaria risk from both survey data [92, 96, 97] and routine HMIS data [98-101] across endemic settings for varied ages.

Additional robust but less common methods used with routine data for risk prediction include: (a) Plotting annual parasite rates from routine reported data at as low spatial resolution as village-level in one district of Sri Lanka between 1991 to 1998 [102]. (b) Generalised linear models (GLM) to predict the effects of environmental predictors in Burundi using province-level monthly estimates of incidence from routinely reported malaria cases between 1996 and 2007 [103]. (c) generalised additive mixed models (GAMM) that provide improved model fitting, with similar results to, though more complex than GLM output, that is demanding to interpret [103]. Despite agreement between these two models, results also indicated that variables other than climate are also very important and should be accounted for. (d) Using the same routine data from Burundi, geo-additive mixed models suggested an improvement on GAMM owing to inclusion of more explicit spatial effects – both correlated and un-correlated at provincial level [104]. (e) Seasonal autoregressive integrated moving average models were used to forecast incidence using key environmental factors, particularly rainfall in Eritrea using monthly incidence estimates from routine data between 2012 and 2016, with recommendations for small area assessments [105].

1.4.1 Mapping malaria burden using routine data

Despite the recent embrace of routine data for generating risk maps in Uganda, there is recent but rather sparse precedent of use of this approach in the region. For instance, a report from Rwanda showed maps of malaria positivity rates as well as incidence for 2010 and 2011 as shown in Figures 7 and 8, respectively [106], and one from Mozambique showed reported inpatient incidence of malaria over the 2010-2012 duration (Figure 9) [107].

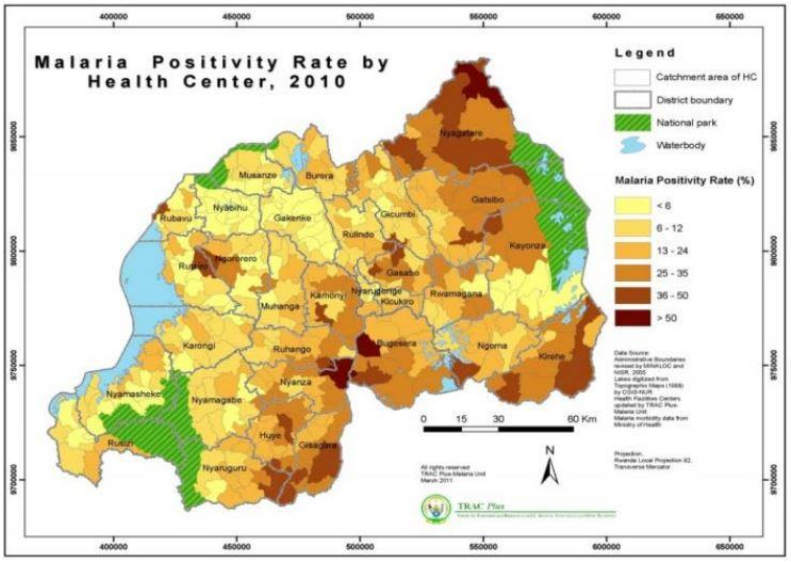


Figure 7. Map of Rwanda malaria burden using test positivity rates for 2010 from PMI evaluation report of 2016

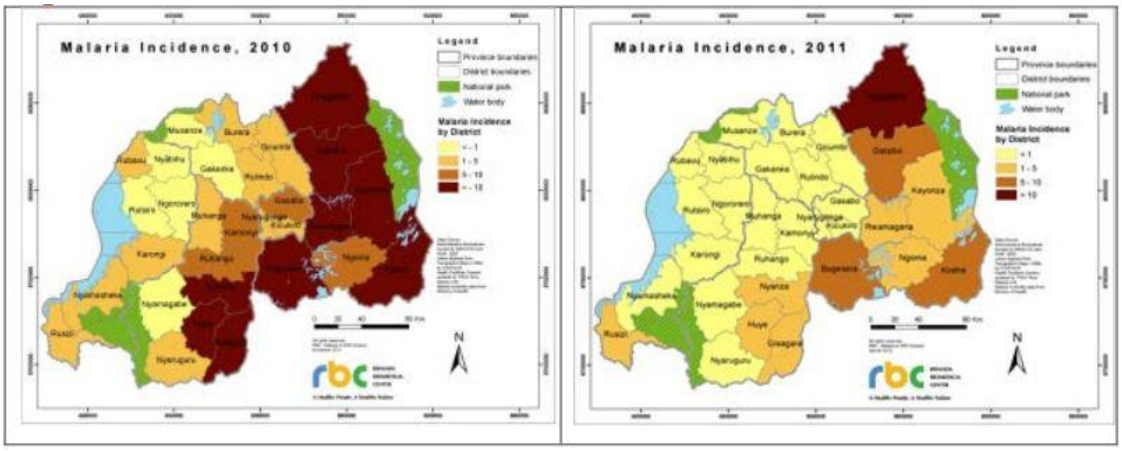


Figure 8. Map of Rwanda malaria burden using incidence rates by district for 2010 and 2011 from the PMI evaluation report of 2016

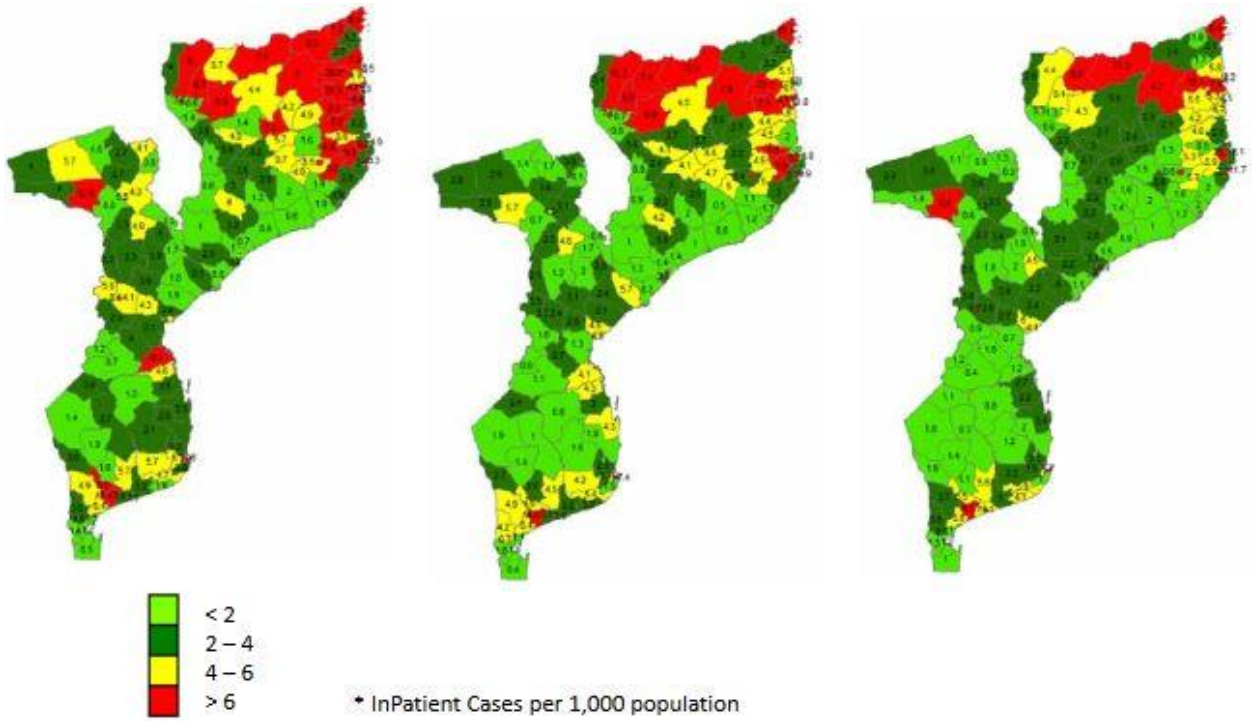


Figure 9. Map of malaria inpatient incidence rates for Mozambique by district for 2010, 2011, and 2012 from the PMI evaluation report of 2016

Unsurprising with minimal utility of HMIS for risk mapping, HMIS-based risk maps have only previously been compared with robust survey-based approaches in very few studies. One study from Malawi investigated the importance of climatic, geographic, and socio-economic determinants of malaria between July-2004 and June-2011 and reported one such methodological comparison [108]. HMIS-based "standardised morbidity ratio (SMR)" of malaria and prevalence from the malaria atlas project (MAP) were compared by visual examination of a map from each. Whilst the spatial distribution of SMR from this study largely reflected the prevalence distribution from MAP for children under 5 years of age, the stark differences found between the two for those 5 years and older may be due to additional effects of age on malaria transmission. These effects potentially remain unexplained and/or unaccounted for in the current survey-based models of burden estimates heavily reliant on data collected primarily from children under 5 years of age [109]. Finding one study that evaluated use of routinely collected data for risk mapping, against more established mapping methods, points to a knowledge gap in fitness-of-purpose of routine data, as a potential low-cost alternative for malaria risk assessment to support optimal resource use.

Regardless of the data used, however, for any spatial temporal distribution of malaria identified to be beneficial, it may need to address some important questions as proposed by Carter et al. These include: "1) Is it operationally possible to reliably distinguish spatial clusters with markedly different malaria case incidence and to determine the locations and extents of all the foci of malaria transmission in a locality? 2) If achieved, can the information be exploited in order to conduct highly effective malaria control by the accurate targeting of an intervention? 3) What

tools for control could be more effective using the generated spatial information? 4) In which situations of endemic malaria is targeting practical and effective and in which is it not?" [110]. To aid disease burden monitoring and control intervention implementation and/or targeting, however, if well understood HMIS data may be a great choice to facilitate assessments that address most of these questions. In the following section therefore, I provide a detailed discussion of HMIS data available in Uganda including the indicators reported, strengths and weaknesses, and its use for impact assessment.

1.5 Routine surveillance and HMIS

1.5.1 Routine reporting of malaria indicators

The WHO has defined routine surveillance as continuous, systematic collection, analysis and interpretation of health-related data for planning, implementation, and evaluation of public health practice [111]. Identified benefits of surveillance include: serving as an early warning system for impending public health emergencies; documentation of impact of intervention, or tracking progress towards specified goals; and, monitoring and clarifying the epidemiology of health problems, to allow priorities to be set and thereby inform public health policy and strategies [111].

Regularly submitted reports to the Ministry of Health that contribute towards malaria routine surveillance emanate from sources such as: implementers of health-related activities like LLIN distribution campaigns; supervision activities conducted by national malaria control programme (NMCP) managers; and, disease surveillance reports from health facilities, all using standardised report formats [112]. Disease surveillance through health facilities in Uganda includes several key activities. First, integrated disease surveillance and response (IDSR), in which data on cases and deaths are reported on a weekly basis to facilitate epidemic detection and/or preparedness [36, 113]. Second, sentinel surveillance programme whose primary objective is to monitor trends using test positivity rates as a key indicator, along with increasing diagnostic testing [68]. Third, demographic surveillance sites (DSS) that include two selected communities for monitoring defined populations on demographic metrics such as births, deaths and migration [114]. Fourth, pharmacovigilance, although this has largely been out of operation [36]. Lastly, outpatient department (OPD) monthly reporting on malaria cases through HMIS form 105 that is central to this research, where malaria reporting primary includes: total monthly reported and confirmed cases, and number of suspected malaria cases tested either by microscopy or mRDT, all categorised into pre-determined age-groups [112].

Whilst there is evidence of use of HMIS data in spatial modelling to identify high burden locations [101, 115] and HMIS data forms the basis of national day-to-day decision making in Uganda, it has not been adopted for national risk mapping, particularly with small area approaches as described above. When considering its utility, it is important to understand both the opportunities and challenges this data source provides.

1.5.2 Contextual framework, Opportunities and Challenges of routine HMIS data

Optimal utility of HMIS data not only requires the identification and harnessing of its strengths as well as identification and mitigation of its weaknesses but also full understanding of contextual factors influencing the records within the HMIS.

Concerning the contextual factors, HMIS records may be assumed to be influenced at three main levels. These levels interact in a predominantly hierarchical flow, though upward influences may also exist. They include: the political system, health system, and community levels. Perceived relationships between these levels of influence, as identified for this study, are presented in a summarised conceptual framework below (Figure 10).

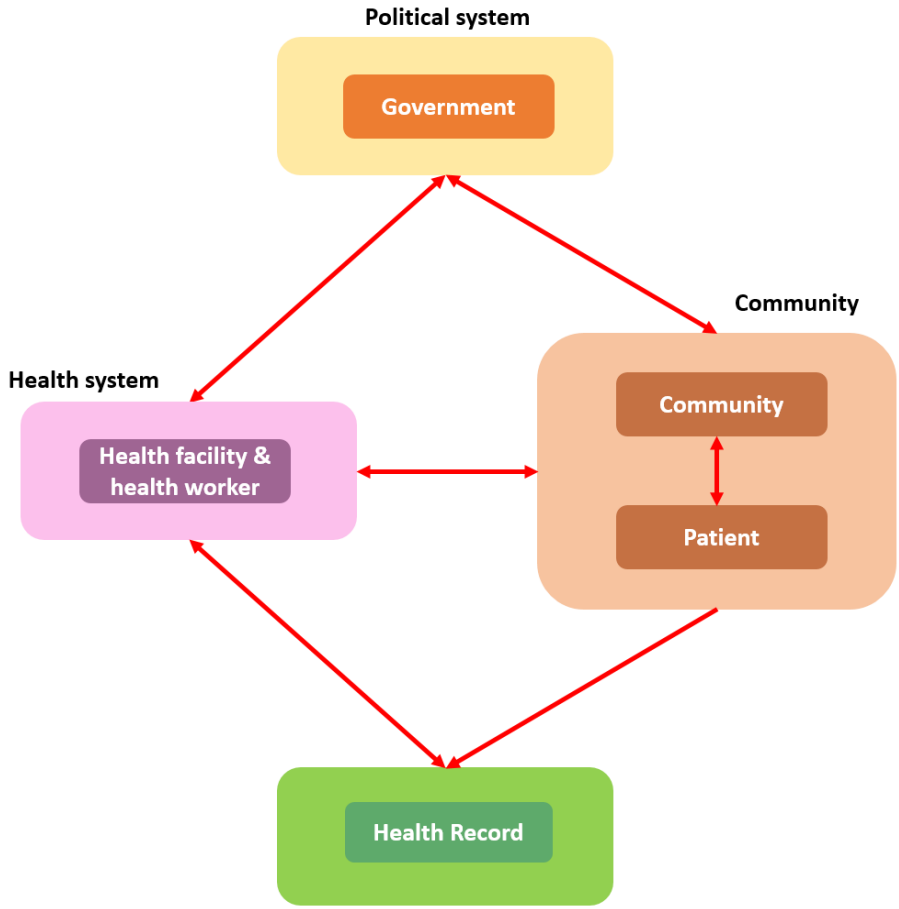


Figure 10. Summary of proposed conceptual framework defining HMIS records.

The major sources of influence that may affect or determine what gets recorded in the OPD or other HMIS registers are broadly categorised into three sources of important factors including: the political, governance, and health financing; health facility, health worker, or localised health system; community or catchment served by the immediate health facility; and, the patients visiting a given health facility.

Additionally, the factors that may determine quality of records at the health facility, which are the basic building blocks of HMIS data, are briefly described below under each of the identified levels of the contextual framework.

The political system (highest level) is characterised by the political environment, healthcare policy, and health financing factors and has overarching influences defining the working environment of the health facilities within the health system. These determine the available services or resources at a given hospital and play a key role in its functionality. At community level, influential factors may be due to the transmission setting, occupational culture within the community, health seeking and community culture surrounding health care that may impact on individual decisions, geographical attributes, and general community context. Proximal to the records, from the community are individual patients that may also directly influence the data recorded, based either on their perceived importance to providing good information to the health workers or their state of illness when they visited. Also, proximal to the records is the health system level that may influence patient, health worker, and the health facility itself. Heywood and Boone classified three levels of influence on health records characterised by demand for and benefit from use of good health records [116]. These include: the beneficiary-level, involving clinicians that need data to follow up patients and monitor their improvement; facility-level, where managers need data for infrastructure and resource improvement; and system-level, where district and national leaders need data to monitor and plan for services delivery. However, these seem to downplay the role of the community which may influence records through community narratives on the available health system, among others.

Collective understanding of (1) the contextual factors influencing HMIS records that need consideration, (2) available opportunities within HMIS data to be harnessed, and (3) prevailing challenges in HMIS to be mitigated, is central to both HMIS improvement efforts and accurate interpretation of indicators of burden derived. This is important for full implementation of the global strategy of transforming surveillance into a core intervention and the ultimate realization of global 2030 malaria targets.

Below, I provide a more detailed breakdown of the opportunities and challenges of routine HMIS data, specifically for malaria surveillance. The opportunities include:

- ***Scalable temporal and spatial resolution:*** Compared to many other sources of malaria case data, HMIS provides unmatched temporal coverage for multiple purposes. For instance, the Uganda NMCP conducts integrated disease surveillance and response (IDSR) using weekly reports to assesses disease epidemics and routine surveillance using monthly OPD HMIS reporting to monitor general trends [36]. However, for any practical purposes, temporal assessments are possible from daily to multi-year scales in HMIS unlike any other study design. Considering spatial scales, HMIS affords both national and regional scales as with indicator surveys. Moreover, given that routine interventions are currently conducted at district level making it the focus in Uganda and elsewhere thus far [36], HMIS has been widely used at this scale [11, 13, 80]. Importantly though, lower spatial scales' assessments of disease burden are also possible with HMIS [102] and I explore this further in Chapter 6 of the thesis.
- ***Comprehensive coverage of age:*** The most common assessments of malaria burden that use small-scale, national cluster-level indicator, and demographic health survey data mainly focus on children under 5 years

of age and seldom on the 5-15 or over 15 years of age if at all. These age-restricted estimates are either largely assumed sufficiently indicative of the scope of malaria burden as seen from indicator surveys [61, 96] and often, population-level estimates are modelled from these [1, 3, 12, 71, 83, 117]. This paradigm seems to downplay any effects of age, particularly older age, in the epidemiology of malaria thereby under-estimating burden [118] or its effects on control efforts and I address this further in Chapter 4. HMIS data, however, covers the full community age distribution making it a richer source, likely to afford more balanced and/or accurate estimates of burden or risk.

- ***Multiplicity of proxy burden measures:*** Whereas malaria incidence rate is primarily defined as number of new cases per duration, divided by total person-time of population observation [119] using cohort studies, these studies are costly. Several proxy measures from routine HMIS data are used for malaria mapping. These have included: (a) Case numbers, either taken as a proportion of estimated population at the time [94, 101, 103, 104, 120-123], or a standardised morbidity ratio [108]; (b) malaria positive fraction (MPF), as a measure that controls for differences in access to care [115, 124, 125]; (c) malaria cases as a proportion of treatment events at a facility [125-128]; (d) case-control analysis of disease clustering defining confirmed malaria as cases, and negatives as controls [129, 130]. Test positivity rate (TPR), though commonly reported in HMIS-based studies has not been widely used for mapping, except in one evaluation report from Rwanda [106]. It has however been used in combination with presumptive cases to generate malaria positive fraction [115, 124, 125] or to adjust for over-estimation when presumptive diagnosis is high [115], as was common practice across sub-Saharan Africa [13]. As a proxy measure of incidence, however, TPR is: (i) inexpensive relative to measuring incidence, (ii) widely used to assess temporal trends, (iii) recommended by WHO [44], and (iv) easy to incorporate and monitor in routine HMIS processes even at peripheral health facilities [44, 45]. The same attributes, however, may hold for all the other commonly used metrics for measuring changes directly like case totals or indicators derived and considered as indirect assessments [131].
- ***Interoperability and systems strengthening:*** There are opportunities within HMIS to link multiple information systems, such as: the patient health records system with logistics information systems to manage stockouts and/or wastage; HMIS with regional or national demographics for health system strengthening; and, conducting multi-disease assessments for enhanced decision making. Importantly, introduction of DHIS-2 in 2012 was associated with 49% increased report completeness and 55.2% increased submission timeliness over the first year, providing greater accessibility to multi-department HMIS data [14]. HMIS data, therefore, provides an evidence base to advance policy proposals from: management, expert opinion, task forces, stakeholder engagements, community dialogues, trainings, investigative research, and field experiences [132]. Evidence exists of triangulation of HMIS with pharmacy and other systems cited as pivotal to monitoring new programs like the anti-retroviral drugs program to inform national HIV response in Kenya [133].

Notwithstanding the great opportunities, several limitations of routine HMIS surveillance data are noteworthy and may affect accuracy in estimates of malaria and/or disease burden derived.

- ***Incompleteness in health facility reporting:*** Nationwide reporting, though improving, may not be absolutely prompt or complete and if completeness is low, assessments may underestimate the burden reported [134]. Contributing factors may include: shortage in staffing, infrequent data checking by in-charges, laborious HMIS documentation along with lack of training, difficulty submitting hard copy reports, and sudden transfer of staff without formal hand-over [135]. Whereas there have been improvements associated with the advent of electronic web-based reporting [14], it remains unclear how factors associated with health care human resources or health worker practices, impact on HMIS data completeness.
- ***Exclusion of close-to-community health services:*** Data from community health services, such as village health teams (VHT) under integrated community case management (iCCM) programmes, are largely excluded from regular reporting. Whilst expected from the entire district health services sector, reporting progress has mostly impacted the formal health centre side. VHT reporting struggles with: inadequate supply of tools, inconsistent and unreliable supervision, shortage of basic required training, and competing demands from multiple implementing partners with a diversity of reporting tools in use [136]. Reports show that training has been poorly attended by a few VHT members and even fewer for any comprehensive course [136]. Deficiency in training, low education levels, and unclear supervision impacts on the quality of VHTs reports, if any.
- ***Health seeking behaviours and the private sector:*** Patient records from the private sector (private-for-profit clinics, drug shops - major players, and pharmacies), said to cater for up to 53.2% of patients in Uganda [137], are dismally captured through HMIS reporting. Preference of the sector is well documented in sub-Saharan Africa citing good service as well as proximal and regular drug supply [138], relative to the public side. Extensive drug shops use may signal high levels of self-medication, since artemisinin combination-based therapy drugs (ACTs) are over the counter drugs [139]. One report indicated that 38% of caregivers first treat fevers at home in Uganda, possibly aided by this drug availability [140]. Moreover, 59% of the children under 60 months of age sought advice or care from private facilities during their most recent fever episode in 2018 [61], an increase from 49% in the 2014 by MIS survey reports [60]. Other reports have indicated 42% versus 16.4% as seeking care from private versus public facilities, respectively, being their first of multiple care options for an illness episode [140]. Moreover, where a single option was used, 68% vs. 27% used private vs. public facilities, respectively [140]. Taken together, the majority [141] of the population seek care from the private sector in Uganda and for effective disease monitoring and control, HMIS-based surveillance needs to critically consider the private sector. Nonetheless, HMIS remains heavily biased to public health facilities to date.
- ***Reliability of diagnosis:*** Testing practices are fairly differential due to health system-related challenges like: disruptive or non-functional facilities, human resource shortages, little or no supervision, and varied health

worker attitudes [142]. Diagnostic testing rates, if low, may reduce confirmed cases realised while encouraging presumptive diagnosis and therefore, compromised accuracy of burden estimate [108]. Moreover, increased use of mRDTs, while curbing the irrational use of antimalarials through reduced presumptive diagnoses, is also associated with: increase in false negative results due to low parasite densities or deletion of target-gene within the parasite; low health worker trust of results; and, antibiotic overuse especially with negative mRDT results [143, 144].

- ***Reliability of population denominators:*** Incidence rates rely on population estimates as the denominator. However, neither the population within an attributed/assumed catchment nor the appropriate catchment of a given health facility or group of facilities can be precisely defined. This may be compounded by: (i) non-alignment of health facility catchments with administrative boundaries though often assumed, (ii) unpredictable trends in population movements, especially with unstable political situations such as areas with rampant refugee activity, or (iii) unreliable frequency of national population census updates and/or restrictive levels of detail of these census data, when available. These factors, individually or collectively, undermine the accuracy of estimates of disease incidence in these low resource settings.

Consequently, the burden of disease reported through routine data is heavily affected by the quality of records generated at the health facilities [145]. As such, large areas of the malaria endemic world, especially sub-Saharan Africa with HMIS classified as poor, still fall short on reporting true measures of disease burden given underutilised routine systems, and alternative model-based sources being used instead [62, 146]. However, this is not the case particularly in the lower transmission settings or where HMIS is reliable [62]. Nevertheless, there are many studies within these high transmission areas that have exemplified the benefit of routine HMIS in mapping malaria, documented from across sub-Saharan Africa [13, 80, 93-95, 101, 103-105, 108, 115, 120-130, 134, 147-157], though minimal compared to other data sources. Therefore, the potential in improved routine reporting through HMIS is great, especially for spatial risk assessment.

1.5.3 HMIS data for malaria impact evaluation

Competing interests on funding that has previously facilitated large-scale declines in malaria burden [158, 159], necessitate renewed data-informed implementation and evaluation of the impact of available control interventions [4, 160], owing to recent stalling in burden declines. Current intervention tools including LLINs, IRS, artemisinin-based combination therapies, and low cost parasitological mRDTs have all long been proved effective. However, following their implementation in routine or real-world settings, assessment of their impacts using cluster randomised trail (CRT) study designs, have often found no impacts [161-163]. One study in Uganda, for instance, successfully implemented a CRT where the intervention trained health workers in fever case management using mRDTs (study introduced) and artemether lumefantrine (AL) but found no differences between arms, in the prevalence of parasitaemia, anaemia, or other outcome [164]. Such designs in routine settings are often overtaken by unexpected competing programs or uncontrolled implementation of other

interventions with diluting effects beyond the confines of CRT design assumptions [161, 162]. Nevertheless, HMIS's spatial and temporal scope may provide the best coverage of real-world contextual changes enabling assessments based on alternative quasi-experimental designs, such as interrupted time-series and dose-response methods, to identify intervention associated impacts [162, 165, 166]. Temporal assessments using HMIS data in Zanzibar – Tanzania, for example, showed declines in malaria incidence following the roll out of ACTs and further declines during expanded vector control (LLIN and IRS), compared to pre-intervention periods [165]. These approaches are fit for purpose because of their capacity to incorporate real-world conditions when carefully applied to contextually comprehensive data such as routine HMIS data. Utilization of the spatial capacity of HMIS data in evaluating impacts of control interventions on malaria burden, however, remains very limited. One study that assessed the effect of case management and vector control on space-time patterns of malaria incidence using HMIS data in Uganda, reported protective effects of ITN coverage among all age-groups, though significant only among children under 5 years [80]. However, these were likely to be predominantly temporal effects, given that no geo-spatial outputs were provided to this effect. Instead, the geo-spatial results reported, only confirmed greater heterogeneity of malaria burden among children under 5 years of age than among those 5 years and older. Taken together, this further highlights the need for improved understanding of the utility of routine HMIS data, for identifying locations at high-risk of malaria in high transmission settings, and thereby its application in evaluating the impact of control interventions in those areas.

1.6 Justification and Rationale

As indicated in previous sections, there are important knowledge gaps surrounding reliability of HMIS as a viable data source, how indicators of malaria burden derived from HMIS relate to each other, their representativeness of burden relative to gold standard estimates, and the potential use of these indicators in identifying high-risk areas across spatial scales. Stalled reduction in malaria burden, coupled with recent strategies of targeted application of well-known effective control interventions informed by surveillance, emphasises an urgent need for improved understanding of routine surveillance systems and better interpretation of indicators of malaria burden from these systems. A stronger understanding of routine surveillance data would improve identification of weaknesses for surveillance system improvement, facilitate increased use of the data generated, foster stronger health systems in low resources settings, and improve the allocation of resources for health in these settings. Moreover, better interpretation of the data and/or indicators of burden from routine surveillance would enable production of stronger evidence or basis for: optimal resource channelling; timely implementation of control interventions; improved assessment of control interventions' impacts; efficient and/or effective decision making; and, sustainable, timely, accurate, and scalable monitoring of malaria burden in the low resource high-burden areas, like Uganda.

My thesis will focus on understanding the HMIS-based indicators of malaria burden. I will particularly focus on malaria incidence rates, both over time and space. As outlined in the previous sections, there are knowledge gaps

surrounding relative magnitudes of these metrics in high circulation and/or frequent use. Extremely few studies have examined the effectiveness, fitness, or utility of HMIS-based indicators, or relationships between and among these or other indicators and none with their gold standard counterparts. Studies of HMIS routine indicators of burden, exploring their inherent sources of bias, examining their representativeness of unbiased or true burden, and assessing their capacity for identification of high-risk locations are needed to address the identified gaps in knowledge on overall utility of routine data. Stronger understanding of relationships between these indicators, their change with age over time, representativeness of unbiased burden and likely sources of bias could provide valuable insights around impact and effectiveness of malaria control strategies. Moreover, increased understanding of the spatial distribution of malaria burden may also inform appropriate scales for optimal implementation and assessment of targeted interventions. Consequently, results will highlight the potential for robust timely map production using HMIS data for target decision making and optimal resource allocation and incentivise improved utility and uptake of risk maps across national malaria control fora. This work is highly timely for the call to transform surveillance into an intervention under the global technical strategy for malaria 2016-2030, and ultimately for the third sustainable development goal to be met [4, 5].

1.7 Thesis aim and objectives

The aim of my thesis is **to investigate the utility of indicators of uncomplicated malaria burden from routinely collected health facility data in describing the changing temporal and spatial distribution of malaria in Uganda**. Addressing this aim will provide evidence to guide strategic use of routine data for malaria control activities. This aim will be reached through the following specific objectives:

1. ***To explore the relationship between alternative measures of uncomplicated malaria incidence generated from sentinel surveillance data.***

While several indicators of malaria burden have been derived from routine public health facility data and used widely to estimate incidence, how they each relate to the other is unclear. Better understanding of this relationship may help with interpretation of burden or risk derived and/or reported through their use. This study objective, therefore, explores the relationship between several indicators of malaria burden (incidence estimates), and will compare them across three transmission settings in Uganda.

2. ***To examine the impact of malaria control interventions on the age distribution of malaria cases using routine sentinel surveillance data in four sites where LLIN and IRS campaigns have been conducted.***

Whereas surveillance has predominantly focused on children under five years of age, a pattern of high-risk of positivity among older children became apparent and raised concerns about the continuation of surveillance as usual. This objective, therefore, explores the possible driver of this changing pattern to provide evidence that supports this apparent trend or shift and highlight the vital role age plays in surveillance considerations.

3. ***To investigate the association between incidence of uncomplicated malaria from routine surveillance data and incidence from cohorts, across sites of different transmission intensities and identify and quantify sources of bias in surveillance incidence to assess its reliability for monitoring burden of malaria.***

Whilst routine HMIS data quality reports range from untimely, incomplete, and unreliable diagnoses to improved, in Uganda and elsewhere [37, 167-169], goodness-of-fit of derivate estimates of malaria incidence to represent unbiased burden, is unknown. This study objective, therefore, compared HMIS-based incidence with incidence from community cohorts in three settings in Uganda, accounting for other associated factors, that are influential on health facility data over time. It then evaluated the potential sources and quantities of bias in routine data to assess reliability of its estimates of malaria burden.

4. ***To explore patterns and determinants of spatial variation of malaria from routine HMIS data at national spatial scales and identify areas at high-risk of malaria***.

Geostatistical analyses of malaria reliant on routine data have been limited to regions, district, and sub-district spatial scales with limited data access. With increasing accessibility given the advent of DHIS-2, more fine-scale assessments of malaria burden and risk may be possible. This study objective, therefore, explored multi-scale spatial temporal patterns of incidence and risk using national routine HMIS data from geolocated health facilities, accounting for known risk factors.

1.8 Thesis outline

To aid interpretation, **Chapter 2** provides a detailed description of the multiple data sets pooled together to address the different components of this research. **Chapter 3** describes the relationship between test positivity and incidence rates from enhanced HMIS surveillance across three sites of varied transmission intensity in Uganda**. Chapter 4** outlines the impacts of effective large-scale community control interventions on the age-specific burden of confirmed malaria across four sites of varied transmission intensity in Uganda, stratified into 'LLIN alone' versus 'LLIN plus IRS' intervention sites**. Chapter 5** evaluates the relationship between HMIS- and cohorts-based incidence of malaria, across three sites of varied transmission intensity around Uganda, and assesses the level of bias from multiple factors of influence to HMIS recorded data**. Chapter 6** presents a concurrent multi-scale assessment of the spatial temporal distribution of incidence of malaria from national routine HMIS reporting, accounting for environmental risk factors, identifying seasonality and high-risk clusters of malaria across the country**.** These chapters have all been published (Chapters 3 to 5) or submitted (Chapter 6) to peer review journals. Finally, **Chapter 7** discusses the findings from this work and the conclusions drawn, limitations identified in this research, and recommendations for policy and/or future research.

Other supportive information towards this work, including: (a) summary of the literature reviewed to assess the use of routine HMIS data in malaria risk or burden mapping has been provided in Appendix 1; (b) Response to reviewers' comments for the published paper in Chapter 3, contained in Appendix 6; (c) Response to reviewers'

comments for the published paper in Chapter 4, contained in Appendix 8; (d) Response to reviewers' comments for the published paper in Chapter 5, contained in Appendix 9.

2 Data Overview

This thesis uses multiple complementary health management information system (HMIS) data sources that are disjointed by study or program design. Consistent across most of these, was that they largely conducted surveillance among the same populations but for independent and/or different study objectives. Together, these datasets provided a unique opportunity to study estimates of malaria burden and factors associated with them. This was possible through leveraging (1) patient-level details from health facilities including dedicated national reference centres and community-based passive cohorts, and (2) a nation-wide network of HMIS reporting health facilities. I thus provide a summary description of the various data sources and how they tie together.

Overall, three separate surveillance projects plus the national routine HMIS data system, provided data for this work. In the following section, I introduce malaria in Uganda's HMIS, after which, I provide a detailed description of all the data sets used to conduct this research.

2.1 Study Data Sources

2.1.1 Malaria in the Uganda's HMIS

Summary of the Uganda health system structure: The health system in Uganda, is a hierarchy comprising of: National referral hospitals, Regional and other referral hospitals, and district health services, that each report to the Ministry of Health's Department of Health Information, through HMIS. The district health services, headed by a district hospital, includes: health centre (HC) IV – providing emergency surgery, in-patient care, maternity, and blood transfusion services; followed by mid-level HC III – providing basic laboratory, maternity, and in-patient care services; then the HC II – providing outpatient and outreach services as the lowest formal care level with premises [170].

Whereas public formal care stops at HC II, other facilities include privately owned and a few government-run special clinics. At the lowest level are community health workers or village health teams (VHT), comprising of volunteers often trained under the integrated community case management (iCCM) strategy to diagnose and treat malaria, pneumonia and diarrhoea in children under five years within communities [171]. Taking advantage of tools like rapid diagnostic test kits for malaria (mRDTs), VHT where operational, provide extended reach of care to communities though these do not consistently perform routine HMIS reporting [172].

Uganda has at least 7000 health facilities and counting to date [170]. Nationally, all public health facilities that include Government owned and private not-for-profit (PNFP) and increasingly private for profit (PFP) health facilities, provide regular (weekly/monthly) HMIS reports on burden of selected diseases and their management to regional authorities, primarily the district medical team [173]. Introduced in 1997 as a paper-based reporting system, HMIS reports are utilised by the Ministry of Health for national level health assessments [15, 173]. They are the primary source of malaria cases data, informing the different Ministry of Health bodies including National Malaria Control Program (NMCP), as an evidence base for decisions on control interventions and wider policy [43,

173]. Since 2012 however, a web-based District Health Information System – version 2 (DHIS-2) was introduced to enable easier access to reports from across the entire national health system, starting with the public sector [14].

Malaria surveillance in Uganda, using standard HMIS, may be considered as conducted at two major levels. The first, broader, and more general level is reporting through the district health services to the NMCP. As in many malaria endemic countries, health facilities provide regular aggregated reports to governments for disease burden assessment and these are entered into the DHIS-2 system, making them readily available to the NMCP(s) [14, 45, 174]. The second and more focal level is through sentinel sites (later known as reference centres) embedded within the HMIS system in epidemiologically diverse settings, to strengthen the collection of high quality data [175]. From these, data are evaluated at patient-level, rather than in aggregates, aiding more robust inferences for control and early warning feedback, for possible epidemics and therefore, action. Reports from the sentinel surveillance are generated monthly by the Uganda malaria surveillance project and made available to the NMCP.

Specific to this study were uncomplicated malaria cases, details of which are recorded in one of many HMIS registers, the outpatient department (OPD) registers – per national policy. Uncomplicated malaria was defined as any episode of malaria where the patient was not hospitalised but treated within the outpatient clinic. OPD registers comprised the main source of data used in this study. In the next sections I describe the two categories of HMIS data used, including patient level or aggregate HMIS data, and two additional data sets including cohorts summarised in Table 1, and explanatory variables data.

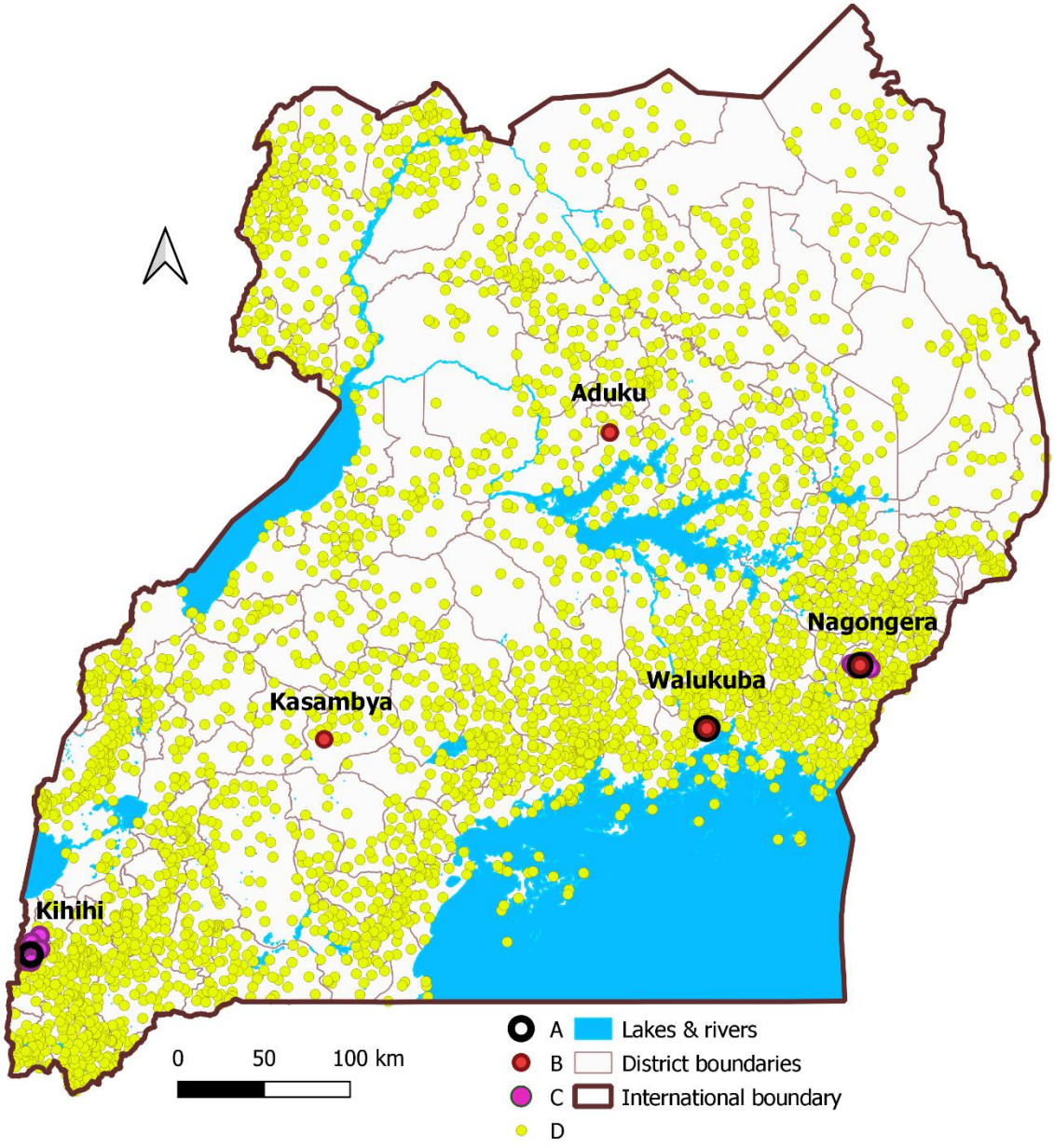


Figure 11. Locations of study health facilities across Uganda, by study objective:

A – Sentinel health facilities included in the objective 1 study, three facilities in total - all being Level IV and Government owned. The sub-county boundary around each was used to define the study area with a varied number of villages per site; Nagongera had 45, Walukuba 21 and Kihihi 117 villages. The three sites were selected due to the concurrent cohorts conducted there, for which epidemiological diversity of the sites was a key consideration in the choice of sites for the cohorts.

B – Sentinel health facilities included in the objective 2 study, four facilities in total and all Government owned.

A & C – Sentinel and lower-level health facilities included in the objective 3 study, 15 facilities in total, with some Government owned and others private not for profit. The sub-county boundary around each was used to define the study area as in objective 1.

D – Nation-wide HMIS reporting health facilities included in the objective 4 study, 3446 facilities, including national and other referral or district and general hospitals, health centres, and clinics both Government and privately owned.

Table 1. Summary of data sources, the respective study populations, and indicators of malaria burden, by study objective.

Data source(s)	Study population age group	Indicator of malaria burden	Study period
1. Exploring the relationship between alternative measures of uncomplicated malaria			
Patient-level HMIS: 3 HCIV's (Malaria reference centres) in 3 sub-Counties including Nagongera, Walukuba, & Kihihi	Children <11 years	Test positivity rate, Malaria incidence rate	Oct-2011 to Jun-2016
2. Examining the impact of malaria control interventions on the case age distributions			
Patient-level HMIS: 3 HCIV's & 1 HCIII (Malaria reference centres) in 4 sub-Counties including Nagongera, Walukuba, Aduku, & Kasambya	3 categories: <5, 5-15, 15-70 years	Test positivity	Jan-2009 to Jul-2018
3. Investigating associations between incidence of uncomplicated malaria from routine surveillance data and cohorts			
Patient-level HMIS: 3 HCIV's (Malaria reference centres), 2 HCIII's, and 7 HCII's in 3 sub-Counties including Nagongera, Walukuba, & Kihihi	Children 0.5-<11 years	Malaria incidence rate	Oct-2011 to Sep-2014
Additional data source - Community cohorts: 3 cohorts involving 100 households from each of the 3 sub-counties of Nagongera, Walukuba, & Kihihi	Children 0.5-<11 years	Malaria incidence rate	Oct-2011 to Sep-2014
4. Exploring patterns and determinants of spatial variation of malaria from routine HMIS data			
National DHIS-2 aggregate HMIS: 3446 health facilities in the national HMIS including (Hospitals, HCIV, HCIII, HCII's, & Clinics)	All	Malaria incidence rate	Jul-2015 to Sep-2019

HC = Health centre

2.1.2 Patient-level HMIS data

2.1.2.1 Sentinel surveillance data

In Uganda, sentinel surveillance for malaria has been conducted since 2006, under the Uganda malaria surveillance project (UMSP) [68]. Against a backdrop of very low capacity for diagnostic testing in Uganda, six sentinel health facilities with operational laboratory facilities and thus, capability to conduct diagnostic testing for malaria using microscopy, were purposefully selected, considering geographical representativeness as determined under the East African Network for Monitoring Antimalarial Treatment (EANMAT) [27]. These sentinel sites were later upgraded to national malaria reference centres for the NMCP [64]. By the start of this PhD in 2017, there were at least 21 operational malaria reference centres in Uganda. Of these, three centres were included in objectives one and three, while four were included in objective two of this research with each centre located in an independent sub-county and district.

At each outpatient (OPD) clinic of these health facilities, for every patient seen, presenting symptoms of illness are assessed by the attending clinician. All suspected malaria cases are sent to the laboratory for a blood test for malaria, by microscopy or mRDT. Based on the test results from the laboratory, appropriate action is then taken by the clinician and all the details pertaining to this patient visit are recorded in the OPD register. These details include age, sex, fever or history of fever, diagnostic test done, test results, diagnosis given, and treatment prescribed, among others. Every month, these data are extracted by a UMSP supported staff at the clinic and entered in a MS Access database (Microsoft Corporation Inc., Redmond WA. USA). The complete monthly data are then sent to the UMSP data centre for cleaning and processing [68]. A detailed description of the data management and processing within this study is provided in section 2.2 below.

2.1.2.2 Additional (non-sentinel) health facility data

To supplement the above sentinel site data and ensure comprehensiveness of HMIS data for the included study sites, 12 non-sentinel health facilities, including level II and III facilities from three sub-counties (each hosting a malaria reference centre) also provided patient-level data in objective three of this study. In keeping with the sentinel facility data collection format, retrospective anonymised individual patient details were collected from OPD registers of each facility, covering a three-year duration. To collect these data, I recruited a team of at least seven research assistants (RA) at a time, per site, and evaluated them with a pre-training test on their basic data and mathematics abilities. I then trained them on the principles of research and the study procedures that were detailed in a standard operating procedure (SOP). Following several days of training, they were all tested using a post-training quiz to evaluate their comprehension of the procedural aspects of the study. The RA's then entered the data from OPD registers into MS access databases, loaded on tablet computers. I provided fulltime supervision of this activity in the field from site to site. On a daily basis, I backed up the data from each tablet and charged the tablets at a central place, making them ready for the next day of work, since our field office – a rented primary school classroom had no power supply.

Also, some data from a cluster-randomised trial (CRT) conducted in several sub-counties of Tororo district including Nagongera, among government-owned lower level health facilities, was included in this study [164]. The CRT study aimed to evaluate the impact of enhanced health facility-based care for malaria and febrile illnesses in children within the study area. With facilities randomised in two arms, the intervention that involved, among others, training health workers on fever case management and use of mRDTs, as well as ensuring adequate supplies of mRDTs and artemether-lumefantrine (AL) was evaluated using HMIS data from OPD registers in both arms [164]. 3/20 facilities including Maundo, Were, and Katajula HCII's were in Nagongera sub-county and data covering the duration between October-2011 and March-2013 for these facilities was obtained from the CRT and included in this study. The primary data collection discussed above, collected the remaining 19 months of data to ensure coverage of the full three-year study duration. Together, these data sets were used to address objective three of this thesis.

2.1.3 National DHIS-2 aggregate HMIS data

From the Department of Health Information within the Uganda Ministry of Health, I obtained nation-wide HMIS data from the DHIS-2 per year for all 128 districts of Uganda (as they were known by 2018) as excel spreadsheet files, formatted as monthly health facility entries. These entries included totals of OPD malaria and OPD malaria confirmed (by microscopy or mRDT) for each health facility, over the duration of January-2014 through September-2019.

Following data review, the study duration was defined to cover July-2015 through September-2019, and these data were compiled into a single database for all the 51 months of the study duration, to address objective 4 of this study.

2.1.4 Additional data source - Community Cohorts

In addition to routine HMIS data, this thesis incorporated data from three enhanced passive cohort studies conducted in Nagongera, Walukuba, and Kihihi sub-counties starting August-2011, under one of ten International Centres of Excellence in Malaria Research (U19AI089674) [176]. The focus for the original project was to describe malaria incidence and prevalence, providing a basis for further analyses on longitudinal trends and risk. For these cohorts, all children aged 0.5-<11years were recruited from a random selection of 100 households, drawn from full enumeration of all households in each sub-county. Being dynamic cohorts, any additional children in this age group within each participating household were all eligible. Clinical assessments happened at enrolment and at 3 monthly scheduled visits using a standardised questionnaire, and a blood sample taken at each to assess for malaria infection by microscopy. However, participants received free medical attention between scheduled assessments throughout the study duration, at the study clinic that was open daily.

For a three-year duration, data was obtained from these three passive community cohorts. Incidence of malaria was estimated monthly, defined as the total number of incident cases of malaria divided by total person-time of follow-up estimated in years, per month, by site.

Whereas the cohorts were used to provide a gold standard estimate of incidence of malaria per site, advantages/strengths, and weaknesses of the data in consideration were identified as detailed in Table 2 below.

Table 2. Strengths and weaknesses of the Cohort data used to derive the 'gold standard' incidence rates against which to evaluate the routine HMIS incidence rates.

Strength	Weakness
Provided standard testing for all suspected malaria: For ill participants at any visit, standard sick visit procedures including measuring temperature and/or recording history of fever in the previous 24 hours; taking a finger prick to obtain smear and filter paper samples and if thick smear positive, the patient was diagnosed with malaria and prescribed artemether-lumefantrine (AL), the recommended first-line therapy per national guidelines [35]. Moreover, the study performed venepuncture on all nonill participants at each clinic visit for a thick blood smear to examine for asymptomatic parasitaemia, among others.	**Only 100 households included across each site:** Though randomly selected, these 100 households accounted for very small proportions of the 9,881, 12,774, and 6,992 households in Walukuba, Kihihi, & Nagongera respectively and therefore, only sufficient to provide a good site-level (sub-county) estimate of incidence but proportions too small for parish or village-level estimates.
Captured cases every day of the week: Unlike standard of care at public health facilities where OPD clinics may sometimes be closed, the cohort clinics were open every day of the week to see participants.	**Excluded children < 6months:** Whilst infants may be assumed to benefit from maternal immunity, sentinel HMIS data showed that high proportions of infants <6 months of age had confirmed malaria including 12.7%, 28.3%, and 23.2% in Walukuba, Kihihi, and Nagongera respectively, over the same study duration. Excluding this group from the cohorts may have limited the understanding of estimates of incidence.
Consistent health worker practice: Given this controlled experiment environment with multiple levels of supervision of study activities, study clinicians followed well documented standard	**Differential loss to follow-up:** Whilst the study purposed to follow-up 100 households, there was considerable loss to follow-up. For instance, by the end of the study-period 21 households had dropped

operating procedures as per ethically approved study procedures for every clinic visit and across all three sites [177].	out of the study in Nagongera due to; relocation, inability to comply, and withdrawn consent, among others [178].
Provided prompt treatment: With a clinic open on every day of the week, participants enjoyed the ideal care provision with the highest likelihood of care availability within 24 hours of symptoms onset unlike under non-study conditions. Under standard care, delay may be caused by multiple limiting factors, especially financial or known unavailability of drugs at facilities.	**Passive follow-up:** Whereas participants were free to come to the clinic for all febrile illness needs and alternative care seeking was minimised, the passive nature of these cohorts could have caused some to choose quicker alternatives. For instance, 0.1% of participants were reported to have sought inappropriate care in the first 24 months of the study [177].
Reimbursed participants' travel costs: Whilst cost or financial challenges have been indicated as inhibiting to appropriate care access, the reimbursement of travel cost for participants provided good motivation for clinic attendance and therefore, improved likelihood of registering incident clinical cases of malaria	On the other hand, the financial motivation through travel cost reimbursements could have inflated case detection rates, to levels unlikely under standard care or routine surveillance. This is especially so in the very high transmissions settings, where minor fevers from other causes that would not have resulted in standard care clinic visits, may be coupled with highly likely asymptomatic parasitaemia leading to confirmed malaria in this incentivised setting [32].

To assess the relationship between HMIS-based and cohort incidence, monthly estimates of cohort incidence were included as an independent variable in the regression models used in the objective three study of this research, results of which are presented in chapter 5 of this thesis. In the following section, I describe the data preparation process by first, explaining the broad data preparation processes undertaken, followed by objective-specific data assessment, with particular focus on the outcomes of interest.

2.2 Outcome Data Management and Processing

2.2.1 Data preparation

2.2.1.1 Data cleaning process

I converted all data sets from the various projects to STATA. The majority of these data sets were already cleaned from the primary analysis projects especially regarding the malaria outcome, diagnostic testing performed, diagnostic test results, and age. This was not the case, however, for villages of residence, especially for the UMSP.

To merge these datasets into a single database, several variables, value definitions and value labels needed cross-checking and alignment, which I conducted in STATA. For villages of residence, I used both my personal experience gained through being involved in the household enumeration exercises, where I led the teams as a research assistant and data officer with the projects during 2009-2011, and also referred to local knowledge. For the three sites, therefore, I created a standard fully coded village names master list against which, to evaluate all incoming data. The remaining list of unresolved named villages (without a match in the master list), I defined as unknown within catchment areas (sub-county), while those with a missing record, I placed in the category of missing.

2.2.1.2 Population at risk of malaria (denominator)

At multiple levels in this research, I needed to generate or define the population at risk, which in turn defined the denominator in estimating village or other defined resolution-level incidence rates per month. As such, the intended resolution-level population estimates were derived using national population gridded surfaces, freely provided by the Worldpop project (http://www.worldpop.org.uk). The main determinant for this choice was the inaccessibility of national housing and population census data for 2014 from UBOS, as well as the unavailability of these estimates at the spatial resolutions of interest in this study, particularly villages and health facility catchment areas, as further discussed in chapters 3, 5, and 6 of this thesis. From the national gridded population surfaces, estimates at the respective spatial resolution, particularly sub-Counties (described fully in Chapter 3) and health facility catchment areas (described fully in chapter 6), were extracted using the ESRI ArcGIS 10.3 *Zonal statistics* tool (ESRI 1995-2014l Redlands, CA. USA) for the objective respective study durations.

From the annual population counts, monthly population estimates were determined using a monthly growth rate generated from national bureau of statistics' (UBOS) 2002-2014 published census reports [179] for each subcounty (in Chapter 3). Moreover, I used linear regression predictions for monthly population estimates within health facility catchment areas, as discussed in Chapter 6.

2.2.1.3 Suspected malaria definition

As no explicit record was made in the HMIS OPD registers of patients with suspected malaria, these were defined as all patients sent to the laboratory for a blood test for malaria, by microscopy or mRDT. Among those not sent to the laboratory, however, suspected malaria cases were identified as those with a clinical diagnosis of malaria. Whilst fever or history of fever in the last 48 hours is a key identifier of cases suspected to be malaria, the recording

of this data or even the temperature taken at the clinic in all the HMIS studies involved was very low. For instance, data from the three sentinel facilities of Walukuba, Kihihi, and Nagongera, between Oct-2011 through June-2016, showed that fever recording among children <11 years of age ranged from 28.4 to 51.9%. Also, whilst a temperature of >=37.5 ^{0}C was considered determinant of fever, this information was predominantly missing in the HMIS databases. For example, in the three sentinel sites discussed above, a maximum of 34 participants had a recorded temperature, as such, these data were not utilised as primary determinants of suspected malaria.

2.2.1.4 Attendance status

For each patient visit recorded in the OPD registers, it is expected that indication is made of whether that patient visit was a new attendance (that is a new episode of illness) or re-attendance (implying a follow-up visit for an illness episode that was previously recorded at the clinic). This was done to avoid possible counting of the same episode of illness more than once as an incident case of malaria and would be applicable for any other illness presented and/or recorded in OPD registers. All re-attendance cases were excluded from any analyses in this study. For instance, though between 44.6 and 51.2% of participants had a missing record of attendance status among the three sentinel facilities during Oct-2011 and June-2016 and were assumed new illness episodes, between 0 and 4.2% of patients <11 years had visits classified as re-attendance, making them ineligible for inclusion. Exclusion of re-attendance visits was not expected to impact on analyses in anyway, given that they had been recorded in the data during their initial clinic visit, for the same illness episode.

In the following section, I provide a detailed description of the data included in addressing each individual study objective, with an emphasis on evaluating those excluded from analysis.

2.2.2 Data description and summary by objective

2.2.2.1 Objective 1: To explore the relationship between alternative measures of uncomplicated malaria incidence generated from sentinel surveillance data

Here, I examined data for the duration between October 2011 to June 2016, from three sites with a sentinel or reference health facility, including Nagongera Health centre IV (HCIV) in Tororo, Walukuba HCIV in Jinja, and Kihihi HCIV in Kanungu districts, as shown in Figure 11 above.

For study participant data preparation, attention was paid to villages of residence, age, test positivity and diagnosis, and attendance status, each contributing to the inclusion criteria.

a) Exclusion from study based on village of residence:

Exclusively village of residence: Among suspected malaria cases under 11 years of age that were new attendance visits, 44,875 (40.6%) were excluded based on a missing (61.9%) or unknown village of residence within the study sites. Interestingly overall, all the excluded patients were suspected to be malaria cases. The majority of these, had a malaria diagnostic test performed (94.9%) with 34.9% confirmed positive for malaria parasites, compared

to 37.5% among those with known villages and therefore, included (*Chi sq.=71.5, p<0.001*). By site, however, two sites showed significant differences including first, Nagongera where 32.0% of those excluded were confirmed positive for malaria parasites compared to 38.2% among those included (*Chi sq.=179.2, p<0.001*). Second, Walukuba where 33.0% of the excluded were confirmed positive cases compared to 24.2% among those included (*Chi sq.=260.2, p<0.001*), but no significant differences in Kihihi (*Chi sq.=1.8, p=0.185*). With a similar distribution of cases among the included and excluded participants in Kihihi, as well as a significantly higher proportion of confirmed malaria cases among the included than the excluded in Nagongera, exclusions due to missing villages of residence may not have impacted findings in this study, for these two sites. For Walukuba, however, the significantly higher proportion of confirmed cases among those excluded may have led to underestimation of indicators generated in this study for this site.

Age by villages of residence status: The majority of patients with recorded age <11 years, were under 5 years of age in all sites, with: Walukuba (66.0%) among those included compared to (70.8%) among those excluded (*p<0.001*); Kihihi (59.2%) among those included compared to 57.7% among those excluded (*p<0.001*); and, Nagongera (79.1%) among those included compared to 81.2% among the excluded (*p<0.001*). Within the highest transmission setting of Nagongera, 50% of the included participants were under 2 years of age compared to 54.9% among those excluded (Table 3).

Table 3. Age distribution of study participants comparing included and excluded patients <11 years, by site.

Site	Age category	Patients Included (%)	Patients Excluded (%)	P-value
Walukuba	<2 years	7,543 (35.9)	4,057 (37.3%)	<0.001
	2-<4 years	4,580 (21.8)	2,679 (24.6%)	
	4-<6 years	3,081 (14.7)	1,671 (15.4%)	
	6-<8 years	2,349 (11.2)	1,102 (10.1%)	
	8-<11 years	3,436 (16.4)	1,363 (12.5%)	
Kihihi	<2 years	7,496 (29.5)	2,135 (29.3%)	<0.001
	2-<4 years	5,207 (20.5)	1,467 (20.1%)	
	4-<6 years	4,425 (17.4)	1,122 (15.4%)	
	6-<8 years	3,597 (14.2)	1,057 (14.5%)	
	8-<11 years	4,652 (18.3)	1,517 (20.8%)	
Nagongera	<2 years	9,671 (50.0)	14,654 (54.9%)	<0.001
	2-<4 years	4,354 (22.5)	5,567 (20.9%)	
	4-<6 years	2,254 (11.7)	2,593 (9.7%)	
	6-<8 years	1,399 (7.2)	1,666 (6.2%)	
	8-<11 years	1,666 (8.6)	2,225 (8.3%)	

The data here showed that in Walukuba and Kihihi, a significantly smaller proportion of participants were excluded due to missing villages by age across these age categories than were included in the study. As such, in these two sites exclusion due to missing villages, may not have significantly impacted the effects of age within the indicators derived. For Nagongera, however, the data showed that a significantly larger proportion of participants were excluded due to missing village by age than were included, implying that participant exclusion due to missing villages may have had a larger impact on age-related effects in the indicators of malaria burden derived for the site.

b) Exclusion based on age

Very few patients (1,203), had a missing record of age from the data collected, and these were excluded from the study. The distribution of these was 72.3, 18.6, and 9.1% in Walukuba, Kihihi, and Nagongera, respectively indicating that Walukuba had the highest occurrence of missing age recording, though all together negligible.

c) Exclusion based on test positivity and diagnosis

Malaria cases were defined as participants that were diagnostically confirmed positive for malaria parasites. Participants with a negative diagnostic test for malaria, but having a diagnosis for malaria given, did not qualify as cases but as suspected malaria cases, and these were very few in the sentinel facilities data - a total of 716 participants with 12.2, 17.3, and 70.5% of them in Walukuba, Kihihi, and Nagongera, respectively. Moreover, 445 participants (37.1, 14.2, and 48.8% of these in Walukuba, Kihihi, and Nagongera) were presumptively diagnosed with malaria and therefore, not considered as cases of malaria but as suspected malaria cases instead. However, 252 participants were registered having a positive diagnostic test for malaria, but without a diagnosis for malaria, and were considered confirmed malaria cases in this study. Among study participants, a large majority of diagnostic testing for malaria was performed using microscopy ranging from 90.9 to 98.7% in Walukuba and Kihihi, respectively. A small proportion of diagnostic tests included rapid diagnostic tests, highest in Nagongera with 954 tests.

2.2.2.2 Objective 2: To examine the impact of malaria control interventions on the age distribution of malaria cases using routine sentinel surveillance data in four sites where LLIN and IRS campaigns have been conducted.

In this study objective, I included data from four sites with a sentinel or reference health facility each (two from objective one above and an additional two), including Walukuba health centre IV (HCIV) in Jinja district of the central Uganda, Kasambya HCIII in Mubende district of mid-western Uganda, Aduku HCIV in Apac district of northern Uganda, and Nagongera HCIV in Tororo district of eastern Uganda, with site locations shown in Figure 11 above, for the duration between January 2009 to July 2018.

The data from these health facilities was prepared and cleaned with particular focus on age, diagnostic tests and test positivity, suspected malaria status, sex, and attendance status.

Exclusion from the study

Age: Given a smaller number of much older patients seeking care, those over the age of 70 years were excluded from the analyses, a total of 1,975 (0.6%), 1,442 (0.9%), 3,673 (1.9%), and 3,833 (1.7%) in Walukuba, Kasambya, Aduku, and Nagongera, respectively. These were not expected to impact on our results in any significant way.

Sex: A very small number of eligible participants (at most 0.03% in a single site) had a missing record of sex in the data and were excluded, given that sex was an important factor included in the analysis for this study objective. However, these were not expected to impact on our findings in any way.

Diagnostic tests and test positivity: All cases confirmed by microscopy or mRDT were considered positive cases and presumptive cases not counted. The exclusion of presumptive cases regardless of being few, is not expected to have a definite effect on our analyses or results, given that the presumptive diagnosis process is highly subjective and therefore indeterminate. Whilst majority of diagnostic testing was performed using microscopy across all the four sites, a slightly larger majority of negative than positive test results were generated using microscopy in Walukuba, Kasambya and Aduku, but not in Nagongera (Table 4). Notably, however, the highest proportion of positive test results generated using mRDT's were observed in Aduku at 28.1%.

Table 4. Proportions of test results, by diagnostic testing method per site.

Test result	Diagnostic method	Walukuba	Kasambya	Aduku	Nagongera
Positive	Microscopy	40,548 (96.7%)	35,273 (84.3%)	26,648 (71.9%)	31,725 (90.2%)
	mRDT	1,403 (3.3%)	6,583 (15.7%)	10,407 (28.1%)	3,438 (9.8%)
Negative	Microscopy	81,295 (97.5%)	55,464 (86.4%)	49,049 (85.4%)	73,175 (87.9%)
	mRDT	2,112 (2.5%)	8,767 (13.7%)	8,393 (14.6%)	10,053 (12.1%)

Assessing potential impacts of diagnostic testing method showed no identifiable pattern, suggesting that the diagnostic method used had very limited influence on the pattern of test results, as further discussed in Chapter 4.

2.2.2.3 Objective 3: To investigate the association between incidence of uncomplicated malaria from routine surveillance data and incidence from cohorts, across sites of different transmission intensities and, identify and quantify sources of bias in surveillance incidence, to assess its reliability for monitoring burden of malaria.

This study objective included HMIS data from which incidence of malaria was estimated, as well as community cohort data that provided the comparative incidence estimates in the three study sites.

HMIS data: This was obtained from all 15 public health facilities located within the geographic administrative boundaries of Nagongera sub-County in Tororo district (5 facilities); Walukuba sub-County in Jinja district (3 facilities); and, Kihihi sub-County in Kanungu district (7 facilities) as shown in Figure 11. The enrolled health facilities included: Nagongera, Walukuba and Kihihi HCIV's, in the respective sub-Counties; Matanda and Nyamwegabira HCIII's, and Bihomborwa, Bushere, Kibimbiri and Nyakashure/Samaria HCII's, in Kihihi sub-County; Were, Katajula, Maundo, and Pokongo HCII's, in Nagongera sub-County; and, Masese Port and Masese 3 HCII's, in Walukuba sub-County. Whereas 16 health facilities were screened for inclusion in this study, one was excluded based on its geo-location falling outside of the study site boundaries, besides the very few residents of the sub-county (study site) who visited this facility for care. These HMIS data were obtained for the three-year duration spanning October-2011 through September-2014.

After extraction from OPD registers, the data was cleaned with particular focus on village of residence, age, diagnostic testing and diagnosis, and attendance status.

Exclusion from the study:

Village of residence: In this study objective, missingness of record of village of residence was corrected for in computing confirmed cases, as later explained in Chapter 5. However, all patient records with villages that were either unknown within the study site or unclear, were excluded. Nevertheless, these were not expected to impact on our estimates of incidence, given that residence within the site boundaries was central to estimating site-specific analysis outcomes. Notably though, there was a higher proportion of patients with unknown villages in the lower-level facilities of Kihihi, which may be attributed to being at the border between Uganda and Democratic republic of Congo (DRC), where sporadic influxes of refugees from DRC have been reported [180]. For instance, Matanda health centre III that is located within a designated refugee transit camp, contributed 41.2% of the patient records with unknown village of residence in Kihihi. Nevertheless, at the time of data collection from this health facility, the camp was unoccupied though the clinic was fully operational, possibly serving the more regular resident users of the facility from nearby villages. Based on this, it can be assumed that the exclusion of participants whose villages of residence were unknown within the site facilitated a more accurate estimate of burden attributable to the site resident population.

Age: To generate HMIS-based incidence estimates that would be comparable to estimates from community cohorts, all patients aged 11 years and older were excluded from this analysis. These were unexpected to impact on our results by virtue of being outside of age-groups of interest. Moreover, all patients with a missing record of age were also excluded and assuming an equal distribution as those with known age and therefore included, exclusion due to missing age was not expected to have considerable impact on incidence estimates.

2.2.2.4 Objective 4: To explore patters and determinants of spatial variation of malaria from routine HMIS data at sub-/national spatial scales and identify areas at high-risk of malaria.

To address this objective, HMIS data was obtained from the national repository for routine HMIS via the DHIS-2 web-based system. Data from all health facilities expected to report through standard surveillance procedures, including total attendance, re-attendance, OPD malaria cases, and confirmed cases (by Microscopy and mRDT) were obtained for at least four years (51 months long).

A total of 3446 health facilities with associated geo-location coordinates, were included in this study (Figure 11). These data were summarised on a monthly time scale for all age-groups combined, for each of the study health facilities.

Exclusion from the study:

HMIS data from January-2014 through September-2019 were extracted and assessed for use in this study. Notably, from January-2014 through June-2015 these data were inconsistent from month to month, with many months of data missing. However, starting July-2015 the format of the data sets was markedly different from the previous duration. The differences included the introduction of additional patient age categories that may have been a consequence of undocumented but evident system revisions or improvements. Given this considerably more complete and consistent data set, the duration of interest in this study was defined as spanning July-2015 through September-2019. Consequently, data from January-2014 through June-2015 was excluded.

For the geo-coding of health facilities, I obtained a database of public health facility geo-coordinates across Africa that was published by the KEMRI-Wellcome Trust Research Programme [181]. Of the 3792 health facilities in the database, 3448 (91.0%) were matched with the health facilities in the HMIS malaria cases database of 2015, excluding all facilities with duplicate geo-coordinates. However, one was geo-located in the lake and another outside the country boundaries, and therefore, both were excluded. A total of 3446 geolocated health facilities that matched with the HMIS malaria cases database consisted the facilities that we defined as study health facilities. Estimated impacts of the exclusion of data reported from health facilities that were not geo-coded per district, are further discussed in Chapter 6 of this thesis.

2.3 Ethical considerations

Two of the programmes that provided data, including Uganda Malaria Surveillance Project (UMSP) and the National HMIS, were not required to have ethical approval as national surveillance programmes. However, the two research projects that provided additional data had independent ethical approvals from the Makerere University School of Medicine Research Ethics Committee (SOM-REC #2010-108 and #2011-167). This was the local Institutional review board at the base of their research activities in Uganda, hosted by the Infectious Diseases Research Collaboration (IDRC). In addition, they each received approval from the Uganda National Council for Science & Technology (UNCST #HS 794 and #HS 1019), which is the national body that oversees research on the Government's behalf. Moreover, ethical approval for each was also obtained from the other collaborating institutions involved in these studies, mainly including the London School of Hygiene & Tropical Medicine (LSHTM #5943 and #5779) and the University of California San Francisco (UCSF).

Specific to the proposed work in this thesis with independent research objectives, separate ethical approval was sought, obtained, and later renewed from SOM-REC, being the local IRB for this study in Uganda (***Appendix 2a & 2b***). Next, approval was sought and obtained from the UNCST for government approval (***Appendix 3***). With these in place, ethical approval was sought, obtained, and later renewed from LSHTM research ethics committee (***Appendix 4a & 4b***).

Also, an amendment was sought and obtained to use publicly available national HMIS data from the national malaria control program to address the fourth objective of this research. For this, permission was sought and obtained from the Ministry of Health (***Appendix 5***) and based on this, ethical approval was sought and obtained for the proposed amendment from SOM-REC (***Appendix 2c***), and ultimately, approval was also obtained from LSHTM research ethics committee (***Appendix 4c***).

Concerning the primary data collection directly from health facilities, other necessary levels of permission were also required, and these included obtaining support letters from the district health officers (DHO) of each of the three districts of Tororo, Jinja, and Kanunugu. With these on hand, permission was then sought and obtained from health facility in-charges to access their stored registers within the respective facility HMIS offices. In two facilities of one district, the in-charges expressed overwhelming reservations to providing access to their registers. These necessitated lengthy explanations as well as additional written permission from other district officials (besides the DHO) before they would permit access to their registers. However, even with these permissions on hand, these two facilities also had both the most disorganised storage and poorest state of registers.

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1513171	**Title**	Mr.
First Name(s)	SIMON PETER		
Surname/Family Name	KIGOZI		
Thesis Title	Exploring the utility of indicators of uncomplicated malaria burden from routine health facility surveillance data in identifying and mapping high-risk areas for malaria in Uganda		
Primary Supervisor	Assoc. Prof. Rachel L. Pullan		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

<u>SECTION B – Paper already published</u>

Where was the work published?	The American Journal of Tropical Medicine and Hygiene		
When was the work published?	May 2019		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**YesYes**	Was the work subject to academic peer review?	**YesYes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I prepared the data, led data analysis, drafted the manuscript. I then led the manuscript submission for peer review, addressed reviewer comments, and managed the final submission for publication

SECTION E

Student Signature	simon peter kigozi
Date	01-September-2020

Supervisor Signature	Rachel Pullan
Date	3rd September 2020

3 Paper 1

Am. J. Trop. Med. Hyg., 101(1), 2019, pp. 137–147
doi:10.4269/ajtmh.18-0901

Malaria Burden through Routine Reporting: Relationship between Incidence and Test Positivity Rates

Simon P. Kigozi,[1,2]* Ruth N. Kigozi,[3] Asadu Sserwanga,[2] Joaniter I. Nankabirwa,[2,4] Sarah G. Staedke,[2,5] Moses R. Kamya,[2,4] and Rachel L. Pullan[1]

[1]*Department of Disease Control, London School of Hygiene & Tropical Medicine, London, United Kingdom;* [2]*Infectious Diseases Research Collaboration, Kampala, Uganda;* [3]*United States Agency for International Development's Malaria Action Program for Districts, Kampala, Uganda;* [4]*School of Medicine, Makerere University College of Health Sciences, Mulago Hospital Complex, Kampala, Uganda;* [5]*Department of Clinical Research, London School of Hygiene & Tropical Medicine, London, United Kingdom*

Abstract. Test positivity rate (TPR)—confirmed cases per 100 suspected cases tested, and test-confirmed malaria case rate (IR)—cases per 1,000 population, are common indicators used routinely for malaria surveillance. However, few studies have explored relationships between these indicators over time and space. We studied the relationship between these indicators in children aged < 11 years presenting with suspected malaria to the outpatient departments of level IV health centers in Nagongera, Kihihi, and Walukuba in Uganda from October 2011 to June 2016. We evaluated trends in indicators over time and space, and explored associations using multivariable regression models. Overall, 65,710 participants visited the three clinics. Pairwise comparisons of TPR and IR by month showed similar trends, particularly for TPRs < 50% and during low-transmission seasons, but by village, the relationship was complex. Village mean annual TPRs remained constant, whereas IRs drastically declined with increasing distance from the health center. Villages that were furthest away from the health centers (fourth quartile for distance) had significantly lower IRs than nearby villages (first quartile), with an incidence rate ratio of 0.40 in Nagongera (95% CI: 0.23–0.63; $P = 0.001$), 0.55 in Kihihi (0.40–0.75; $P < 0.001$), and 0.25 in Walukuba (0.12–0.51; $P < 0.001$). Regression analysis results emphasized a nonlinear (cubic) relationship between TPR and IR, after accounting for month, village, season, and demographic factors. Results show that the two indicators are highly relevant for monitoring malaria burden. However, interpretation differs with TPR primarily indicating demand for malaria treatment resources and IR indicating malaria risk among health facility catchment populations.

INTRODUCTION

National strategies for malaria control and intervention planning typically rely on the existence and strengths of national health management information systems (HMIS's).[1] Strong systems that incorporate complete case notification and accurate population data provide a firm basis for monitoring the present burden and trends, and making future projections. However, in much of sub-Saharan Africa, HMIS data miss a substantial number of malaria cases from the community who either do not seek diagnosis and/or treatment through the health system, or do but are not correctly reported.[1,2] In addition, catchment population data are not readily available at the facility level.[3] As a result, determining the true burden and interpreting trends at local levels remain a challenge. Test positivity rate (TPR) defined as the proportion of tested suspected malaria cases that return a positive malaria result is an indicator generated from facility records. Suspected malaria cases are patients presenting with fever or a history of fever in the last 48 hours, without any other recognizable cause, sent to the laboratory for a malaria diagnostic test.[4] Those who have a positive test result are, therefore, diagnosed with malaria. The test positivity rate overcomes the challenges of both denominators and numerators by assuming that those seeking diagnosis and treatment with suspected malaria make a representative sentinel population, and is also recommended by the WHO as a key surveillance indicator.[5,6] Interpretation of TPR is, however, affected by the incidence or number of non-malaria fevers by potentially inflating the sentinel population (denominator for TPR).[4] When health facility catchment populations are available, however, the test-confirmed malaria care rate or incidence rate (IR), defined as the number of malaria cases per 1,000 population at risk per unit time, may provide a better indication of the burden of malaria than TPR. Whereas TPR, owing to easier accessibility, is more commonly reported than IR,[7–10] the two indicators have been complimentarily reported but with no expressed implication of one indicator on the other.[11,12] The relationship between these indicators, therefore, remains unclear, and to our knowledge, very few studies have addressed it.

Health facility–based surveillance forms the basis of malaria burden reporting in Uganda.[5,6] Data are reported through the district health services to the Ministry of Health using regular aggregated reports including suspected malaria cases; cases tested and cases confirmed by age category (by microscopy or rapid diagnostic test for malaria [mRDT]); and confirmed or presumed cases treated with antimalarial medicine; among others, and the indicators derived are used for disease burden assessment.[5,6] In addition, multiple reference centers that were selected to cover diverse transmission intensities across the country have been embedded within the HMIS to strengthen the collection of high-quality data.[13] Using the data from three of these malaria reference centers, we investigated the relationship between IR and TPR as indicators of malaria burden at the facility level, accounting for important factors in this relationship. The few previous studies that explored relationships between incidence measures have predominantly been limited to a single site—either at the village level[5] or provincial level,[14] with a general focus on a single dimension of time. Here, we evaluate the relationship between the two indicators and explore the relationship in both time and space, providing additional detail to our understanding of the utility and representativeness of HMIS data for understanding the changing malaria burden in endemic settings.

*Address correspondence to Simon P. Kigozi, London School of Hygiene & Tropical Medicine, Keppel St., London WC1E 7HT, United Kingdom. E-mail: skigozi@yahoo.com

METHODS

Study setting. This study used routine surveillance data collected from three level IV health facilities located in different malaria transmission settings in Uganda (Figure 1). These facilities were 1) Nagongera in Tororo district with an annual entomological inoculation rate (aEIR) of 562 infective mosquito bites, 2) Walukuba in Jinja district with an aEIR of 6, and Kihihi in Kanungu district with an aEIR of 6 as per early 2000's assessments.[15] The three facilities formed part of a cohort of malaria reference centers established by the National Malaria Control Programme and led by the Uganda Malaria Surveillance Project (UMSP). Level IV health facilities typically serve a health subdistrict with an estimated population of 100,000 people and are run by a medical doctor. Each of the three sites has benefited from malaria control activities, including 1) the use of artemisinin-based combination therapies—specifically artemether–lumefantrine, which is the first-line treatment for uncomplicated malaria across the country, and 2) distribution of long-lasting insecticidal nets aiming to achieve universal coverage, conducted during September to November 2013 in Nagongera and Walukuba, and during December 2013 to February 2014 in Kihihi. In addition, Nagongera received three rounds of indoor residual spraying (IRS) with bendiocarb during December 2014 to February 2015, June to July 2015, and November to December 2015, and at least one round of IRS with pirimiphos-methyl (Actellic) during June to July 2016. The three subcounties of Nagongera, Walukuba, and Kihihi were made up of 45, 21, and 117 villages, respectively, each identified by names as known by district and subcounty officials. All these villages were located, mapped, and uniquely identified during enumeration surveys conducted at each site over 2009/2010, and are fully described elsewhere.[16,17]

Study population. All patients aged less than 11 years, presenting with suspected malaria, who were seen at the outpatient department clinic of each facility during October 2011 to June 2016, comprised the study population. This age group includes both the age range most at risk of the effects of and most sensitive to changes in malaria transmission (< 5 years),[18,19] and the age range across which naturally acquired immunity to malaria is seen to develop (5–11 years).[20] Monthly TPR trends were similar between under-fives and those aged five to less than 11 years (over-fives). However, TPR in

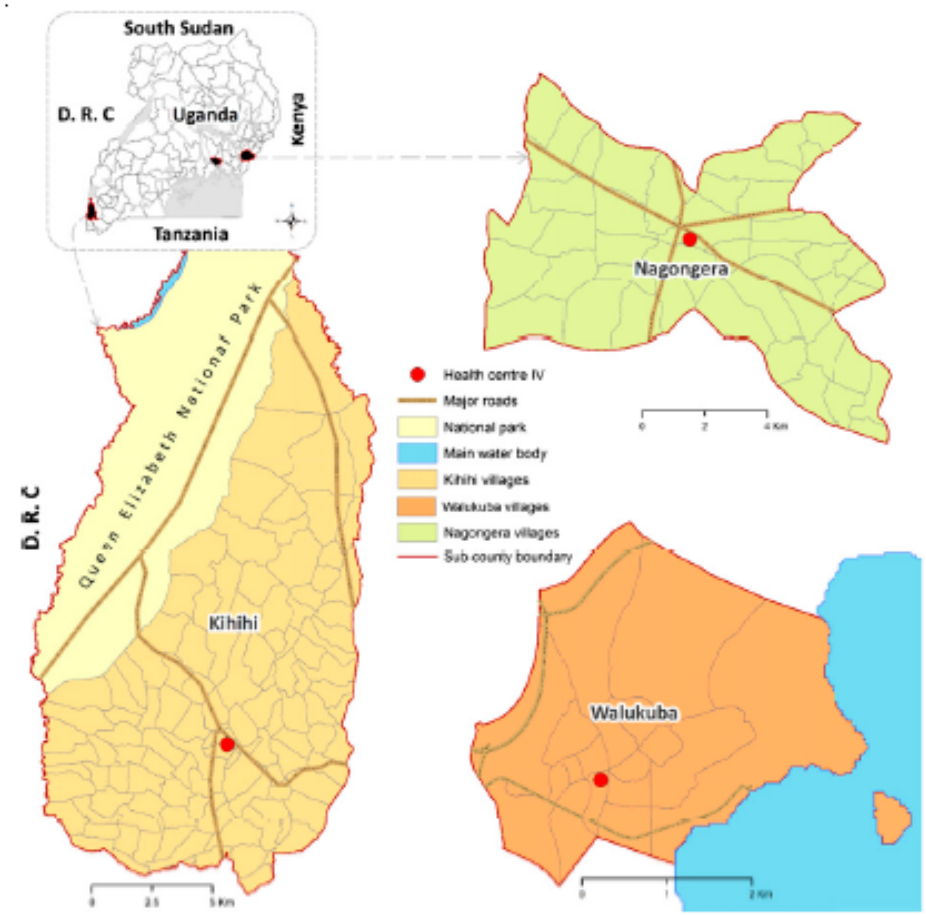


FIGURE 1. Location of the three sites of Nagongera (farthest east), Walukuba (central), and Kihihi (southwest) in Uganda, with the respective locations of the health facilities included. Red dots represent the site study health facilities, each being a level IV health center for the health subdistrict to which the respective subcounty (our catchment area) belongs—all shown on the inset map of Uganda in black. Kihihi (in shades of yellow) is located southwest of Uganda with a 2002 estimated annual entomological inoculation rate (aEIR) of 6 infective bites per person; it is home to Queen Elizabeth National Park at the border between Uganda and Democratic Republic of Congo. Walukuba (orange area) is located in the central part of Uganda at the shores of Lake Victoria with an aEIR of 6, whereas Nagongera (light-green area) is located far east close to the border between Uganda and Kenya with an aEIR of 562. This figure appears in color at www.ajtmh.org.

under-fives was generally lower in the low-transmission sites (Walukuba and Kihihi) than in over-fives, and vice versa in the high-transmission site of Nagongera. The pattern in Nagongera was consistent with findings previously reported from a high-transmission site of Aduku in Uganda.[9] Members of the study population whose clinic visit was classified as "re-attendance," referring to follow-up visits for a previously recorded episode, were excluded. All members of the study population with a new illness episode were included in temporal trend assessments, whereas those whose village of residence (VOR) was known to be within the facility–host subcounty boundaries (catchment area) were included in spatial–temporal assessments. Those whose VOR was either located outside the catchment area or unknown were excluded from the space–time evaluations (Figure 2).

Outcome measures. Outcome data for all outpatients were extracted from health facility outpatients' registers, which recorded the presence of fever, diagnostic test results, diagnosis made, and treatment prescribed, as well as demographic data, height, weight, and VOR. These data were entered from the register into an MS Access database (Microsoft Corporation, Redmond, WA) at the facility and regularly submitted to the UMSP data center, where they were checked for inconsistencies and cleaned before transfer to STATA (Stata Corporation, College Station, TX) for analysis. Detailed methods are further explained elsewhere.[10,21]

The study primary outcomes were the test-confirmed IR, defined as the number of confirmed primary malaria cases per 1,000 catchment study population of children aged < 11 years (by mRDT or microscopy), and TPR, defined as the proportion of patients (children aged < 11 years) presenting to the health facility who were suspected to have malaria and tested with a positive malaria test result (by either mRDT or microscopy). Suspected malaria was defined as patients with fever or a history of fever in the past 48 hours, without any other apparent cause. These are considered eligible for a malaria test from which positive results determine the confirmed malaria cases. Because of inaccessible population estimates of our study area from the most recent census, catchment study populations were estimated using cartographic boundaries and a gridded population surface using the ArcMap geographic information system (ESRI, Inc., Redlands, CA), with cartographic boundaries defined as the 2006 subcounty boundary for each health facility. Subcounty total population estimates for 2010 and 2015 were obtained from the worldpop.org project data portal.[22] The population of children aged less than 11 years for each year from 2010 to 2016 was estimated from the overall estimates using population pyramids provided in the 2002 national census results, available in 5-year age bins at the district level. It was assumed that the proportion of the population aged less than 11 years was the same for all subcounties within each district, the proportion of population aged 10 years was one-fifth of the population aged 5–9 years, and the intercensal population growth rate was 3.0%, as published by the National Bureau of Statistics from the period of 2002 to 2014.[23] Per village study population estimates were then derived based on the proportion of village households enumerated per village in each subcounty.

Analysis. The two outcome indicators were summarized by month and village, and examined for temporal and spatial

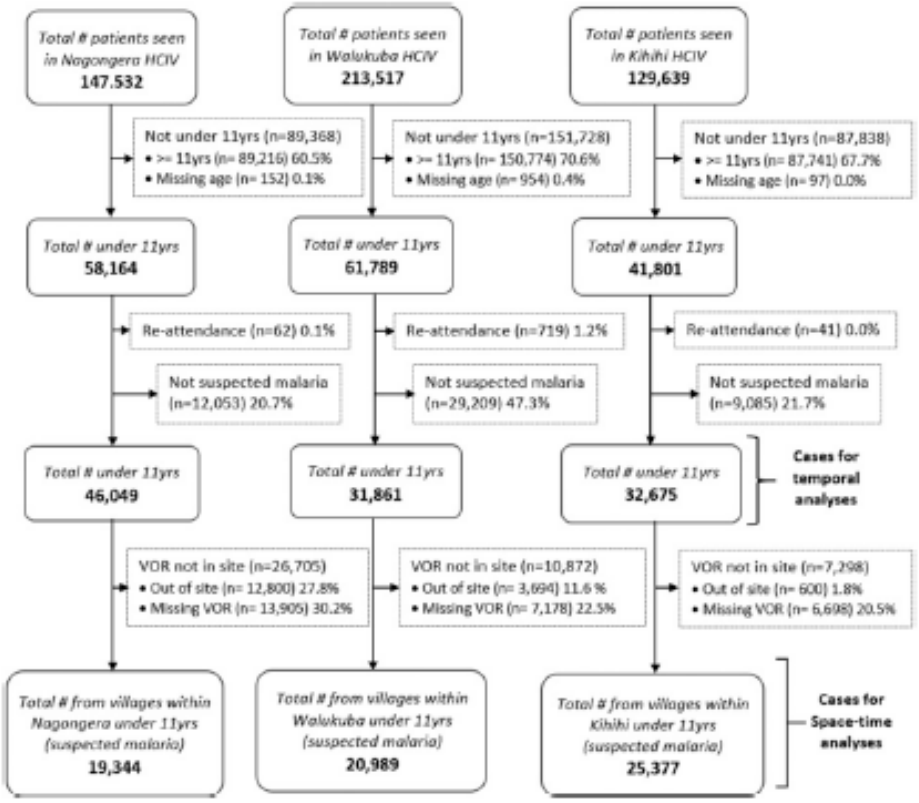


FIGURE 2. Trial profile indicating the participants included in the study and the exclusion criteria at the two levels of time and space–time evaluations. VOR = patients' village of residence.

trends. Straight-line distance to the health facility from the VOR was estimated using the ArcGIS point distance tool (ESRI, Inc.) between the centroid point of the village and the coordinate point of the health facility. Regression analyses were conducted using STATA versions 14 and 15 (Stata Corporation).

The pairwise relationship between IR and TPR was visualized using scatter plots and approximated using locally weighted scatterplot smoothing and quadratic prediction plots. As a measure of agreement between the indicators, concordance analysis was performed through evaluation of the Bland–Altman diagrams by site, with no predetermined threshold for agreement.[24] Two of the three sites were found to be eligible for concordance evaluation based on an approximately normal distribution of the differences between TPR and IR, thereby excluding Walukuba.

Factors influencing the association between TPR and IR per village per month were explored using mixed-effects Poisson regression models, with random effects at the village and month levels, to account for clustering at both levels. Model selection was performed using Akaike's information criteria (AIC) from a list of linear, quadratic, cubic, and exponential fits. Explanatory variables considered were age (proportion of the presenting population aged 5 years and more), gender (proportion of the presenting population that was male), distance to the health facility (by quartile), and season (determined using the predominant annual patterns of rainy [March–May and September–November] and dry [rest of the year] seasons in the southern parts of Uganda).[25] Explanatory variables with $P \leq 0.05$ in the unadjusted analysis were considered for inclusion in a fully adjusted multivariable model using a step-down process and evaluated using likelihood ratio tests.

RESULTS

Characteristics of the study population. A total of 161,754 patient visits by children aged less than 11 years were recorded for the three health facilities between October 2011 and June 2016, accounting for 33% of all patient visits. Among these, 110,585 (68%) were suspected malaria cases. A monthly mean of 346, 382, and 445 suspected malaria patients aged less than 11 years were seen in Nagongera, Walukuba, and Kihihi, respectively, over the study duration (Figure 2). The respective site mean (median) age in years among these participants was 2.8 (2), 3.8 (3), and 4.2 (4), respectively. In addition, the mean age of patients less than 11 years was significantly higher for females than males, both overall and within each of the three sites ($P < 0.001$). No differences were observed in trends of TPR between participants included or excluded based on the availability of VOR data, hence the assumption of no considerable impact on our results due to the exclusion. Additional participant characteristics are detailed in Table 1 for those included in the space–time assessments.

Trends in indicators of malaria. Figure 3 shows the trends in TPR and IR stratified by site over the study duration, on a monthly time scale, suggesting that the two show similar trends. In addition, there is evidence of good agreement between these indicators through concordance analysis using Bland–Altman's diagrams shown in Supplemental Figures 1 and 2 in the Supplemental Information, for the two sites that fulfilled Bland–Altman's criteria of approximately normally distributed differences.

Figure 4 shows the relationship between TPR and IR per month, suggesting a tendency for months with low case detection rates to also have low TPRs for all three sites. Given the strong seasonal trends in transmission, this confirmed TPR as a reasonable indicator for transmission intensity. However, when plotting TPR and IR by village, this relationship was lost as seen in Supplemental Figure 3 in the Supplemental Information. This suggests that factors other than intensity of transmission were influencing the number of malaria cases from each village that attend the health center for diagnosis and/or treatment. This was also confirmed through the spatial distribution of IR compared with that of TPR by village shown in Figure 5. Here, the highest IR is seen to cluster closer to the

TABLE 1

Characteristics of study population and distribution of suspected and clinically confirmed malaria cases for the duration of October 2011 through June 2016, evaluated in the space–time analysis for this study

Site	Nagongera, *N* (%)	Walukuba, *N* (%)	Kihihi, *N* (%)
Participants	19,344	20,989	25,377
Gender			
Male	9,407 (49%)	9,367 (45%)	12,003 (47%)
Female	9,937 (51%)	11,622 (55%)	13,371 (53%)
Age (years), mean (SD)			
All	2.79 (2.61)	3.78 (3.01)	4.17 (3.02)
Male	2.62 (2.49)	3.54 (2.93)	3.98 (2.93)
Female	2.94 (2.71)	3.97 (3.07)	4.34 (3.09)
Testing rates			
Diagnostic tests with recorded results (proportion of all participants who were tested)			
Tested and results recorded	18,655 (96%)	19,140 (91%)	25,048 (99%)
Testing rates by the diagnostic method (proportion of all participants tested by the method)			
Microscopy	17,701 (95%)	19,073 (99%)	25,043 (99%)
Rapid diagnostic test	954 (5%)	67 (0%)	5 (0%)
TPR	15,295 (38%)	7,910 (24%)	15,212 (47%)
TPR by gender, i.e., proportion of tested blood slides that were positive for malaria			
Male	3,479 (38%)	2,120 (25%)	5,699 (48%)
Female	3,637 (38%)	2,506 (24%)	6,107 (46%)
Malaria diagnosis			
Malaria diagnosed with a confirmatory test (proportion of all malaria diagnoses made that had a confirmatory test)			
Confirmed positive	7,604 (97%)	4,554 (97%)	11,856 (99%)

TPR = test positivity rate.

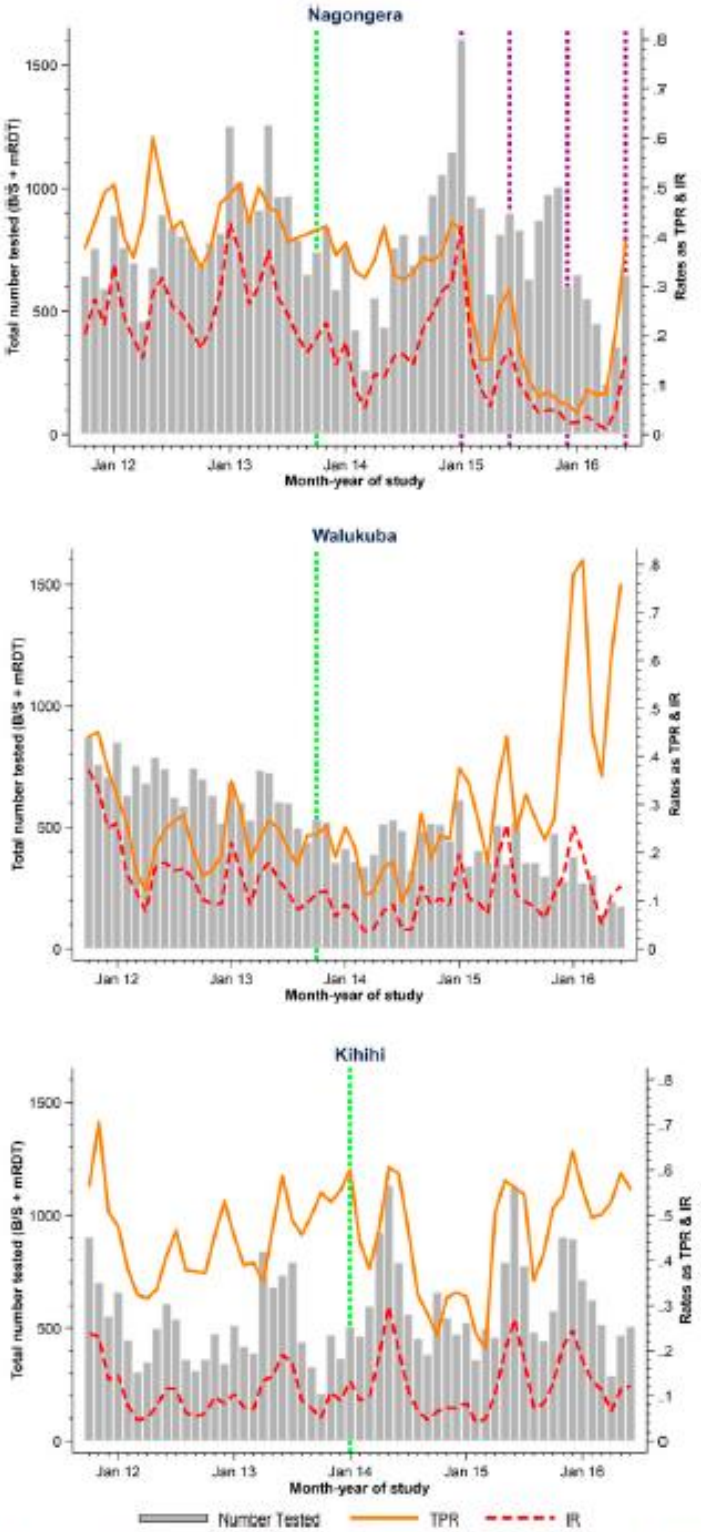


FIGURE 3. Trends in incidence and test positivity rates (TPRs) over time (monthly), by site. In Figure 3, the solid orange line represents TPR, whereas the dashed red line represents incidence rate, both overlaid on bar plots of total number of children < 11 years tested for malaria per month. A scale of monthly numbers tested is shown on the left *y* axis, whereas that for the two incidence measures (indicators) is on the right *y* axis. The green dotted line represents the time of a universal long-lasting insecticidal-treated net distribution campaign in each site, whereas the purple dotted lines represent the timing of indoor residual spraying in Nagongera. This figure appears in color at www.ajtmh.org.

health facility and less so further away, whereas TPR has no distinguishable spatial variation. Moreover, results in Figure 6 showed that there was no change in TPR with distance, whereas in Figure 7, IR was considerably reduced with increasing distance from the health facility in all three settings. The number of patients tested for malaria was not as highly varied by month as it was by village, although Walukuba recorded steady decline across the study duration and

Nagongera recorded the highest mean (arithmetic) monthly number of patients tested for malaria, as seen in Figure 3. Kihihi recorded the highest mean number of patients tested for malaria, by village, compared with the other sites, and Walukuba, the lowest for the same. The mean TPR per month ranged from 28.9% to 45.7%, whereas the village mean TPR per month ranged from 23.9% to 47.1% excluding outliers. In addition, the mean IR ranged from 11.8 to 18.6 cases per 1,000 children per month, whereas the village mean IR ranged from 21.8 to 40.3 cases per 1,000 children per month.

Association between the TPR and IR. Univariable analysis of the association between IR and explanatory variables revealed significant associations with the proportion of the presenting population aged more than 5 years, distance from the health facility, and TPR in all three settings; and for season in Nagongera (see Supplemental Information). Model selection revealed a cubic fit for TPR as best in all settings, based on AIC, compared with linear, quadratic, and exponential fits (see Supplemental Information).

In the final multivariable models, IR was significantly lower in villages further from the health facility (Table 2). Associations with seasonal and population factors also varied by site. Incidence rate was significantly lower during the rainy season in Nagongera only, whereas in Kihihi and Walukuba, when an increasing proportion presenting were aged more than 5 years, IR stayed borderline significantly associated with age (Table 2).

Comparison of village and temporal random effects suggested considerably more unexplained heterogeneity between villages than over time (month). Average variability between villages was lowest in Kihihi and highest in Nagongera, although this relationship was reversed when examining unexplained temporal variation, with Nagongera being lowest.

DISCUSSION

This study took advantage of enhanced malaria surveillance, conducted at three health facilities in varied malaria transmission settings in Uganda. We investigated the relationship between two standard indicators of malaria burden: TPR and IR. Findings showed that these indicators strongly agreed over time (month) in all three settings, most notably during periods of low transmission. However, TPR remained unchanged, whereas IR was drastically reduced with increased distance from the health facility. This further points to TPR as an appropriate proxy of transmission intensity regardless of the setting, but suggested that despite enhanced reporting, IR is strongly influenced by facility accessibility and/or use by the catchment populations using distance to the health facility as a proxy.[26] We demonstrated here that after

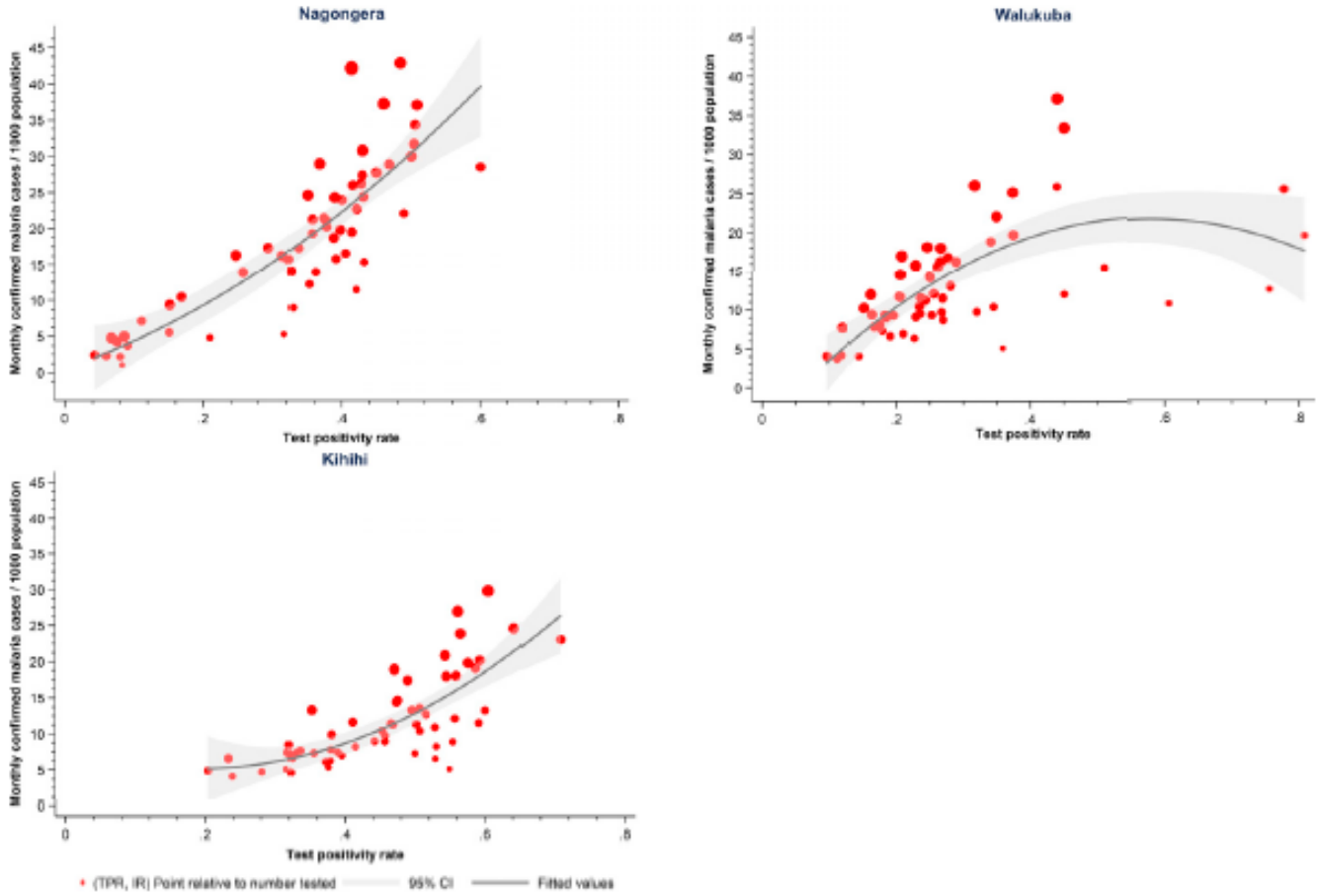


FIGURE 4. Plots of test positivity rate (TPR) against test-confirmed malaria case rate with points as months and point sizes accounting for the number tested for malaria by month. Each red dot here represents a month during the study duration, and the size of dots is relative to the number of suspected cases for independent episodes tested for malaria within each month. The gray curve is the fitted curve of the estimated relationship between the two measures of TPR and incidence rate. This figure appears in color at www.ajtmh.org.

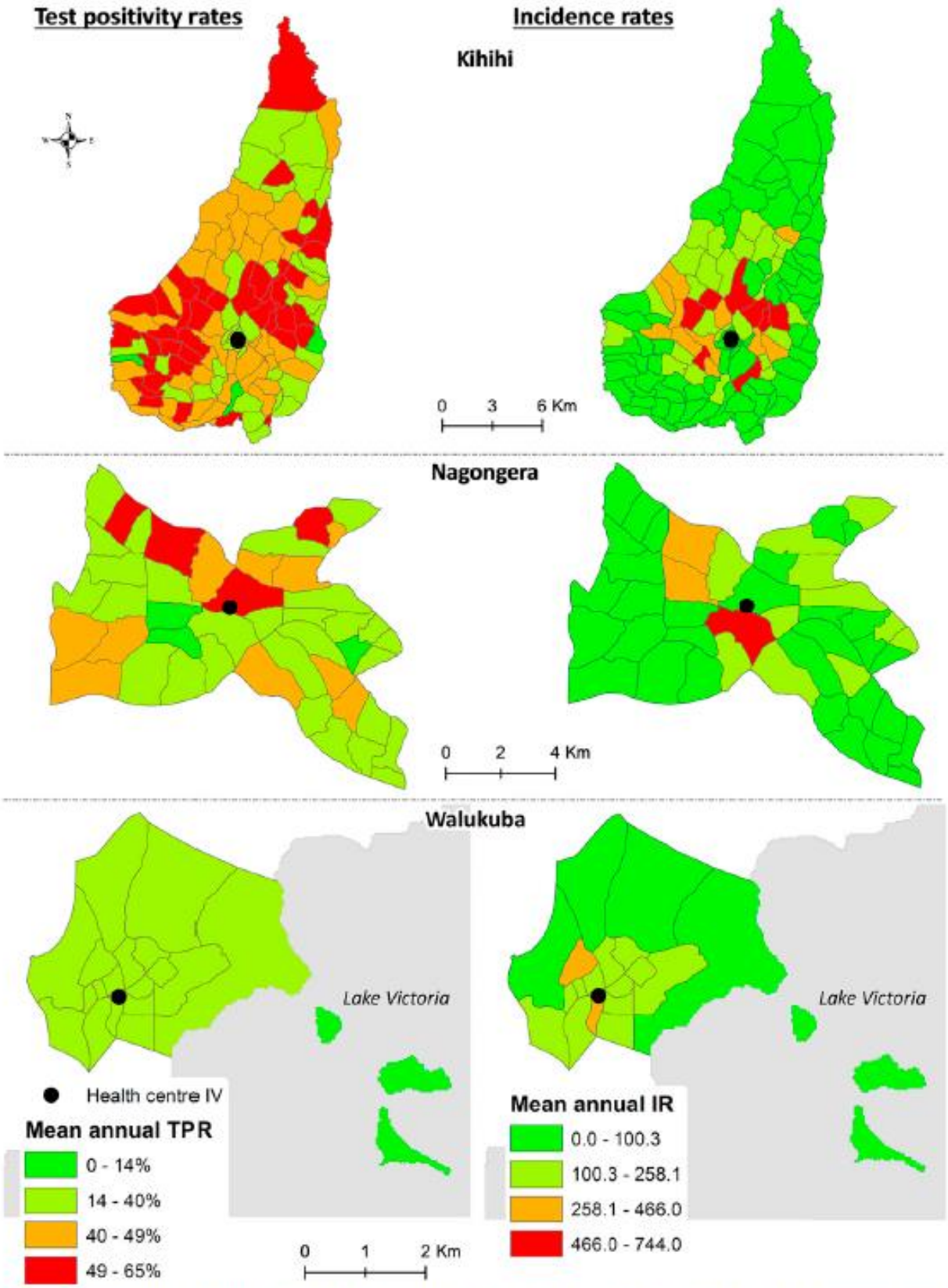


FIGURE 5. Comparison of the spatial distribution of incidence rate (IR) and test positivity rate (TPR) based on the village-level annual mean of each indicator in the three sites of Kihihi, Nagongera, and Walukuba. Three sites of Kihihi, Nagongera, and Walukuba shown side by side to depict a comparison between the spatial distribution of the village mean annual TPR (left) against IR (right). Village boundaries for each site are represented with the gray lines, whereas the respective site study health facility location is represented by the black dot. A single legend per indicator is placed at the bottom on respective sides. This figure appears in color at www.ajtmh.org.

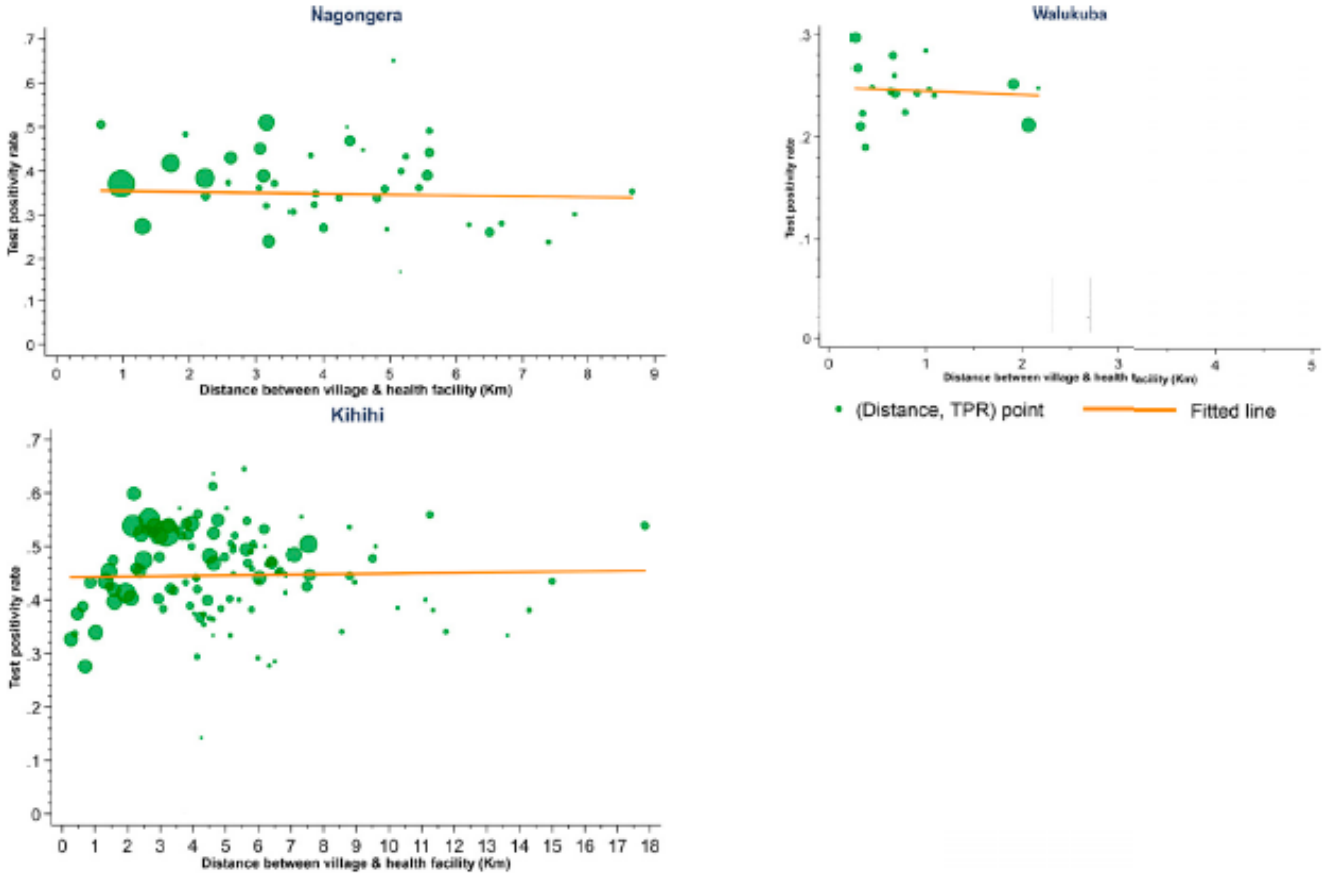


FIGURE 6. Plot of the distance between the village and health facility against the average annual test positivity rate (TPR), by site. Each point or circle corresponds to a village located in the site catchment area. The *y* axis is the average annual TPR, whereas the *x* axis is the distance of each village from the health facility in kilometers. The size of points is relative to the number of suspected cases tested for independent episodes of illness from the village overall. This figure appears in color at www.ajtmh.org.

accounting for transmission via TPR, some variability in IR remained unexplained both between villages and months. This pointed to important systemic differences in access to diagnosis and treatment within health facility catchments.

Trends in and concordance between these indicators over time (month) showed patterns suggestive of stronger agreement during low-transmission seasons. This was particularly evident in Nagongera during the period of intense IRS activity. This suggests that assessment of the trends over time in the two measures together may serve as a more sensitive marker of drastic changes in transmission within low-transmission settings and, thus, jointly provide an important surveillance tool. Consistent with other studies, including one in western Uganda,[27] we demonstrated that the relationship between TPR and IR was nonlinear, with both indicators increasing steadily until TPR reached around 50%. This suggested that either an inherent saturation point in TPR or that dynamics in treatment-seeking and care practices masked any further increase in TPR. It also reiterated the presence of malaria heterogeneity within each site.

The three-level mixed-effects Poisson model results from this study revealed that after controlling for the transmission intensity (via TPR), IR remained significantly influenced by the age of the patients, distance to the health facility from the VOR, and climatic conditions represented by a wet or dry season of the year. Studies have previously reported the influence of village or "area of residence" and age on TPR[27] as well as season[25] on malaria through routine reporting. Consequently, age, distance, village, and time (month) or season (rain versus dry) were important factors in the assessment of malaria burden through routine health facility–based surveillance. Concerning age, the small but strongly significant association of increased IRs with age in Kihihi and Walukuba but not Nagongera could be explained by Nagongera having significantly lower mean age, coupled with its historically significantly higher parasite prevalence in children than the other two sites.[17] Distance to the health facility, the increase of which was associated with a significant decline in IR, is known to influence the care-seeking behavior[28] for reasons such as cost[29] and, thus, may be considered a proxy for accessibility.[30] There were significant associations between distance and IR in all three settings. These highlight a strong influence of access to health facility and factors associated with access such as the use of drug shops as the first action, reported to be anywhere between 25% in patients of all ages[31] and 62.7% in children aged less than 5 years,[32] in parts of Uganda. Accessibility may also provide an explanation for the reduced IR observed in Nagongera during the rainy season, given that floods, intense farming activity, and inaccessible roads are common in this area during this time. It should be noted that distance did not explain all between-village variation, suggesting that other factors play an important role in health facility accessibility and use.

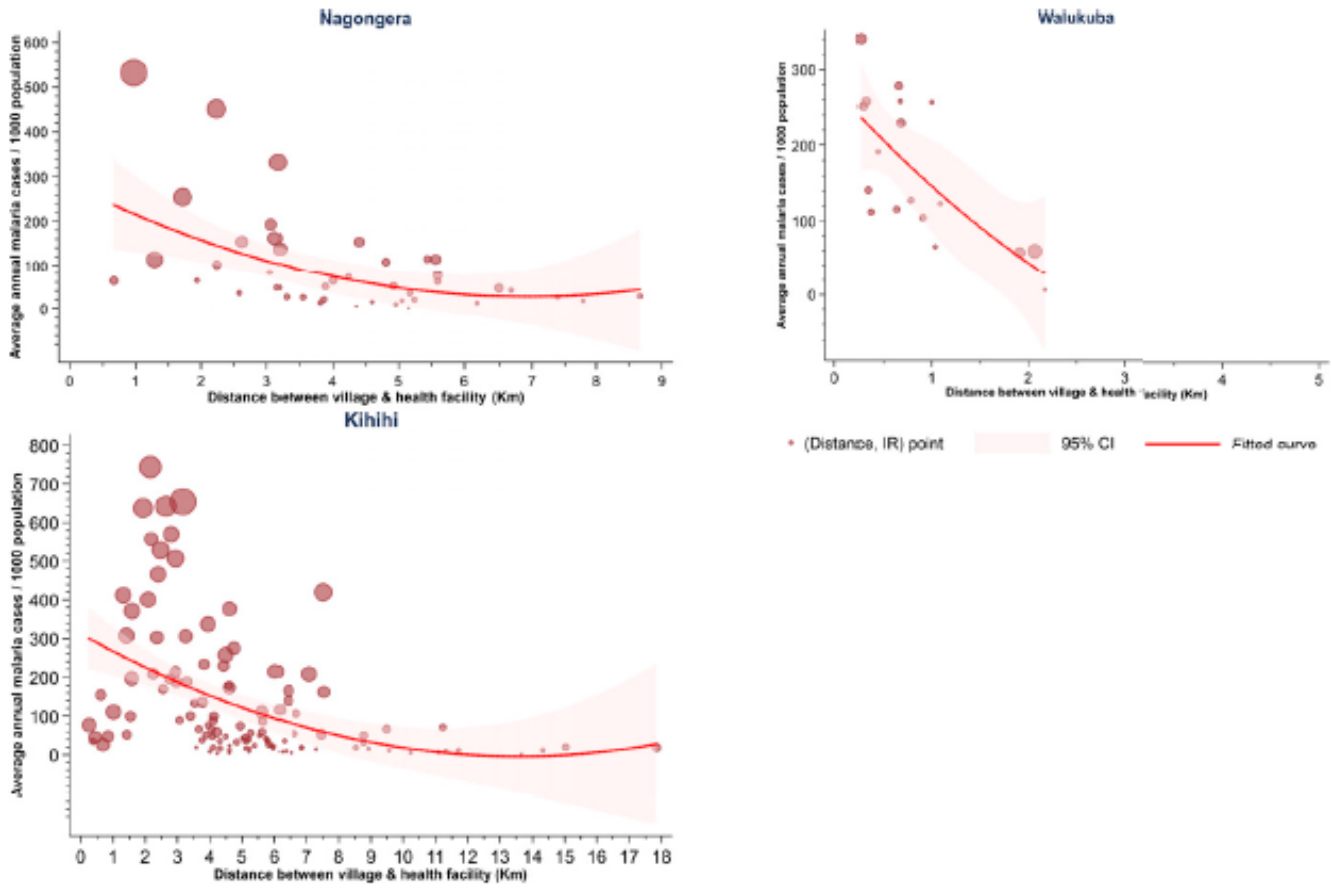


FIGURE 7. Plot of the distance between the village and health facility against the average annual incidence rate (IR), by site. Each point or circle corresponds to a village located in the site catchment area. The *y* axis is the average annual IR, whereas the *x* axis is the distance of each village from the health facility in kilometers. The size of points is relative to the number of suspected cases tested for independent episodes from the village overall. This figure appears in color at www.ajtmh.org.

There were a number of limitations that should be acknowledged. Incomplete records represent a major challenge when interpreting health facility data. However, an evaluation of the greatest source of missing data (VOR) suggested no systematic differences between the two populations. Distance to the health facility could only be measured as a straight-line distance between the health facility and the centroid of the village as more detailed information was unavailable within the confines of routine passive surveillance. In addition, given possible repeated measures on the individual

TABLE 2

Mixed-effects Poisson model results (adjusted) assessing association between incidence rate and TPR in Nagongera, Kihihi, and Walukuba controlling for age, distance to the health facility in all sites, gender, and season in Nagongera and Kihihi as fixed effects and including random effects of village of residence and month of study year

		Nagongera		Kihihi		Walukuba	
		Fixed effects					
Exposure		IRR (95% CI)	*P*-value	IRR (95% CI)	*P*-value	IRR (95% CI)	*P*-value
Case positivity	Cubic TPR term	1.00 (1.00–1.00)	< 0.001	1.00 (1.00–1.00)	< 0.001	1.00 (1.00–1.00)	< 0.001
	Quadratic TPR term	0.98 (0.98–0.98)	< 0.001	0.98 (0.98–0.98)	< 0.001	0.98 (0.98–0.98)	< 0.001
	Linear TPR term	1.11 (1.09–1.13)	< 0.001	1.18 (1.17–1.19)	< 0.001	1.02 (1.00–1.05)	0.111
Age	Increasing proportion of ≥ 5 years	1.00 (0.97–1.02)	0.727	1.01 (1.00–1.02)	< 0.043	1.03 (1.00–1.06)	0.068
Distance to health facility	1st quartile	1	Reference	1	Reference	1	Reference
	2nd quartile	0.48 (0.28–0.83)	0.009	1.06 (0.76–1.49)	0.723	0.70 (0.31–1.60)	0.395
	3rd quartile	0.38 (0.22–0.67)	0.001	0.68 (0.49–0.94)	0.018	0.88 (0.43–1.79)	0.716
	4th quartile	0.40 (0.23–0.68)	0.001	0.55 (0.40–0.75)	< 0.001	0.25 (0.12–0.51)	< 0.001
Season	Dry/sunny	1	Reference	1	Reference	1	Reference
	Wet/rain	0.86 (0.80–0.91)	< 0.001	N/A	N/A	N/A	N/A
Random effects	Village (standard error)	0.3*75 (0.088)		0.305 (0.045)		0.325 (0.114)	
	Month (standard error)	0.075 (0.009)		0.104 (0.008)		0.105 (0.014)	

IRR = incidence rate ratio; TPR = test positivity rate; N/A = not applicable.

and household levels, there could be clustering at those levels that remained unaccounted for in this analysis. Nevertheless, repeated measures within the same episode of malaria were minimized using visit classification. Last, there are several lower level facilities within the same catchment of each of our study facilities that may influence the results observed; however, these data were not available at the same level of quality for inclusion in this study. Nevertheless, although lower level facilities absorb many malaria cases in their proximity, they too see patients from much farther away in a relatively similar form to the higher level facilities (such as our study facilities) for myriad reasons, which may undermine their influence on results in this study.

CONCLUSION

Strong nonlinear relationships between the two indicators of TPR and IR emphasize their distinct relevance to monitor malaria; however, caution is necessary in their interpretation. Given the strong impact of the distance of patients' residence from the health facility, a good proxy for the care accessibility, on IR and none on TPR, burden estimates in the assumed health facility catchment differs from one indicator to the other. The influence of access to health facility on IR depicts it as a good indicator for malaria burden in the health facility catchment, whereas the absence of the same effect of access on TPR suggests TPR as a good indicator for resource planning within the health facility system. More information is needed, however, on how well IR reflects the true burden on well-characterized catchments.

Received November 12, 2018. Accepted for publication March 30, 2019.

Published online May 6, 2019.

Note: Supplemental information appears at www.ajtmh.org.

Acknowledgments: We would like to thank the leadership at the National Malaria Control Programme; district health leadership in Tororo, Jinja, and Kanungu districts; staff at the three health center IV's of Nagongera, Walukuba, and Kihihi; the Uganda malaria surveillance project data officers at these sites; members of the UMSP project, especially Arthur Mpimbaza and Geoff Lavoy; and the wider IDRC team that have been instrumental to the smooth running of this surveillance work over the years.

Financial support: We acknowledge, with gratitude, funding from the President's Malaria Initiative, U.S. Agency for International Development under the terms of Interagency Agreement (1U51CK000117) with the CDC for the parent study and the Fogarty International Centers training grant (D43TW7375) for this substudy.

Availability of data and materials: The dataset used and/or analyzed during the present study are available from the corresponding author on reasonable request.

Authors' addresses: Simon P. Kigozi and Rachel L. Pullan, Department of Disease Control, London School of Hygiene & Tropical Medicine, London, United Kingdom, E-mails: skigozi@yahoo.com and rachel.pullan@lshtm.ac.uk. Ruth N. Kigozi, USAID's Malaria Action Program for Districts, Kampala, Uganda, E-mail: kigoziruth@gmail.com. Asadu Sserwanga, Infectious Diseases Research Collaboration, Kampala, Uganda, E-mail: asserwanga@idrc-uganda.org. Joaniter I. Nankabirwa and Moses R. Kamya, School of Medicine, Makerere University College of Health Sciences, Mulago Hospital Complex, Kampala, Uganda, E-mails: jnankabirwa@yahoo.co.uk and mkamya@infocom.co.ug. Sarah G. Staedke, Department of Clinical Research, London School of Hygiene & Tropical Medicine, London, United Kingdom, E-mail: sarah.staedke@lshtm.ac.uk.

REFERENCES

1. WHO, 2016. *World Malaria Report 2016*. Geneva, Switzerland: World Health Organization, 38–44.
2. WHO, 2014. *World Malaria Report 2014*. Geneva, Switzerland: World Health Organization, xii.
3. Zu Erbach-Schoenberg E et al., 2016. Dynamic denominators: the impact of seasonally varying population numbers on disease incidence estimates. *Popul Health Metr 14:* 35.
4. WHO, 2015. *Guidelines for the Treatment of Malaria*, 3rd edition. Geneva, Switzerland: World Health Organization.
5. Boyce RM, Reyes R, Matte M, Ntaro M, Mulogo E, Lin FC, Siedner MJ, 2016. Practical implications of the non-linear relationship between the test positivity rate and malaria incidence. *PLoS One 11:* e0152410.
6. WHO, 2012. *Diseases Surveillance for Malaria Control: An Operational Manual*. Geneva, Switzerland: World Health Organization.
7. Argaw MD, Woldegiorgis AGY, Abate DT, Abebe ME, 2016. Improved malaria case management in formal private sector through public private partnership in Ethiopia: retrospective descriptive study. *Malar J 15:* 352.
8. Githinji S, Noor AM, Malinga J, Macharia PM, Kiptui R, Omar A, Njagi K, Waqo E, Snow RW, 2016. A national health facility survey of malaria infection among febrile patients in Kenya, 2014. *Malar J 15:* 591.
9. Raouf S et al., 2017. Resurgence of malaria following discontinuation of indoor residual spraying of insecticide in an area of Uganda with previously high-transmission intensity. *Clin Infect Dis 65:* 453–460.
10. Sserwanga A et al., 2011. Improved malaria case management through the implementation of a health facility-based sentinel site surveillance system in Uganda. *PLoS One 6:* e16316.
11. Aregawi M et al., 2017. Effect of anti-malarial interventions on trends of malaria cases, hospital admissions and deaths, 2005–2015, Ghana. *Malar J 16:* 177.
12. Oguttu DW, Matovu JKB, Okumu DC, Ario AR, Okullo AE, Opigo J, Nankabirwa V, 2017. Rapid reduction of malaria following introduction of vector control interventions in Tororo District, Uganda: a descriptive study. *Malar J 16:* 227.
13. UMSP, 2016. *UMSP: Malaria Reference Centres Surveillance Report. January 2016. Kampala, Uganda: Uganda Malaria Surveillance Project*.
14. Bi Y, Hu W, Liu H, Xiao Y, Guo Y, Chen S, Zhao L, Tong S, 2012. Can slide positivity rates predict malaria transmission? *Malar J 11:* 117.
15. Okello PE, Van Bortel W, Byaruhanga AM, Correwyn A, Roelants P, Talisuna A, D'Alessandro U, Coosemans M, 2006. Variation in malaria transmission intensity in seven sites throughout Uganda. *Am J Trop Med Hyg 75:* 219–225.
16. Staedke SG et al., 2016. The impact of an intervention to improve malaria care in public health centers on health indicators of children in Tororo, Uganda (PRIME): a cluster-randomized trial. *Am J Trop Med Hyg 95:* 358–367.
17. Kamya MR et al., 2015. Malaria transmission, infection, and disease at three sites with varied transmission intensity in Uganda: implications for malaria control. *Am J Trop Med Hyg 92:* 903–912.
18. Mabaso ML, Craig M, Vounatsou P, Smith T, 2005. Towards empirical description of malaria seasonality in southern Africa: the example of Zimbabwe. *Trop Med Int Health 10:* 909–918.
19. Mabaso ML, Vounatsou P, Midzi S, Da Silva J, Smith T, 2006. Spatio-temporal analysis of the role of climate in inter-annual variation of malaria incidence in Zimbabwe. *Int J Health Geogr 5:* 20.
20. Doolan DL, Dobano C, Baird JK, 2009. Acquired immunity to malaria. *Clin Microbiol Rev 22:* 13–36.
21. Kigozi R et al., 2012. Indoor residual spraying of insecticide and malaria morbidity in a high transmission intensity area of Uganda. *PLoS One 7:* e42857.
22. Linard C, Gilbert M, Snow RW, Noor AM, Tatem AJ, 2012. Population distribution, settlement patterns and accessibility across Africa in 2010. *PLoS One 7:* e31743.

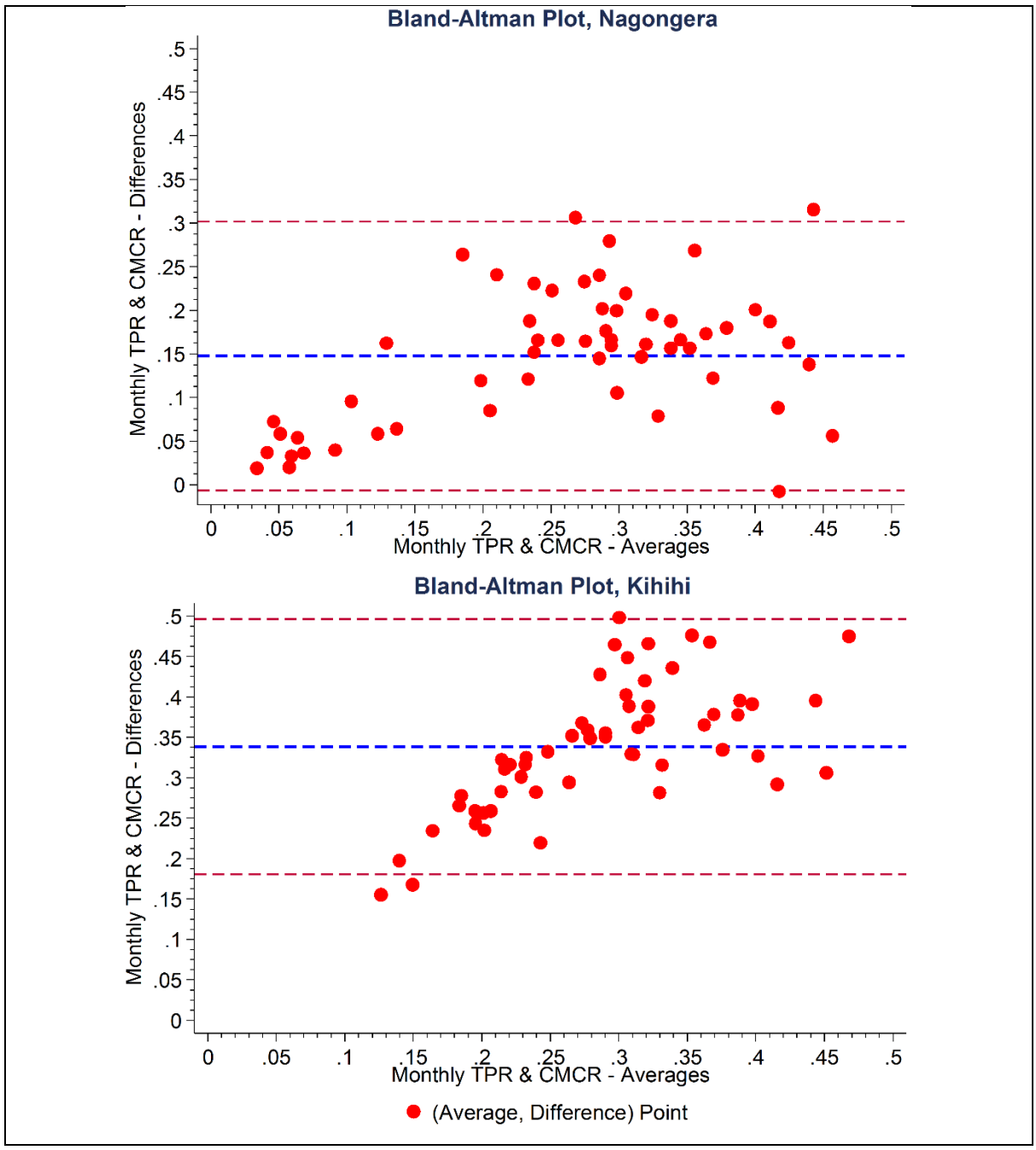
23. UBOS, 2016. *The National Population and Housing Census 2014–Main Report*. Kampala, Uganda: Uganda Bureau of Statistics.
24. Kwiecien R, Kopp-Schneider A, Blettner M, 2011. Concordance analysis: part 16 of a series on evaluation of scientific publications. *Dtsch Arztebl Int 108:* 515–521.
25. Kigozi R, Zinszer K, Mpimbaza A, Sserwanga A, Kigozi SP, Kamya M, 2016. Assessing temporal associations between environmental factors and malaria morbidity at varying transmission settings in Uganda. *Malar J 15:* 511.
26. Githinji S, Kigen S, Memusi D, Nyandigisi A, Wamari A, Muturi A, Jagoe G, Ziegler R, Snow RW, Zurovac D, 2014. Using mobile phone text messaging for malaria surveillance in rural Kenya. *Malar J 13:* 107.
27. Francis D, Gasasira A, Kigozi R, Kigozi S, Nasr S, Kamya MR, Dorsey G, 2012. Health facility-based malaria surveillance: the effects of age, area of residence and diagnostics on test positivity rates. *Malar J 11:* 229.
28. Amuge B, Wabwire-Mangen F, Puta C, Pariyo GW, Bakyaita N, Staedke S, Kamya M, Olico-Okui, 2004. Health-seeking behavior for malaria among child and adult headed households in Rakai district, Uganda. *Afr Health Sci 4:* 119–124.
29. Mbonye AK, Neema S, Magnussen P, 2006. Malaria in pregnancy, risk perceptions and care seeking practices among adolescents in Mukono district Uganda. *Int J Adolesc Med Health 18:* 561–573.
30. Muhumuza G, Mutesi C, Mutamba F, Ampuriire P, Nangai C, 2015. Acceptability and utilization of community health workers after the adoption of the integrated community case management policy in Kabarole district in Uganda. *Health Syst Policy Res 2:* 13.
31. Ndyomugyenyi R, Magnussen P, Clarke S, 2007. Malaria treatment-seeking behaviour and drug prescription practices in an area of low transmission in Uganda: implications for prevention and control. *Trans R Soc Trop Med Hyg 101:* 209–215.
32. Rutebemberwa E, Pariyo G, Peterson S, Tomson G, Kallander K, 2009. Utilization of public or private health care providers by febrile children after user fee removal in Uganda. *Malar J 8:* 45.

3.1 Additional information for Paper 1

3.1.1 Concordance analysis

Concordance analysis results are presented by month in Figure 1 and by village in Figure 2 below. The included two (Nagongera and Kihihi) of three sites are they that met the 'normal distribution of differences' criteria required in Bland-Altman's method, whereas Walukuba did not qualify and was therefore, excluded.

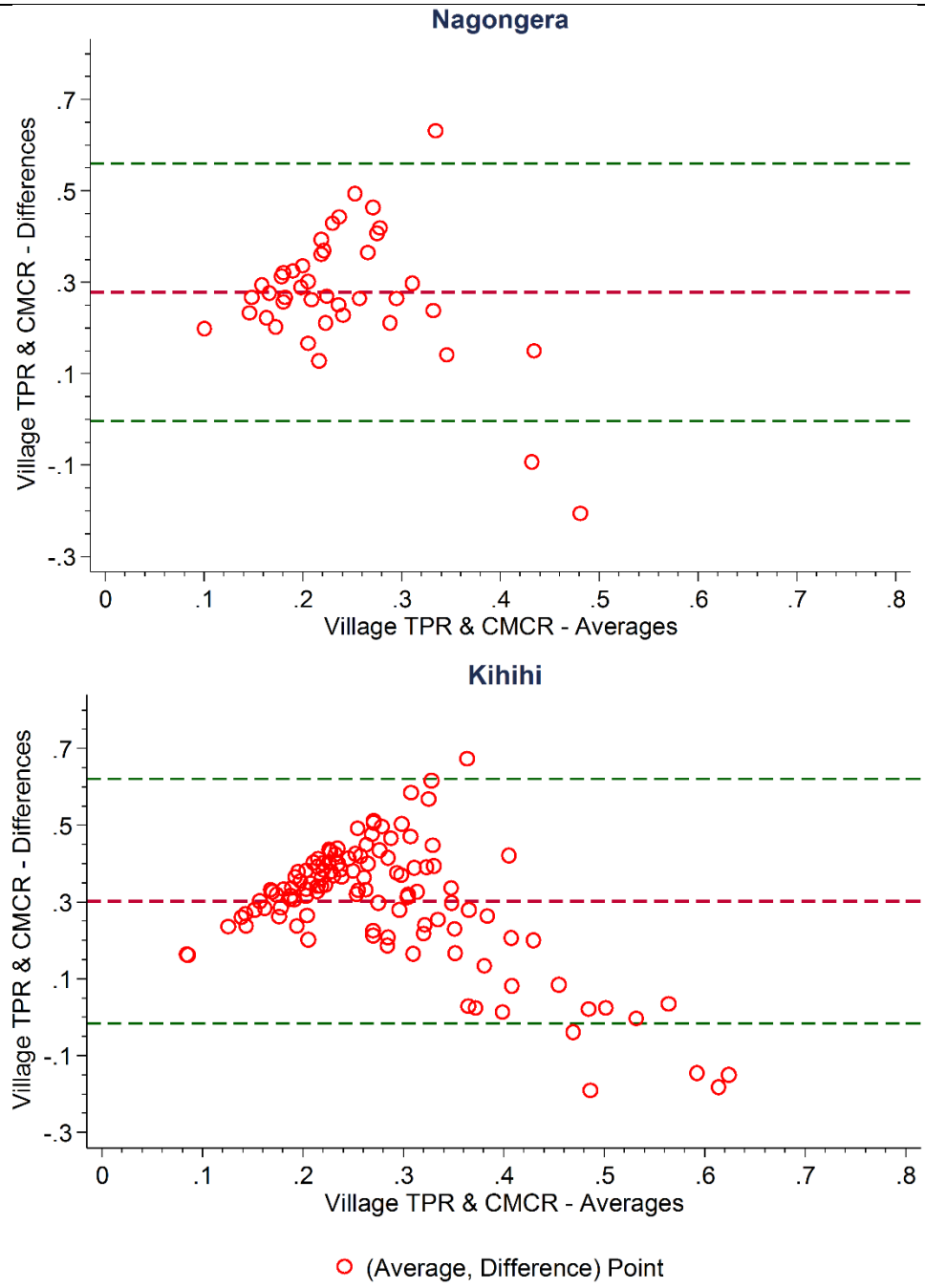
Figure 1. Bland-Altman diagram for Nagongera and Kihihi, assessing incidence estimates of TPR and IR at the level of time (month).



Each red dot represents a month of study year within each site, the dashed blue lines – the mean of differences, and the dashed red lines – the 95% agreement limits at approximately two standard deviations away from the mean.

The mean of differences for these monthly assessments (Figure 1) was much lower in Nagongera than Kihihi, being 0.148 and 0.338 respectively, a higher than two-fold and significant difference ($p<0.001$) with the means represented by the blue dashed line. However, the spread of limits of agreement was nearly the same for both sites i.e. 0.154 and 0.158 respectively, indicated by the dark-red dashed lines. Thus, difference between TPR and IR per month was less than 0.08 at both sites within 95% confidence bounds.

Figure 2. Bland-Altman diagram for Nagongera and Kihihi, assessing TPR against IR at by village, stratified by year of study.



The mean of differences by village is represented by the red dashed line for both sites. Also, spread of the limits of agreement by village is indicated by the area between green dashed lines, which mark the 95% confidence interval limits, at approximately two standard deviations from the mean. Each circle represents a four-year village average for each indicator for the respective site.

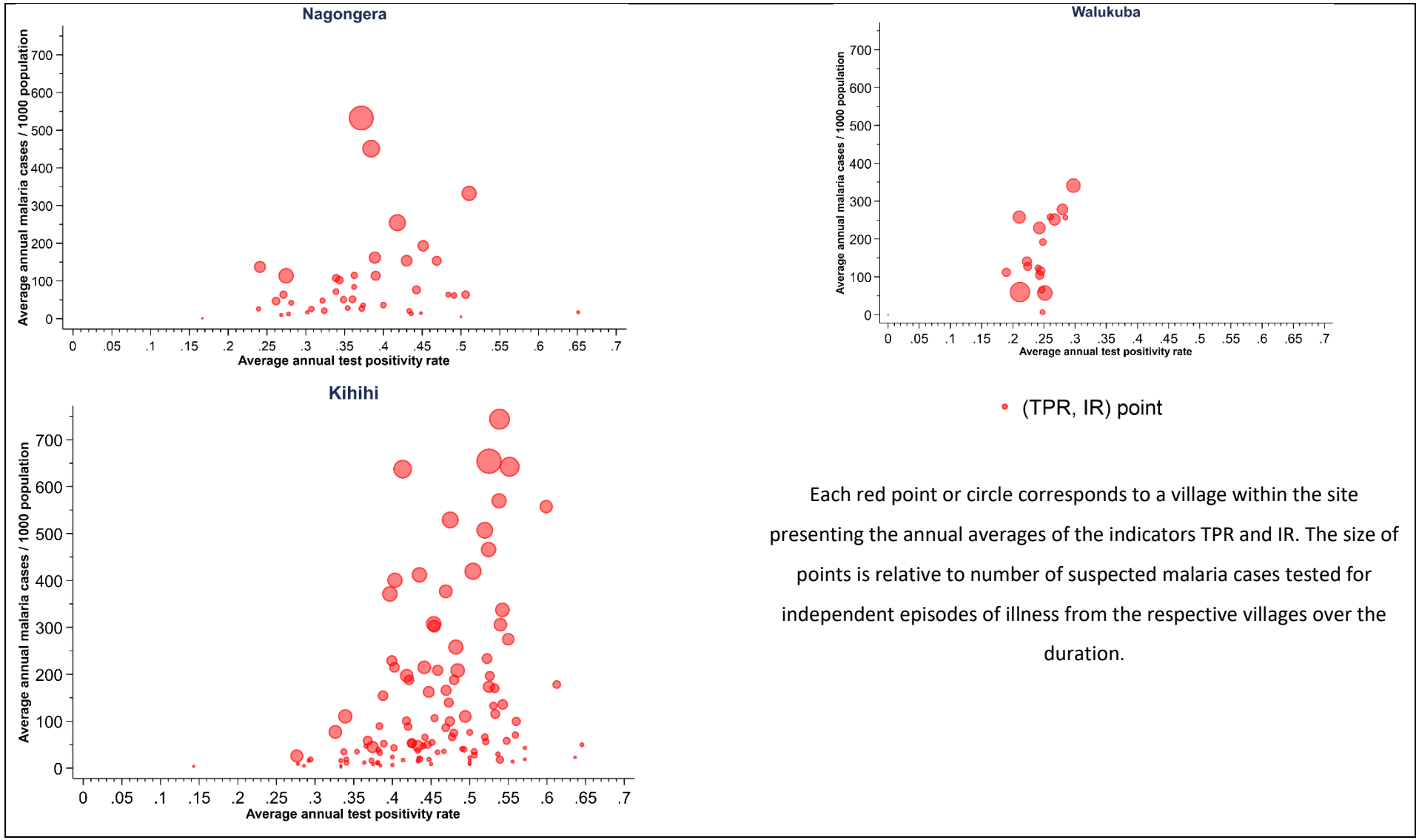
Concordance results (shown in Figures 1 and 2) revealed higher mean of differences between TPR and IR in Kihihi, 0.33 than Nagongera, 0.17 within 95% CI, suggesting a similarly large average difference between the sites. Furthermore, differences between TPR and IR by village were limited to 0.16 in either site and greater than

differences by month that were limited to 0.08 within 95% CI, pointing to greater heterogeneity between villages than months. Consistency in the differences between TPR and IR for either site on the two dimensions of month and village, provides further evidence in support of agreement between these indicators [182] regardless of transmission setting. By village, TPR was on average 30% higher than IR for both sites and there was an apparent relationship between variability in the two indicators and the quantity of each, with smaller differences observed at lower quantities in each of the indicators and greater differences as well as uncertainty when these are larger, that implies that there is greater agreement at lower transmission levels.

3.1.2 Relationship between TPR and IR

The relationship between the two indicators of TPR and IR by village, was explored using the mean annual value of each indicator, for each village. Here, unlike the case of the same examination by month presented in Figure 4 in the paper, the relationship is unclear as seen in Figure 3 below.

Figure 3. Scatter plot of village-level four-year average test positivity rate against annual test-confirmed malaria case rate, by site. Point sizes account for number tested for malaria by village



3.1.3 Univariable analysis

For each of the sites, explanatory variables including: sex or gender as 10% increments in the proportion of males among the study participants; distance to the health facility first determined in kilometres and then transformed to site specific quartiles; and, season determined using the predominant annual patterns of rain (March-May and September-November) and dry (rest of the year) seasons in the southern parts of Uganda, [183] were evaluated. For each site, age as 5% increments in the proportion of children 5 to under 11 years of age, was considered a default variable for inclusion, given that the existence of significant association between age with risk of infection is well known. [118, 184] Results from the univariate analysis are presented in Tables 1, 2, and 3 for Nagongera, Kihihi, and Walukuba, respectively.

Table 1. Mixed effects Poisson model results (crude) assessing associations in Nagongera between IR and TPR, age, gender, distance to health facility, and season as fixed effects; and, including random effects of village of residence and month of study year.

Exposure		Un-adjusted			
		Fixed effects		Random effects	
		IRR (95% CI)	p-value	Village (Std. Err.)	Month (Std. Err.)
Case positivity	TPR	1.14 (1.13-1.15)	<0.001	0.759 (0.175)	0.172 (0.016)
Age	Increasing proportion of >=5yrs	0.97 (0.94-1.00)	0.042	0.748 (0.173)	0.398 (0.030)
Gender	Increasing proportion of Males	1.01 (0.98-1.03)	0.595	0.721 (0.167)	0.395 (0.030)
Distance to health facility	1st Quartile	*1*	*Reference*	0.548 (0.128)	0.398 (0.030)
	2nd Quartile	0.45 (0.23-0.86)	0.016		
	3rd Quartile	0.34 (0.17-0.65)	0.001		
	4th Quartile	0.33 (0.17-0.62)	0.001		
Season	Dry / Sunny	*1*	*Reference*	0.743 (0.172)	0.387 (0.029)
	Wet / Rain	0.83 (0.75-0.91)	<0.001		

Table 2. Mixed effects Poisson model results (crude) assessing associations in Kihihi between IR and TPR, age, and distance to health facility as fixed effects; and, including random effects of village of residence and month of study year.

Exposure		Un-adjusted			
		Fixed effects		Random effects	
		IRR (95% CI)	p-value	Village (Std. Err.)	Month (Std. Err.)
Case positivity	TPR	1.12 (1.11-1.13)	<0.001	0.545 (0.079)	0.225 (0.013)
Age	Increasing proportion of >=5yrs	1.06 (1.04-1.07)	<0.001	0.534 (0.079)	0.428 (0.021)
Gender	Increasing proportion of Males	1.01 (0.99-1.02)	0.298	0.565 (0.082)	0.442 (0.022)
Distance to health facility	1st Quartile	*1*	*Reference*	0.459 (0.067)	0.442 (0.022)
	2nd Quartile	1.16 (0.78-1.74)	0.467		
	3rd Quartile	0.66 (0.45-0.98)	0.041		
	4th Quartile	0.49 (0.33-0.72)	<0.001		
Season	Dry / Sunny	*1*	*Reference*	0.565 (0.082)	0.442 (0.022)
	Wet / Rain	0.93 (0.88-0.99)	0.034		

Table 3. Mixed effects Poisson model results (crude) assessing associations in Walukuba between IR and TPR, age, and distance to health facility as fixed effects; and, including random effects of village of residence and month of study year.

Exposure		Un-adjusted			
		Fixed effects		Random effects	
		IRR (95% CI)	p-value	Village (Std. Err.)	Month (Std. Err.)
Case positivity	TPR	1.14 (1.12-1.15)	<0.001	0.671 (0.230)	0.285 (0.027)
Age	Increasing proportion of >=5yrs	1.09 (1.05-1.13)	<0.001	0.594 (0.205)	0.481 (0.041)
Gender	Increasing proportion of Males	1.03 (0.99-1.08)	0.088	0.645 (0.222)	0.494 (0.042)
Distance to health facility	1st Quartile	*1*	*Reference*	0.320 (0.115)	0.495 (0.042)
	2nd Quartile	0.72 (0.32-1.62)	0.423		
	3rd Quartile	0.84 (0.45-0.98)	0.623		
	4th Quartile	0.25 (0.33-0.72)	<0.001		
Season	Dry / Sunny	*1*	*Reference*	0.646 (0.222)	0.494 (0.042)
	Wet / Rain	0.92 (0.81-1.04)	0.202		

3.1.4 Model selection

The best model fit was selected using the Akaike's information criteria where the model with the lowest value is considered better than others with higher values. This model can be considered as the model with maximum precision using all the important covariates accounted for. In this study, four models were considered including the linear, the quadratic, the exponential and the cubic. Results for each of these models considered are presented in Table 4 below, indicating that the cubic was preferable.

Table 4. Akaike's information criteria values for the models each compared to the linear model to determine significant improvement of the linear model to fit the relationship between TPR and IR

Site	Model			
	Linear	Quadratic	Exponential *	Cubic
Nagongera	5847.68	5363.82	5650.67	**5317.82**
Kihihi	13298.93	11878.39	12399.11	**11857.46**
Walukuba	3828.50	3510.93	3710.12	**3452.89**

**The exponential model considered here was one that included a linear term of TPR given it was better than model that was purely exponential and excluded a linear term*

3.1.5 Multi-variable analysis

The cubic fit of the model, as compared to the linear, quadratic, and exponential models was selected as best based on AIC (Table 4). This fitted relationship from the multi-variable model was presented as a predicted plot using values of all covariates in the model, fixed at their mean values in each of the three sites (Figure 4).

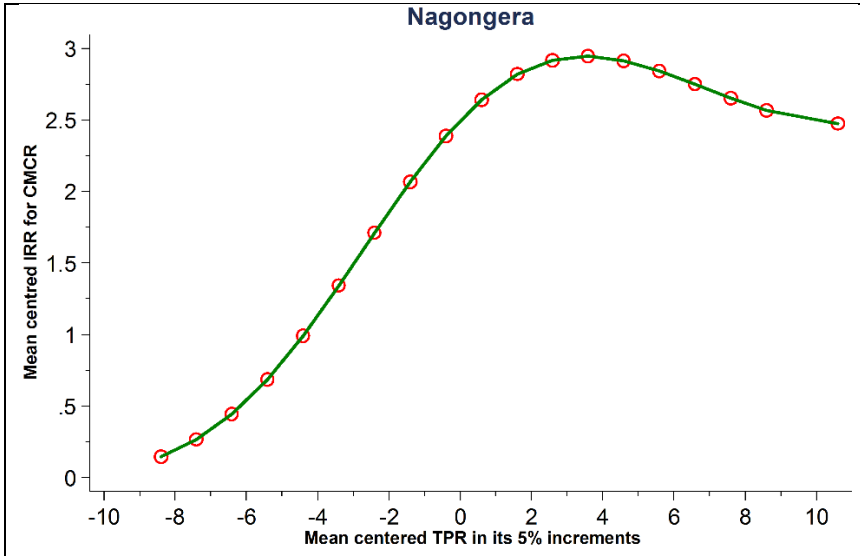
This relationship takes on the form of

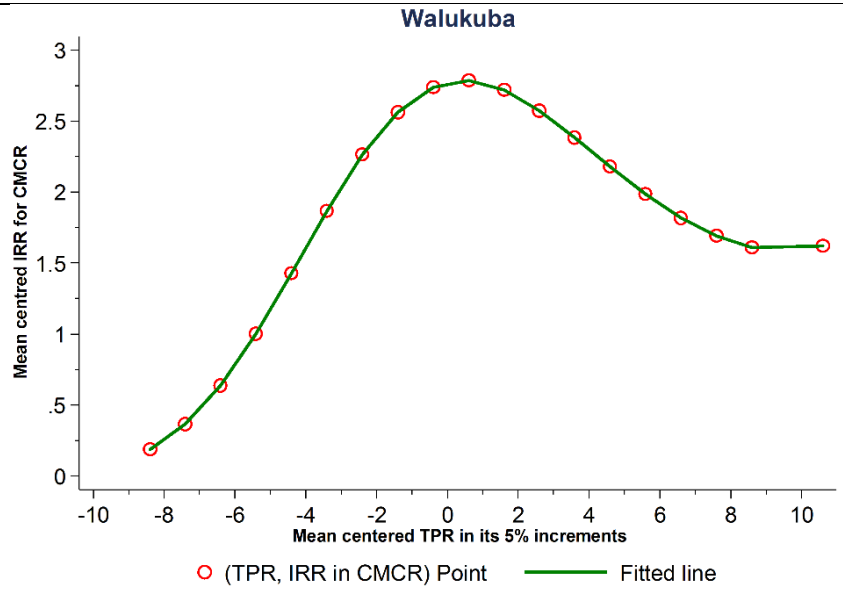
$$y = ax^3 + bx^2 + cx + \beta$$

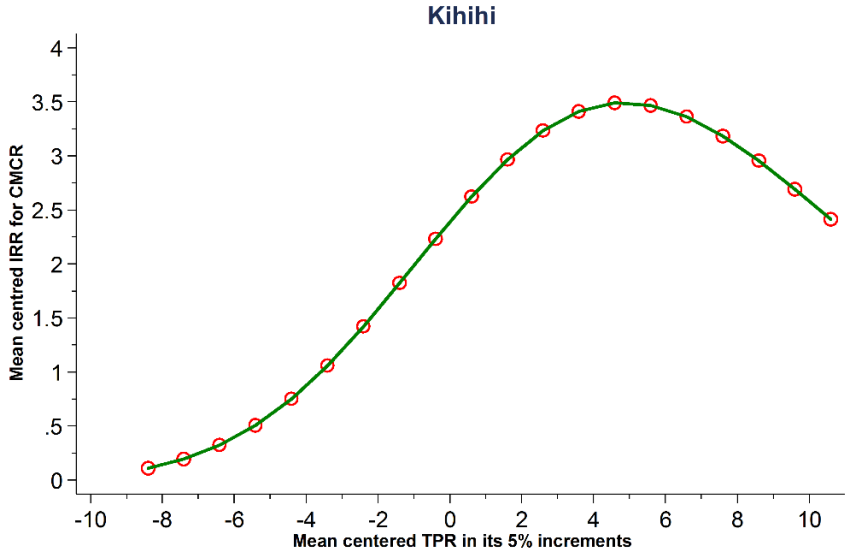
Where y = village IR per month, x = village TPR per month, and a, b, & c are coefficients, while β is an error term.

The same relationship between TPR and IR was sustained at all three settings with one exception in Walukuba where the linear term does not hold a significant effect. In all three settings, the fitted relationships between TPR and IR suggested that observed IR were highest when TPR was above the site mean, although the nature of the relationship had slight variations by site: in Nagongera, fitted IR peaked at 25% above the mean of TPR, whilst in Walukuba this was at 10% above and in Kihihi at 50% above mean of TPR (Figure 4).

Figure 4. Prediction plots for the relationship (cubic) between TPR and IR from the multi-variable mixed effects model for the sites of Nagongera, Walukuba, and Kihihi.







Predicted relationship between TPR and IR, with TPR centered around the mean of its 5% increments as fitted in the multi-variable model, and IR centered around its mean incidence rate ratio (IRR), by site. All three settings maintained a cubic relationship with slightly varied slopes, as well as peaks of IR.

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1513171	**Title**	Mr.
First Name(s)	SIMON PETER		
Surname/Family Name	KIGOZI		
Thesis Title	Exploring the utility of indicators of uncomplicated malaria burden from routine health facility surveillance data in identifying and mapping high-risk areas for malaria in Uganda		
Primary Supervisor	Assoc. Prof. Rachel L. Pullan		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

<u>SECTION B – Paper already published</u>

Where was the work published?	Malaria Journal		
When was the work published?	MARCH 2020		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I prepared the data, led data analysis, drafted the manuscript. I then led the manuscript submission for peer review, addressed reviewer comments, and managed the final submission for publication

SECTION E

Student Signature	simon peter kigozi
Date	01-September-2020

Supervisor Signature	Rachel Pullan
Date	3rd September 2020

RESEARCH Open Access

Rapid shifts in the age-specific burden of malaria following successful control interventions in four regions of Uganda

Simon P. Kigozi[1,2*], Ruth N. Kigozi[3], Adrienne Epstein[4], Arthur Mpimbaza[5], Asadu Sserwanga[2], Adoke Yeka[6], Joaniter I. Nankabirwa[2,7], Katherine Halliday[1], Rachel L. Pullan[1], Damian Rutazaana[8], Catherine M. Sebuguzi[2,8], Jimmy Opigo[8], Moses R. Kamya[2,7], Sarah G. Staedke[2,9], Grant Dorsey[2,10], Bryan Greenhouse[11,12] and Isabel Rodriguez-Barraquer[11]

Abstract

Background: Malaria control using long-lasting insecticidal nets (LLINs) and indoor residual spraying of insecticide (IRS) has been associated with reduced transmission throughout Africa. However, the impact of transmission reduction on the age distribution of malaria cases remains unclear.

Methods: Over a 10-year period (January 2009 to July 2018), outpatient surveillance data from four health facilities in Uganda were used to estimate the impact of control interventions on temporal changes in the age distribution of malaria cases using multinomial regression. Interventions included mass distribution of LLINs at all sites and IRS at two sites.

Results: Overall, 896,550 patient visits were included in the study; 211,632 aged < 5 years, 171,166 aged 5–15 years and 513,752 > 15 years. Over time, the age distribution of patients not suspected of malaria and those malaria negative either declined or remained the same across all sites. In contrast, the age distribution of suspected and confirmed malaria cases increased across all four sites. In the two LLINs-only sites, the proportion of malaria cases in < 5 years decreased from 31 to 16% and 35 to 25%, respectively. In the two sites receiving LLINs plus IRS, these proportions decreased from 58 to 30% and 64 to 47%, respectively. Similarly, in the LLINs-only sites, the proportion of malaria cases > 15 years increased from 40 to 61% and 29 to 39%, respectively. In the sites receiving LLINs plus IRS, these proportions increased from 19 to 44% and 18 to 31%, respectively.

Conclusions: These findings demonstrate a shift in the burden of malaria from younger to older individuals following implementation of successful control interventions, which has important implications for malaria prevention, surveillance, case management and control strategies.

Keywords: Malaria, Routine surveillance, Age distribution, Reduced transmission, Burden shift

Background

In Africa, the burden of malaria has decreased significantly, primarily through the scale-up of vector control interventions, including long-lasting insecticidal nets (LLINs), and indoor residual spraying of insecticide (IRS) [1–3]. These interventions were coupled with improved case management using artemisinin-based combination

*Correspondence: skigozi@yahoo.com
[1] Department of Disease Control, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK
Full list of author information is available at the end of the article

therapy (ACT), and intermittent preventive therapy with sulfadoxine-pyrimethamine for pregnant women [4, 5]. Whereas the impact of malaria control interventions is generally measured in terms of changes in *Plasmodium falciparum* infection prevalence and case numbers [1], more evidence is needed on how interventions influence the age distribution of malaria cases, a vital marker of progress in malaria control [6]. In high transmission settings, younger children bear the brunt of the malaria burden [7, 8], particularly for severe malaria and malaria deaths [9, 10]. However, it is unclear how quickly and to what extent the age distribution of uncomplicated malaria cases may shift with changes in transmission, following the successful implementation of control interventions.

Malaria surveillance efforts have generally focused on children under 5 years of age, a group that contributes the majority of reported cases [11] and are thus the focus of control measures and research, in areas of stable malaria transmission in sub-Saharan Africa. School-aged children (5–15 years) have received less attention due to the lower morbidity and occurrence of severe outcomes in this group [12]. However, older children often experience low-density asymptomatic infections and have been identified as important contributors to the infection reservoir for onward transmission [13]. Even less attention is given to adults (over 15 years), except for pregnant women [14], despite documented high prevalence of asymptomatic infections with high parasitaemia in adults [15, 16].

Commitments to malaria control made at the Abuja Summit in 2000 [17] led to World Health Organization (WHO) recommendations of universal coverage with LLINs, beginning in 2007 [18]. In Uganda, access to LLINs increased gradually, first among pregnant women and children under 5 years of age [19, 20]. This later culminated in universal LLIN campaigns aimed at one LLIN for every two household residents [21], with nationwide distributions in 2013–2014 [22] and 2017–2018. Furthermore, IRS was implemented in 10 districts effective 2010, then shifted to 14 districts from 2014 onwards [23, 24]. In northern Uganda, integrated community case management (iCCM) of malaria began from 2016 to 2017 and the 10 districts plus one also resumed IRS [25].

Following the scale-up of malaria control efforts in Uganda, confirmed malaria cases reported from health facilities declined by an average of 10.8% per year between 2013 and 2015 before increasing in 2016 [26]. Moreover, malaria indicator surveys (MIS) showed that parasite prevalence in children under 5 years of age declined from 42% in 2009 to 19% in 2014–2015 [27, 28], whilst malaria mortality reportedly decreased from 59 to 23 deaths per 100,000 between 2010 and 2017 [25].

This study aimed to investigate the impact of control interventions on the age distribution of malaria cases, using high-quality malaria surveillance data from four sites in Uganda, where mass LLIN distribution was conducted at all four sites and IRS implemented at two, one of which also received iCCM.

Methods

Malaria surveillance

The National Malaria Control Division (NMCD) of the Ministry of Health has conducted surveillance through the health management information system (HMIS) since 2007. However, more detailed malaria surveillance including collection of individual-level data has been conducted in selected malaria reference centres (MRCs) since 2006. MRCs are level III or IV health facilities located across the country, representing varied transmission settings as previously described [29, 30]. In summary, patients visiting these centres are assessed at the outpatient department (OPD) and basic information recorded using an OPD registry, including history of fever, age, gender, malaria diagnostic test results, and treatments prescribed. Individual visit-level OPD records are then entered into an MS Access database (Microsoft Corporation, Redmond, WA, USA) and sent to the Uganda Malaria Surveillance Programme (UMSP) data centre and cleaned before transfer to STATA (Stata Corporation, College Station, TX, USA) for analysis. Data from these sites are used to generate monthly reports that are shared and reviewed by the NMCD and other stakeholders.

Study sites

Four MRCs were selected for this study including: Walukuba in Jinja District, Kasambya in Mubende District, Aduku in Apac District, and Nagongera in Tororo Distict (Fig. 1). These sites were purposively selected because of their temporal representation of the malaria indicators of interest, with data covering the period of interest (January 2009 and July 2018), and have had malaria control intervention activities implemented in each of them. By transmission settings, Walukuba and Kasambya had an estimated annual entomological inoculation rate (aEIR) of less than 10 infective bites per person per year in the early 2000s, while in Aduku and Nagongera aEIR was in excess of 1500 and 550, respectively [31].

Variable description

Records in the database were classified either as 'suspected malaria' or 'no suspected malaria'. Suspected malaria was defined as patients who: (a) had a laboratory test done for malaria (microscopy or rapid diagnostic test

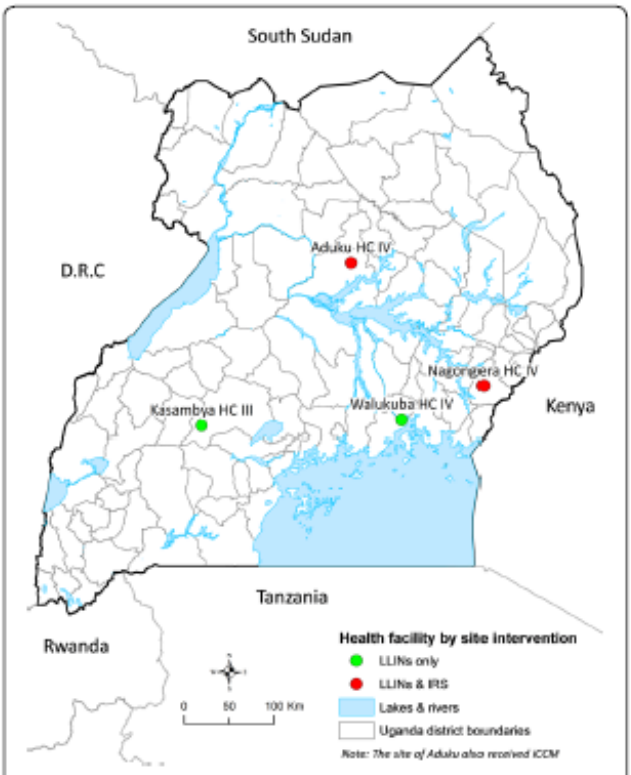


Fig. 1 Site locations of the four study health facilities categorized by the main intervention activity used

(RDT)); or, (b) were given a clinical diagnosis of malaria in the absence of laboratory testing. All records that did not meet this definition were classified as 'no suspected malaria'.

Among suspected malaria cases, 'malaria cases' were defined as all those patients with positive malaria diagnostic test results (microscopy or RDT). Moreover, 'malaria negative cases' were those who were tested for malaria (microscopy or RDT) but had negative test results.

Description of malaria control interventions at the study sites

Using available information from the NMCD on when specific interventions were implemented and/or interrupted at each study site, calendar time was divided into three to four different intervention periods per site (Fig. 2). The first period for each site, before large-scale interventions were implemented, is referred to as baseline. During baseline periods, control activities were largely limited to targeted distribution of insecticide-treated nets (ITNs) to vulnerable populations such as children under 5 years of age and pregnant mothers [32, 33]. The subsequent intervention periods were then used to quantify the impact of interventions on the age-distribution of test-confirmed malaria cases.

The main control interventions implemented in Walukuba and Kasambya were two mass LLIN distribution campaigns conducted in 2013 and 2017 at both sites. In addition to LLINs in the same time periods, Aduku and Nagongera also received IRS. The baseline period in Aduku included one round of IRS with the pyrethroid alphacypermethrin insecticide, which was not considered to be effective due to insecticide resistance [30]. This period in Aduku, was followed by nine rounds of IRS with the carbamate bendiocarb insecticide approximately every 6 months, as well as the first mass LLIN distribution campaign. IRS was stopped for 3 years and then a single round of IRS with the organophosphate Actellic insecticide was conducted, immediately followed by the second mass LLIN distribution campaign. In Nagongera, the baseline period was followed by the first mass LLIN distribution campaign and then a sustained period of IRS including three rounds with the bendiocarb insecticide, approximately every 6 months, followed by three rounds with Actellic approximately every 12 months. A second mass LLIN distribution campaign was also conducted in Nagongera during the sustained period of IRS [34].

Statistical analysis

Data from patients with age missing (0.3%), age over 70 years (1.2%), and follow-up visits for any previously recorded illness episode (0.7%) were excluded from the analyses. First, data were explored by calculating, for each site and intervention period, the total number of patients seen and the proportion of those that were suspected, tested and classified as confirmed malaria cases. To characterize changes in testing practices, proportion of cases that were tested by microscopy vs RDT were calculated. Test positivity rate (TPR) was defined as the proportion of patients tested for malaria that tested positive, a metric often considered a viable proxy for transmission intensity [35].

Trends in the gender distribution of patients with or without suspected malaria and malaria cases were also explored across intervention periods using cross-tabulation with Chi squared tests.

To characterize the impact of control interventions on the age distribution of malaria cases, first, violin density plots were used to visualize changes over the intervention periods and compare the distributions of patients not suspected of malaria to those with confirmed malaria. Second, scatter plots of age (as a continuous variable) and test positivity, stratified by intervention periods, per site were examined. Third, multinomial logistic regression models were fit, by site. The outcome in these models was 'the age category of

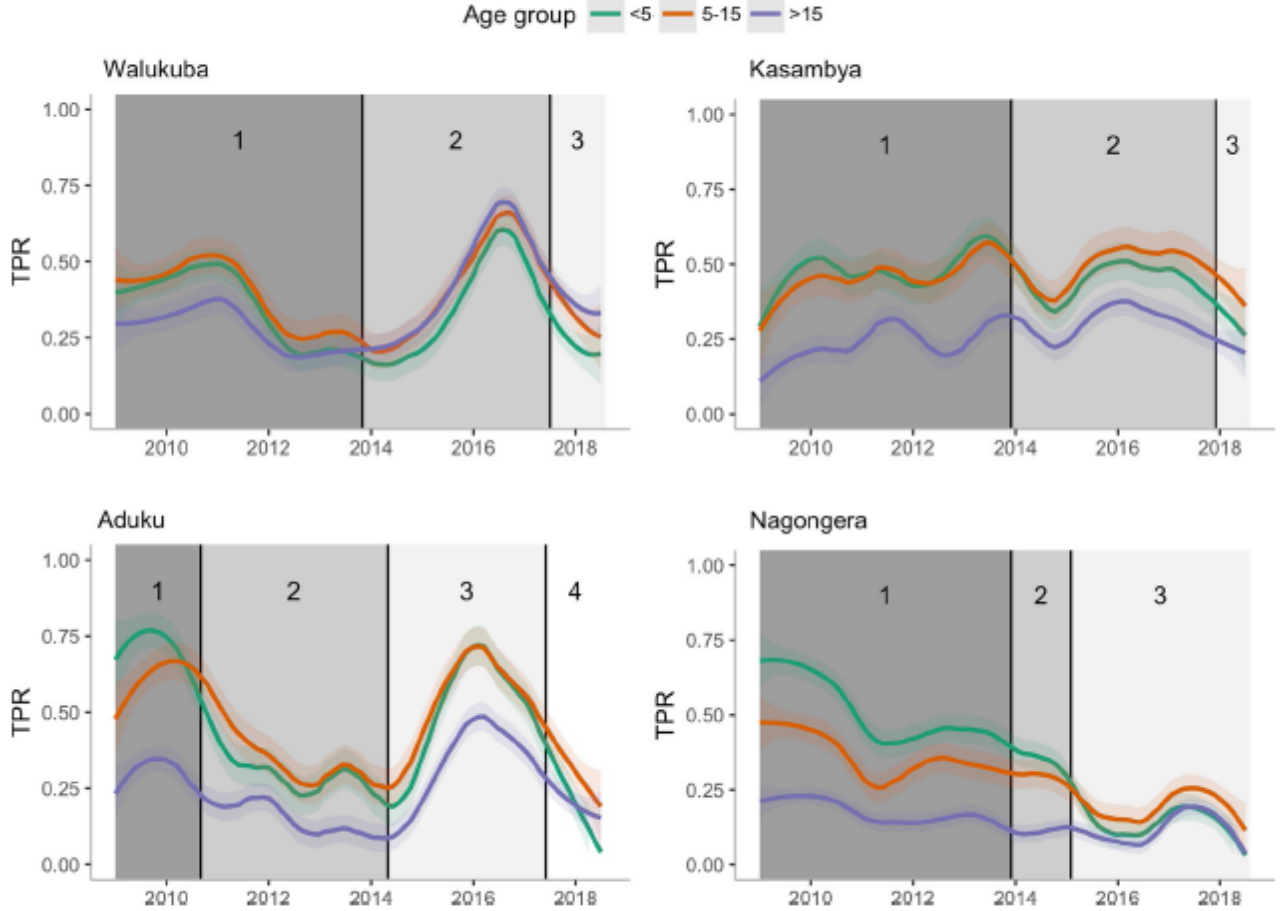


Fig. 2 Intervention periods, and TPR trends by age category (< 5, 5–15 and > 15 years) and site. Intervention periods, defined based on malaria control activities per site: Walukuba: LLINs alone. Period 1 (Jan 2009–Oct 2013): Prior to the first mass LLIN distribution campaign conducted during October 2013. Period 2 (Nov 2013–May 2017): Post first mass LLIN distribution and prior to second distribution campaign conducted during May 2017. Period 3 (Jun 2017–Jul 2018): Post second mass LLIN campaign to study endline. Kasambya: LLINs alone. Period 1 (Jan 2009–Nov 2013): Prior to first mass LLIN distribution campaign conducted during November 2013. Period 2 (Dec 2013–Nov 2017): Post first mass LLIN distribution campaign and prior to second distribution campaign conducted during November 2017. Period 3 (Dec 2017–Jul 2018): Post second mass distribution campaign to study endline. Aduku: LLINs, IRS and ICCM. Period 1 (Jan 2009–Aug 2010): One IRS round with the pyrethroid alpha-cypermethrin insecticide was implemented in March 2010. Period 2 (Sep 2010–Apr 2014): Received intense IRS campaign with nine rounds of carbamate bendiocarb approximately 6-monthly, the last of which was conducted during May 2014. The first mass LLIN distribution campaign was also conducted in July 2014. Period 3 (May 2014–May 2017): IRS withdrawal for 3 years till one round of IRS with the organophosphate Actellic insecticide conducted during May 2017. The second mass LLIN distribution campaign was also conducted over this period and ICCM implemented effective 2016. Period 4 (Jun 2017–Jul 2018): Post last round of IRS (May 2017) to study endline. Nagongera: LLINs and IRS. Period 1 (Jan 2009–Nov 2013): Prior to first mass LLIN distribution campaign conducted during November 2013. Period 2 (Dec 2013–Jan 2015): Post first mass distribution and prior to first IRS campaign conducted during January 2015. Period 3 (Feb 2015–Jul 2018): Period where six rounds of IRS were conducted, three rounds with bendiocarb approximately 6-monthly followed by three rounds with Actellic approximately 12-monthly, as well as the second mass LLIN distribution campaign conducted in May 2017

confirmed malaria cases' (under 5 years, 5–15 years, and over 15 years; three age categories being conveniently defined), while the main predictor was the intervention period. Models were adjusted for diagnostic test used (microscopy vs RDT), as well as patient gender. To validate the potential impact of increasing use of RDTs across sites over time on these findings, the multinomial regression models were fit with an interaction between diagnostic test and time. Using the full models, adjusted proportions of malaria cases in each of the age categories per intervention period were predicted. These marginal predictions were made at the mean values of the variables included in the model and analyses performed in R [36] and STATA 15 (Stata Corporation 2017, College Station, TX, USA).

Results

Patient composition and attendance by site

Between January 2009 and July 2018, the four health facilities of Walukuba, Kasambya, Aduku, and Nagongera recorded 896,550 patient visits in their outpatient department clinic registers, with over half of them (53%) suspected to be malaria cases. Walukuba recorded the highest number of both patient attendance (323,856) and suspected malaria cases (130,296), while Kasambya had the lowest attendance at 153,811 (Table 1).

The highest annual mean number of patients seen was in Walukuba (33,053) followed by Nagongera (22,701), then Aduku (20,079), and least in Kasambya (15,897). Mean monthly patient attendance per year remained fairly constant at all sites except Walukuba where this value peaked in 2011 at 3400 and steadily declined to 1952 by 2018 (Fig. S1, Additional file 1). Mean monthly attendance of patients not suspected of malaria per year increased slowly over time at all sites with cyclic variations (Fig. S2, Additional file 1). From 2009 to 2018, these increases were significant in Kasambya and Aduku though not in Walukuba or Nagongera by Wilcoxon rank-sum test (Table S1, Additional file 1). Conversely, mean monthly attendance of suspected malaria cases per year followed a general decline over time at all sites with cyclic variations (Fig. S3, Additional file 1). From 2009 to 2018, these declines were significant in Walukuba, Aduku and Nagongera, but not in Kasambya (Table S1, Additional file 1).

The majority of non-suspected and suspected malaria cases were female (67 and 63%, respectively). The difference between gender among confirmed malaria cases was smallest among children under 5 years of age (per cent female: Aduku 49%, Nagongera 49%, Walukuba 54%, and Kasambya 55%) and largest among patients over 15 years of age (per cent female: Aduku 79%, Nagongera 75%, Walukuba 63%, and Kasambya 66%). No observable trend in gender overall was found across intervention periods.

Trends in diagnostic testing over time

Throughout the study period, the majority of laboratory testing for malaria was by microscopy (89%) and the highest and lowest overall testing rates (percentage of suspected malaria cases that received a diagnostic test for malaria) were observed in Walukuba (97%) and Aduku (93%), respectively. Across intervention periods however, the proportion tested by RDT increased at all four sites mostly in the last 4 years of study duration. By the last period, 77% of cases were tested by RDT in Kasambya, 72% in Aduku, 40% in Nagongera, and 25% in Walukuba (Fig. S4, Additional file 1).

Test positivity rates

Although Aduku and Nagongera were historically the highest transmission settings, the highest TPR was observed in Aduku followed by Kasambya, Nagongera and Walukuba (Table 1). When considering only children under 5 years of age, however, baseline TPR levels were reflective of the historical transmission intensities. Baseline TPR in this group was highest at Nagongera (64%) and Aduku (63%), and lower in Kasambya (37%) and Walukuba (31%).

In all four sites, control interventions were associated with moderate reduction in overall TPR with

Table 1 Malaria-associated demographics of study participants for each site, by intervention period

Site	Intervention period	Total observations[a]	Suspected malaria (%)	Tested for malaria (%)[b]	Microscopy (%)	Positive test result (%)[b]
Walukuba	Jan 2009–Oct 2013	183,327	95,080 (51.9%)	92,784 (97.6%)	92,738 (99.9%)	30,785 (33.2%)
	Nov 2013–May 2017	111,506	27,483 (24.7%)	26,051 (94.8%)	24,360 (93.5%)	9239 (35.5%)
	Jun 2017–Jul 2018	29,023	7733 (26.6%)	7024 (90.8%)	5227 (74.4%)	2037 (29.0%)
Kasambya	Jan 2009–Nov 2013	85,200	65,768 (77.2%)	63,479 (96.5%)	60,070 (94.6%)	24,538 (38.7%)
	Dec 2013–Nov 2017	60,488	41,823 (69.1%)	38,328 (91.6%)	30,199 (78.8%)	15,996 (41.7%)
	Dec 2017–Jul 2018	8123	5272 (64.9%)	5083 (96.4%)	1156 (22.7%)	1529 (30.1%)
Aduku	Jan 2009–Aug 2010	34,596	19,024 (55.0%)	17,877 (94.0%)	17,877 (100%)	9921 (55.5%)
	Sep 2010–Apr 2014	71,329	40,428 (56.7%)	40,277 (99.6%)	39,056 (97.0%)	9825 (24.4%)
	May 2014–May 2017	66,092	34,467 (52.2%)	29,410 (85.3%)	17,466 (59.4%)	15,311 (52.1%)
	Jun 2017–Jul 2018	25,389	9013 (35.5%)	8158 (90.5%)	2280 (28.0%)	2189 (26.8%)
Nagongera	Jan 2009–Nov 2013	124,711	82,762 (66.4%)	76,915 (92.9%)	74,959 (97.5%)	27,007 (35.1%)
	Dec 2013–Jan 2015	24,530	15,893 (64.8%)	15,723 (98.9%)	14,874 (94.6%)	3977 (25.3%)
	Feb 2015–Jul 2018	72,236	27,297 (37.8%)	27,064 (99.2%)	16,213 (59.9%)	4309 (15.9%)

[a] Patients seen at the health facility excluding those with a missing record of age

[b] Testing for malaria includes both microscopy and RDT

acyclic secular trends in between. Larger reductions were observed in the two sites where both LLINs and IRS were implemented (Fig. 2). In these sites, the declining trend in TPR is consistent except in Aduku, during the 3 years of withdrawn IRS, characterized by a sharp increase. Between baseline and the last intervention periods, TPR declined in Aduku from 56 to 27% and in Nagongera from 35 to 16%. In the two sites that received LLINs only, a similarly (with acyclic secular but less notably) reducing trend was observed. Between baseline and the last intervention periods, overall TPR decreased in Walukuba from 33 to 29% and in Kasambya from 39 to 30% (Table 1).

Over time and in all four sites, test positivity was seen to decline among the younger children while increasing among older participants. In all sites, a shift in the peak age of test positivity from the youngest to the older ages, was observed between baseline and last intervention period (Figs. S5 and S6, Additional file 1). Interestingly, this pattern was reversed in Aduku during the 3 years when IRS was withdrawn, further confirming the effect of control interventions on test positivity with age (Fig. S6, Additional file 1).

Differences in age distribution of malaria cases between sites at baseline

Although the duration of baseline periods varied between the sites due to the different timing of intervention activities, the age distribution of patients not suspected of malaria was very similar between all four sites at baseline, with median ages ranging from 23 to 25 years (Fig. 3). In contrast, the age distributions of malaria cases varied significantly between sites. These distributions were similar between the highest transmission sites of Nagongera and Aduku, with median ages of 2 and 3 years, respectively. The distributions were also similar, but higher, between the lower transmission sites, with median ages of 8 and 11 years for Kasambya and Walukuba, respectively.

Consistent with unadjusted analyses, results from the final adjusted multinomial regression models (adjusting for diagnostic test and gender of patient; evaluated using Akaike's information criteria and likelihood ratio tests (Table S2, Additional file 1) for model selection; and, using scatter plots with fitted lines for goodness of fit (Fig. S7, Additional file 1) showed that the proportion of malaria cases per age-group were significantly different between sites at baseline. The majority of malaria cases were among children under 5 years of age in the highest transmission sites (Aduku 58% and Nagongera 64%), while the highest proportion of malaria cases was among patients 5–15 years of age in Kasambya (35%), and over 15 years of age in Walukuba (31%) (Fig. 4).

Changes in age distribution of non-suspected, test-negative, and laboratory confirmed malaria cases over time

The age of patients not suspected of malaria decreased slightly over the study duration at all four sites. Moreover, for malaria negative patients, the age distribution slightly shifted downwards and then upwards at all sites except Nagongera, where it shifted downwards across the intervention periods. In contrast, the age distribution of patients with laboratory confirmed malaria shifted upwards over time at all four sites. Comparing the last observation period to baseline, the median age of patients with malaria increased from 8 (IQR: 2.5–19) to 11 (IQR: 5–21) in Kasambya; 11 (IQR: 3.5–24) to 22 (IQR: 8–32) in Walukuba; 2 (IQR: 1.1–10) to 6 (IQR: 2–18) in Nagongera; and 3 (IQR: 1.2–13) to 14 (IQR: 5–22) in Aduku (Fig. 3).

Across all sites, a progressive decline in the proportion of malaria cases from the youngest age group and a progressive increase in the proportion of cases from the oldest age group were observed. Comparing the last intervention period to baseline, the adjusted proportion of malaria cases among children under 5 years of age decreased from 58% (95% CI 57–59%) to 30% (95% CI 28–31%) in Aduku; 31% (95% CI 30–31%) to 16% (95% CI 15–17%) in Walukuba; 64% (95% CI 63–65%) to 47% (95% CI 45–48%) in Nagongera; and 35% (95% CI 34–36%) to 25% (95% CI 23–27%) in Kasambya. Comparing the same periods, the proportion of malaria cases among patients over 15 years of age increased from 19% (95% CI 19–20%) to 44% (95% CI 42–46%) in Aduku; 40% (95% CI 40–41%) to 61% (95% CI 59–64%) in Walukuba; 18% (95% CI 17–18%) to 31% (95% CI 29–32%) in Nagongera; and, 29% (95% CI 28–29%) to 39% (95% CI 37–41%) in Kasambya.

The upward shift in the age distribution of malaria cases occurred gradually throughout the study periods in all sites except Aduku, where IRS was withdrawn in 2014 for three years (defining the 3rd intervention period) before another round was implemented in 2017. In Aduku, during the intervals from the 2nd to the 3rd intervention periods, the proportion of malaria cases among children under 5 years of age increased from 38% (95% CI 37–39%) to 44% (95% CI 43–44%), followed by a decrease to 30% (95% CI 28–31%) following the last round of IRS. At this site, the proportion of malaria cases among patients over 15 years decreased from 35% (95% CI 34–36%) during the 2nd to 30% (95% CI 29–31%) during the 3rd intervention period, before increasing to 44% (95% CI 42–46%) in the last period (Fig. 4).

The upward shift in age distribution of confirmed malaria occurred consistently in both males and females (Fig. S8, Additional file 1). Whilst the majority of all

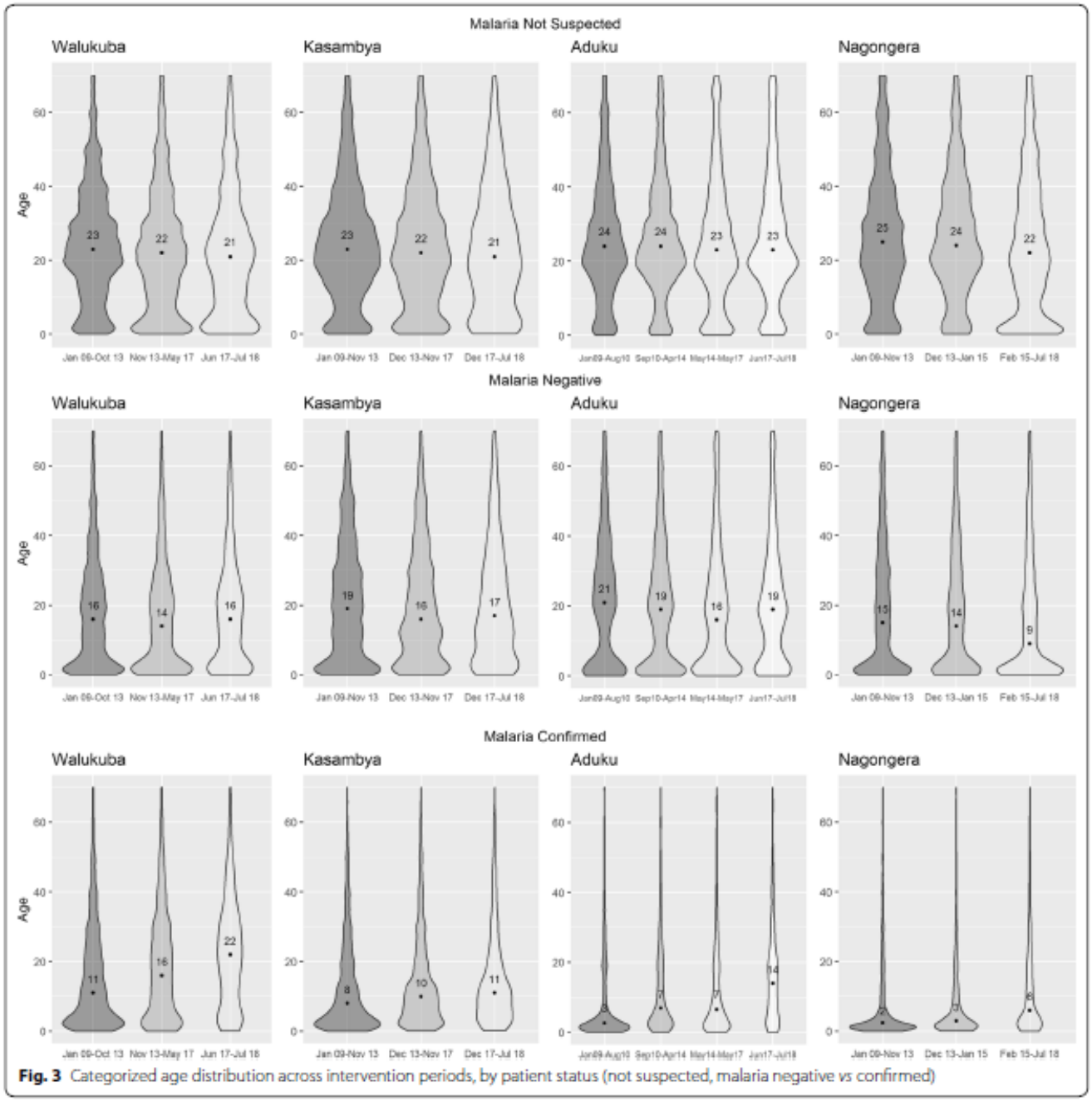


Fig. 3 Categorized age distribution across intervention periods, by patient status (not suspected, malaria negative *vs* confirmed)

patients, non-suspected and suspected malaria cases were female across the study durations, small differences were observed in age distribution between males and females that were not suspected to be malaria cases (Fig. S9, Additional file 1), but the age distribution of females was older than that of males among malaria-negative patients across all sites (Fig. S10, Additional file 1). Moreover, models allowing an interaction between gender and intervention period suggest greater increase in proportion of males than females among confirmed malaria cases over time (Fig. S11 and Table S3, Additional file 1).

Overall, Aduku experienced the largest change in the age distribution of malaria cases throughout the study period. The odds of an upward shift in the age category of confirmed malaria cases in the last relative to the baseline intervention periods were 3.27 (95% CI 2.97–3.61) in Aduku, 2.35 (95% CI 2.14–2.58) in Walukuba, 2.03 (95% CI 1.90–2.17) in Nagongera, and 1.59 (95% CI 1.44–1.76) in Kasambya (Table 2). Whereas the interaction between

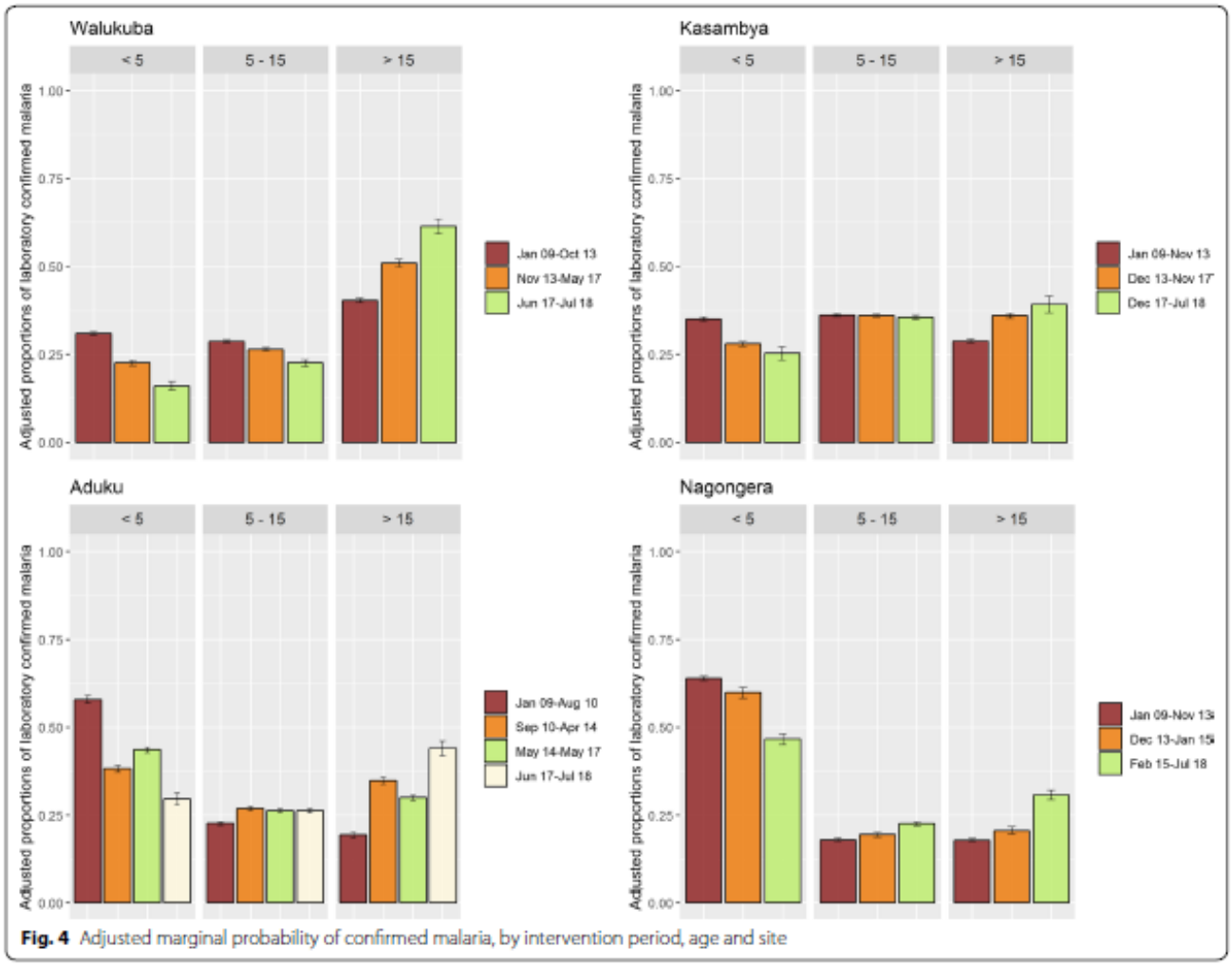


Fig. 4 Adjusted marginal probability of confirmed malaria, by intervention period, age and site

intervention time and diagnostic test used was significant in Kasambya and Nagongera, the same did not notably impact the overall effect observed (Table S4, Additional file 1). Also, RDT use increased gradually at all sites, reaching 20% only in the last three to four years of the study except Aduku (Fig. S4, Additional file 1).

Discussion

The impact of reductions in malaria transmission following the scale-up of malaria control interventions on the age distribution of malaria cases, was investigated using routine surveillance data from four sentinel health facilities in Uganda. The study included data from sites with historically varied transmission intensity where large-scale programmatic control interventions of either LLINs alone or LLINs plus IRS were implemented over the approximately 10-year study period.

Study findings provide empirical evidence of rapid shifts in the malaria burden to older individuals, following implementation of effective malaria control interventions. Over time, the proportion of test confirmed malaria cases progressively decreased in children under 5 years of age while it progressively increased in those over 15 years of age, irrespective of transmission settings. This is also reflected by subtle but greater decline in TPR among children than adults. The absence of similar changes in age patterns among patients not suspected of malaria suggests that the primary drivers of this shift were declines in malaria transmission intensity associated with effective control interventions and not changes in patient demographics. The reverse shift observed in Aduku during a period of malaria resurgence after three years of interrupted IRS [37] provides further support for this conclusion.

Many factors may have contributed to the shift in age distribution of malaria cases in this study. For many endemic infectious diseases, decreases in transmission are expected to result in increased age of infection and cases [38]. For pathogens like *P. falciparum*, where partial immunity develops gradually as a consequence of

Table 2 Association between being malaria confirmed within ages (< 5, 5–15, > 15 years) and control intervention period

Factor	Categories	Odds	95% CI	P value
Walukuba				
Gender	Male	1	Ref	
	Female	1.34	1.29–1.39	< 0.001
Malaria test done	Microscopy	1	Ref	
	RDT	2.11	0.61–7.26	0.237
Intervention period	Baseline	1	Ref	
	First period	1.55	1.48–1.62	< 0.001
	Last period	2.27	2.05–2.51	< 0.001
Kasambya				
Gender	Male	1	Ref	
	Female	1.39	1.34–1.45	< 0.001
Malaria test done	Microscopy	1	Ref	
	RDT	0.98	0.93–1.03	0.368
Intervention period	Baseline	1	Ref	
	First period	1.39	1.34–1.44	< 0.001
	Last period	1.59	1.44–1.76	< 0.001
Aduku				
Gender	Male	1	Ref	
	Female	2.64	2.54–2.75	< 0.001
Malaria test done	Microscopy	1	Ref	
	RDT	1.26	1.19–1.32	< 0.001
Intervention period	Baseline	1	Ref	
	First period	2.23	2.11–2.35	< 0.001
	Second period	1.78	1.68–1.89	< 0.001
	Last period	3.27	2.97–3.61	< 0.001
Nagongera				
Gender	Male	1		
	Female	2.18	2.09–2.28	< 0.001
Malaria test done	Microscopy	1		
	RDT	1.24	1.15–1.34	< 0.001
Intervention period	Baseline	1		
	First period	1.19	1.11–1.27	< 0.001
	Last period	2.03	1.90–2.17	< 0.001

repeated exposure, waning immunity due to decreased infection rates may lead to reduced ability to control parasites [39]. This waning immunity may result in a relative increase in the disease burden among adolescents and adults [12, 40]. Concentration of the malaria burden among older age groups following reduced transmission has been predicted in other studies [8], and seen in children for both severe and non-severe outcomes [10, 41].

Behavioural factors may have also contributed to rapid shifts in the age distribution of malaria. A reduced proportion of cases among children under 5 years of age may have been due to an increased use of LLINs among this age group relative to older age groups [22, 41, 42]. In adults, behavioural factors including travel, leisure and social activities, and occupational activities such as agriculture or night-time work may have increased the risk of exposure outside the household as compared to children. Travel has been reported as a risk factor for *P. falciparum* infection in East and Southern Africa [43, 44] and for cases of imported malaria being older than those not imported in Southern Africa [45]. Whereas occupational hazards were not evaluated in this study, considerable occupational risk of malaria has been documented among mobile male workers in Asia [46, 47] and populations involved in agriculture in Africa [48, 49].

Results suggest that implementation of LLINs plus IRS was associated with larger decreases in transmission and larger shifts in age distribution of cases than LLINs alone. However, among LLIN-only sites, Walukuba recorded a much larger shift than Kasambya and in Walukuba the magnitude of the shift was comparable to that of Nagongera, a site with both LLINs and IRS. This suggests that other non-intervention factors for example urbanization [50] may have contributed to the shifts, consistent with other reported findings [38, 51].

Findings from this study showed that a significantly higher proportion of females, as compared to males, seem to seek care for febrile illnesses among the school aged (5–15 years) and adults (over 15 years old) at all sites. This is consistent with the generally older age distribution of females among both the confirmed and the negative malaria cases, but similar age distribution of males and females among those not suspected of malaria. However, shifts in the age distribution of malaria cases observed after control interventions seem to have disproportionately affected males, suggesting a role of gender-based occupational or behavioural differences. Nevertheless, reported greater involvement of females in agriculture in the region [52] than males, as well as documented associations between occupation and education status of mothers and malaria infection risk of families [53] may explain the prevailing significantly high proportion of older females and hence the need for continued interventions that address this vulnerable group.

This study had limitations. First, without well-defined catchment populations for the study sites, it was not possible to estimate changes in malaria incidence over time or effects of environmental factors within the catchment areas and limiting the study to describing temporal shifts in the age distribution of laboratory confirmed cases of malaria. Additional research is necessary to determine impacts of reduced transmission on incidence of malaria. Second, the increased use of RDTs over time, may have influenced results from this study through varied sensitivity [54]. Nevertheless, in regression analyses,

diagnostic test used was accounted for. Moreover, patterns seen among suspected and confirmed cases were absent in the cases not suspected of malaria, suggesting that the impact of diagnostic method would be minimal. Third, inability to account for levels of uptake of the interventions implemented at the four study sites could have masked some important variations. However, the community-wide benefit of reduced vector populations [55], coupled with the multiple rounds of interventions, was believed to have mitigated this effect and thus its impact on these results. Fourth, whereas fever or history of fever in the past 48 h is the main indicator of suspected malaria, this specific measure was not consistently recorded in the health facilities, and so inconsistency in consideration of the diagnosis may have impacted results from this study. Nevertheless, other proxy indicators of suspected malaria were also considered to ensure a fairly complete capture of these cases. Lastly, having analysed health facility surveillance data, this would only include participants that seek care in the public health facilities. As such community level interventions, especially iCCM that target young children could have had an impact on these findings. However, iCCM was limited to one of the four sites and for less than 2 years.

Conclusions

The findings from this study have important implications regarding targeted malaria control interventions that have historically focused on children under 5 years of age. Whereas these efforts need to continue, new strategies may be necessary to address the shift in burden to older age groups following the implementation of successful malaria control interventions. For instance, extending iCCM for malaria to older age groups after intensive IRS, with the post-IRS duration often associated with malaria resurgence. These findings also have implications for the optimal allocation of health care resources for the diagnosis, treatment and control of malaria amidst changing transmission. Further, they highlight the usefulness of considering age distribution in surveillance and monitoring change, and the role of surveillance in understanding the epidemiology of malaria.

Supplementary Information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12936-020-03196-7.

Additional file 1: Figure S1. Trends in mean monthly overall patient attendance per year, stratified by site. **Figure S2.** Trends in mean monthly attendance of patients not suspected of malaria per year, by site. **Figure S3.** Trends in mean monthly suspected malaria patients per year, stratified by site. **Table S1.** Changes in attendance of patients suspected versus not suspected of malaria over-time, comparing mean monthly attendance between first and last calendar years of study duration. **Figure S4.** Age distribution of test confirmed malaria cases, by sex and site. **Figure S5.** Age distribution of patients not suspected of malaria by sex and site. **Figure S6.** Age distribution of patients that tested negative for malaria, by sex and site. **Figure S7.** Adjusted marginal probability of test confirmed malaria, by sex, intervention period, age, and site. **Table S2.** Multivariable association between age (in three categories) and covariates of interest among malaria confirmed cases, accounting for effect modification of intervention periods on sex. **Table S3.** Association between age (in three categories) and covariates of interest among malaria confirmed cases, fitting an interaction between diagnostic test used (B/S vs. RDT) and intervention duration. **Figure S8.** Trends in the annual proportion of RDT use among tested participants, stratified by site. **Figure S9.** Scatter plot of age with test positivity for LLINs only sites, stratified by intervention period. **Figure S10.** Scatter plot of age with test positivity for (LLIN plus IRS) sites, by intervention period.

Abbreviations

ACT: Artemisinin-based combination therapy; aEIR: Annual entomological inoculation rate; CI: Confidence interval; HMIS: Health Management Information System; iCCM: Integrated Community Case Management; IQR: Inter quartile range; IRS: Indoor residual spraying of insecticide; ITN: Insecticide treated net; LLIN: Long-lasting insecticidal net; MIS: Malaria indicator survey; MRC: Malaria Reference Centre; NMCD: National Malaria Control Division; OPD: Outpatients department; RDT: Rapid diagnostic test; TPR: Test positivity rate; UMSP: Uganda Malaria Surveillance Project; WHO: World Health Organization.

Acknowledgements

We would like to especially thank God in this work. We also would like to thank the administration and staff at the Infectious Diseases Research Collaboration, and HMIS officers and in-charges at each of the clinics for all their hard work and great contributions to the success of this project. We also acknowledge the district health leadership in Jinja, Tororo, Mubende, and Apac districts for all their cooperation with and support toward this study.

Authors' contributions

SPK AND RK conceived of the study with input from RP, KH, JIN, GD and IR-B. AM, AS, AY, and RK led the data collection activities with input from GD, SGS, DR, CS, JO and MK. RK led the data management with input from GD. SPK, IR-B, GD and RK led the data analysis, with support from AE, BG. All authors, with all the data in this study at their disposal, reviewed the manuscript and gave permission for publication. SPK, the corresponding author, takes full responsibility for submission towards publication. All authors read and approved the final manuscript.

Funding statement

Funding for parent study was provided by the President's Malaria Initiative, U.S. Agency for International Development under the terms of Interagency Agreement (1US1CK000117) with the CDC. For the sub-study however, research reported in this publication was supported by the Fogarty International Center of the National Institutes of Health under Award Numbers (D43TW7375) and (D43TW010526). JIN is supported by the Fogarty International Center (Emerging Global Leader Award) grant number (K43TW010365). The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the President's Malaria Initiative, U.S. Agency for International Development or the National Institutes of Health.

Availability of data and materials

The datasets used and/or analyzed for this study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

This study was approved by the Uganda National Council for Science and Technology (UNCST Ref SS 4455), the Makerere University School of Medicine Research and Ethics Committee (REC Ref. 2017-119), and the London School of Hygiene & Tropical Medicine Ethics Committee (LSHTM Ref. 13902).Written informed consent of participants included in this study was not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1] Department of Disease Control, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK. [2] Infectious Diseases Research Collaboration, PO Box 7475, Kampala, Uganda. [3] USAID's Malaria Action Program for Districts, PO Box 8045, Kampala, Uganda. [4] Department of Epidemiology & Biostatistics, University of California, San Francisco, 550 16th Street, San Francisco, CA 94158, USA. [5] Child Health and Development Centre, Makerere University College of Health Sciences, Mulago Hospital Complex, PO Box 7072, Kampala, Uganda. [6] School of Public Health, Makerere University College of Health Sciences, Mulago Hospital Complex, PO Box 7072, Kampala, Uganda. [7] School of Medicine, Makerere University College of Health Sciences, Mulago Hospital Complex, PO Box 7072, Kampala, Uganda. [8] National Malaria Control Division, Uganda Ministry of Health, Kampala, Uganda. [9] Department of Clinical Research, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK. [10] Department of Medicine, University of California, San Francisco, 1001 Potrero Ave, SFGH Building 30, San Francisco, CA 94110, USA. [11] Division of HIV, ID, and Global Medicine, University of California, San Francisco, 1001 Potrero Ave, SFGH, Building 3, San Francisco, CA 94110, USA. [12] Chan Zuckerberg Biohub, San Francisco, CA, USA.

Received: 20 November 2019 Accepted: 20 March 2020
Published online: 30 March 2020

References

1. Bhatt S, Weiss DJ, Cameron E, Bisanzio D, Mappin B, Dalrymple U, et al. The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. Nature. 2015;526:207–11.
2. Murray CJ, Ortblad KF, Guinovart C, Lim SS, Wolock TM, Roberts DA, et al. Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2014;384:1005–70.
3. Noor AM, Kinyoki DK, Mundia CW, Kabaria CW, Mutua JW, Alegana VA, et al. The changing risk of *Plasmodium falciparum* malaria infection in Africa: 2000-10: a spatial and temporal analysis of transmission intensity. Lancet. 2014;383:1739–47.
4. Nkumama IN, O'Meara WP, Osier FHA. Changes in malaria epidemiology in Africa and new challenges for elimination. Trends Parasitol. 2017;33:128–40.
5. WHO. World Malaria Report 2008. Geneva, World Health Organization, 2008.
6. WHO. Malaria surveillance, monitoring & evaluation: a reference manual. Geneva, World Health Organization, 2018:8.
7. Molineaux L, Gramiccia G. The Garki Project: Research on the epidemiology and control of malaria in the Sudan Savanna of West Africa. Geneva: World Health Organization; 1980.
8. Carneiro I, Roca-Feltrer A, Griffin JT, Smith L, Tanner M, Schellenberg JA, et al. Age-patterns of malaria vary with severity, transmission intensity and seasonality in sub-Saharan Africa: a systematic review and pooled analysis. PLoS ONE. 2010;5:e8988.
9. Schellenberg D, Menendez C, Aponte J, Guinovart C, Mshinda H, Tanner M, et al. The changing epidemiology of malaria in Ifakara Town, southern Tanzania. Trop Med Int Health. 2004;9:68–76.
10. Ceesay SJ, Casals-Pascual C, Erskine J, Anya SE, Duah NO, Fulford AJ, et al. Changes in malaria indices between 1999 and 2007 in The Gambia: a retrospective analysis. Lancet. 2008;372:1545–54.
11. WHO. World Malaria Report 2017. Geneva, World Health Organization, 2017.
12. Nguyen HV, van den Eede P, van Overmeir C, Thang ND, le Hung X, D'Alessandro U, et al. Marked age-dependent prevalence of symptomatic and patent infections and complexity of distribution of human *Plasmodium* species in central Vietnam. Am J Trop Med Hyg. 2012;87:989–95.
13. Nankabirwa J, Wandera B, Kiwanuka N, Staedke SG, Kamya MR, Brooker SJ. Asymptomatic *Plasmodium* infection and cognition among primary schoolchildren in a high malaria transmission setting in Uganda. Am J Trop Med Hyg. 2013;88:1102–8.
14. WHO. Global technical strategy for malaria 2016-2030. Geneva, World Health Organization, 2015.
15. Fekadu M, Yenit MK, Lakew AM. The prevalence of asymptomatic malaria parasitemia and associated factors among adults in Dembia district, northwest Ethiopia, 2017. Arch Public Health. 2018;76:74.
16. Oladeinde BH, Omoregie R, Osakue EO, Onaiwu TO. Asymptomatic Malaria among Blood Donors in Benin City Nigeria. Iran J Parasitol. 2014;9:415–22.
17. WHO. World Malaria Report 2005. Geneva, World Health Organization and UNICEF; 2005.
18. WHO. World Malaria Report 2010. Geneva, World Health Organization, 2010.
19. Ministry of Health. The Uganda Malaria Reduction Strategic Plan 2014-2020. Uganda: Kampala; 2014.
20. Uganda Bureau of Statistics (UBOS) and Macro International Inc. Uganda Demographic and Health Survey 2006. Calverton, Maryland, USA: 2007.
21. Ministry of Health. Uganda National Malaria Control Strategyic Plan 2010/11 - 2014/15. Uganda: Kampala; 2011.
22. Wanzira H, Katamba H, Rubahika D. Use of long-lasting insecticide-treated bed nets in a population with universal coverage following a mass distribution campaign in Uganda. Malar J. 2016;15:311.
23. Ministry of Health. Annual Health Sector Performance Report - FY 2010/2011. Uganda: Kampala; 2011.
24. Uganda Bureau of Statistics, NMCP. Uganda malaria indicator survey 2014-15. Kampala, Uganda, and Rockville, Maryland, 2015.
25. Ministry of Health. Annual Health Sector Performance Report - FY 2016/2017. Uganda: Kampala; 2017.
26. Ssempiira J, Kissa J, Nambuusi B, Kyozira C, Rutazaana D, Mukooyo E, et al. The effect of case management and vector-control interventions on space-time patterns of malaria incidence in Uganda. Malar J. 2018;17:162.
27. Ministry of Health. Annual Health Sector Performance Report - FY 2014/2015. Uganda: Kampala; 2015.
28. Uganda Bureau of Statistics, NMCP. Uganda Malaria Indicator Survey. Calverton, Maryland. USA: UBOS and ICF Macro; 2009. p. 2010.
29. Sserwanga A, Harris JC, Kigozi R, Menon M, Bukirwa H, Gasasira A, et al. Improved malaria case management through the implementation of a health facility-based sentinel site surveillance system in Uganda. PLoS ONE. 2011;6:e16316.
30. Kigozi R, Baxi SM, Gasasira A, Sserwanga A, Kakeeto S, Nasr S, et al. Indoor residual spraying of insecticide and malaria morbidity in a high transmission intensity area of Uganda. PLoS ONE. 2012;7:e42857.
31. Okello PE, Van Bortel W, Byaruhanga AM, Correwyn A, Roelants P, Talisuna A, et al. Variation in malaria transmission intensity in seven sites throughout Uganda. Am J Trop Med Hyg. 2006;75:219–25.
32. Sexton AR. Best practices for an insecticide-treated bed net distribution programme in sub-Saharan eastern Africa. Malar J. 2011;10:157.
33. Teklehaimanot A, Sachs JD, Curtis C. Malaria control needs mass distribution of insecticidal bednets. Lancet. 2007;369:2143–6.
34. Katureebe A, Zinszer K, Arinaitwe E, Rek J, Kakande E, Charland K, et al. Measures of malaria burden after long-lasting insecticidal net distribution and indoor residual spraying at three sites in Uganda: a prospective observational study. PLoS Med. 2016;13:e1002167.
35. Bi Y, Hu W, Liu H, Xiao Y, Guo Y, Chen S, Zhao L, Tong S. Can slide positivity rates predict malaria transmission? Malar J. 2012;11:117.
36. Team RC. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2018.
37. Raouf S, Mpimbaza A, Kigozi R, Sserwanga A, Rubahika D, Katamba H, et al. Resurgence of malaria following discontinuation of indoor residual spraying of insecticide in an area of Uganda with previously high-transmission intensity. Clin Infect Dis. 2017;65:453–60.
38. O'Meara WP, Mangeni JN, Steketee R, Greenwood B. Changes in the burden of malaria in sub-Saharan Africa. Lancet Infect Dis. 2010;10:545–55.
39. Farnert A, Yman V, Homann MV, Wandell G, Mhoja L, Johansson M, et al. Epidemiology of malaria in a village in the Rufiji River Delta, Tanzania: declining transmission over 25 years revealed by different parasitological metrics. Malar J. 2014;13:459.
40. Vliegenthart-Jongbloed K, de Mendonca Melo M, van Wolfswinkel ME, Koelewijn R, van Hellemond JJ, van Genderen PJ. Severity of imported

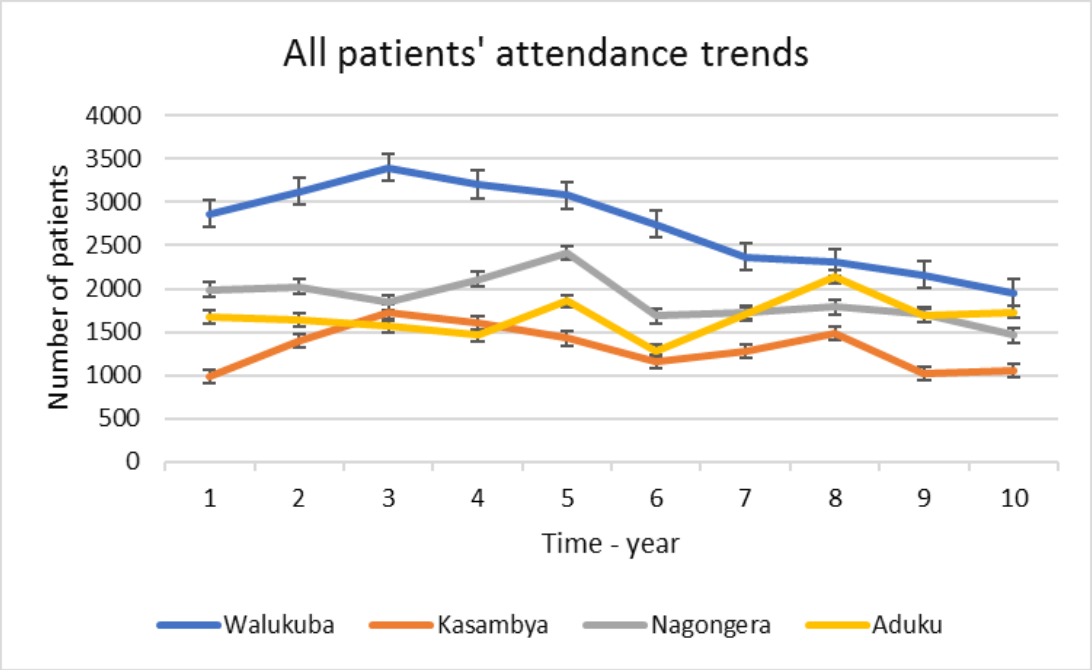
malaria: protective effect of taking malaria chemoprophylaxis. Malar J. 2013;12:265.
41. O'Meara WP, Bejon P, Mwangi TW, Okiro EA, Peshu N, Snow RW, et al. Effect of a fall in malaria transmission on morbidity and mortality in Kilifi, Kenya. Lancet. 2008;372:1555–62.
42. Ferrari G, Ntuku HM, Ross A, Schmidlin S, Kalemwa DM, Tshefu AK, et al. Identifying risk factors for *Plasmodium* infection and anaemia in Kinshasa, Democratic Republic of Congo. Malar J. 2016;15:362.
43. Arinaitwe E, Dorsey G, Nankabirwa JI, Kigozi SP, Katureebe A, Kakande E, et al. Association between recent overnight travel and risk of malaria: a prospective cohort study at three sites in Uganda. Clin Infect Dis. 2018;68:313–20.
44. Mutegeki E, Chimbari MJ, Mukaratirwa S. Assessment of individual and household malaria risk factors among women in a South African village. Acta Trop. 2017;175:71–7.
45. Lowa M, Sitali L, Siame M, Musonda P. Human mobility and factors associated with malaria importation in Lusaka district, Zambia: a descriptive cross sectional study. Malar J. 2018;17:404.
46. Erhart A, Ngo DT, Phan VK, Ta TT, Van Overmeir C, Speybroeck N, et al. Epidemiology of forest malaria in central Vietnam: a large scale cross-sectional survey. Malar J. 2005;4:58.
47. Yang Y, Liu Y, Xie Z, Wu S, Yang L, Li W, et al. Epidemiology of malaria in Yulin, South China 1999-2016: imported malaria threatens zero local case status. Vector Borne Zoonotic Dis. 2018;18:533–8.
48. Janko MM, Irish SR, Reich BJ, Peterson M, Doctor SM, Mwandagalirwa MK, et al. The links between agriculture, *Anopheles* mosquitoes, and malaria risk in children younger than 5 years in the Democratic Republic of the Congo: a population-based, cross-sectional, spatial study. Lancet Planet Health. 2018;2:e74–82.
49. Naidoo S, London L, Burdorf A, Naidoo RN, Kromhout H. Occupational activities associated with a reported history of malaria among women working in small-scale agriculture in South Africa. Am J Trop Med Hyg. 2011;85:805–10.
50. Kigozi SP, Pindolia DK, Smith DL, Arinaitwe E, Katureebe A, Kilama M, et al. Associations between urbanicity and malaria at local scales in Uganda. Malar J. 2015;14:374.
51. Marsh K. Africa in transition: the case of malaria. Int Health. 2016;8:155–6.
52. Woldu DO, Haile ZT. Gender roles and perceptions of malaria risk in agricultural communities of Mwea Division in Central Kenya. Women Health. 2015;55:227–43.
53. Essendi WM, Vardo-Zalik AM, Lo E, Machani MG, Zhou G, Githeko AK, et al. Epidemiological risk factors for clinical malaria infection in the highlands of Western Kenya. Malar J. 2019;18:211.
54. Laban NM, Kobayashi T, Hamapumbu H, Sullivan D, Mharakurwa S, Thuma PE, et al. Comparison of a PfHRP2-based rapid diagnostic test and PCR for malaria in a low prevalence setting in rural southern Zambia: implications for elimination. Malar J. 2015;14:25.
55. Hawley WA, Phillips-Howard PA, ter Kuile FO, Terlouw DJ, Vulule JM, Ombok M, et al. Community-wide effects of permethrin-treated bed nets on child mortality and malaria morbidity in western Kenya. Am J Trop Med Hyg. 2003;68:121–7.

Publisher's Note

4.1 Additional information for Paper 2

4.1.1 Consideration of trends in attendance

Figure S1. Trends in mean monthly overall patient attendance per year, stratified by site.



The years on the x-axes in Figures S1 to S4 are represented as 1 to 10 corresponding to the years 2009 to 2018 while the number of patients and/or cases on the y-axis represent monthly average number per year in the study duration.

Figure S2. Trends in mean monthly attendance of patients not suspected of malaria per year, by site

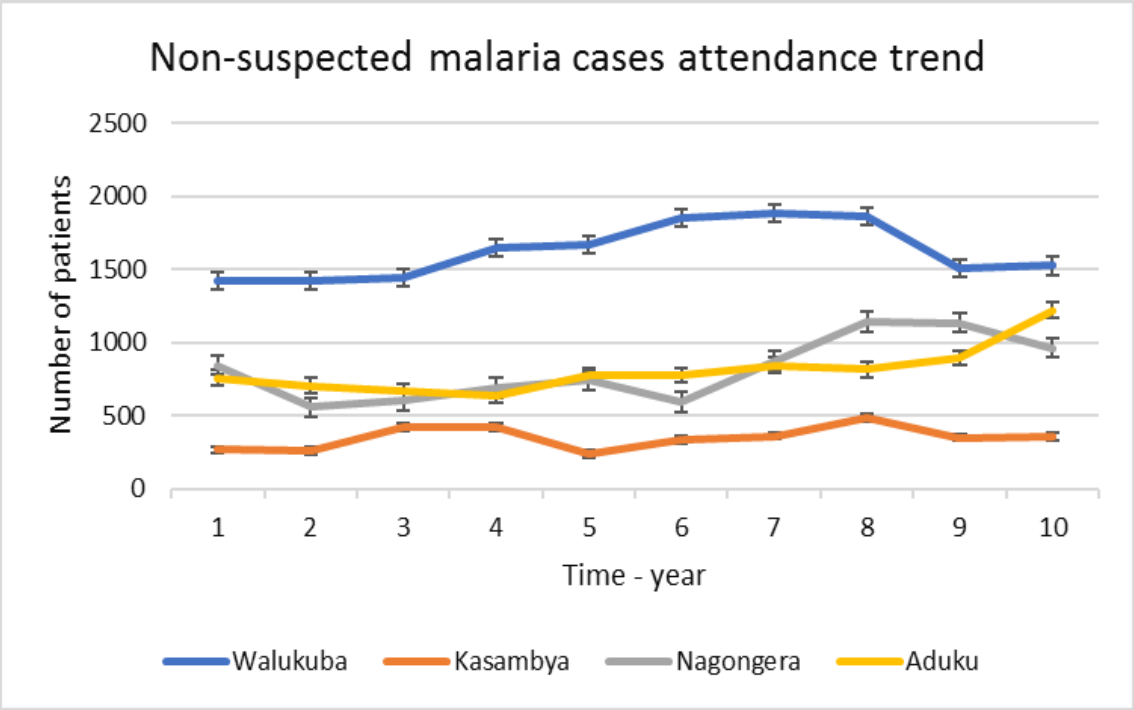


Figure S3. Trends in mean monthly suspected malaria patients per year, stratified by site

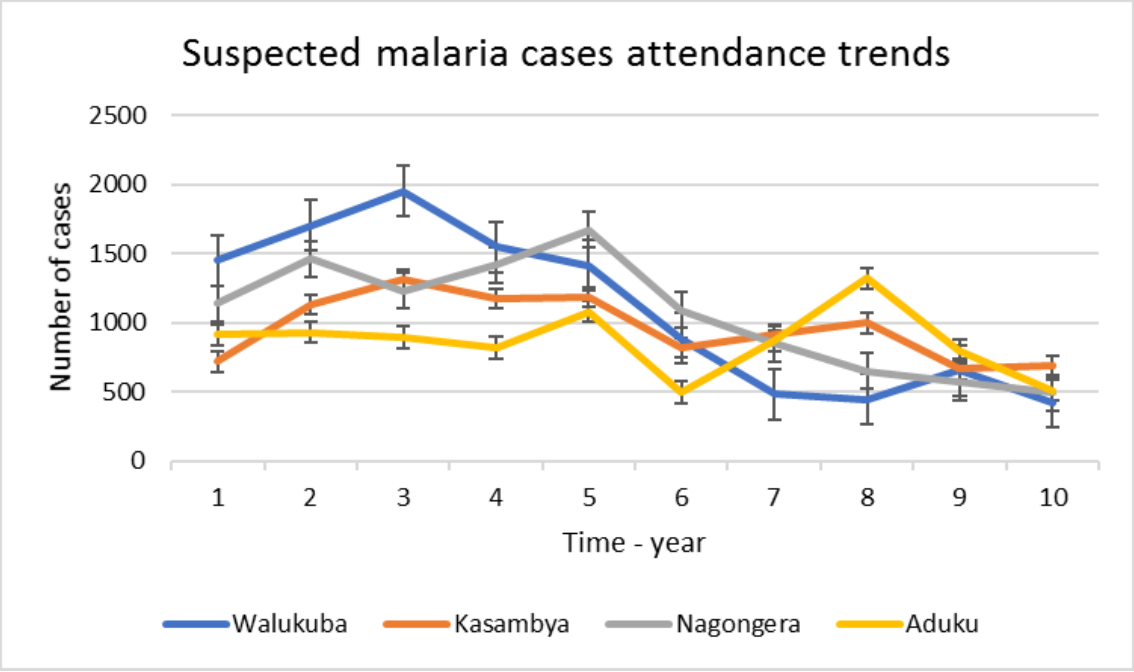
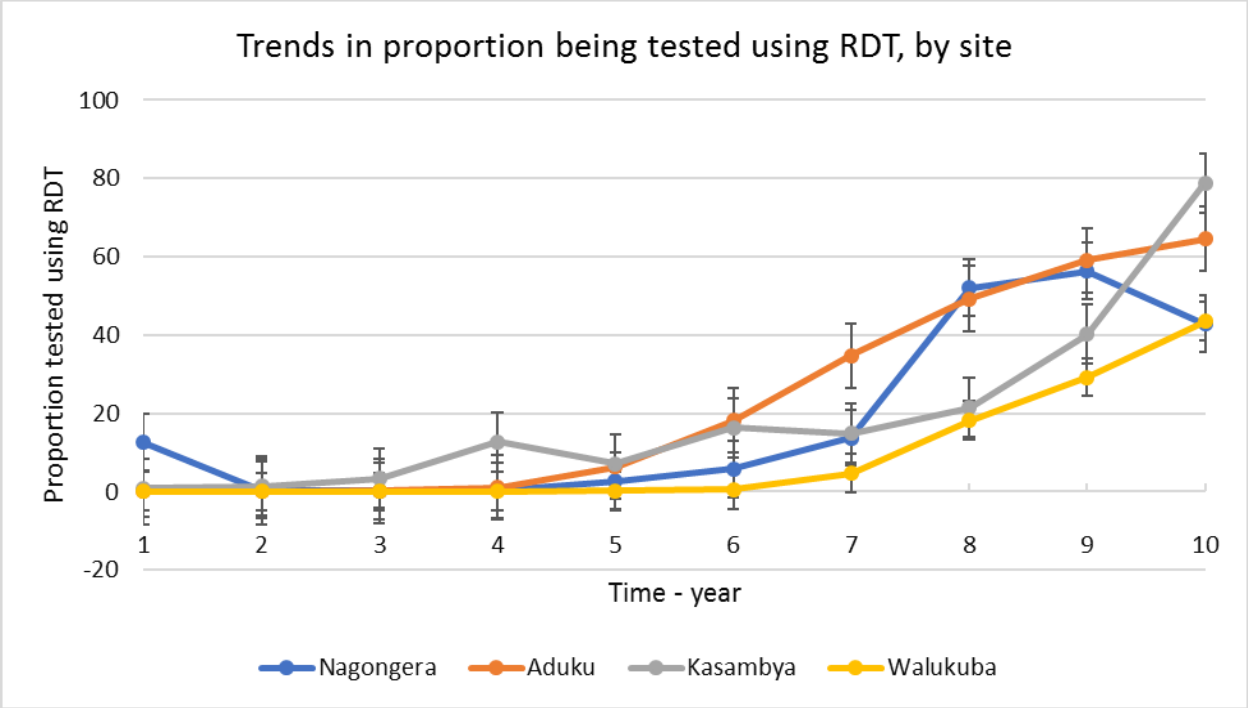


Table S1. Changes in attendance of patients suspected versus not suspected of malaria over-time, comparing mean monthly attendance between first and last calendar years of study duration.

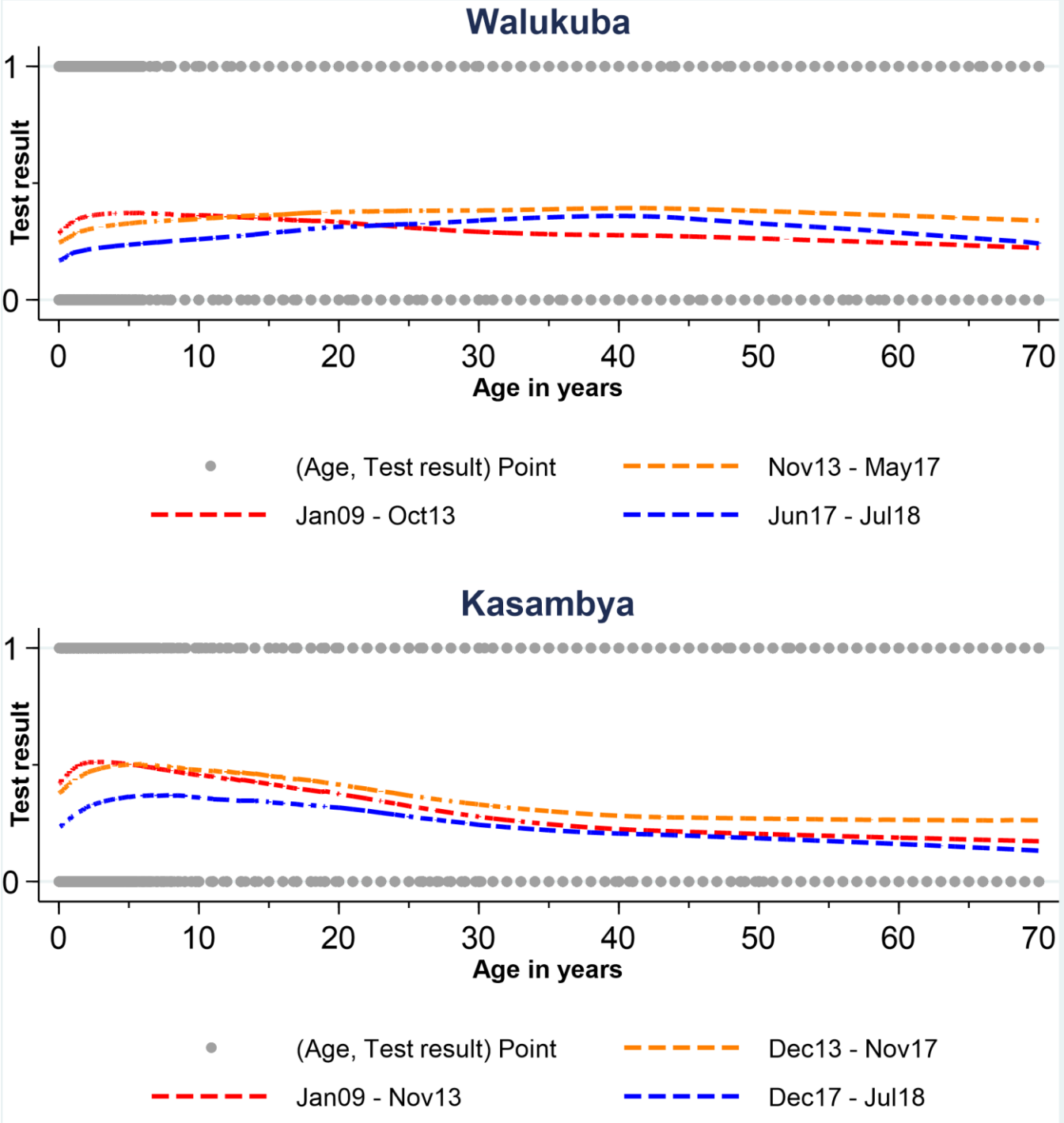
Site	Patient category	Mean monthly attendance per year (SD)		Wilcoxon rank-sum test
		2009	2018	P value
Walukuba	Not suspected of malaria	1418 (178)	1525 (257)	0.353
	Suspected malaria	1452 (372)	427 (84)	<0.001
Kasambya	Not suspected of malaria	268 (73)	360 (73)	0.023
	Suspected malaria	722 (186)	692 (280)	0.866
Aduku	Not suspected of malaria	761 (175)	1222 (156)	<0.001
	Suspected malaria	915 (286)	512 (99)	0.003
Nagongera	Not suspected of malaria	846 (140)	965 (134)	0.108
	Suspected malaria	1139 (157)	496 (183)	<0.001

Figure S4. Trends in the annual proportion of RDT use among tested participants, stratified by site.



There was little to no RDT use in the first five years of this study duration and most sites did not get to 20% use of RDTs till after 2015 (year number 7 in Figure S4). The predominant diagnostic test used in this study therefore, was microscopy.

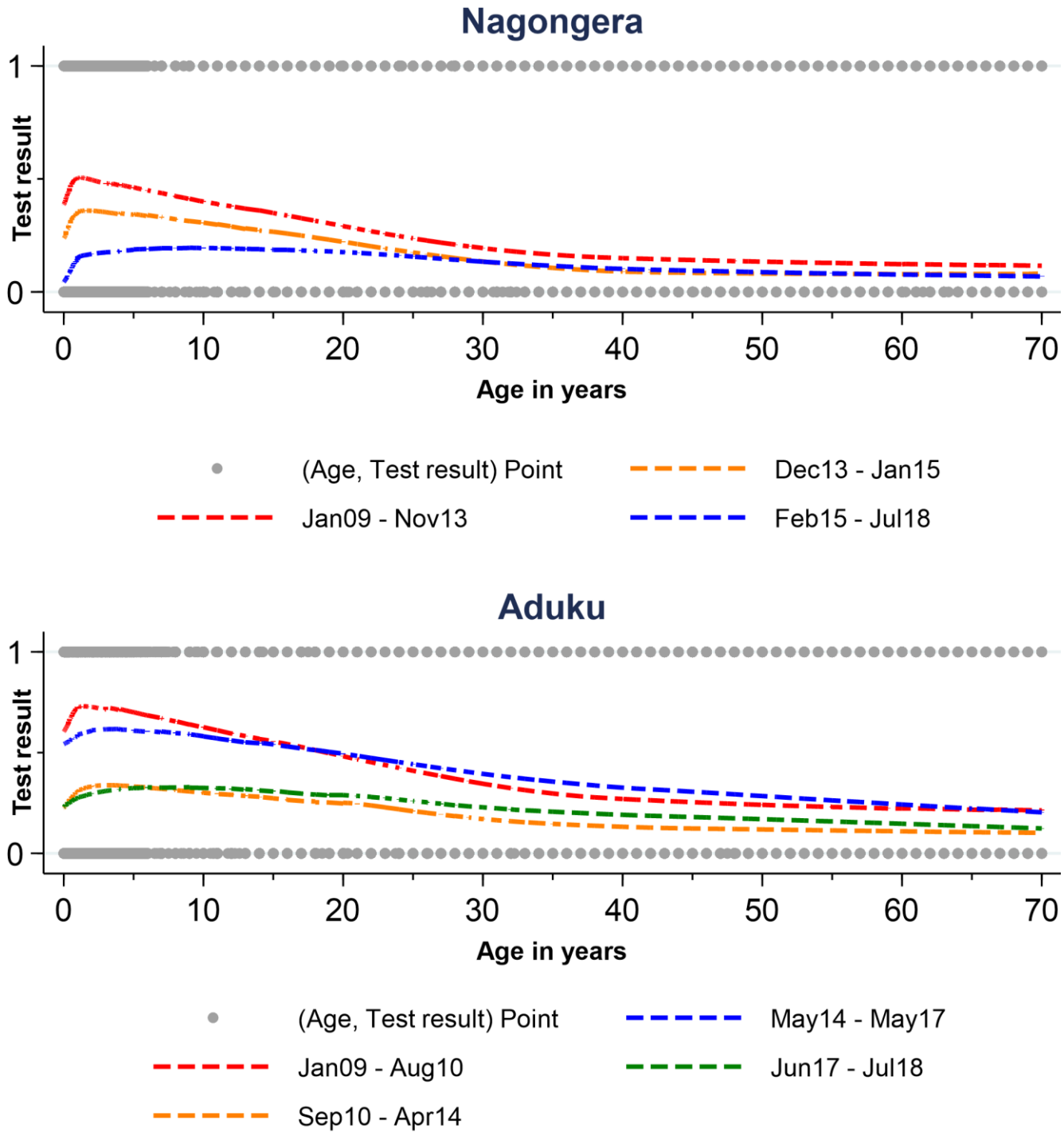
Figure S5. Scatter plot of age with test positivity for LLINs only sites, stratified by intervention period.



In Figure S5, the x-axis represents the age of the participants (70 years and younger) and the y-axis, the test result from malaria diagnostic tests performed. From these tests, 0 corresponds to a negative result while 1 represents a positive result. The grey points are the (age, test result) coordinates of the scatter plot and the dashed curves the relationship fitted using the Lowess smoother function. The red dashed curve represents the relationship of the baseline period, the orange dashed curve – the first intervention period, and the blue dashed curve – the last intervention period of the study duration. By the last intervention period, positivity among the youngest participants was lower than during baseline and the largest shift was observed in Walukuba where in the last

intervention period the peak age of malaria positivity was over 40 years compared to among under 5 years at baseline.

Figure S6. Scatter plot of age with test positivity for (LLIN plus IRS) sites, by intervention period.



In this case (Figure S6), the x-axis represents the age of participants and the y-axis, the test result from malaria diagnostic test performed. From these tests, 0 on the x-axis corresponds to a negative result while 1 represents a positive result. The gray points are the (age, test result) coordinates of the scatter plot and the dashed curves the relationship fitted using the Lowess smoother function. The red dashed curve represents the relationship for the baseline period, the orange dashed curve – the first intervention period, and the blue dashed curve – the last intervention period in Nagongera, but the second intervention period in Aduku. For Aduku, the green dashed curve represents the last intervention period of the study duration.

For all sites in Figure S6, larger decreases in test positivity among the younger children were observed compared to the sites in Figure S5 above. When IRS was withdrawn in Aduku, however, a pattern similar to that during baseline was observed (represented by the blue dashed curve). During the last intervention period, once IRS was resumed and partly supplemented by integrated community case management iCCM for malaria, the pattern (represented by the green dashed curve) was comparable to the first intervention period when intense IRS was implemented (represented by the orange dashed curve).

4.1.2 Model evaluation for the adjusted multinomial regression

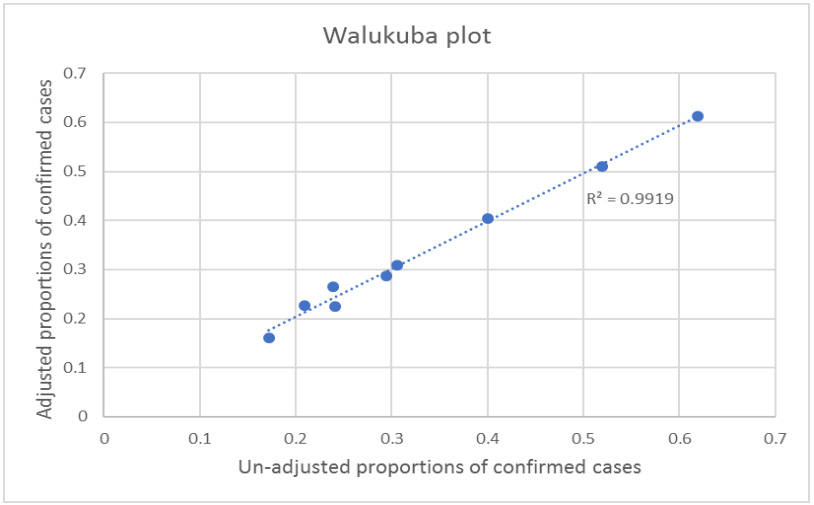
Table S2. Model selection for the final model based on performance with inclusion of the main exposure metric of the intervention over time.

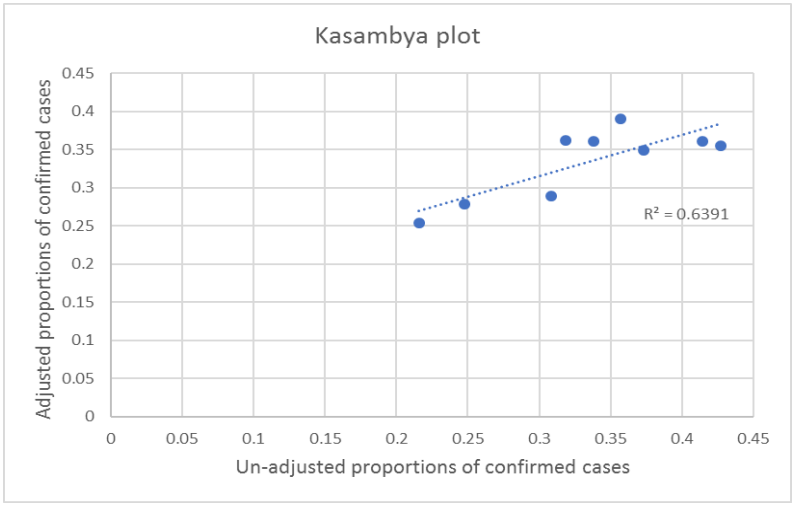
Site	Model	AIC	Model 1 is nested in the final model
Walukuba	Model 1	89998.26	Chi-sq. = 582.45; P<0.001
	Model 2 (final)	89419.82	
Kasambya	Model 1	91870.95	Chi-sq. = 323.16; P<0.001
	Model 2 (final)	91551.79	
Aduku	Model 1	76239.23	Chi-sq. = 1051.01; P<0.001
	Model 2 (final)	75194.22	
Nagongera	Model 1	64666.48	Chi-sq. = 417.42; P<0.001
	Model 2 (final)	64253.06	

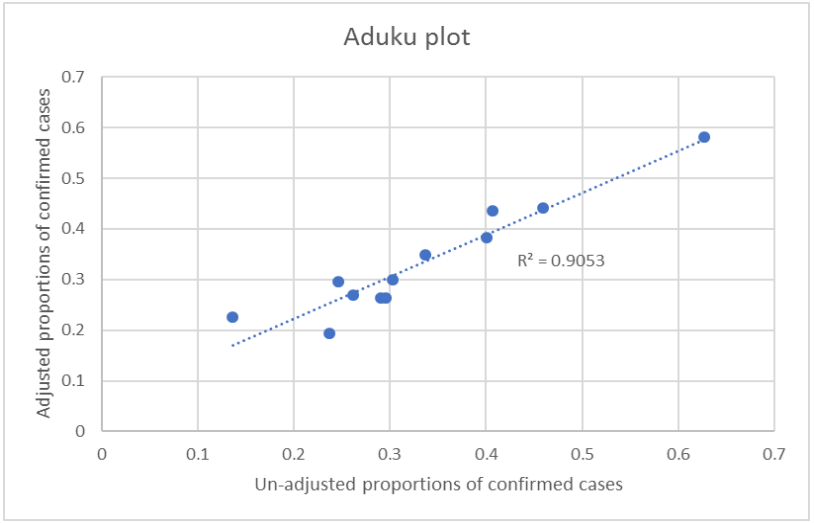
Model 1 = The model adjusted for gender (male vs. female) and diagnostic test used (microscopy vs. RDT) only

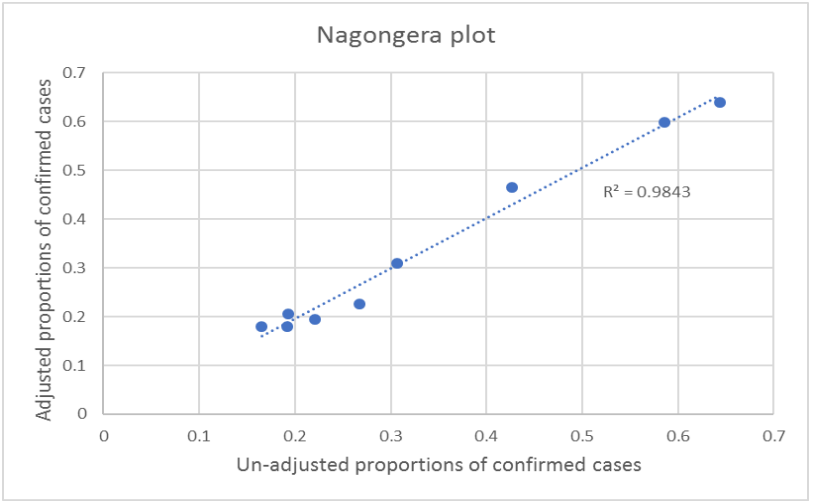
Model 2 = Final model that was adjusted for gender and diagnostic test used, as well as intervention period. This model was found to improve model 1 and therefore the final one based on both AIC and likelihood ratio test evaluations.

Figure S7. Evaluation of model goodness of fit by examining relationship between model predicted proportions of confirmed malaria cases by age category (<5, 5-15, & >15years) adjusted for gender and diagnostic test used, and crude proportion of confirmed malaria cases across intervention periods, by site.









At all sites, multinomial models are seen to fit the data very well, best in Walukuba, Nagongera and Aduku and a little less so in Kasambya.

4.1.3 Consideration of age distribution by gender of patients.

Figure S8. Age distribution of test confirmed malaria cases, by gender and site across intervention periods.

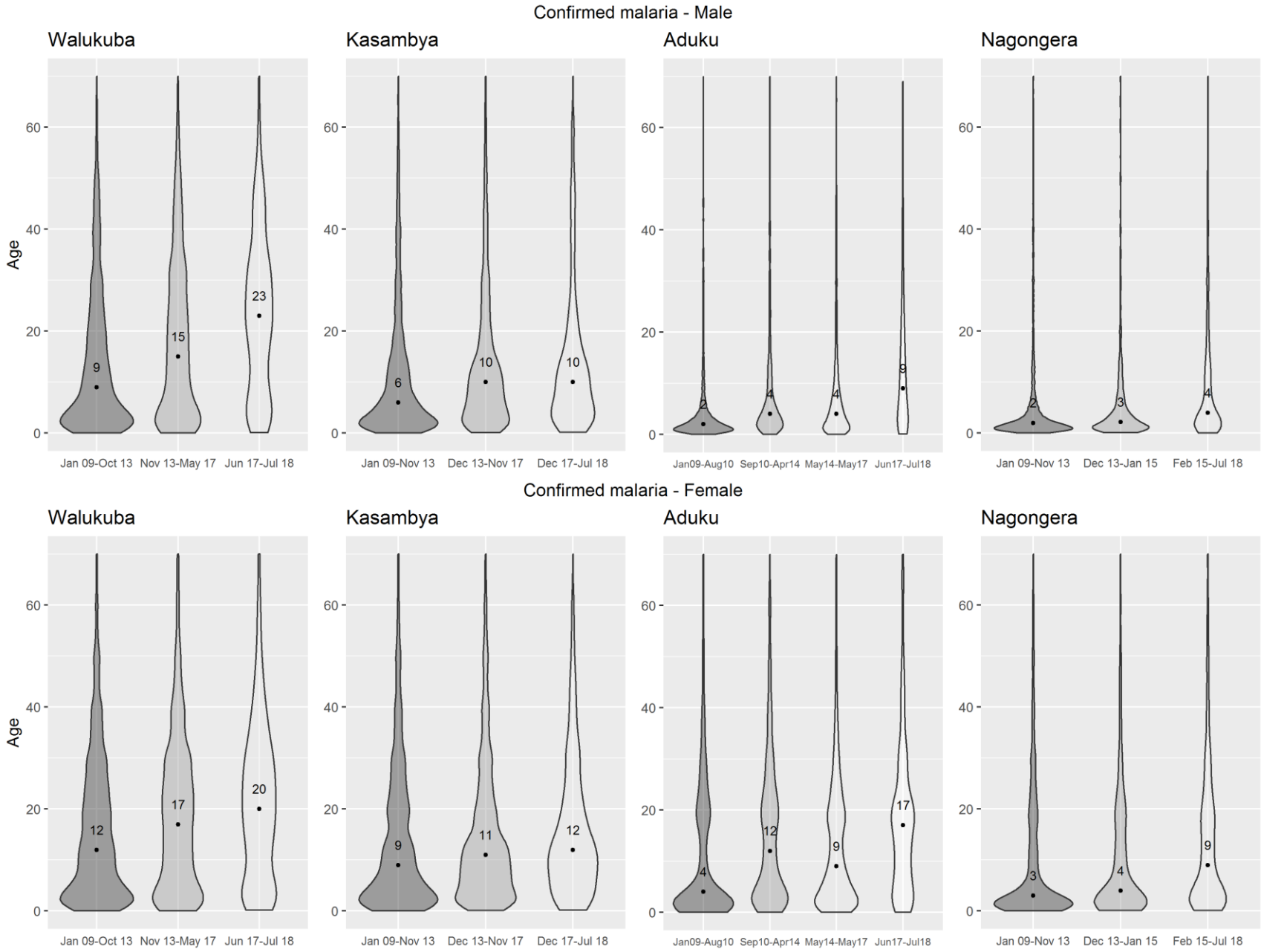


Figure S9. Age distribution of patients not suspected of malaria, by gender and site across intervention periods.

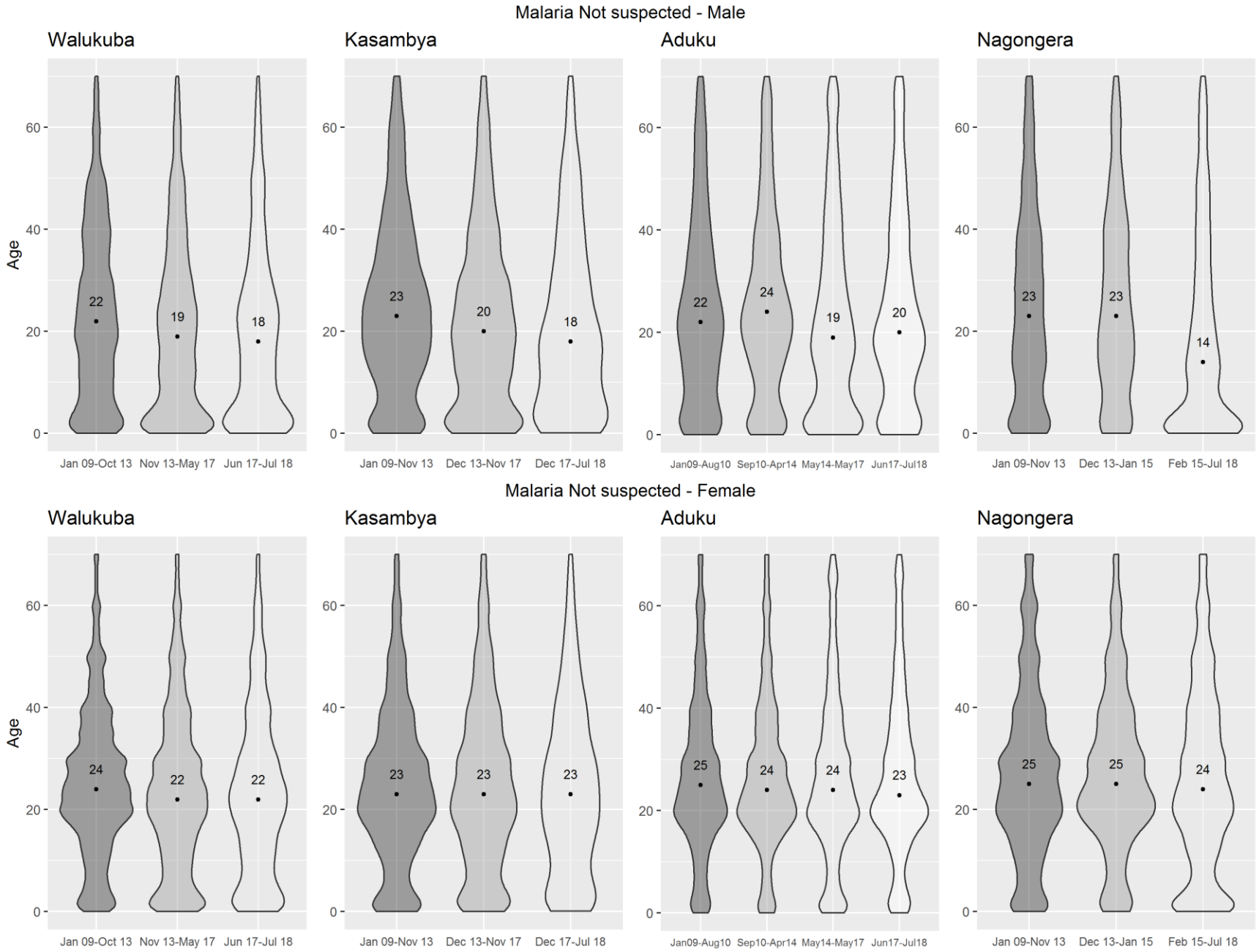


Figure S10. Age distribution of patients that tested negative for malaria, by gender and site across intervention periods.

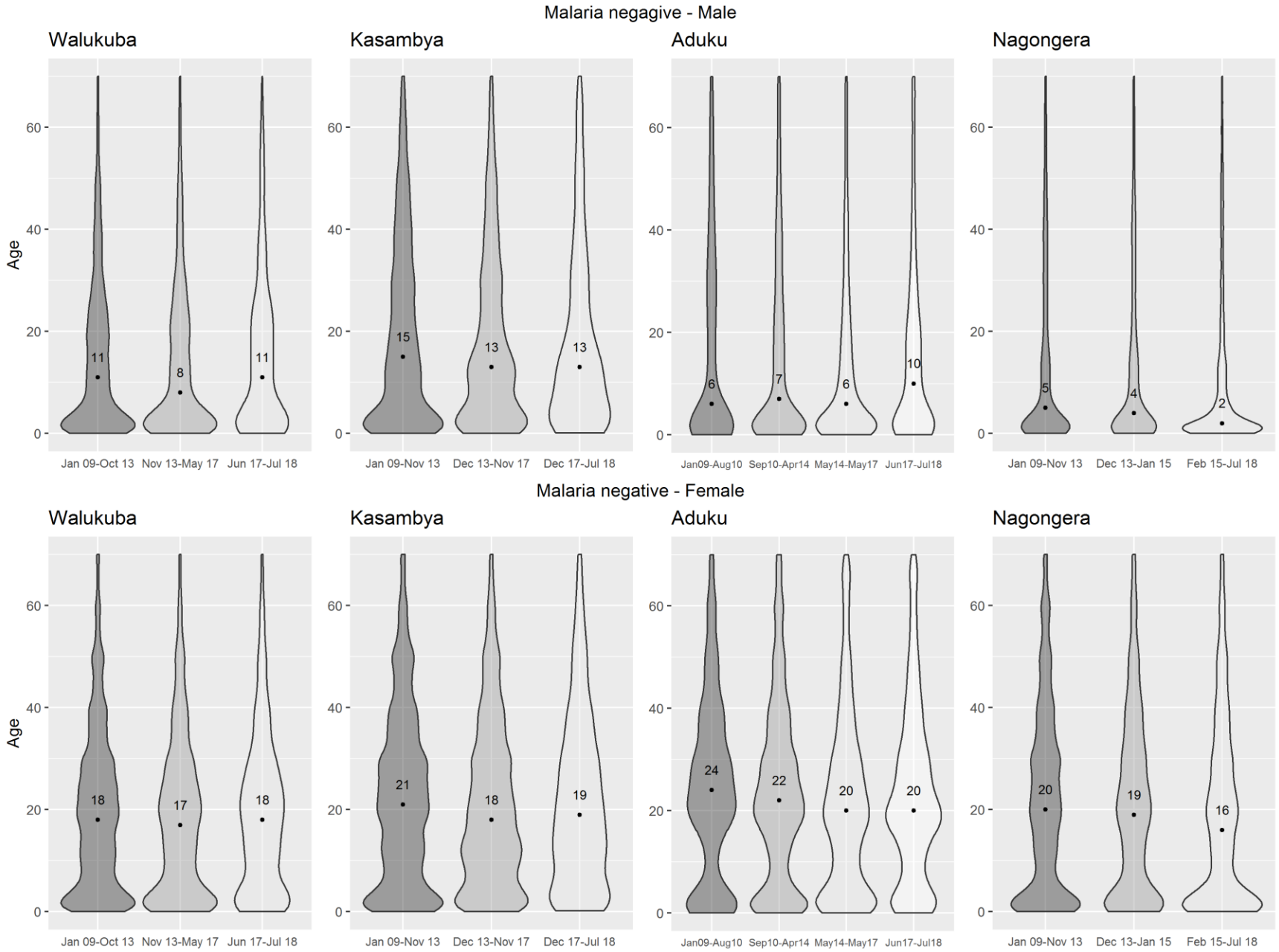
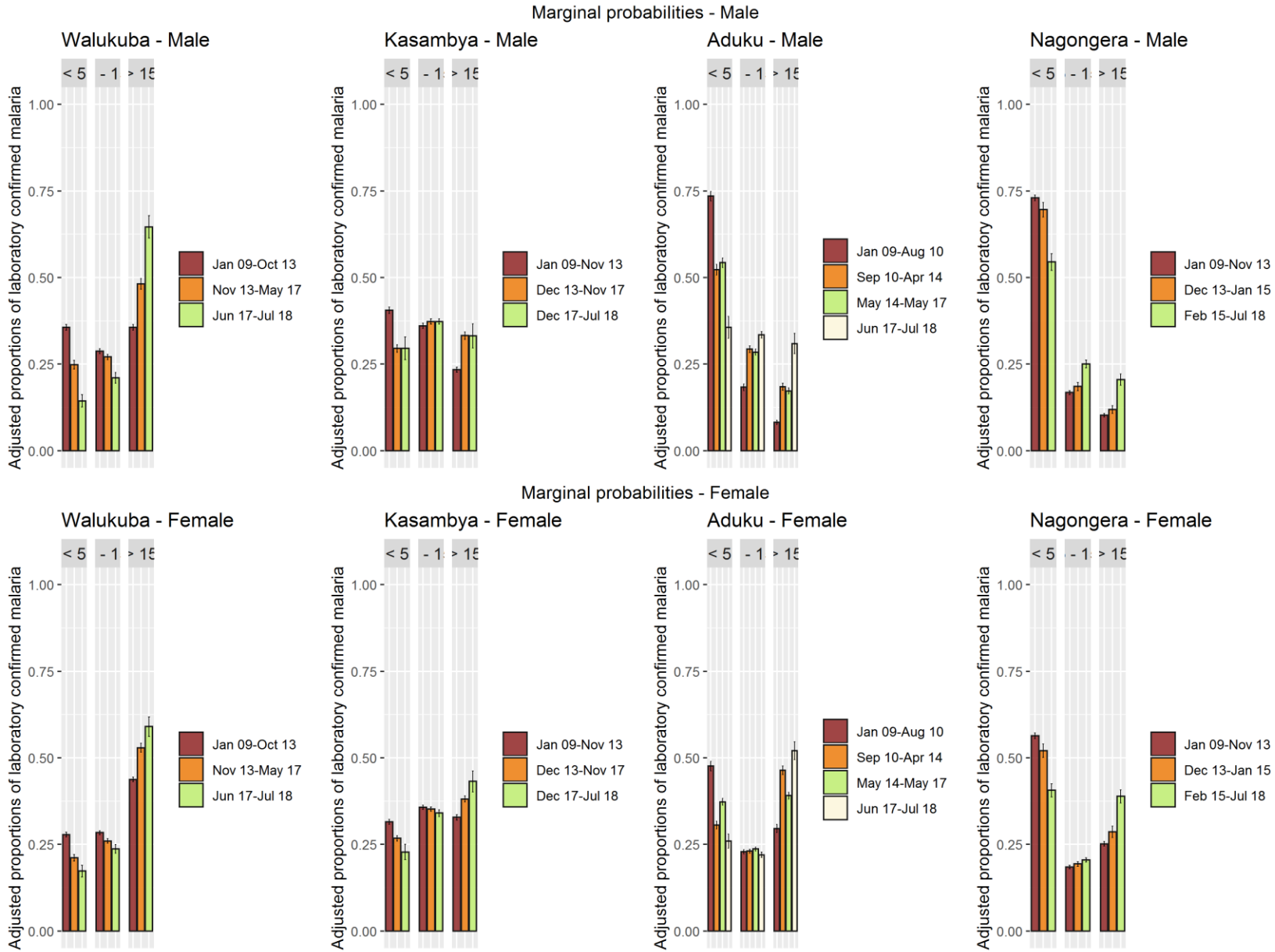


Figure S11. Adjusted marginal probability of test confirmed malaria, by gender, intervention period, age, and site.



The three age categories include: under 5 years, 5-15 years, and over 15 years in each site while intervention periods are arranged by dates.

Table S3. Multivariable association between covariates of interest and age category of confirmed malaria cases (<5, 5 – 15, and >15 years) , accounting for effect modification of intervention periods on gender.

	Covariate category	Multi-variable OR	95% CI	P - value
Walukuba				
Diagnostic test done	B/S	1	*Ref*	
	RDT	1.01	0.91 - 1.13	0.844
Gender	Male	1	*Ref*	
	Female	1.41	1.36 - 1.47	<0.001
Intervention period	Jan 2009 - Oct 2013	1	*Ref*	
	Nov 2013 - May 2017	1.66	1.55 - 1.78	<0.001
	Jun 2017 - Jul 2018	3.27	2.83 - 3.77	<0.001
Effect of gender by intervention period	(Jan 2009 - Oct 2013) x Female	1	*Ref*	
	(Nov 2013 - May 2017) x Female	0.87	0.79 - 0.95	0.002
	(Jun 2017 - Jul 2018) x Female	0.57	0.47 - 0.68	<0.001
Kasambya				
Diagnostic test done	B/S	1	*Ref*	
	RDT	0.98	0.93 - 1.04	0.54
Gender	Male	1	*Ref*	
	Female	1.53	1.46 - 1.61	<0.001
Intervention period	Jan 2009 - Nov 2013	1	*Ref*	
	Dec 2013 - Nov 2017	1.61	1.51 - 1.71	<0.001
	Dec 2017 - Jul 2018	1.61	1.38 - 1.88	<0.001
Effect of gender by intervention period	(Jan 2009 - Nov 2013) x Female	1	*Ref*	
	(Dec 2013 - Nov 2017) x Female	0.78	0.73 - 0.84	<0.001
	(Dec 2017 - Jul 2018) x Female	0.97	0.80 - 1.17	0.75
Aduku				
Diagnostic test done	B/S	1	*Ref*	
	RDT	1.25	1.19 - 1.32	<0.001
Gender	Male	1	*Ref*	
	Female	3.26	2.99 - 3.56	<0.001
Intervention period	Jan 2009 - Aug 2010	1	*Ref*	
	Sep 2010 - Apr 2014	2.4	2.19 - 2.63	<0.001

	May 2014 - May 2017	2.19	2.00 - 2.39	<0.001
	Jun 2017 - Jul 2018	4.38	3.77 - 5.09	<0.001
Effect of gender by intervention period	(Jan 2009 - Aug 2010) x Female	1	*Ref*	
	(Sep 2010 - Apr 2014) x Female	0.89	0.79 - 0.99	0.038
	(May 2014 - May 2017) x Female	0.72	0.65 - 0.80	<0.001
	(Jun 2017 - Jul 2018) x Female	0.63	0.53 - 0.76	<0.001
Nagongera				
Diagnostic test done	B/S	1	*Ref*	
	RDT	1.25	1.16 - 1.34	<0.001
Gender	Male	1	*Ref*	
	Female	2.29	2.18 - 2.41	<0.001
Intervention period	Jan 2009 - Nov 2013	1	*Ref*	
	Dec 2013 - Jan 2015	1.19	1.07 - 1.33	0.001
	Feb 2015 - Jul 2018	2.35	2.13 - 2.60	<0.001
Effect of gender by intervention period	(Jan 2009 - Nov 2013) x Female	1	*Ref*	
	(Dec 2013 - Jan 2015) x Female	1	0.87 - 1.15	0.995
	(Feb 2015 - Jul 2018) x Female	0.79	0.70 - 0.89	<0.001

Results in Table S3 showed that after accounting intervention period and for the effect of gender, the same being modified by intervention periods at all sites, diagnostic test used was only significantly associated with age category of confirmed malaria cases (the outcome in the regression model) in Aduku and Nagongera but not in Walukuba or Kasambya. Importantly however, the effect of gender across all sites is seen to significantly increase in males, given its reduction in females by intervention periods relative to the baseline.

4.1.4 Consideration of possible effect of changes in diagnostic testing methods

Table S4. Association between age (in three categories) and covariates of interest among malaria confirmed cases, fitting an interaction between diagnostic test used (B/S vs. RDT) and intervention duration.

Factor	Categories	Coefficient	95% CI	P - Value
Walukuba				
Gender	Male	1	*Ref*	
	Female	1.34	1.29 - 1.39	<0.001
Malaria test done	Microscopy	1	*Ref*	
	RDT	2.11	0.61 - 7.26	0.237
Intervention period	Jan 2009 - Oct 2013	1	*Ref*	
	Nov 2013 - May 2017	1.55	1.48 - 1.62	<0.001
	Jun 2017 - Jul 2018	2.27	2.05 - 2.51	<0.001
Interaction term	(Jan 2009 - Oct 2013) x RDT	1	*Ref*	
	(Nov 2013 - May 2017) x RDT	0.45	0.13 - 1.56	0.208
	(Jun 2017 - Jul 2018) x RDT	0.56	0.16 - 1.98	0.372
Kasambya				
Gender	Male	1	*Ref*	
	Female	1.4	1.35 - 1.45	<0.001
Malaria test done	Microscopy	1	Ref	
	RDT	1.08	0.99 - 1.18	0.101
Intervention period	Jan 2009 - Nov 2013	1	*Ref*	
	Dec 2013 - Nov 2017	1.41	1.36 - 1.47	<0.001
	Dec 2017 - Jul 2018	1.79	1.46 - 2.20	<0.001
Interaction term	(Jan 2009 - Nov 2013) x RDT	1	*Ref*	
	(Dec 2013 - Nov 2017) x RDT	0.87	0.78 - 0.97	0.013
	(Dec 2017 - Jul 2018) x RDT	0.79	0.62 - 1.00	0.055
Aduku				
Gender	Male	1	Ref	
	Female	2.64	2.54 - 2.75	<0.001
Malaria test done	Microscopy	1	Ref	
	RDT	1.05	0.87 - 1.28	0.603
Intervention period	Jan 2009 - Aug 2010	1	Ref	

	Sep 2010 - Apr 2014	2.2	2.08 - 2.33	<0.001
	May 2014 - May 2017	1.79	1.69 - 1.90	<0.001
	Jun 2017 - Jul 2018	3.77	3.15 - 4.52	<0.001
Interaction term	(Jan 2009 - Aug 2010) x RDT	1	Ref	
	(Sep 2010 - Apr 2014) x RDT	1.46	1.13 - 1.87	0.003
	(May 2014 - May 2017) x RDT	1.18	0.96 - 1.44	0.121
	(Jun 2017 - Jul 2018) x RDT	omitted	N/A	
Nagongera				
Gender	Male	1	Ref	
	Female	2.18	2.09 - 2.28	<0.001
Malaria test done	Microscopy	1	Ref	
	RDT	1.03	0.91 - 1.16	0.649
Intervention period	Jan 2009 - Nov 2013	1	Ref	
	Dec 2013 - Jan 2015	1.15	1.08 - 1.24	<0.001
	Feb 2015 - Jul 2018	1.92	1.77 - 2.08	<0.001
Interaction term	(Jan 2009 - Nov 2013) x RDT	1	Ref	
	(Dec 2013 - Jan 2015) x RDT	1.48	1.17 - 1.86	0.001
	(Feb 2015 - Jul 2018) x RDT	1.35	1.14 - 1.59	<0.001

Whereas diagnostic testing increasingly (in the last three years of the study) included RDT use, with the highest increase observed in Kasambya, Aduku, and Nagongera and least in Walukuba (Figure S4), the potential impact of this change in diagnostic testing method did not generally change the effect identified as due to control intervention activities. The significant interaction in Kasambya and Nagongera provides some evidence of an effect of change in diagnostic testing approach, however, after accounting for this effect, the impact of control interventions on age distribution of confirmed malaria cases persists and remains strongly statistically significant (Table S2). This, therefore, provides further evidence that given other factors at play, the upward shift from younger to older age-groups of malaria cases following successful malaria control interventions is significantly attributable to impacts of control interventions on malaria transmission.

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1513171	**Title**	Mr.
First Name(s)	SIMON PETER		
Surname/Family Name	KIGOZI		
Thesis Title	Exploring the utility of indicators of uncomplicated malaria burden from routine health facility surveillance data in identifying and mapping high-risk areas for malaria in Uganda		
Primary Supervisor	Assoc. Prof. Rachel L. Pullan		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

<u>SECTION B – Paper already published</u>

Where was the work published?	The American Journal of Tropical Medicine and Hygiene		
When was the work published?	APRIL 2020		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**YesYes**	Was the work subject to academic peer review?	**YesYes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I prepared the data, led data analysis, drafted the manuscript. I then led the manuscript submission for peer review, addressed reviewer comments, and managed the final submission for publication

SECTION E

Student Signature	simon peter kigozi
Date	01-September-2020

Supervisor Signature	Rachel Pullan
Date	3rd September 2020

5 Paper 3

Am. J. Trop. Med. Hyg., 00(0), 2020, pp. 1–11
doi:10.4269/ajtmh.19-0950

Practical Implications of a Relationship between Health Management Information System and Community Cohort–Based Malaria Incidence Rates

Simon P. Kigozi,[1,2] Emanuele Giorgi,[3] Arthur Mpimbaza,[4] Ruth N. Kigozi,[5]* Teun Bousema,[6] Emmanuel Arinaitwe,[2] Joaniter I. Nankabirwa,[2,7] Catherine M. Sebuguzi,[2,8] Moses R. Kamya,[2,7] Sarah G. Staedke,[2,9] Grant Dorsey,[2,10] and Rachel L. Pullan[1]

[1]*Department of Disease Control, London School of Hygiene & Tropical Medicine, London, United Kingdom;* [2]*Infectious Diseases Research Collaboration, Kampala, Uganda;* [3]*CHICAS, Lancaster Medical School, Lancaster University, Lancaster, United Kingdom;* [4]*Child Health and Development Centre, Makerere University College of Health Sciences, Kampala, Uganda;* [5]*USAID's Malaria Action Program for Districts, Kampala, Uganda;* [6]*Department of Medical Microbiology, Radboud University, Nijmegen, Netherlands;* [7]*School of Medicine, Makerere University College of Health Sciences, Kampala, Uganda;* [8]*National Malaria Control Division, Uganda Ministry of Health, Kampala, Uganda;* [9]*Department of Clinical Research, London School of Hygiene & Tropical Medicine, London, United Kingdom;* [10]*Department of Medicine, University of California, San Francisco, San Francisco, California*

Abstract. Global malaria burden is reducing with effective control interventions, and surveillance is vital to maintain progress. Health management information system (HMIS) data provide a powerful surveillance tool; however, its estimates of burden need to be better understood for effectiveness. We aimed to investigate the relationship between HMIS and cohort incidence rates and identify sources of bias in HMIS-based incidence. Malaria incidence was estimated using HMIS data from 15 health facilities in three subcounties in Uganda. This was compared with a gold standard of representative cohort studies conducted in children aged 0.5 to < 11 years, followed concurrently in these sites. Between October 2011 and September 2014, 153,079 children were captured through HMISs and 995 followed up through enhanced community cohorts in Walukuba, Kihihi, and Nagongera subcounties. Although HMISs substantially underestimated malaria incidence in all sites compared with data from the cohort studies, there was a strong linear relationship between these rates in the lower transmission settings (Walukuba and Kihihi), but not the lowest HMIS performance highest transmission site (Nagongera), with calendar year as a significant modifier. Although health facility accessibility, availability, and recording completeness were associated with HMIS incidence, they were not significantly associated with bias in estimates from any site. Health management information systems still require improvements; however, its strong predictive power of unbiased malaria burden when improved highlights the important role it could play as a cost-effective tool for monitoring trends and estimating impact of control interventions. This has important implications for malaria control in low-resource, high-burden countries.

INTRODUCTION

Global burden of malaria has declined over the past 20 years because of implementation of wide-scale control interventions and the effective treatment of cases using artemisinin-based combination therapy (ACT).[1,2] Nevertheless, malaria remains a global public health challenge, with more than 200 million cases and 400,000 deaths occurring in 2018 alone.[3] Despite this high burden, case incidence has reduced substantially in at least 31 malaria-endemic countries, including Uganda, and these are on track to reduce the incidence by 40% or more by 2020.[3] To sustain these current gains in control and prevent future epidemics, malaria surveillance is a core intervention as was proposed in the 2016–2030 Global Technical Strategy for Malaria.[4] To be effective, this strategy demands improved understanding of surveillance data.

National malaria control programs rely on data from health management information systems (HMISs) for malaria surveillance and day-to-day decision-making.[5] Although there are extensive HMIS improvements through standardized data formats and quality assessment tools,[6] among others, HMIS data are still underused to provide rigorous evidence of program effectiveness because of concerns about incompleteness, delayed reporting, data quality, and often low rates of definitive diagnostic testing.[7,8] Consequently, other sources of burden estimate, such as Malaria Indicator Surveys and Demographic Health Surveys, are often used as alternatives.[8–11] Although these surveys are valuable to estimate intervention coverage and enable detailed subnational comparisons, they cannot provide continuous information to support ongoing decision-making as can HMISs.[9]

A national HMIS was introduced in Uganda in 1997 to enable priority disease surveillance at national levels[12,13] and has since expanded with the introduction of the District Health Information System version 2 web-based system in 2012.[14] Specific to malaria, the HMIS was additionally supplemented by routine sentinel surveillance through malaria reference centers in different endemicity settings to enable a more detailed understanding of transmission intensity and its temporal trends and improve HMIS data quality.[15] These sentinel site data have been used to evaluate the impact of control interventions and inform decisions on control and prevention in Uganda, and notable improvements in health service delivery and utilization have also been reported as a result of improved use of HMIS data elsewhere.[16,17] Although the precision of HMIS estimates requires further investigation, HMIS data are highly attractive for monitoring malaria trends and supporting intervention programs, considering their temporal and spatial coverage, and cost-effective community representativeness.

To our knowledge, no study in sub-Saharan Africa has investigated how malaria incidence rates from HMIS data compare with incidence rates from rigorous cohort studies. Although cohorts are uncommon and normally small, they provide a gold standard method for measuring the incidence or risk of malaria in defined populations. Here, we quantified the relationship between HMIS incidence and cohort

* Address correspondence to Simon P. Kigozi, London School of Hygiene & Tropical Medicine, Keppel St., London WC1E 7HT, United Kingdom. E-mail: skigozi@yahoo.com

incidence rates in three contrasting settings in Uganda. The analysis sought to then use these data to explore potential sources of bias in HMIS incidence relative to cohort incidence and explain discrepancies between the two, providing important insight into the utility and representativeness of HMIS estimates of malaria burden compared with true population burden.

METHODS

Study design. This was an evaluation study comparing two longitudinal estimates of malaria burden. These estimates were generated using 1) reports of uncomplicated malaria from the HMISs in outpatient departments (OPD) of all public health facilities in three sites and 2) community cohort studies conducted over the same duration in these sites with passive case detection in a dedicated clinic for each site.

Study setting and population. The study used data collected between October 2011 and September 2014 at three varied transmission intensity sites, described in detail elsewhere.[18,19] In brief, these include the following. 1) Walukuba subcounty in Jinja district, approximately 12.0 sq. km in size (Figure 1). Walukuba had approximately 9,800 households according to our enumeration and mapping survey of 2011. It was a peri-urban site of moderate-to-low transmission intensity, with an estimated annual entomological inoculation rate (aEIR) of approximately three infective bites per person per year during this study duration.[18] 2) Kihihi subcounty in Kanungu district, ≈ 186.0 sq. km with approximately 12,700 households, was a rural site of moderate-to-low transmission intensity, with an aEIR of 32. 3) Nagongera subcounty in Tororo district, ≈ 81.0 sq. km with approximately 6,900 households, was a rural site of high transmission intensity, with an aEIR of 310.[18]

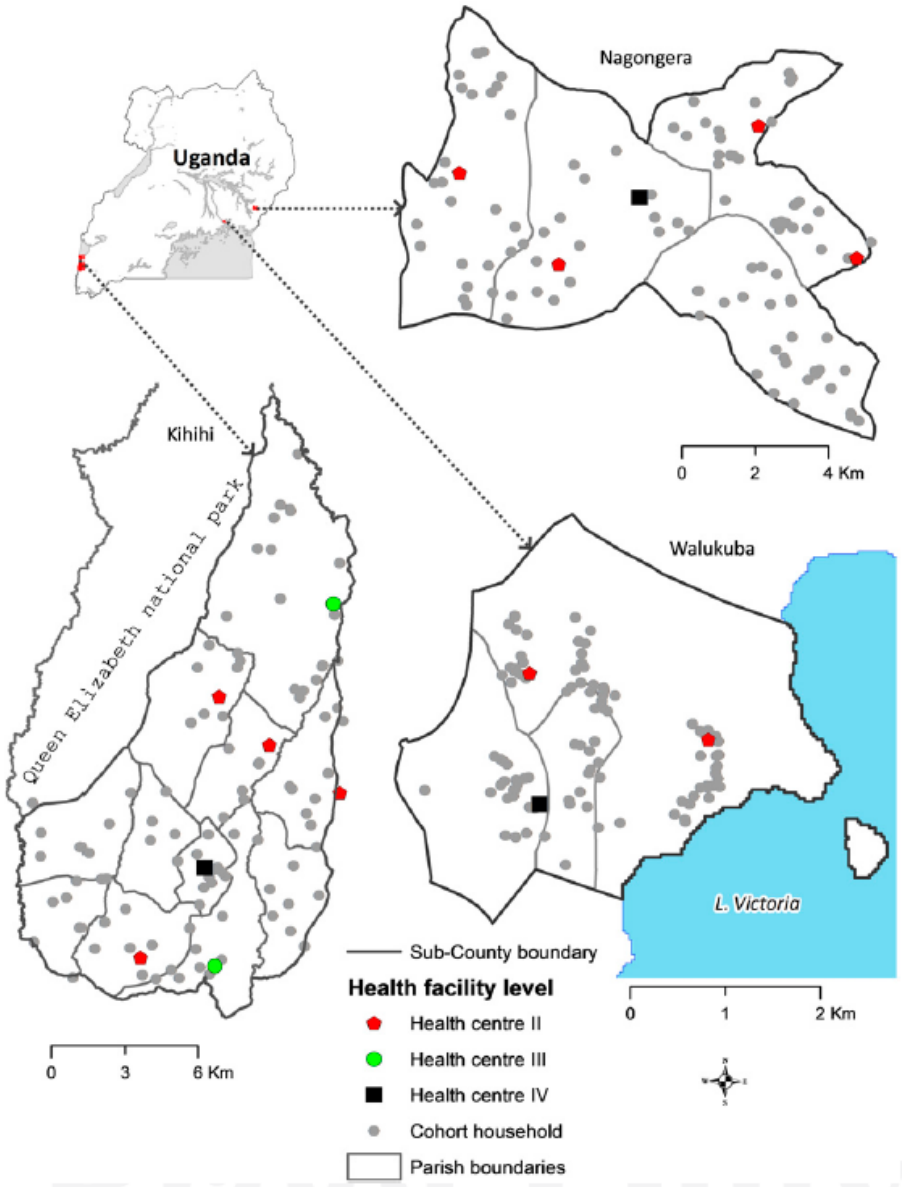


FIGURE 1. Site map showing the sites and health facilities included. This figure appears in color at www.ajtmh.org.

This study encompassed two populations. The first comprised all children aged between 6 months and younger than 11 years living within the three sites considered for HMIS estimates. The second included all children aged between 6 months and younger than 11 years recruited into an enhanced passive community cohort in each of the three sites. Cohort data were interpreted as the unbiased monthly malaria incidence in the three communities.

Data sources. *Health facilities' HMISs.* In Uganda, health facilities receive and assess patients with uncomplicated illness through their OPD clinics. A full record is made, in a standardized OPD register (HMIS Form 031), of patient demographics, place of residence, presenting symptoms, diagnostic test performed, diagnosis made, and treatment prescribed for each patient seen. Here, we extracted data from these registers, entered and cleaned it in a MS Access database (Microsoft Corporation, Redmond, WA), and analyzed it using Stata 15 (Stata Corporation, College Station, TX). This involved three independent projects: the primary project in 12 facilities, and the Uganda Malaria Surveillance Program[15] and the ACT-PRIME project[20] in three facilities each, to provide additional HMIS data. All public health facilities (including three level VI, two level III, and 10 level II) located within the three subcounties provided the HMIS data for this study. Level IV facilities serve ≈ 50,000 people providing inpatient, laboratory, and maternity services, whereas level III facilities serve ≈ 20,000 with inpatient and laboratory services, and level II ≈ 5,000 with basically outpatient and community outreach services.[21]

As a denominator in estimating HMIS incidence rates, estimates of participant population at risk per month were derived, as explained elsewhere.[19] In summary, gridded population images for 2010 and 2015, from the world Pop (www.worldpop.org) project, were used together with intercensual growth rates between the 2002 and 2014 censuses, provided by the national bureau of statistics of Uganda.

Enhanced passive community cohorts. Part of the International Centers of Excellence for Malaria Research, the Program for Resistance, Immunology, Surveillance and Modeling of malaria (PRISM) project, was conducted in Uganda between October 2011 and June 2016. One key component of this project was cohorts conducted in Nagongera, Walukuba, and Kihihi subcounties that comprised study sites. Although details about the PRISM project are fully explained elsewhere,[18] the cohorts involved 100 randomly selected households from each subcounty in which all resident children aged between 6 months and younger than 11 years, with due consent, were enrolled.

Over time, follow-up was terminated for all children who became 11 years old, whereas new children born into or joining study households were considered for enrollment. At enrollment, participants agreed to visit the cohort clinic for all their treatment needs, thereby minimizing clinical visits to other places. Participants were followed up at a dedicated clinic at each site, open 7 days athe week, through regular visits scheduled for once every 3 months, in addition to interim visits if they had illness and treatment needs. At each visit, a blood specimen was assessed for malaria, and transport to the clinic was reimbursed. If a participant was symptomatic and tested positive for malaria parasites, treatment with artemether–lumefantrine was administered and data recorded in a database. By these symptomatic diagnostically confirmed infections, incident malaria cases were identified, and person-time of follow-up per participant per month was computed.

Outcome and explanatory variables. The primary outcome in this study was the monthly malaria incidence rate derived from health facility HMIS OPD incident malaria cases data for children aged between 6 months and younger than 11 years, hereafter referred to as HMIS incidence. An incident case of malaria was defined as an independent symptomatic episode of malaria among participants seen at any study health facility OPD clinic. Level II and III health facilities had very low testing rates and predominantly diagnosed malaria presumptively. Assuming that the risk of malaria for children aged between 6 months and younger than 11 years seen at each reference facility was the same as for similar age children seen at the respective lower level facilities each month, total monthly presumptive malaria cases from lower level facility sites were corrected for test positivity using the monthly test positivity rates from the respective site reference facility (level IV). This approach was also supported by a linear relationship between the cohort fever incidence and HMIS clinical malaria incidence suggesting case identification at facilities as a major factor (Figure 1 in the additional file). Moreover, to optimize catchment-sourced malaria cases, cases with a missing village of residence were corrected for nonresidence within study site boundaries, using the facility monthly proportion of participants with recorded villages of residence that were located or known within the study sites. Notably, cases from villages unknown within site boundaries were excluded. To generate monthly HMIS incidence rates, the site-level sum of incident cases of malaria among children aged between 6 months and younger than 11 years after accounting for nontesting among presumptive cases, and for nonresidence among those with missing data on village of residence, was divided by the site estimated monthly population of children aged younger than 11 years at risk of malaria.

Data from health facility clinic visits that were classified as "reattendance" in the registers, referring to follow-up visits for a previously recorded episode, were excluded from the analyses.

The main explanatory variable was the monthly cohort incidence rate, providing an indication of the level of community transmission, and, thus, true burden. To generate monthly cohort incidence rates, site total incident malaria cases per month were divided by the total person-time of follow-up of the respective site participants over that month. Incident cases of malaria in the cohort were determined as any symptomatic participant visits, at least 2 weeks apart, that were each diagnostically confirmed positive for malaria parasites using blood slide microscopy. Participants with asymptomatic parasitemia at the time of assessment were not included in case estimates. Preliminary analyses showed evidence of bias in incidence because of age, which we attribute to a growing difference in the age structure as the cohorts aged over time. Consequently, we age-standardized incidence estimates in the cohorts using six age categories defined between 6 months and < 11 years, based on the initial recruitment age distribution in these categories as the standard (as explained in section E in the additional file). Initial recruitment into the cohorts was conducted primarily during August and September 2011 for each site.

Regression model. After accounting for community transmission using cohort incidence, we selected factors

quantifiable from our data that may influence the reliability and representativeness of the routine data, including both health-seeking and health facility characteristics. Health seeking was measured by health facility accessibility and health facility availability, whereas health facility characteristics were estimated using health facility performance in recording patients' diagnoses.

Accessibility between residence and health facility was measured by proxy as monthly rainfall estimates (which influence agricultural demands and ease of road use to travel) obtained from TAMSAT raster data described elsewhere[22,23] and extracted as site monthly mean estimates. On examining monthly trends in rainfall and attendance, we observed a general pattern of peaks in rainfall corresponding to troughs in attendance for the same month and vice versa, suggesting associations between rainfall and attendance and supporting its use as proxy for accessibility (Figure 2 in the additional file). These also served as a proxy for seasonality, which is an important factor in malaria transmission dynamics.[24] It was assumed that the higher the mean rainfall received per month the less physically accessible the health facility is for the population that month. Mean rainfall estimates varied by site, with Walukuba receiving 98.9 mm (SD = 45.2), Kihihi 82.2 mm (SD = 40.3), and Nagongera 105.2 mm (SD = 55.8) of rainfall per month. For use in the analysis, however, this metric was standardized around its mean to a mean of zero and SD of one, to generate its *z*-score.

To account for health facility availability, measured as ease of care availability at the facility, we generated the average proportion of days per month that health facilities within each site were open to see patients. The average proportion was defined as the mean number of days a site's facilities were open in a month divided by the respective calendar month's total number of days. In this case, level IV facilities performed best, with mean days open per month 25.8 (SD = 4.0), followed by level III with 24.8 days (SD = 4.1) and level II with 18.8 days (SD = 4.7). However, there was limited variation between sites, with Walukuba recording the highest mean number of days open per month 22.3 (SD = 2.4), followed by Kihihi with 21.5 days (SD = 2.3) and Nagongera with 21.0 days (SD = 1.9). For inclusion in the regression as a covariate, the *z*-score of this metric was generated.

To measure health facility performance at recording vital patient information (recording completeness or performance), we generated the average site health facility proportion of patients seen per month that had no diagnosis recorded in the OPD register. Here, level II facilities performed the best, with the lowest mean number of patients missing a diagnosis per month of 6.5 (SD = 24.6), followed by level IV with 19.7 patients (SD = 25.4) and then level III with 20.9 patients (SD = 40.7). The reciprocal of the proportion was derived so as to enable intuitive interpretation of its trend as performance (high values correspond to high performance and vice versa). This measure was varied by site: Walukuba scored lowest with a median of 89.9 (IQR, 38.3–125.5) points and then Kihihi with 124.1 (IQR, 64.3–178.0) points. Nagongera recorded both highest performance and variability over time, with 310.5 (IQR, 193.2–496.2) points. This score was then standardized to its *z*-score for inclusion in the analysis.

Overall, although Nagongera had the highest recording completeness, it had the lowest availability and least accessibility, making it the lowest HMIS performance site of the three.

Statistical analysis. *Relationship between HMIS and cohort incidence.* We explored the relationship between HMIS and cohort incidence among children aged between 6 months and younger than 11 years on a monthly timescale, stratified by site. First, we examined trends in monthly raw incidence rates as cases per person-year at risk for both HMIS and cohort incidence, using line plots. Second, using monthly incidence rates (standardized to their *z*-scores and hereafter known as incidence rate *z*-scores), we examined the relationship between HMIS and cohort incidence using scatterplots with fitted lines. Third, we examined trends in incidence rate *z*-scores using line plots of calendar month against incidence. And finally, we evaluated agreement between these rates using the Bland–Altman diagram approach without a predetermined threshold of agreement.

Moreover, using test positivity and nonresidence-corrected HMIS incidence as primary outcome and age-standardized cohort incidence of malaria as primary exposure, site-specific Poisson regression models (with incident case counts as outcome and population at risk as offset) were fit with health facility accessibility, health facility availability, and health facility recording performance as the included explanatory variables.

Potential drivers of the differences between HMIS and cohort incidence (sources of bias). Significant factors in the relationship between HMIS and cohort incidence were assessed for viability as potential contributors to differences between HMIS and cohort incidence (potential sources of bias) using linear regression. Factors assessed included monthly estimates of health facility accessibility, health facility availability, and health facility recording performance estimated as explained earlier. To investigate bias due to these factors, we fit a simple linear regression model with the difference between corresponding pairs of incidence rate *z*-scores per month as the primary outcome and the identified factors as explanatory variables.

RESULTS

Study participants. *Health management information systems.* From the 15 OPD clinics in public health facilities located in Walukuba, Kihihi, and Nagongera subcounties enrolled in this study, 153,079 participants aged between 6 months and younger than 11 years were seen between October 2011 and September 2014, with a mean age of 4.14, 4.52, and 3.52 years, respectively (Table 1). Of these participants, 114,919 (75.1%) had their village of residence known to be within site boundaries, whereas 1,840 (1.2%) had an unclear or no village record, and 36,320 (23.7%) were from a village unknown in the study site. A small majority of participants seen at OPD clinics were female (53%). Suspected malaria cases comprised the majority of participants seen (70%), with Nagongera recording the highest proportion of suspected malaria participants (83%), followed by Kihihi (69%) and Walukuba (56%).

Cohorts. From the three sites, 995 children aged between 6 months and younger than 11 years were recruited (304, 357, and 334 from Walukuba, Kihihi, and Nagongera, respectively) in the study, 755 (76%) of them at the start of the study (August–September 2011) (Table 2). Overall, 19,911 clinic visits were made, of which 9,109 (46%) were by female participants. Among all clinic visits, 8,954 (45%) of participants

TABLE 1
Summary of HMIS participant data from health facilities from October 2011 through September 2014

Site	Health facility	Total OPD visits	Visits with suspected malaria*, *n* (%)	Suspected malaria with laboratory test done, *n* (%)	Cases of laboratory-confirmed malaria	Test positivity rate† (%)	Clinical cases of malaria without laboratory testing	Test-adjusted clinical cases‡	Residence-adjusted confirmed cases§	Total estimated cases of malaria‖
Nagongera	Nagongera HCIV	20,611	18,748 (91)	18,380 (98)	8,014	43.60	332	N/A		
	Were HCII	8,054	7,138 (89)	612 (9)	494	N/A	6,499			
	Maundo HCII	8,311	7,089 (85)	1,143 (16)	999		5,931			
	Katajula HCII	10,983	9,179 (84)	5,520 (60)	3,724		3,655			
	Pokongo HCII	12,582	7,891 (63)	1 (0)	1		7,890			
	All	60,541	50,045 (83)	25,656 (51)	13,232	51.60	24,307	8,770	10,190	18,960
Walukuba	Walukuba HCIV	26,018	17,003 (65)	16,890 (99)	4,062	24.00	113	N/A		
	Masese port HCII	5,423	1,765 (33)	134 (8)	34	N/A	1,631			
	Masese three HCII	9,630	4,111 (43)	552 (13)	357		3,559			
	All	41,071	22,879 (56)	17,576 (77)	4,453	25.30	5,303	912	3,972	4,884
Kihihi	Kihihi HCIV	20,510	18,310 (89)	18,300 (99)	8,875	48.50	10	N/A		
	Matanda HCIII	12,293	7,020 (57)	0 (0)	0	N/A	7,020			
	Nyamwegabira HCIII	5,236	2,628 (50)	0 (0)	0		2,628			
	Nyakashure HCII	4,930	2,320 (47)	276 (12)	203		2,044			
	Bihomborwa HCII	6,334	3,448 (54)	0 (0)	0		3,448			
	Bushere HCII	1,168	959 (82)	648 (68)	560		311			
	Kibimbiri HCII	996	847 (85)	0 (0)	0		847			
	All	51,467	35,532 (69)	19,224 (54)	9,638	50.10	16,308	3,079	8,979	12,058

OPD = outpatient departments.
* Patients sent to the laboratory for malaria testing or those given a clinical diagnosis of malaria without laboratory testing.
† Cases of laboratory-confirmed malaria/suspected malaria with laboratory test performed.
‡ Total monthly (clinical cases of malaria without laboratory testing × HCIV test positivity rate).
§ Cases of laboratory-confirmed malaria from known villages + residence adjusted laboratory-confirmed cases from missing villages.
‖ Test-adjusted clinical cases (*c*) + cases of laboratory-confirmed malaria (*d*).

were febrile, with the majority of them being male (55%). Whereas Kihihi had the highest number of participants recruited overall (36%) and Walukuba the lowest at 31%, the majority of clinic visits when febrile (53%) were recorded in Nagongera, the highest transmission site, followed by Kihihi (45%) and Walukuba (29%). That being the case, Nagongera had the highest mean monthly febrile visit frequency at 1.35 visits per individual per month, whereas Kihihi and Walukuba were similar and lower at 1.21 and 1.20 visits per individual per month, respectively.

Incidence rates. *Health management information system incidence rates.* Across the study duration, a total of 4,884, 12,058, and 18,960 symptomatic test and residence-corrected incident malaria cases (Table 1) were generated among participants in Walukuba, Kihihi, and Nagongera, respectively. The majority of incident cases of malaria were registered from lower level (level II) facilities in the highest transmission site of Nagongera (79%), whereas in the lower transmission settings of Walukuba and Kihihi, the majority were recorded from the level IV facilities (53% compared with level II and 43% compared with levels II and III, respectively). Consistent with transmission strata of the sites, mean monthly HMIS incidence rates were lowest in Walukuba, followed by Kihihi, and highest in Nagongera (Table 2).

Cohort incidence rates. Between October 2011 and September 2014, a total of 189, 1,477, and 2,325 incident malaria cases were recorded in the Walukuba, Kihihi, and Nagongera cohorts, respectively (Table 2). From these, site mean monthly incidence rates were derived and age-standardized. These cohort incidence rates were found to be lowest in Walukuba, followed by Kihihi, and highest in Nagongera, a pattern consistent with HMIS incidence rates.

Trends in HMIS and cohort incidence rates. The monthly HMIS and cohort incidence, expressed as raw estimates, showed similar trends in Walukuba but were unclear in the

TABLE 2
Summary of malaria incidence rates

	HMIS				Community cohort*			
Site	Population	Person time in years†	Total cases of malaria‡	Mean monthly incidence rate (range)§	Number of participants	Person time in years	Total cases of malaria	Mean monthly incidence rate (range)
Walukuba	41,071	11,089	4,884	0.19 (0.05–0.41)	304	600.80	189	0.33 (0.02–1.15)
Kihihi	51,467	22,428	12,058	0.26 (0.07–0.57)	357	847.34	1,477	1.71 (0.29–4.70)
Nagongera	60,541	15,907	18,960	0.46 (0.14–0.81)	334	781.68	2,325	3.26 (1.59–5.31)

HMIS = health management information systems.
* Age standardized to age structure of cohort participants at recruitment.
† Population of children younger than 11 years per site at mid-point of study duration, (average for months of March and April 2013).
‡ Cases corrected for nonresidence within site boundaries among those with missing villages, as well as for test positivity among presumptive malaria cases.
§ Incidence rates were corrected for test positivity among presumptive cases and site nonresidence among missing villages.

other sites (Figure 2). As the respective mean-standardized scores, however, results showed similar trends in Walukuba and Kihihi, except for the first two months of study duration in Kihihi (Figure 3, right column). In Nagongera, there was a tendency for HMIS incidence to increase and peak during February 2013, followed by a generally downward trend. The cohort incidence, on the other hand, followed an increasing trend before peaking during May 2014, and then sharply declined through to September 2014.

Relationship between HMIS and cohort incidence rates. Health management information system incidence estimates were seen to underestimate the true (cohort) incidence rates in all three sites, both as mean monthly estimates, by magnitudes of between 2 and 10 (Table 2), and from trends in raw incidence estimates (Figure 2). Nevertheless, mean-standardized rates (*z*-scores) showed the same trends over time, particularly in Walukuba and Kihihi, but not clearly in Nagongera (Figure 3, right column). The overall concordance

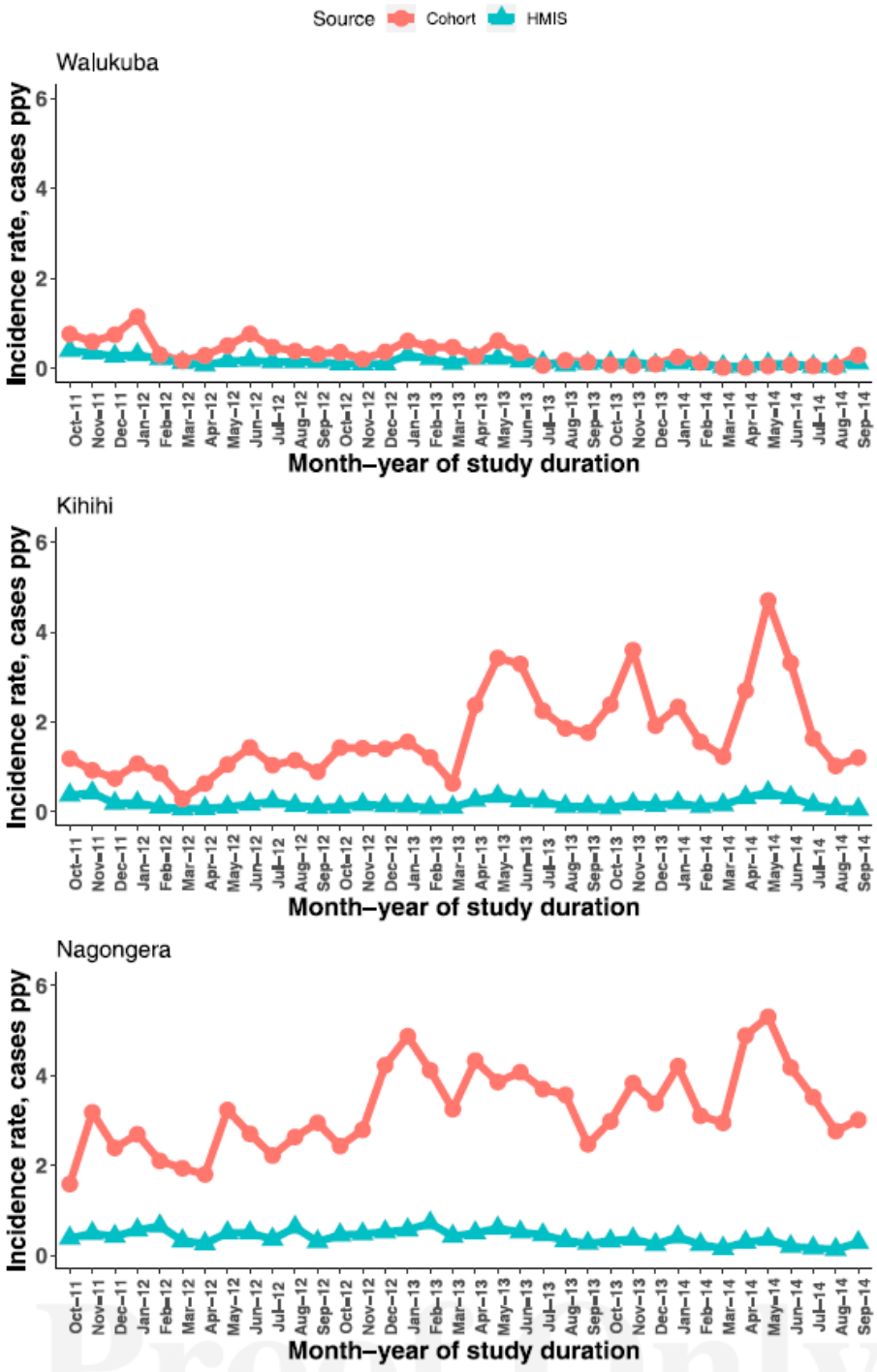


FIGURE 2. Trends in raw incidence rates per month. This figure appears in color at www.ajtmh.org.

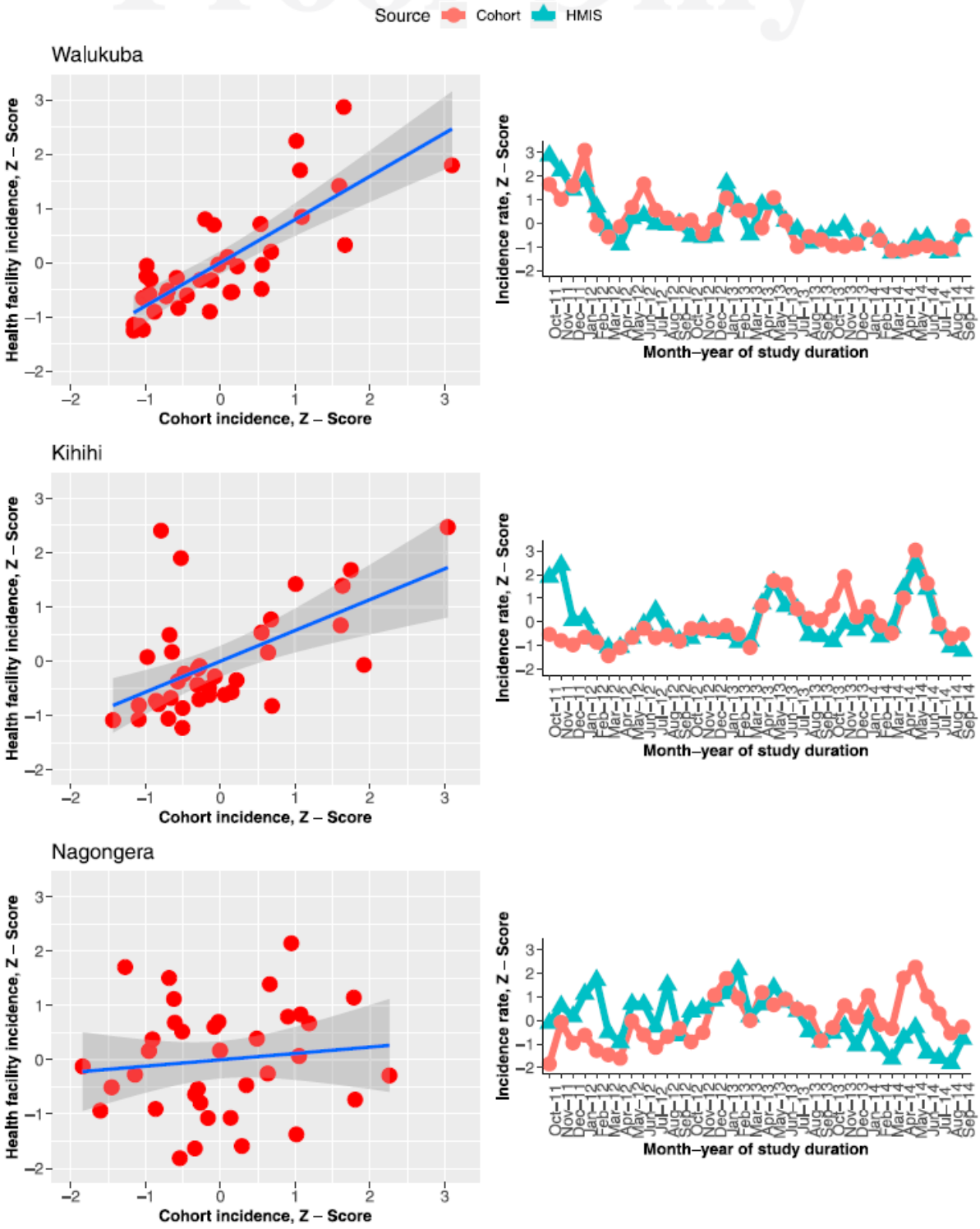


FIGURE 3. Scatterplots and trends in health management information systems and cohort incidence rates on respective mean-standardized scales (*z*-scores). This figure appears in color at www.ajtmh.org.

between HMIS and cohort incidence rates was very good and well within the 95% confidence bounds, by the Bland–Altman criteria at all three sites (Figure 4). Only Kihihi had 2 months as major outliers, consistent with other results observed, including the trends.

Relationships explored using scatterplots of age-standardized cohort incidence *z*-scores against HMIS incidence *z*-scores with fitted lines showed a linear relationship for two of the three sites. This relationship was strongest in Walukuba (adjusted *R*-square = 0.6235), weaker in Kihihi

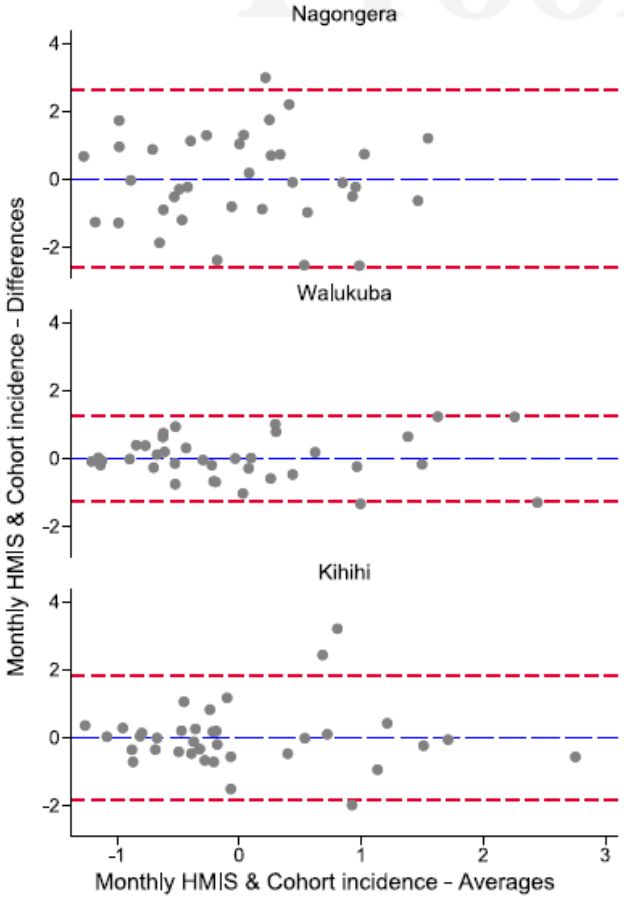


FIGURE 4. Concordance between health management information system incidence and cohort incidence rates using Bland–Altman diagrams by site, based on incidence z-scores. This figure appears in color at www.ajtmh.org.

(Adj. R-sq. = 0.4066), and not present in Nagongera (Adj. R-sq. = 0.0922), as shown in the first column of Figure 3 (and Table 1 in the additional file).

Consistent with exploratory results, adjusted analysis results revealed a significant association between mean monthly HMIS incidence and cohort incidence rates, with varied but significant modifications to the associations by year. In Walukuba, changes in cohort incidence had an increasingly greater effect on HMIS incidence, with the adjusted incidence rate ratio (aIRR) increasing from 0.39 to 14.13 for 2011 to 2014, respectively. This was followed by Kihihi, with reduction in the aIRR from 7.18 to 1.39, respectively, and Nagongera, with the aIRR increasing from 1.09 to 1.34, respectively (Table 3). Factors controlled for were health facility accessibility (rainfall estimates), health facility availability, and health information recording completeness at each site. Moreover, after accounting for the underlying disease burden, the HMIS incidence rate was significantly associated with health facility accessibility, health facility availability, and health information recording completeness in all sites, except recording completeness in Nagongera. Results showed that after accounting for the community burden and year of study, recording completeness was significantly associated with a minimal increase in the HMIS incidence rate for Walukuba, and a sizable decline in Kihihi (aIRR = 0.85, 95% CI: 0.83–0.87, $P <$ 0.001), but not in Nagongera (Table 3). By contrast, health facility availability was significantly associated with a minimal increase in the HMIS incidence rate in Nagongera and Kihihi, but a decrease for Walukuba. Moreover, reduced health facility accessibility was significantly associated with declines in the HMIS incidence rates at all sites, highest in Kihihi (aIRR = 0.88, 95% CI: 0.86–0.90, $P < 0.001$), and moderate in Nagongera and Walukuba.

Notably, calendar year of study was a significant modifier of this relationship, with statistically significant ($P < 0.001$)

TABLE 3

Association between health management information systems and cohort incidence rates, accounting for potential confounders in the relationship

			Unadjusted			Adjusted		
Site	Covariate	Year	IRR	95% CI	P-value	IRR	95% CI	P-value
Walukuba	Effect of cohort incidence* by year	2011	1.01	0.42–2.40	< 0.001	0.39	0.13–1.19	< 0.001
		2012	2.75	2.35–3.21		2.67	2.23–3.20	
		2013	4.26	3.40–5.34		4.37	3.45–5.55	
		2014	21.26	10.62–42.56		14.13	6.70–29.78	
	Recording		–	–	–	1.07	1.02–1.13	0.005
	HF availability		–	–	–	0.95	0.91–0.99	0.017
	HF accessibility		–	–	–	0.97	0.93–0.99	0.042
Kihihi	Effect of cohort incidence* by year	2011	3.07	2.38–3.97	0.017	7.18	5.38–9.58	< 0.001
		2012	1.83	1.63–2.05		1.71	1.51–1.95	
		2013	1.45	1.40–1.50		1.50	1.44–1.55	
		2014	1.51	1.47–1.54		1.39	1.35–1.43	
	Recording		–	–	–	0.85	0.83–0.87	< 0.001
	HF availability		–	–	–	1.03	1.00–1.07	0.023
	HF accessibility		–	–	–	0.88	0.86–0.90	< 0.001
Nagongera	Effect of cohort incidence* by year	2011	1.17	1.09–1.26	< 0.001	1.09	1.01–1.17	< 0.001
		2012	1.14	1.10–1.18		1.22	1.17–1.27	
		2013	1.46	1.40–1.52		1.33	1.27–1.39	
		2014	1.30	1.25–1.35		1.34	1.28–1.39	
	Recording		–	–	–	1.01	0.99–1.03	0.256
	HF availability		–	–	–	1.09	1.07–1.11	< 0.001
	HF accessibility		–	–	–	0.93	0.91–0.94	< 0.001

IRR = incidence rate ratio.

* Cohort incidence rates were age-standardized using the site-specific recruitment population structure as standard population. Other covariates are standardized around their means to obtain the corresponding z-score for each.

interactions of calendar year with cohort incidence in all three sites.

Factors that potentially influence differences (bias) between HMIS and cohort incidence rates. The three factors evaluated were found to be significant confounders in the association between HMIS and cohort incidence rates in all three sites. Given these independent associations with HMIS incidence rates, health facility accessibility and availability, and health information recording completeness were assessed as suspected sources of bias in HMIS incidence relative to the gold standard.

After accounting for calendar year of study, none of the three factors was found to be significantly associated with the difference between the incidence rates (Table 2 in the additional file).

DISCUSSION

In this analysis, we investigated the relationship between routine HMIS incidence (health facility case detection and reporting) and cohort incidence (burden in the community) rates of malaria among children aged between 6 months and younger than 11 years in three diverse sites in Uganda. We found a strong relationship between the two measures in the higher HMIS performance moderate-to-low endemicity sites of Walukuba and Kihihi, with both data sources revealing similar monthly trends, although this was not observed in the lowest HMIS performance high-endemicity setting of Nagongera. Although these findings suggest HMIS incidence as a viable proxy for the true incidence of malaria in moderate-to-low transmission settings,[19,25] results also highlighted how HMIS incidence substantially underestimates cohort incidence at all sites, with subtle differences in cohort incidence reflected by much larger changes in HMIS incidence after controlling for confounders.

These findings are consistent with previous work that reported HMIS incidence as a good measure for evaluating trends in malaria burden within facility catchments, compared with test positivity rates.[19] Taken together, our findings suggest that HMISs will become increasingly relevant as a robust, cost-effective means of monitoring changes in malaria burden, particularly as effective control interventions continue to drive wide-scale reductions in malaria burden.[1,3]

Our findings contrasted substantially between low-to-moderate and high-transmission settings, suggesting important differences in the proportion of malaria infected persons recorded through passive surveillance. This is likely due to variations both in treatment seeking among facility catchment populations, driven by caregiver characteristics, accessibility, and availability of healthcare resources,[26,27] and by the quality of diagnosis and reporting at facilities. It is not uncommon, however, for people to move outside of their closest facility catchment for care, for example, when traveling or because of perceived better care. In those cases, clinical malaria cases may not be registered for surveillance within the facility catchment of origin or not at all, within the private sector. Previous studies have suggested that most care seeking is conducted in private health facilities in Uganda, including private-for-profit hospitals/clinics, pharmacies, and drug shops[9,28]; however, private facilities were very few in our study sites and therefore excluded. Also, prompt treatment is not yet attainable outside of research settings, owing to care-seeking characteristics and diversity. In the lower transmission settings with better HMIS performance, most malaria cases were collected from facilities with high diagnostic testing rates, further improving the confirmation rates in those sites compared with the high-transmission site.[7] This suggests that these findings strongly depend on improved surveillance systems and can be reliable in all transmission settings.

In addition, although treatment was free of monitory cost at public health facilities, regular patient visits to the health facility still costed them in the form of transport cost and long waiting times. Within the cohorts, however, transport was reimbursed for every clinic visit made and waiting times minimized because of the dedicated clinics. This status quo may have limited potential HMIS clinic visits and, thereby, contributed to HMIS underestimating cohort incidence estimates.

To further explore the extent by which other factors may systematically influence the representativeness of HMIS data, we identified indicators that were quantifiable from our data and explored their association with HMIS incidence. Our findings are consistent with previous reports on reliability of HMIS measures of burden being dependent on completeness and accuracy of HMIS records, as well as healthcare access,[10,11] although results were not always consistent across sites. For example, health facility accessibility, health facility availability, and recording completeness were each significantly associated with heterogeneous changes in HMIS incidence after accounting for community burden. In Walukuba, the site with highest availability according to the site mean number of days facilities were open, improved recording completeness was associated with the increased HMIS incidence. For Kihihi, with the second highest availability, however, increased recording completeness was associated with the reduction of HMIS incidence, suggesting heterogeneous effects of the same factor in different locations. We believe these effects may be due to variations in factors such as resource availability and staffing or workload, but these were not evaluated in this study.

For all three study sites, reduced health facility accessibility (represented by increasing monthly rainfall) was associated with reduction in HMIS incidence, the largest impact being in Kihihi, a hilly site with roads that can become impassable when it rains. Moreover, increased rainfall in a given month is also known to be inversely associated with mosquito abundance in that month because of flushing of long-term breeding sites.[29,30] A lagged increase in rainfall that may not influence access and, therefore, not considered here, however, is associated with increased mosquito abundance.[31] Although these findings confirm the perpetuation of errors in routine HMIS data through health system factors, when assessed for likelihood to be sources of bias in HMIS incidence relative to cohort incidence, none qualified. Although health facility accessibility, availability, and recording completeness are vital for accurate estimates of HMIS incidence and its reliability,[8] the absence of their significant impact on the difference between HMIS and cohort incidence rates provides evidence against claims of systematic bias in HMIS burden estimates through these HMIS data drawbacks.[7]

Results revealed the calendar year of study to be a significant modifier of the relationship between the rate of change in HMIS incidence and cohort incidence. This effect may have been due to seasonal or temporal variation in health system factors, such as drugs or diagnostics stocks and staffing that

were reportedly changing over this study duration.[32] It may also be attributable to community campaign–based activities, such as the distribution of long-lasting insecticidal nets in 2013–14 or urbanization.[33,34] Moreover, variability in care seeking for alternatives such as self-medication, use of herbal medicines or drug shops,[35] and availability of diagnostics or drugs[36] could also explain year-to-year impacts on this relationship. These results provide some evidence that the HMIS level of performance and treatment seeking are key factors that drive the strength of the relationship between HMIS incidence and cohort incidence.

This study had several limitations that could be categorized as HMIS data and analytical limitations. First, from the data side, low testing rates in lower level facilities could have impact on the number of cases observed and, thus, on incidence estimates. This may have been exemplified by the lack of clear relationship and absence of clear trends in the highest transmission site, where the majority of malaria cases were recorded from lower level, low testing rate facilities. We did, however, adjust these presumptive cases by reference facility positivity rates, which acted to reduce this impact on incidence estimates. Second, not all facilities had complete data for every month, with some registers missing at the facilities, reducing the number of cases registered. Third, the difference in diagnostic testing methods between the cohort (microscopy) and health facilities (microscopy or rapid diagnostic test) may have introduced some disparities due to sensitivity and specificity differences in methods. Fourth, being unable to access patient records from private facilities may have impact on the true number of malaria cases in each community. However, private facilities were very few in each site, and results from unpublished data showed that most private facilities that were expected to be in the sites had closed. Fifth, being unable to account for malaria commodity stock levels in health facilities may have undermined HMIS incidence through reduced HMIS attendance whenever there was a considerable stock-out season. Sixth, health facility availability, accessibility, and recording performance are more complex than this study proposed to estimate them. This could have masked any potentially observable associations otherwise not found. Last, it is unclear what proportion of patients with missing age would have been participants in this study, and this too could have had an impact on our results.

On the side of analytical limitations, first, we were unable to obtain census estimates of the population for the health facility catchments considered. Consequently, we estimated the denominator for HMIS incidence, which may not have been a precise measure of the catchment population at risk. Second was the inability to directly account for bias due to gender that was identified. However, the overall effect of this bias was not expected to vary much over time, given that gender does not change over time. Third, being limited to three sites may imply that comparisons between them were limited; however, the diversity of settings and transmission provides important contributions of benefit to surveillance and considerably generalizable findings in Uganda. Last, there were limited data on covariates describing treatment seeking, hospital characteristics, and governance factors, which may have left several influences unaccounted for. However, using the available data, considerable scope was taken into account.

CONCLUSION

Findings from this study show that although HMIS substantially underestimates the cohort-based malaria incidence rate, there is empirical evidence of a strong linear relationship between these incidence rates in children living in high HMIS performance moderate-to-low endemicity settings that deteriorates in a low HMIS performance very high–transmission setting. This, coupled with similar trends in these rates and good concordance, suggests that HMIS incidence rates may constitute a reliable estimate of malaria burden and its trends with improved HMISs. These findings have important implications for malaria risk assessment in low-resource settings that bear most of the burden of malaria, given improved information systems. Coupled with successful control interventions reported globally, the stronger predictive power of HMIS incidence for the true burden with improving HMISs, despite transmission settings, suggests a cost-effective means of evaluating malaria risk for effective control.

Received December 20, 2019. Accepted for publication March 1, 2020.

Note: Supplemental file appears at www.ajtmh.org.

Acknowledgments: We would like to thank Katonda Omukulu, Joy Wanigine, Mary Nyaketcho, Julius Moi, Cosmas Ofwono, Isaac Onyango, Fatuma Adikini, Stanley Ogolla, Deborah Ekusai, Rhoda Namubiru, Faith Bagatya, Sylvia Kamisha, Charles Kavuma, Hajara Kyokisa, John Rek, Nicholas Wendo, Bryan Greenhouse, Asadu Sserwanga, Geoff Lavoy, and the administration and staff at the Infectious Diseases Research Collaboration for their contributions. We would also like to thank the HMIS officers and in charges at each of the clinics for all their hard work and great contributions to the success of this project, as well as the district health leadership in Jinja, Tororo, and Kanungu districts for all their cooperation with and support toward this study.

Financial support: Funding for the research reported in this publication was supported by the Fogarty International Center of the National Institutes of Health under award numbers D43TW7375 and D43TW010526. For the Uganda malaria surveillance project that provided secondary data, funding was provided by the President's Malaria Initiative, U.S. Agency for International Development under the terms of Interagency Agreement (1U51CK000117) with the CDC. For the PRISM study that also provided secondary data, funding was provided by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health as part of the International Centers of Excellence in Malaria Research program [U19AI089674]. For the ACT PRIME study that also provided secondary data, funding was provided by the ACT Consortium (http://www.actconsortium.org/) through a grant from the Bill & Melinda Gates Foundation to the London School of Hygiene and Tropical Medicine. J. I. N. is supported by the Fogarty International Center (Emerging Global Leader Award) grant number K43TW010365. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the President's Malaria Initiative, U.S. Agency for International Development, the National Institutes of Health, and the Bill & Melinda Gates Foundation.

Disclosure: This study was approved by the Uganda National Council for Science and Technology (UNCST Ref SS 4455), the Makerere University School of Medicine Research and Ethics Committee (REC Ref. 2017–119), and the London School of Hygiene & Tropical Medicine Ethics Committee (LSHTM Ref. 13902).

The datasets used and/or analyzed for this study are available from the corresponding author on reasonable request.

Authors' addresses: Simon P. Kigozi and Rachel L. Pullan, Department of Disease Control, London School of Hygiene & Tropical Medicine, London, United Kingdom, E-mails: skigozi@yahoo.com and rachel.pullan@lshtm.ac.uk. Emanuele Giorgi, CHICAS, Lancaster Medical School, Lancaster University, Lancaster, United Kingdom,

E-mail: e.giorgi@lancaster.ac.uk. Arthur Mpimbaza, Child health and development centre, Makerere University College of Health Sciences, Kampala, Uganda, E-mail: arthurwakg@yahoo.com. Ruth N. Kigozi, USAID's Malaria Action Program for Districts, Kampala, Uganda, E-mail: kigoziruth@gmail.com. Teun Bousema, Department of Medical Microbiology, Radboud University, Nijmegen, Netherlands, E-mail: teun.bousema@radboudumc.nl. Emmanuel Arinaitwe, Infectious Diseases Research Collaboration, Kampala, Uganda, E-mail: earinaitwe@idrc-uganda.org. Joaniter I. Nankabirwa and Moses R. Kamya, School of Medicine, Makerere University College of Health Sciences, Kampala, Uganda, E-mails: jnankabirwa@yahoo.co.uk and mkamya@infocom.co.ug. Catherine M. Sebuguzi, National Malaria Control Division, Uganda Ministry of Health, Kampala, Uganda, E-mail: cmaiteki@yahoo.com. Sarah G. Staedke, Department of Clinical Research, London School of Hygiene & Tropical Medicine, London, United Kingdom, E-mail: sarah.staedke@lshtm.ac.uk. Grant Dorsey, Department of Medicine, University of California, San Francisco, San Francisco, CA, E-mail: grant.dorsey@ucsf.edu.

REFERENCES

1. Bhatt S et al., 2015. The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. *Nature 526:* 207–211.
2. Murray CJ et al., 2014. Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet 384:* 1005–1070.
3. WHO, 2019. *World Malaria Report 2019*. Geneva, Switzerland: World Health Organization.
4. WHO, 2015. *Global Technical Strategy for Malaria 2016–2030*. Geneva, Switzerland: World Health Organization.
5. WHO, 2003. *Improving Data Quality: A Guide for Developing Countries*. Manila: World Health Organization.
6. Yourkavitch J, Prosnitz D, Herrera S, 2019. Data quality assessments stimulate improvements to health management information systems: evidence from five African countries. *J Glob Health 9:* 010806.
7. WHO, 2011. *World Malaria Report 2011*. Geneva, Switzerland: World Health Organization.
8. WHO, 2018. *World Malaria Report 2018*. Geneva, Switzerland: World Health Organization.
9. Uganda Bureau of Statistics (UBOS), 2015. *Uganda Malaria Indicator Survey 2014–15*. Kampala, Uganda, and Rockville, MD.
10. Maokola W, Willey BA, Shirima K, Chemba M, Armstrong Schellenberg JRM, Mshinda H, Alonso P, Tanner M, Schellenberg D, 2011. Enhancing the routine health information system in rural southern Tanzania: successes, challenges and lessons learned. *Trop Med Int Health 16:* 721–730.
11. Mate KS, Bennett B, Mphatswe W, Barker P, Rollins N, 2009. Challenges for routine health system data management in a large public programme to prevent mother-to-child HIV transmission in South Africa. *PLoS One 4:* e5483.
12. Gladwin J, Dixon RA, Wilson TD, 2003. Implementing a new health management information system in Uganda. *Health Policy Plan 18:* 214–224.
13. Hotchkiss DR, Aqil A, Lippeveld T, Mukooyo E, 2010. Evaluation of the performance of routine information system management (PRISM) framework: evidence from Uganda. *BMC Health Serv Res 10:* 188.
14. Kiberu VM, Matovu JKB, Makumbi F, Kyozira C, Mukooyo E, Wanyenze RK, 2014. Strengthening district-based health reporting through the district health management information software system: the Ugandan experience. *BMC Med Inform Decis Mak 14:* 40.
15. Sserwanga A et al., 2011. Improved malaria case management through the implementation of a health facility-based sentinel site surveillance system in Uganda. *PLoS One 6:* e16316.
16. Nisingizwe MP et al., 2014. Toward utilization of data for program management and evaluation: quality assessment of five years of health management information system data in Rwanda. *Glob Health Action 7:* 25829.
17. Wagenaar BH et al., 2015. Effects of a health information system data quality intervention on concordance in Mozambique: time-series analyses from 2009–2012. *Popul Health Metr 13:* 9.
18. Kamya MR et al., 2015. Malaria transmission, infection, and disease at three sites with varied transmission intensity in Uganda: implications for malaria control. *Am J Trop Med Hyg 92:* 903–912.
19. Kigozi SP, Kigozi RN, Sserwanga A, Nankabirwa JI, Staedke SG, Kamya MR, Pullan RL, 2019. Malaria burden through routine reporting: relationship between incidence and test positivity rates. *Am J Trop Med Hyg 101:* 137–147.
20. Staedke SG, Chandler CIR, DiLiberto D, Maiteki-Sebuguzi C, Nankya F, Webb E, Dorsey G, Kamya MR, 2013. The PRIME trial protocol: evaluating the impact of an intervention implemented in public health centres on management of malaria and health outcomes of children using a cluster-randomised design in Tororo, Uganda. *Implement Sci 8:* 114.
21. MoH, 2018. *National Health Facility Master List 2018*. ■: Ministry of Health, Division of Health Information.
22. Maidment R, Allan RP, Tarnavsky E, Stringer M, Hewison T, Roebeling R, Black E, 2014. The 30 year TAMSAT African rainfall climatology and time series (TARCAT) data set. *J Geophys Res Atmos 119:* 10619–10644.
23. Tarnavsky E, Grimes D, Maidment R, Black E, Allan RP, Stringer M, Chadwick R, Kayitakire F, 2014. Extension of the TAMSAT satellite-based rainfall monitoring over Africa and from 1983 to present. *Appl Meteorol Climatol 53:* 2805–2822.
24. Cairns ME et al., 2015. Seasonality in malaria transmission: implications for case-management with long-acting artemisinin combination therapy in sub-Saharan Africa. *Malar J 14:* 321.
25. Boyce RM, Reyes R, Matte M, Ntaro M, Mulogo E, Lin F-C, Siedner MJ, 2016. Practical implications of the non-linear relationship between the test positivity rate and malaria incidence. *PLoS One 11:* e0152410.
26. Kadobera D, Sartorius B, Masanja H, Mathew A, Waiswa P, 2012. The effect of distance to formal health facility on childhood mortality in rural Tanzania, 2005–2007. *Glob Health Action 5:* 19099.
27. Burton DC, Flannery B, Onyango B, Larson C, Alaii J, Zhang X, Hamel MJ, Breiman RF, Feikin DR, 2011. Healthcare-seeking behaviour for common infectious disease-related illnesses in rural Kenya: a community-based house-to-house survey. *J Health Popul Nutr 29:* 61–70.
28. Ndyomugyenyi R, Magnussen P, 2004. Trends in malaria-attributable morbidity and mortality among young children admitted to Ugandan hospitals, for the period 1990–2001. *Ann Trop Med Parasitol 98:* 315–327.
29. Okorie PN, Popoola KOK, Awobifa OM, Ibrahim KT, Ademowo GO, 2014. Species composition and temporal distribution of mosquito populations in Ibadan, southwestern Nigeria. *J Entomol Zool Stud 2:* 164–169.
30. Benedum CM, Seidahmed OME, Eltahir EAB, Markuzon N, 2018. Statistical modeling of the effect of rainfall flushing on dengue transmission in Singapore. *PLoS Negl Trop Dis 12:* e0006935.
31. Ssempiira J, Kissa J, Nambuusi B, Kyozira C, Rutazaana D, Mukooyo E, Opigo J, Makumbi F, Kasasa S, Vounatsou P, 2018. The effect of case management and vector-control interventions on space-time patterns of malaria incidence in Uganda. *Malar J 17:* 162.
32. MoH, 2015. *Annual Health Sector Performance Report–FY 2014/2015*. Kampala, Uganda: Ministry of Health - Uganda.
33. Kigozi SP et al., 2015. Associations between urbanicity and malaria at local scales in Uganda. *Malar J 14:* 374.
34. Wanzira H, Katamba H, Rubahika D, 2016. Use of long-lasting insecticide-treated bed nets in a population with universal coverage following a mass distribution campaign in Uganda. *Malar J 15:* 311.
35. Buabeng KO, Duwiejua M, Dodoo ANO, Matowe LK, Enlund H, 2007. Self-reported use of anti-malarial drugs and health facility management of malaria in Ghana. *Malar J 6:* 85.
36. Johansson EW, Gething PW, Hildenwall H, Mappin B, Petzold M, Peterson SS, Selling KE, 2015. Effect of diagnostic testing on medicines used by febrile children less than five years in 12 malaria-endemic African countries: a mixed-methods study. *Malar J 14:* 194.

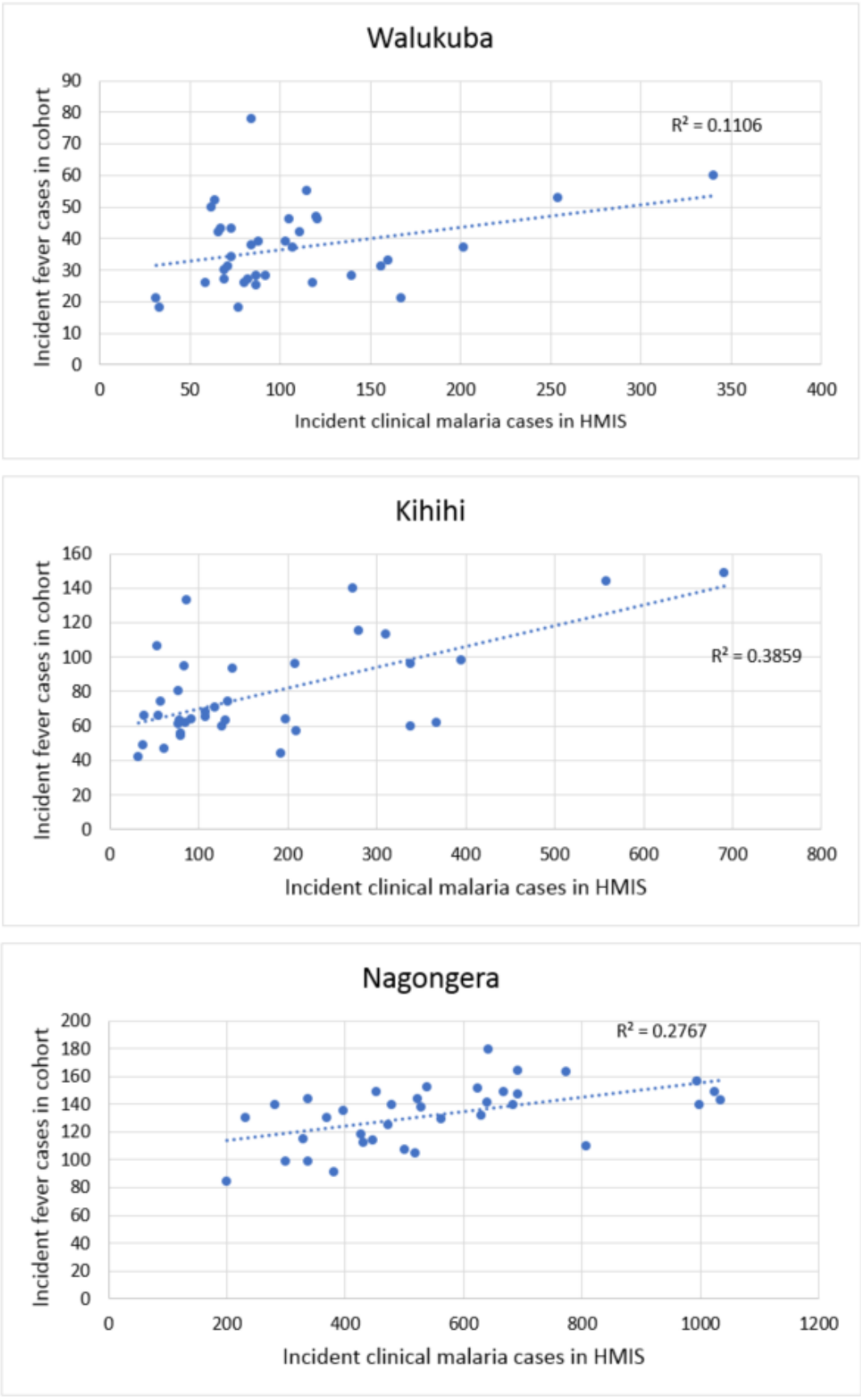
5.1 Additional information for Paper 3

5.1.1 Cohort fever and HMIS clinical malaria

We examined the relationship between cohort fever incidence per month and monthly clinical malaria case incidence to evaluate the viability of our approach of using reference health facility test positivity rates to correct for non-testing among lower-level facilities.

Scatter plots of HMIS clinically diagnosed malaria cases against cohort incident cases per month (Figure 1) indicate a linear relationship between these metrics.

Figure 1. Relationship between HMIS clinical case incidence and fever case incidence in the cohorts

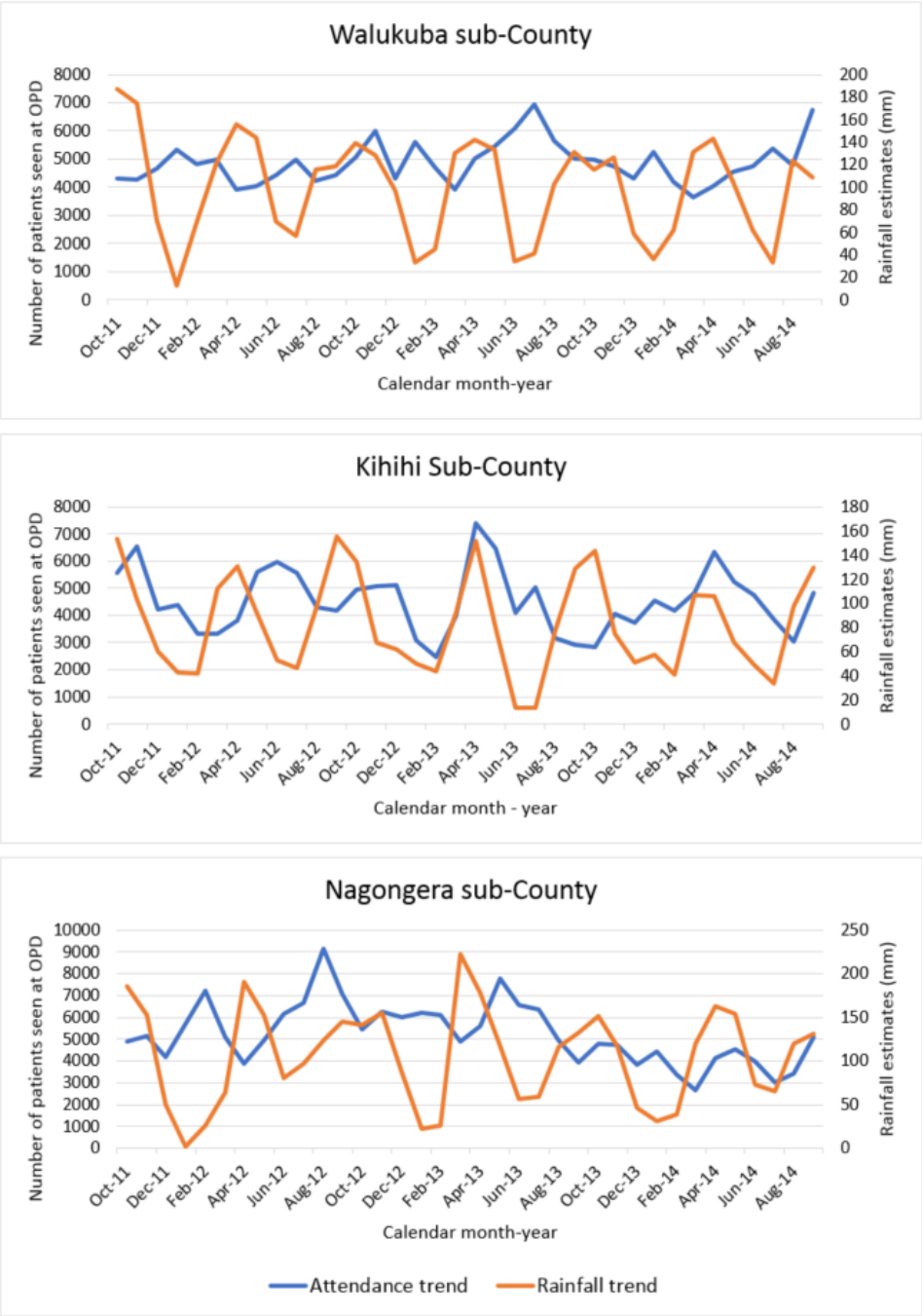


From the observed relationship in all three sites, there is evidence that the rate of change in HMIS clinical malaria incidence is higher than the rate of change in cohort fever incidence (gentle slope). This suggests that within the HMIS, identification of malaria cases would be exaggerated if it were based on clinical symptoms to be higher than expected within a well characterised sample of the study population (cohorts) in the three sites.

5.1.2 Rainfall as a proxy for accessibility

We examined the viability of rainfall as a proxy for health facility accessibility using trends lines of rainfall and total number of patients seen at all facilities within each site, which are total patients visiting public health facilities within out study sites (Figure 2).

Figure 2. Monthly trends in HMIS patient attendance compaared to rainfall estimates per site



The trends in Figure 2 were estimated using site monthly total patient attendance on the left y-axis and rainfall estimates (mm) of the right y-axis, with calendar month of study year on the x-axis by site. Orange trend-lines represent rainfall while blue trendlines represent attendance. We observed in these trends that while they do not track each other directly, for a large number of months in each site, rainfall peaks correspond to low patient attendance and vice-versa. This provides some evidence of negative associations between monthly rainfall amounts and monthly patient attendance at public health facilities. We assume therefore, that increased rainfall

amounts facilitate reduced attendance at public health facilities, and this may be through agriculture involvement or travel difficulties, among others.

5.1.3 Fitted relationship between cohort and HMIS incidence rates

Scrutinizing the linear relationship between HMIS-based incidence and cohort incidence, we used multi-variable regression and examined each explanatory variable for the potential effect in improving model fit using the coefficient of determinations, R-squared and R-squared adjusted. Results of this evaluation are summarised in Table 1 below.

Table 1. Evaluation of linear relationship fit between cohort incidence and HMIS incidence rates by site, involving three vital explanatory variables.

Site	Coefficients of determination	Explanatory variables in model					
		Cohort incidence	Rainfall	HF Availability	Recording performance	Recording + Availability	Recording + Availability + Rainfall
Walukuba	R-sq	0.6322	0.645	0.6322	0.6329	N/A	N/A
	R-sq (adj)	0.6214	**0.6235**	0.61	0.6107		
Kihihi	R-sq	0.3222	0.323	0.3249	0.4405	N/A	N/A
	R-sq (adj)	0.3022	0.282	0.284	**0.4066**		
Nagongera	R-sq	0.0137	0.0577	0.0701	0.0991	0.1334	0.1959
	R-sq (adj)	-0.0153	**0.0006**	**0.0138**	**0.0445**	**0.0521**	**0.0922**

HF = Health facility

Owing to the fact that adjusted R-squared provides a measure of importance of any additional explanatory variable in improving a model fit whereby once inclusion of the variable in the model leads to an increase in the adjusted R-squared, the variable is considered vital in the association evaluated. This outcome of increased adjusted R-squared implies that the added explanatory variable improved the model fit well enough above the increase being an occurrence due to chance.

Walukuba: Results here showed that in Walukuba, adding rainfall (the proxy for accessibility) to the basic model Improved the model fit, given the increase in adjusted R-squared from 0.6214 to 0.6235 (Table 1). However, the same was not true for the addition of either health facility availability or recording performance in Walukuba. As a result, it was not necessary to proceed with more complex models that would include health facility availability and/or recording performance in addition to rainfall. Our results showed that after accounting for rainfall, the

best fit for the linear relationship between HMIS and cohort incidence rates was such that variation in cohort incidence explained up to 65% of the variation HMIS incidence rates in Walukuba.

Kihihi: Here, adding recording performance to the basic model Improved the model fit, given the increase in adjusted R-squared from 0.3022 to 0.4066 (Table 1). However, the same was not true for the addition of either health facility availability or rainfall in Kihihi. As a result, it was not necessary to proceed with more complex models that would include rainfall and/or health facility availability in addition to recording performance. Our results here showed that after accounting for health facility recording performance, the best fit model for the linear relationship was such that at least 44% of variation in HMIS incidence rates was explained by variation in cohort incidence rates in Kihihi.

Nagongera: Here, unlike the other two sites, results showed that all three explanatory variables of rainfall, health facility availability, and health facility recording performance when added independently to the basic model, improved the model fit as shown in Table 1 above. This, therefore, implied that there was need to proceed with adding all explanatory variables and with every additional covariate, the model fit was improved as indicated by progressive increase in adjusted R-squared. The model with all three covariates included provided the best fit for this weak linear relationship with up to 20% variation in HMIS incidence explained by cohort incidence after accounting for rainfall as well as health facility recording performance and availability.

5.1.4 Potential sources of bias in HMIS relative to cohort incidence rates

Evaluating the sources of bias in HMIS-based incidence of malaria (HMIS incidence) estimates relative to the gold standard (cohort incidence), we considered factors that could be influential to the number of malaria cases that end up being recorded at the health facility. These factors, estimated at site level on a monthly timescale, included; health information recording performance, health facility availability, and rainfall estimates that represent both seasonality and ease of access to the health facility from the patients' viewpoint.

Table 2. Regression results evaluating association between "difference between HMIS and Cohort incidence" and potential sources of bias (important factor in recording malaria cases in health facilities) in HMIS incidence

Site	Covariate	Coefficient	95% CI	P-value
Nagongera	HF Recording Performance	0.02480	-0.28632 – 0.33591	0.872
	HF Availability	0.26672	-0.00903 – 0.54247	0.057
	Rainfall estimate	-0.21206	-0.48522 – 0.06110	0.124
	Year of study	-1.17198	-(1.49468 – 0.84927)	<0.001
Walukuba	HF Recording Performance	0.04405	-0.206637 – 0.29447	0.722
	HF Availability	-0.05617	-0.31409 – 0.20174	0.660
	Rainfall estimate	0.15120	-0.08453 – 0.38692	0.200
	Year of study	0.05992	-0.20033 – 0.32016	0.642
Kihihi	HF Recording Performance	-0.21759	-0.55307 – 0.11789	0.196
	HF Availability	0.18247	-0.12444 – 0.48938	0.234
	Rainfall estimate	-0.02498	-0.30566 – 0.25569	0.857
	Year of study	-0.57000	-(0.95522 – 0.18478)	0.005

HF = Health facility

Results here showed that the effect of the explanatory variables on the difference between HMIS and cohort incidence were varied with recording performance, health facility availability and rainfall having no significant association in Nagongera (Table 2) with or without controlling for calendar year of study. Similarly, for Walukuba, neither of the three explanatory variables had a significant association with the difference between HMIS and cohort incidence. In Kihihi, without accounting for calendar year of study, health facility recording performance was significantly associated with the difference between HMIS and cohort incidence (p=0.004). In this case improved recording performance was associated with reduced difference between the incidence estimates. That being the case however, after accounting for calendar year of study, an important modifier of the relationship between HMIS and cohort incidence, none of the three factors was associated the difference between incidence estimates in Kihihi. While these factors are associated with HMIS incidence, these findings show that there is no evidence that they contribute to biased HMIS incidence estimates relative to the true incidence.

5.1.5 Age standardizing the cohort incidence rates

Procedures

It was observed that there was bias due to age ostensibly resulting from differences between the health facility and the cohort populations. Given further that age is a significant risk factor for malaria, in order to evaluate relationship between health facility and cohort incidence, it is important to account for this bias.

Two possible ways that I identified of doing this were: 1) adjusting for age in a binary outcome model at patient-level & 2) Age standardization of the cohort incidence, given that the cohort is the more stringently held population that differed from the normal population over time due to aging of participants, be it so slowly and minimally over time.

The major drawback to approach 1 is that the patient-level analysis is limited to diagnostically tested cases, which situation implies that one has much lower number of cases of malaria from the health facility system, given that lower level facilities were not testing the majority of their suspected malaria cases. The benefit to this approach, however, is that we can account for other known or identifiable sources of bias such as gender (that was identified as a source of bias) and ultimately generate a model-based incidence that is corrected for bias.

The drawback to approach 2 is that there is no known standard population structure to fit the small rage of ages included in this study as compared to known standard age structures that span a much wider age-range. However, even with a small age-range, the ability to define finer age-categories provides the detailed effect of age that is not obtainable when international standard populations with larger categories, are used.

The other drawback to approach 2 is being unable to account for other sources of bias as is the case in approach 1.

Taking approach 2

Having examined standard population structures, both international and national and not obtaining any that fits a small range of ages i.e. 6months to under 11years old, we chose to use the age structure of the cohort participants at the time of enrolment – a two months duration of August to September, 2011. Whereas the cohort continued to recruit participants over the entire duration, the cohort generally grew older while at the same time taking in a considerable number of children under 1year of age (children born into the cohort).

After obtaining the site specific 'standard population' we generated the age category-specific incidence rate, for the six age categories created including [0.5-1yr], [>1-2yrs), [3-4yrs), [5-6yrs), [7-8yrs), [9-<11yrs], where the notation for brackets implies that [is bounded at the value, while) implies unbounded at the value, we generated the age-category or stratum incidence rates i.e. total age-category incident cases/ age category person time for the month.

As explained in https://www.statcan.gc.ca/eng/dai/btd/asr as well as in Chapter 2 of Introduction to Epidemiology p.27, age standardization is then achieved by multiplying each age specific incidence rate by the proportion for each age-category from the recruitment age structure, known as the standard population weight for each age category. These products are summed up for all the age categories per month to obtain site monthly age-standardized incidence rates from the cohort. Note: Whereas there are two approaches to standardization i.e. direct & indirect, we used direct standardization because we could generate age-specific

rates for the age categories defined. Otherwise the indirect approach is applicable for situations where age-specific rates are unknown or where the population under study is small (unlike in our case for both conditions)

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1513171	**Title**	Mr.
First Name(s)	SIMON PETER		
Surname/Family Name	KIGOZI		
Thesis Title	Exploring the utility of indicators of uncomplicated malaria burden from routine health facility surveillance data in identifying and mapping high-risk areas for malaria in Uganda		
Primary Supervisor	Assoc. Prof. Rachel L. Pullan		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

<u>SECTION B – Paper already published</u>

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	BMC Public Health
Please list the paper's authors in the intended authorship order:	Simon P. Kigozi, Ruth N. Kigozi, Catherine M Sebuguzi, Jorge Cano, Damian Rutazaana, Jimmy Opigo, Teun Bousema, Adoke Yeka, Anne Gasasira, Benn Sartorius, & Rachel L Pullan
Stage of publication	**SubmittedSubmitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I prepared the data, led data analysis, and drafted the manuscript. I then led the manuscript submission for peer review, I will address reviewer comments, and lead the final submission for publication

SECTION E

Student Signature	simon peter kigozi
Date	01-September-2020

Supervisor Signature	Rachel Pullan
Date	3rd September 2020

6 Paper 4

Spatial-temporal patterns of malaria incidence in Uganda using HMIS data from 2015 to 2019

Simon P Kigozi[1,2], Ruth N Kigozi[3], Catherine M Sebuguzi[2,4], Jorge Cano[1], Damian Rutazaana[4], Jimmy Opigo[4], Teun Bousema[5], Adoke Yeka[6], Anne Gasasira[7], Benn Sartorius[1], Rachel L Pullan[1]

1 *Department of Disease Control, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK,*
2 *Infectious Diseases Research Collaboration, PO Box 7475, Kampala, Uganda,*
3 *USAID's Malaria Action Program for Districts, PO Box 8045, Kampala, Uganda,*
4 *National Malaria Control Division, Uganda Ministry of Health, Kampala, Uganda*
5 *Department of Medical Microbiology, Radboud University, Nijmegen, Netherlands*
6 *Department of Disease Control and Environmental Health, College of Health Sciences, School of Public Health, Makerere University. PO Box 7072, Kampala, Uganda*
7 *African Leaders Malaria Alliance (ALMA), Kampala, Uganda*

Corresponding author: Simon P Kigozi, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK; Email: skigozi@yahoo.com; Tel: +256 (0) 751 680244

Abstract

Background. As global progress to reduce malaria transmission continues, it is increasingly important to track changes in malaria incidence rather than prevalence. Risk estimates for Africa have largely underutilized available health management information systems (HMIS) data to monitor trends. This study uses national HMIS data together with environmental and geographical data, to assess spatial temporal patterns of malaria incidence at facility catchment level in Uganda over a recent 5-year period.

Methods. Data reported by 3446 health facilities in Uganda, between July 2015 and September 2019 was analysed. To assess the geographic accessibility of the health facilities network, a WHO tool for modelling accessibility (AccessMod) was employed to determine a three-hour cost-distance catchment around each facility. Using confirmed malaria cases and total catchment population by facility, an ecological Bayesian conditional autoregressive spatial temporal Poisson model was fitted to generate monthly posterior incidence rate estimates, adjusted for caregiver education, rainfall, land surface temperature, night-time light (an indicator of urbanicity), and vegetation index.

Results. An estimated 38.8 million (95% Credible Interval [CI]: 37.9 – 40.9) confirmed cases of malaria occurred over the period, with a national mean monthly incidence rate of 20.4 (95% CI: 19.9 - 21.5) cases per 1000, ranging from 8.9 (95% CI: 8.7 – 9.4) to 36.6 (95% CI: 35.7 – 38.5) across the study period. Strong seasonality was observed, with June-July experiencing highest peaks and February-March the lowest peaks. There was also considerable geographic heterogeneity in incidence, with health facility catchment relative risk during peak transmission months ranging from 0 to 50.5 (95% CI: 49.0 – 50.8) times higher than national average. Both districts and health facility catchments showed significant positive spatial autocorrelation; health facility catchments had global Moran's I = 0.3 (p<0.001) and districts Moran's I = 0.4 (p<0.001). Notably, significant clusters of high-risk health facility catchments were concentrated in Acholi, West Nile, Karamoja, and East Central – Busoga regions.

Conclusion. Findings showed clear countrywide spatial temporal patterns with clustering of malaria risk across districts and health facility catchments within high-risk regions, which can facilitate targeting highest risk areas with interventions. Moreover, despite high and perennial transmission, seasonality for malaria incidence highlights the potential for optimal and timely implementation of targeted interventions.

Key words: Uganda, Malaria, Incidence, Relative risk, Routine surveillance, HMIS, Seasonality

Background

The global burden of malaria has declined since 2000 primarily due to the scale up of control interventions including long-lasting insecticidal nets (LLINs), indoor residual spraying with insecticide (IRS), and use of artemisinin-based combination therapy (ACT) [1, 8, 71]. Nevertheless, incidence rates in sub-Saharan Africa remained high at an estimated 219 cases per 1000 in 2017 – 2018 [71]. The incidence estimates used to monitor trends across sub-Saharan Africa are typically generated using parasite prevalence in children 2-10 years fitted in prevalence-to-incidence models [71]. Though informative, the surveys included happen infrequently [60] and may be limited in scale. Derived burden estimates, therefore, cannot adequately support day-to-day monitoring for decision making at national or sub-national levels [64].

National malaria control programmes typically depend on routine health management information systems (HMIS) data to guide programme decisions in control and elimination efforts. With the advent and extended access to web-based health information systems, such as the District Health Information System - version 2 (DHIS-2), timely access to nation-wide HMIS data and quality of these data have been shown to have greatly improved in sub-Saharan Africa [185, 186]. As such, the WHO has reiterated that timely and high-quality HMIS-based burden estimates are achievable, and can be used to inform on-going decision making [4]. Despite this, HMIS remains underutilized, especially for risk mapping, due to concerns over incompleteness and delayed reporting [71, 187, 188]. Whilst HMIS has had, and still needs, further improvement, substantial discrepancies between estimates of burden from the current prevalence-to-incidence model approach and HMIS reports have been identified among at least 30 high burden countries [71]. Thus, questions remain as to the reliability of HMIS estimates and their corresponding representation of fine-scale spatial distribution of risk to support evidence-based decision making by country-level programme managers.

Small area space-time disease models fitted to routinely reported data have been widely implemented to accurately identify contextually important risk factors and unpack spatial temporal patterns of infectious diseases, including tuberculosis and malaria [156, 189-192]. These models have the capacity to explain the spatial autocorrelation in disease data, and can provide robust means of understanding ecological connectivity and relationships [193] that are critical for control processes in high malaria or other disease burden countries. Moreover, foci of high malaria risk or burden are pertinent to the principle of strategic information to drive impact under the global “high burden to high impact” initiative, for effective targeting of interventions [160]. This study, therefore, aims to investigate a pragmatic novel small-area space-time approach using a nationwide network of health facilities in estimating malaria incidence from HMIS data, in order to identify areas of high malaria burden and risk across Uganda and assess malaria seasonality.

Methods

Study setting

Uganda was estimated to be the 3rd highest contributor of *Plasmodium falciparum* malaria cases globally in 2018, with incidence rates of >250 cases per 1000 population at risk within a perennial transmission setting [3]. Located between -1^0 and 4^0 latitudes, it covers a total area of ≈241,500 square kilometres that was divided into 15 non-administrative regions (comprised of between one to 13 districts each) considered to be the malaria endemicity zones under the Uganda Demographic and Health Survey (UDHS) Program by 2018 [61]. Nested within these regions were 128 districts (as they were known in 2018), representing the second administrative level of government.

Data and population

Health management information systems data: In Uganda, all health facilities are required to submit monthly reports from their out-patients department (OPD) registers on all reported diseases to the Department of Health information of the Ministry of Health (MoH). Health facilities are either private-for-profit (PFP) or public comprised of the government owned and private-not-for-profit (PNFP) facilities. HMIS was introduced in 1997 as a paper-based reporting system from each health facility to the Ministry of Health. In 2012, however, a web-based reporting version, the DHIS-2, was implemented with full roll-out across the country in 2013 [14]. In this system, health facility data is either entered directly among high-level facilities or sent as paper reports from lower-level facilities to the districts for entry into the online system.

For this study, HMIS data consisted of monthly counts of all reported and confirmed malaria cases from study facilities defined here as reporting facilities with available geo-coordinates. Reported malaria cases are defined as all cases reported regardless of confirmation status while confirmed malaria are laboratory confirmed cases using either blood slide microscopy (B/S) or rapid diagnostic test for malaria (RDT) – per national guidelines. Whereas the recruited reporting facilities with available geo-coordinates represented 3453/7029 (49.1%) of all facilities included within the DHIS-2, 2656/7029 (37.8%) neither reported nor were geolocated and were therefore not recruited (Fig. S1, Additional file 1). Whilst majority of reporting geolocated facilities were publicly owned, the majority of non-geolocated health facilities were private for profit (PFP) commonly located in urban areas and these were excluded. Notably, the two districts of Kampala and Wakiso that together formerly comprised the capital city, contributed 49% of these excluded facilities. All reporting facilities that were not geolocated or geolocated facilities without a matching reporting health facility were excluded from this study. A total of 3446 geo-located health facilities constituted the study facilities for this work (Fig. 1).

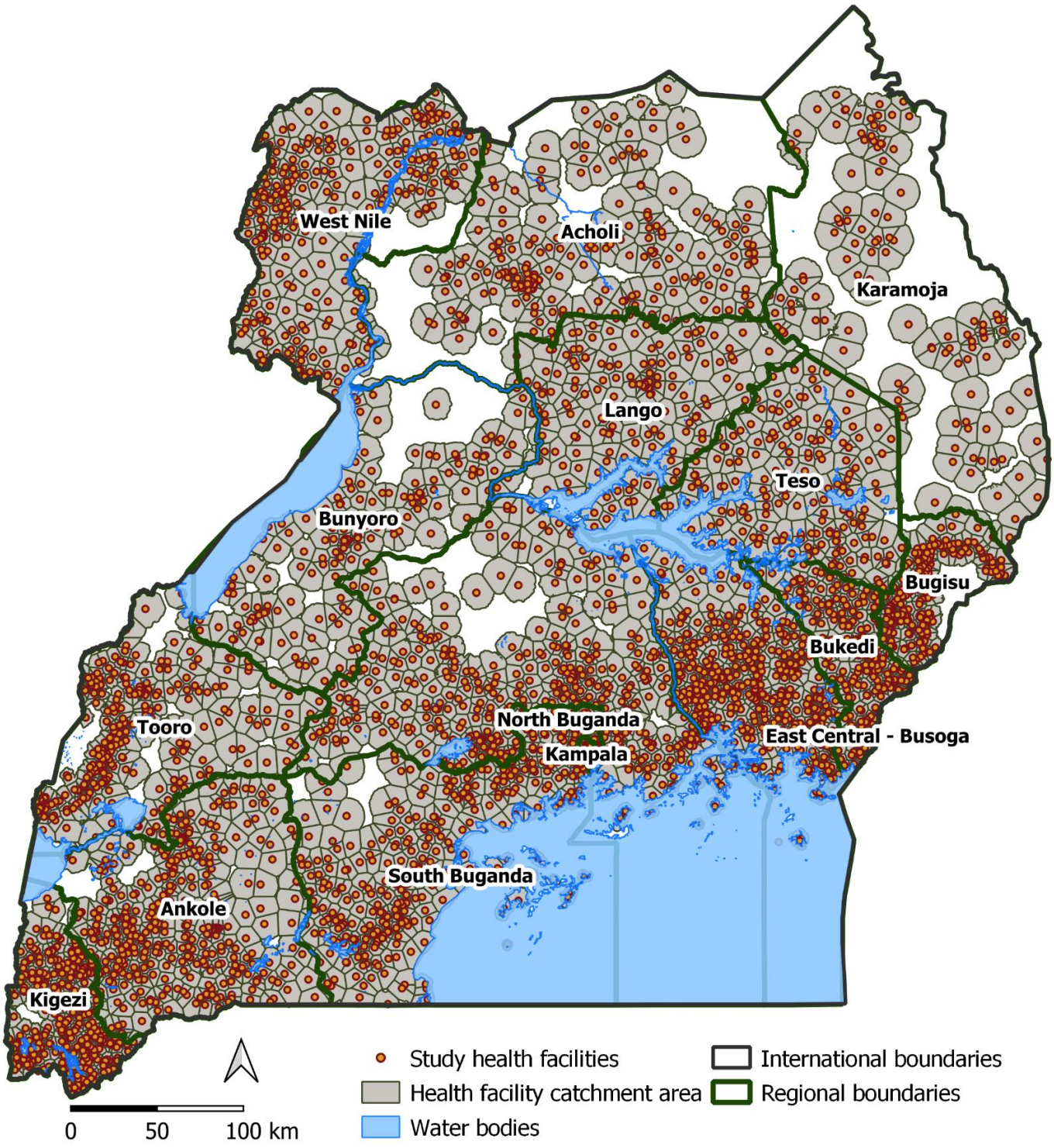


Fig 1. Map of Uganda showing locations of study health facilities within their defined catchment areas

Ancillary data: To define accessibility to health facilities, four categories of single timepoint ancillary data were incorporated to develop a raster surface within which each pixel was assigned a time cost of travel across it and is herein referred to as a cost-distance surface (Table 1). First, a digital elevation model (DEM) provided a measure of penalty on travel speed depending on direction of travel along the elevation. Second, a land use and land cover

raster data set from 2016 was used to define diversity of land cover across which, travel speed would be affected. Third, wetlands, lakes, and rivers were identified as barriers for travel. Lastly, road networks were incorporated categorized by feasible travel speed class.

Table 5. Description of ancillary data sets and the sources of these covariates

Data set	Data type	Data source
Single time point data sets		
National geo-located health facilities	Vector	https://figshare.com/articles/Public_health_facilities_in_sub_Saharan_Africa/7725374 Accessed September-2019.
Digital elevation model	Gridded raster	https://www.rcmrd.org/ Accessed October-2019.
Land use and land cover	Gridded raster	http://geoportal.rcmrd.org/layers/servir%3Auganda_sentinel2_lulc2016 Accessed October-2019.
National wetlands	Vector	http://maps.nema.go.ug/layers/geonode%3Augandawetlands2008 Accessed September-2019.
Lakes and rivers	Vector	https://geodata.lib.berkeley.edu/catalog/stanford-fh022bz4757 Accessed September-2019.
Road network	Vector	http://cod.humanitarianresponse.info/sites/default/files/uganda_roads_feb2009.zip Accessed September-2019 and from KEMRI.
Multi-time point data sets		
Land surface temperature	Gridded raster	https://earlywarning.usgs.gov/fews/ewx/index.html?region=af Accessed October-2019.
Normalized difference vegetation index (NDVI)	Gridded raster	
Rainfall	Gridded raster	https://www.tamsat.org.uk/data/archive Accessed September-2019.
Night-light emissivity	Gridded raster	https://earthobservatory.nasa.gov/features/NightLights Accessed November-2019.
Mean years of education for women of childbearing age over 2000-2015	Gridded raster	http://ghdx.healthdata.org/record/africa-educational-attainment-geospatial-estimates-2000-2015 Accessed November-2019.

To generate predicted incidence rates accounting for spatially variable risk factors, ancillary data sets at multi-time points were considered and utilized (Table 1). Notably, whilst vegetation quantities (NDVI) were quantified as the first ten days (dekad) per month and rainfall as monthly estimates, monthly night-light emissivity was projected using 2012 and 2016 data sets, and the mean number of years in of attending school among childbearing women published in [194] as a single estimate.

Health facility catchments: Currently, the HMIS is used to report malaria burden down to the district level, limiting the ability to observe and act upon heterogeneity at finer spatial scales. In part, this is because of limited information on health facility catchments. Considering proximity as the most important determinant of health facility access and utility [138, 141], health facility catchments were defined based on a cost-distance surface generated using a WHO supported tool known as AccessMod [195] as described in (Section E, Additional file 1). This tool has been widely used in assessments for general and emergency care accessibility, and estimation of care utilization for febrile illnesses, among others [196-198].

Using the cost-distance surface generated based on anisotropic (direction dependent) analysis with direction of travel considered as 'towards the health facility' in the geographic accessibility model, three-hour travel catchment buffers were generated for each health facility included in the study. To delineate each facility's catchment area, the intersection polygon between the three-hour travel buffer and a Thiessen polygon around each health facility, generated using ESRI ArcGIS 10.5 *Thiessen polygon* tool (ESRI 1995-2016; Redlands, CA, USA), was derived. This intersection polygon constituted the catchment area for each health facility covering majority of the country.

Population data: Population estimates for the country were obtained from gridded population surfaces generated by the WorldPop project whose estimates are based on national census estimates and other factors, accessible from www.worldpop.org. Annual gridded population surfaces were obtained for the duration between 2014 and 2019 and population estimates per year extracted as summary statistics for each calendar year of the study duration 2015 to 2019. These estimates were extracted using ESRI ArcGIS 10.5 *Zonal Statistics* tool at the level of the defined catchment area for each study health facility, regardless of administrative boundaries, given that care seeking is not restricted by these boundaries in Uganda.

Spatial, temporal, and spatial temporal analyses

The primary outcome in this analysis was monthly cumulative malaria incidence rate, derived from HMIS data as the number of new confirmed cases per facility catchment divided by the total population of the catchment per month.

Inherent spatial correlation of malaria infections is unexplained within classical regression approaches though remains in the residuals and induces spatial autocorrelation in the response even after known available risk

factors are accounted for [199]. Using spatial conditional autoregressive models, however, explains this autocorrelation in the outcome using random effects within a Bayesian framework that uses prior distribution, maximum likelihood, and neighbourhood to predict a more reliable outcome [200, 201].

A Beyesian space-time model employed here consisted of three segments, including: a data model that accounts for data distribution; process model that accounts for spatial structure and trends; and, the parameter model that accounts for prior distribution estimated and utilised [202].

The data model for incidence data in this study assumed a Poisson distribution and was defined by

$$Y_{it} \sim Pois(E_{it}\mu_{it}),$$

Where, Y_{it} was the incidence at time t for area i whose expected incidence was E_{it} and relative risk μ_{it}. This outcome variable Y_{it} is assumed to be conditionally independent across the spatial process defined by health facility catchments in this study.

With the process model, a spatial temporal fit of a BYM (Besag, York and Mollie) conditional autoregressive model with two random effects was implemented using integrated nested Laplace approximation (INLA) (www.r-inla.org), fit to the monthly crude confirmed case rates in R (code presented in Appendix 10) [203]. Random effects included a structured and an unstructured spatial effect, as well as a structured and an unstructured temporal effect. This spatial convolution model takes the form,

$$Log(\mu_i) = \beta_1 + \beta_2 x_{i2} + \ldots + \beta_p x_{ip} + s_i + u_i$$

Where s_i is spatially structured and modelled using an inverse gamma process to enable smoothing among neighbouring locations and u_i is spatially unstructured and modelled using a Gaussian process to allow for increased heterogeneity due to included covariates/risk factors [202, 204]. Within the parameter model, structured random effects were assigned an inverse Gamma prior, while unstructured random effects were assigned a Gaussian prior under the assumption of a normal distribution of noise. β_1 denotes the overall risk represented as a fixed intercept. x's are explanatory spatial covariates including rainfall, land-surface temperature, night-time light – proxy for social economic status, level of education for women of child-bearing age – proxy for treatment seeking behaviour, and … β_2 to β_p are regression coefficients estimated to be constant across catchment areas for fixed effects [205].

In this study, posterior estimates of incidence rates and the crude incidence rates were shown to be correlated (Fig. S6 and Fig. S7, Additional file 1)). Moreover, to avoid overfitting, a time restriction using a random walk of the first order was included.

Candidate covariates had been used in other studies, given their association with malaria transmission including rainfall, temperature, vegetation index, night-time lights (proxy for urbanicity), and caregiver education [191, 207-209]. For inclusion in the final model, covariates quantities were evaluated for impact on a linear regression model of crude incidence rates using Akaike's information criteria values (Table S2, Additional file 1). The final covariate list included catchments estimates of: mean years of education for women of childbearing age, mean of current and three months' lags for both rainfall and land surface temperature estimates, mean monthly night-time light emissivity, and mean of current and one month's lag of vegetation amounts. All these were significantly associated with crude incidence estimates (Table S3, Additional file 1). Both β and b were assigned monthly informative Gaussian distributions over the full 51 months length of the study duration. The full model was validated by withholding 20% of data points at random and comparing the model predicted values with the actual observed values using scatter plots and spearman's correlation coefficients (Section G, Additional file 1).

Relative risk of malaria at district and health facility catchment levels was derived as the respective predicted incidence rate divided by the overall predicted mean incidence rate at national level per month in the study duration. All maps of the posterior estimates of incidence rates and relative risk of malaria were generated using R.

Spatial clustering in the modelled outcome was further investigated using the global Moran's Index statistic within the spatial dependence (spdep) package of R and visually examined Moran's scatter plots of incidence and risk estimates at both district and health facility catchment resolutions. To identify cluster locations, the local Moran's Index using ESRI ArcGIS 10.5 *Cluster and Outlier Analysis* (*Anselin Local Moran's I)* tool was used, set for first order queen contiguity, running 999 permutations and clusters evaluated at 0.01 level of significance.

Also, study model estimates of confirmed malaria cases were compared with estimates from both the WHO's recent reports [3, 6, 71] and Malaria Atlas Project (MAP) estimates for the same period from https://malariaatlas.org/trends/country/UGA (Section J, Additional file 1) and relationship between MIS regional estimates of prevalence of malaria in children under five years [61] and estimated relative risk of malaria are regional-level, examined using visual inspection of scatter-plots with results presented in supplementary information (Fig. S13, Additional file 1).

Results

Study population

The total population identified within the health facility catchment, considered at risk of malaria infection and likely to seek care from the associated geo-located publicly reporting health facility, were considered the study population of interest. The total population was estimated at 34.9 and 39.6 million in 2015 and 2019 respectively, with the ≈2.8% located outside of the defined catchments (Section D, Additional file 1).

HMIS data summary

Between 62.2 and 88.7% of nationally reported cases of malaria annually were diagnostically confirmed cases in 2015 and 2019, respectively (Fig. S2, Additional file 1). Whilst these proportions increased across the 15 regions of the country over time, Kampala recorded marginal improvements. Moreover, the majority of confirmed malaria cases in Kampala (ranging from 61.8 to 81.0% in 2015 and 2018) were unaccounted for due to exclusion of facilities, leaving only up to 38% of the burden in this metropolitan district estimated (Table S1, Additional file 1). Excluding Kampala, however, results showed that estimates accounted for between 67 to 96% of the routine HMIS-based burden of malaria among the remaining 14 regions, over the study duration. Moreover, in these regions, average annual proportion of reported confirmed cases excluded from the study ranged from 5.3 to 19.8% in Karamoja and Tooro, respectively. Diagnostic testing of suspected malaria cases across the country was conducted either by microscopy or rapid diagnostic tests and reported as a single total.

Mean incidence rates, seasonality, and risk of malaria

Highest burden regions and districts also hosted health facilities with the highest number of confirmed malaria cases reported. For instance, Bala health centre (HC) III in Kole district of the Lango region reported 3,317 cases during November 2015, while Bira HCII in Adjumani district of the West Nile region reported 6,697 cases during June 2016. Moreover, Barakala HCIII (highest for two consecutive years) also from West Nile in Yumbe district, reported 9,654 cases during October 2017 and 9,246 cases during July 2018. Lastly, Matany hospital in Napak district of Karamoja region reported 8,089 confirmed cases during September 2019.

This study showed spatial and temporal variation in incidence rates between regions and districts in any given region, as well as between health facility catchments within districts, both during the low (Fig. S9, Additional file 1) and high burden seasons (Fig. 2).

Fig 2. Spatial distribution of malaria incidence rates during high burden months of study duration

Maps in column A show the regional boundaries (regional stratification of malaria in Uganda per 2018 MIS), Column B show district boundaries (the second government administrative level) and column C maps show study defined health facility catchment area boundaries for the study health facilities.

National incidence rates: The model estimated 38.8 (95% CI: 37.9 – 40.9) million confirmed malaria cases over the study period of July, 2015 to September, 2019, highest in 2016 with 10.3 (95% CI: 9.9 – 10.7) million cases and lowest in 2018 with 6.5 (95% CI: 6.4 – 6.9) million cases among complete calendar years (Table S4, Additional file

1). Annual incidence rates reduced from 281.7 (95% CI: 274.9 – 296.7) in 2016 to 170.0 (95% CI: 165.9 – 178.8) cases per 1000 in 2018.

Monthly incidence rates showed a general declining trend in the burden of malaria from 2015 to 2019, strongest through 2018 followed by an increase in 2019 (Fig. 3). In all the years of the study, the incidence rates consistently peaked in June and July, reaching a maximum of 36.6 (95% CI: 35.7 – 38.5) cases per 1000 in June 2017 (Table S5, Additional file 1). Conversely, low risk periods were less consistent, although often lowest in February and March, reaching a minimum of 8.9 (95% CI: 8.7 – 9.4) in February 2018.

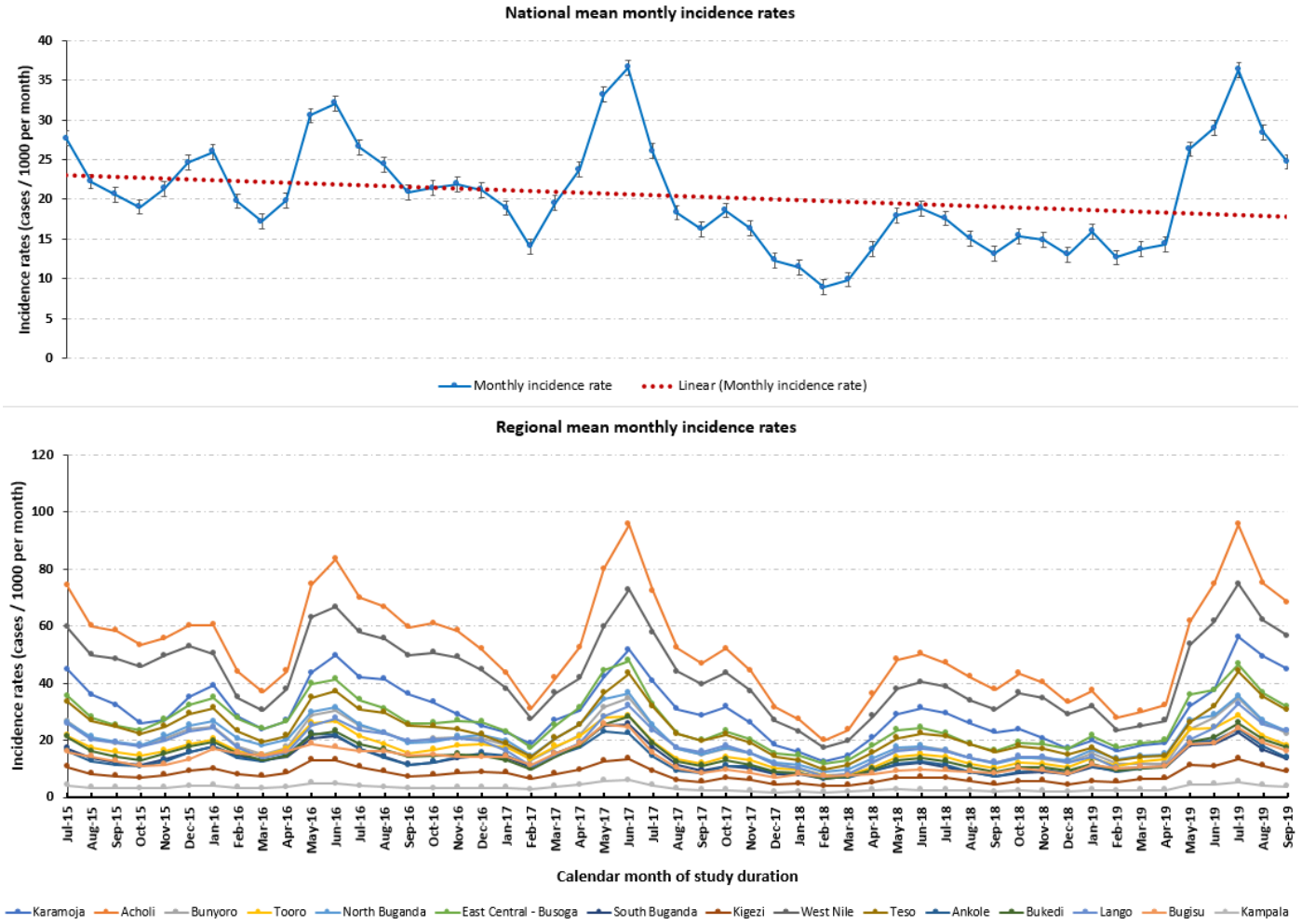


Fig 3. Trends in the national and regional monthly malaria incidence rates between July, 2015 – September, 2019

Spatial distribution of incidence rates across the country: Overall, mean monthly regional incidence rates were highest in Acholi region (Northern Uganda) at 52.3 (95% CI: 50.3 – 59.6) cases per 1000 per month and lowest in Kigezi region (South Western Uganda) at 7.9 (95% CI: 7.6 – 8.2) cases per 1000 per month (besides Kampala).

Consistent with national trend assessments, monthly trends in regional incidence rates showed the highest peaks in June-July, highest in June, 2017 (Range: 13.4 – 95.6 cases per 1000) and July, 2019 (Range: 13.5 – 95.5 cases per 1000 in Kigezi and Acholi, respectively) and the lowest troughs in February-March of each calendar year (Fig. 3).

These trends showed that Acholi, West Nile, Karamoja, East Central – Busoga, and Teso persistently recorded the highest monthly incidence rates across the entire study duration. Moreover, the greatest variability in incidence rates was also observed among these five highest burden regions of with respective estimated mean monthly incidence rates of 52.3 (SD: 17.8), 43.3 (13.9), 30.3 (10.4), 26.3 (8.6), and 23.5 (8.0) cases per 1000 per month.

Within these regions, high burden and risk districts were also identified, both during the highest and lowest burden months. During June 2017 district monthly incidence reached the maximum in Lamwo of Acholi, Moyo of West Nile, Kaabong of Karamoja, Namayingo of East Central - Busoga, and Katakwi of Teso regions, at 167.6 (95% CI: 165.6 – 169.8), 192.5 (95% CI: 189.9 – 195.1), 81.1 (95% CI: 79.6 – 82.5), 73.1 (95% CI: 71.9 – 75.0), 72.0 (95% CI: 70.9 – 73.1), cases per 1000 per month, respectively (Table S6, Additional file 1).

Monthly incidence rate trends among districts showed that Moyo, Lamwo, Adjumani, Pader, Nwoya, and Maracha persistently recorded the highest monthly incidence rates across the study duration (Fig. 3). Moreover, higher incidence rates were also associated with higher variability in monthly incidence rates with the mean monthly estimate in Moyo at 115.8 (SD: 36.5) and lower rates less variability with Rubanda at 1.6 (SD: 0.5) cases per 1000 (Figs. S10 and S11, Additional file 1).

Within individual districts, a wide distribution of incidence rates was estimated among health facility catchments both during the lowest and highest burden months. From the 3446 catchment areas identified across the country, mean monthly incidence rate reached a maximum of 569.8 (95% CI: 555.2 – 584.3) cases per 1000 per month in Namayingo district of East Central – Busoga region and minimum of 0.13 (95% CI: 0.10 – 0.17) cases per 1000 per month in Rukungiri district of Kigezi region, excluding Kampala. Also, higher incidence rates within catchments were associated with higher viability in monthly incidence rates and lower incidence rates with less variability (Fig. S10, Additional file 1). Among health facility catchments, variability in incidence rates reached a maximum standard deviation (SD)= 142.4 cases per 1000 in highest incidence rate catchment located in Namayingo and a minimum SD= 0.1 among the lowest burden catchments in Arua and Kasese districts.

Spatial distribution of relative risk across the country: Consistent with incidence rates, relative risk of malaria was highest among the highest burden regions of Acholi, West Nile, Karamoja, East Central – Busoga, and Teso, both during the lowest and highest burden months, maintaining their rank of risk at both times (Table S7, Additional file 1). During the highest burden month of June 2017, the relative risk of malaria among these regions ranged from 1.18 (95% CI: 1.17 – 1.19) to 2.6 (95% CI: 2.6 – 2.8)-times higher than national average in Teso and Acholi, respectively. Moreover, while mean relative risk among districts within these regions was higher during the highest burden month at 1.8 (95% Confidence Interval:1.5 – 2.1) than the lowest at 1.7 (95% Conf. I:1.4 – 2.0), the difference was not significant (p= 0.676) by a two-sample t-test.

Spatial and temporal variation in relative risk observed between regions, and districts within regions (largely informative at programmatic or NMCP levels), was also present between catchments within districts (informative for district health managers). Relative risk remained consistent among the 15 regions, between low and high burden seasons, but showed additional variability among districts and health facility catchments across the two seasons (Fig. 4).

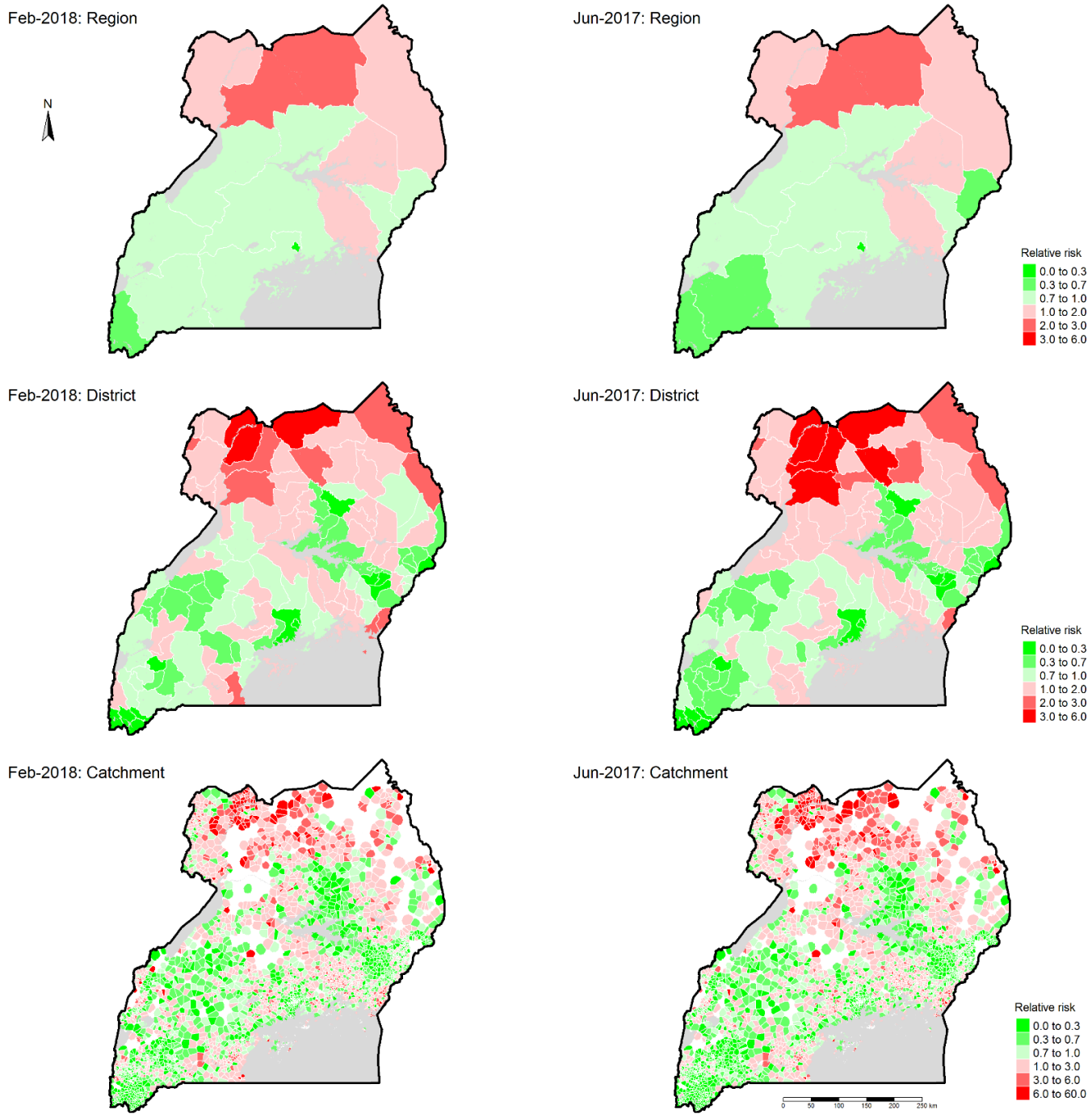


Fig 4. Spatial distribution of the relative risk of malaria during lowest and highest burden months of the study duration

Results showed that catchment risk ranged from 0 to 24.9 (95% CI: 24.4 – 24.9) times higher than national average during the highest burden month and from 0 to 50.5 (95% CI: 49.0 – 50.8) during the lowest burden month. Moreover, a non-linear association of catchment risk was observed between the lowest and highest burden months further confirming this rising risk during lower burden months (Fig. S16, Additional file 1). However, the highest risk catchments at the two time points were neither identical nor located in the same district or region.

Spatial clustering of risk: Assessment for spatial autocorrelation of incidence and/or risk showed consistent levels of moderate global autocorrelation between both districts (Moran's I range by month: 0.4 to 0.6, $p<0.001$) and health facility catchments (0.3 to 0.5, $p<0.001$). Both during the highest (June-2017) and lowest (February-2018) burden months, global autocorrelation between districts was very similar (Moran's I = 0.5, $p<0.001$) (Figs. 18 and 19, Additional file 1) but slight difference between health facility catchments (Moran's I = 0.4 and 0.3, $p<0.001$, respectively) (Figs. S20 and S21, Additional file 1).

Analysis of local spatial autocorrelation at two levels of significance ($p<=0.05$ and $p<=0.01$) identified substantial significant high-high clustering in Acholi and West Nile regions in the North, as well as East Central – Busoga region in the South East of the country, both during the highest and lowest burden seasons (Fig. 5). Similarly, large low-low clustering was identified in the Southern regions of the country. Moreover, outlier catchments typically had significantly lower risk than their neighbours in the north, and higher risk than their neighbours in the rest of the country. Significant monthly high-high clusters were comprised of between 191 health facility catchments during February 2018 and 236 during June 2017 and 2019 (Fig. S22, Additional file 1).

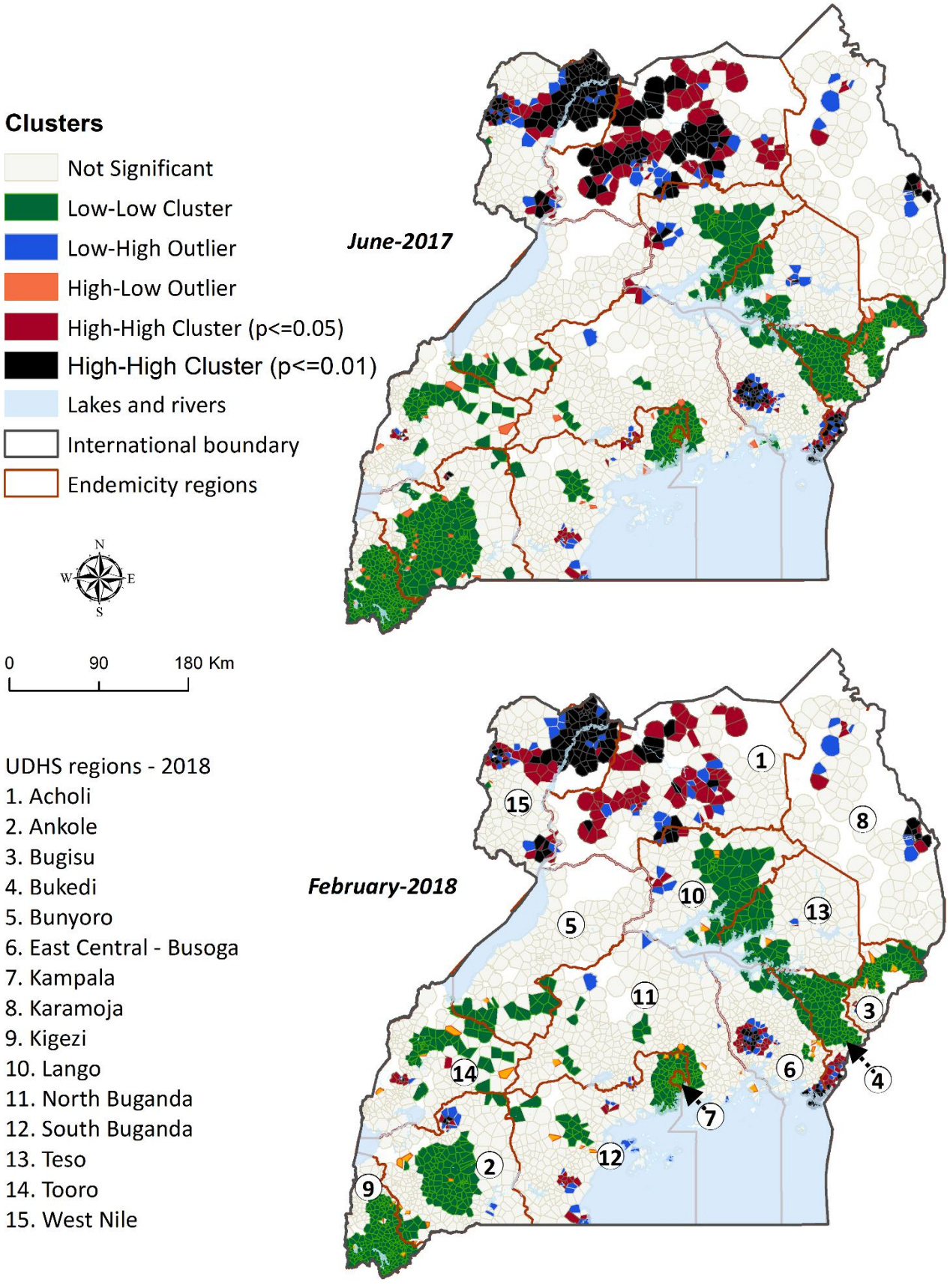


Fig 5. Spatially significant clusters of malaria risk for the highest and lowest burden months between 2015 and 2019, across Uganda

Discussion

Results from this innovative, large-scale, longitudinal observational study suggest that with improved HMIS reporting, credible high-risk areas at both high and low spatial scales are identifiable. The study revealed distinct monthly spatial distribution of malaria incidence across the fifteen regions of Uganda, in a concurrent multi-resolution assessment including coarse (regional) down to fine (health facility catchment) spatial resolutions. Moreover, whilst Uganda is considered a perennial transmission setting, this study revealed a nation-wide seasonal pattern in incidence rates with two peaks (major and minor), the highest during June-July and the minor peak during October. This approach may facilitate efficient implementation and optimization of targeted control activities that can leverage existing health facility systems [210]. It may also improve managers' understanding of the heterogeneity and/or clustering of malaria burden within districts that currently form the lowest level of malaria burden assessments, though acknowledged as difficult to use or unusable for planning control [64].

This study showed that the risk of malaria by regional rank among the highest and lowest risk regions had minimal temporal variability, with these regions maintaining their status both during low and high burden seasons. These findings were consistent with extant UDHS regional stratification of Uganda where Acholi, West Nile, and Karamoja are among the highest transmission regions, and Ankole and Kigezi among the lowest. This stratification supports tailored approaches for long-term malaria control efforts aiming at elimination, as advocated in the global 'high burden to high impact' initiative [17] that was recently adopted as central to onward national malaria control strategies for Uganda [38]. Whilst targeted interventions including IRS [211] and larval source management [37] have been used, further emphasis is necessary [81, 160] with implementation taking greater account of local context. Importantly however, temporal variability of risk among many regions highlights the continued vital role of routine surveillance for planning and timely action towards control. Moreover, higher risk among high burden locations during the lowest than highest burden seasons suggests persistent high-risk in these locations, the identification of which could facilitate high precision targeted actions for effective control.

This study also identified several distinct clusters of high-risk health facility catchments, which were consistent over time though largest during the highest burden seasons and smallest at the lowest. The largest high-risk clusters were concentrated in the West Nile and Acholi regions in Northern Uganda, although smaller clusters were noted in the recognised high transmission regions of Karamoja and East-Central Busoga [61]. Conversely, the most notable low-risk health facility catchment clusters could be grouped into three categories: highland regions (e.g. Kigezi, Ankole and Bugisu) [152, 212]; regions with recent intense targeted multi-year IRS activity associated with high impacts on transmission (e.g. Bukedi, Teso, and Lango) [60, 61, 213, 214]; and, large urban municipalities (e.g. Southern Buganda) with urbanization associated with reduced transmission [215, 216]. These findings provide further evidence of identifiable candidate locations for targeted control interventions among the high-risk clusters and an approach for assessment of possible impacts of previous interventions.

Trends in annual confirmed malaria cases in Uganda declined between 2016 and 2018, despite increased reporting and proportions of confirmed cases over time, consistent with MIS findings between 2014 and 2018 [60, 61], before a sharp increase in 2019. Moreover, the relationship between regional relative risk and prevalence of malaria (among children under five years of age from the 2018 MIS) showed that small changes in parasite prevalence were associated with sharp increases in relative risk among regions at lower than national average risk. However, large changes in parasite prevalence were associated with small changes in relative risk among regions at higher than national average risk. This further confirms the variability of risk among many regions while pointing to strong effects of age on malaria [217]. In addition to estimated confirmed cases being lower than estimates reported by WHO and MAP per year (possibly due to study design of excluding some facilities), trends were dissimilar with WHO and MAP cases increasing between 2016 and 2017 [3], unlike in this present study. Nevertheless, such dissimilarities have been documented [71] and likely explained by the use in global assessment for sub-Saharan Africa of prevalence surveys that are predominantly conducted among children [82]. With estimates for the whole population generated from these surveys, despite shifts in malaria burden from children to the older population following effective control interventions [217], the dynamic effects on burden may not be adequately accounted for in the prevalence-to-incidence models used.

The observed seasonality with June-July peaks and February-March troughs was consistent with reports from south western Uganda, where epidemics followed a regular July pattern except during El-nino in 1998 [64, 218] and in Gulu district (Northern Uganda) where between 2006 and 2015 biannual peaks of malaria were reported during June-July and October-November [219]. One study however, reported two peaks of malaria during April-May and September-November in Northern Uganda following the rain seasons, though unsubstantiated [220]. Findings from this present study may inform optimal timing for control activities including IRS, mass drug administration (MDA), or community mobilization campaigns towards increased malaria risk awareness for control vigilance.

PFP facilities, a small majority of which do not report to the HMIS and were therefore excluded from this study, limit the utility of focal analyses such as presented here. This highlights an important missed surveillance opportunity. The limited capacity to detect outbreaks in settings largely served by PFP may exacerbate the severity of malaria outcomes among their most vulnerable residents with increased case management costs [221]. There are several possible initiatives to increase reporting in these facilities where a small majority seek care for febrile illnesses [59-61]. First, provision of guarantees on exclusive use of data for public health not revenue monitoring, may improve confidence and alleviate any fears of punitive intensions in their reporting. Second, ensured availability of standardized reporting tools, may offset running costs of stationery in the private facilities while it enables improved documentation of health records. Third, training of PFP managers and owners on the benefits

of surveillance and/or reporting. Lastly, implementation of regular feedback mechanisms may provide a means of continued evaluation that fosters risk and other assessments that are mutually beneficial.

Given that policymakers' remediating responses as well as policy formulation processes are informed by pooled information from diverse sources, including but not limited to research, political, and funding provisions, it is unrealistic to expect these technocrats to be expert generators of the evidence from these multi-disciplinary sources. Whilst there are no simple solutions to the implementation of analyses such as in this present study, interpretation of contemporary outputs is nowhere nearly as demanding, highlighting the criticality of partnerships between policy and research dimensions for malaria and other disease control efforts.

This study had limitations. First, the disproportionately low proportion of geolocated reporting private facilities impacted on the estimates of malaria burden, especially among highly urban locations including Kampala and Wakiso districts and others across the country. Results for the Kampala region (and Wakiso district) in this study, represent only a small proportion of the burden and were excluded from results discussions. Moreover, exclusion of non-geolocated reporting public health facilities (such as in Kitgum district), impacted on the estimates of incidence due to unidentified catchments in those places. Nevertheless, there was wide coverage of health facilities across the country with a small proportion of districts under-represented, minimizing effects of this constraint. Second, the study did not account for level of health facility and other population level factors that impact on differential health seeking behaviour, which may have inflated incidence rates and risk where a given level or type of facility is preferred. However, in this analysis it was assumed that for uncomplicated malaria, people attend the closest health facility and some important factors such as urbanicity and primary care giver education were accounted for, though further research may be required to better understand impacts of level of health facility on care seeking for uncomplicated malaria. Third, the study did not account for stock levels of antimalarials or test kits, variations of which may impact on the number of cases recorded between seasons of full stock versus stockouts. A better understanding of the linkage between logistics management and HMIS may be required, given known associations between stockouts and increased under-five mortality or compromised treatment practices like dosage rationing and use of less effective remedies [222]. Fourth, given that health facility recruitment into the study was not dynamic, any increase in number of facilities reporting could have had impacts on study findings. Moreover, the systematic exclusion of non-geolocated facilities, may have biased study results towards more long-term established than newer health facilities, but duration of facility existence was beyond the scope of this study.

Conclusion

Assessment of malaria burden and/or risk in high burden countries using routine surveillance data is highly achievable. Using national routine data, this study provided needed evidence of vital concurrent assessment of

malaria risk and burden among regions, districts, and health facility catchments with identifiable significant spatial clustering of risk. Targeting hotspots as an intervention approach has been shown to yield modest and transient impacts on malaria prevalence [223]. However, locations with persistently high-risk of malaria that are potential candidates for health facility-based interventions such as community outreaches, provision of LLINs, mass drug administration and enhanced case management were identified, an approach that may be beneficial beyond isolated health facility catchments. Furthermore, whilst extensive geo-spatial analytical output with scales either too large (region or district) or too fine (pixel or neighbourhood) may be challenging for control programmes to use [224], this study provides evidence of HMIS-based assessments at practical scales for districts to implement and assess intervention impacts. Moreover, in perennial settings, the identifiable strong seasonal patterns as seen with June-July highest peaks and February-March lowest troughs in Uganda, provide vital information for intervention timing. Taken together, these results show the potential in routine HMIS surveillance data for pragmatic timely identification of high-risk areas and with further research assessments for optimal implementation of targeted control activities and their impacts.

List of Abbreviations

ACTs – Artemisinin-based combination therapies

AL – Artemether-lumefantrine

B/S – Blood slide microscopy

BYM – Besag, York and Mollie modelling approach

CI – Credible Interval

DEM – Digital elevation model

DHIS-2 – District health information system – version 2

DHS – Demographic health survey

HMIS – Health management information system

INLA – Integrated Nested Laplace Approximation

IRS – Indoor residual spraying with insecticide

KEMRI – Kenya Medical Research Institute

LLIN – Long-lasting insecticidal net

MAP – Malaria Atlas Project

MIS – Malaria indicator survey

MoH – Uganda Ministry of Health

NDVI – Normalized Differences Vegetation index

NMCP – National Malaria Control Programme

OPD – Outpatient department

PFP – Private for profit

PNFP – Private not for profit

RDT – Rapid diagnostic test for malaria

UBOS – Uganda bureau of statistics

UTM – Universal Transverse Mercator coordinate system

WGS – World Geodetic System of modelling earth

WHO – World Health Organization

Declarations

Acknowledgements

We would like to thank Katonda Omukulu for his support for this project. We would also like to acknowledge Deborah Ekusai, Rhoda Namubiru, Faith Bagatya, Dr. Charles Olara, Dr. Joaniter Nankabirwa and the administration and staff at the National Malaria Control Division in charge of the national HMIS and the DHIS-2 reporting for their contributions, as well as the administration and staff at the Infectious Diseases Research Collaboration and the Makerere University College of health sciences for their support.

Funding statement

Funding for the research reported in this publication was supported by the Fogarty International Center of the National Institutes of Health under Award Numbers (D43TW7375) and (D43TW010526). The content is solely the responsibility of the authors and does not necessarily represent the official view of the National Institutes of Health.

Authors' contributions

SPK and RP conceived of the study. SPK led the data collation with support from RK, JC, CMS, DR, and JO. SPK led the data analysis with support from BS and RP. SPK led the manuscript drafting with support from TB, AY, AG, BS, and RP. SPK and BS had full access to all the data in this study, but all authors reviewed the manuscript and gave permission for publication. SPK, the corresponding author, had final responsibility for the decision to submit for publication.

Ethics approval

This study was approved by the Uganda National Council for Science and Technology (UNCST Ref SS 4455), the Uganda Ministry of Health (ADM.386/01), the Makerere University School of Medicine Research and Ethics Committee (REC Ref. 2017-119), and the London School of Hygiene & Tropical Medicine Ethics Committee (LSHTM Ref. 13902 - 1)

Written informed consent of participants included in this study was not applicable.

Consent to participate

Not applicable

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Availability of data and material

The datasets used and/or analysed for this study are available from the corresponding author on reasonable request and with permission from the Department of Health Information, Uganda Ministry of Health.

Figure legend

Fig. 1 Map of Uganda showing locations of study health facilities within their defined catchment areas

The orange points are the relative geo-locations of the study health facilities recruited from across the country, each situated in a grey background representative of the exclusive catchment area for each facility. The catchment areas were constituted using a three-hour cost distance surface towards each health facility. These are overlaid with the regional boundaries (dark green) defining the 15 endemicity regions across the country.

Fig. 2 Spatial distribution of malaria incidence rates during high burden months of study duration

Columns A, B, and C represent regions, districts, and heath facility catchments respectively, while the rows correspond to the respective highest burden month of each year. The lighter the shade of colour, the lower the incidence rates within a region, district, or catchment and the darker the colour, the higher the incidence rates.

Fig. 3 National and regional trends in mean monthly malaria incidence rates July 2015 – September 2019

Trend plots of incidence rates (confirmed malaria cases / 1000) over study time (x-axis) – monthly. The top plot shows the national mean incidence rates per month (blue line) with a linear trend-line (dotted red). The bottom plot shows the trends for the 15 endemicity regions that comprise the country.

Fig. 4 Spatial distribution of the relative risk of malaria during lowest and highest burden months of the study duration

The left column shows, from top to bottom, relative risk by region, district, and health facility catchment for the lowest risk month of February 2018 while the right column shows a similar arrangement for the highest risk month of June 2017. For each row, the same levels (region, district, or health facility catchment) are side-by-side.

Green areas are locations with relative risk of malaria lower than the national average where the darker the colour the lower levels of risk below national average.

Red coloured areas are locations with relative risk of malaria higher than national average, where the darker the colour the high the risk

Fig. 5 Spatially significant clusters of malaria risk for the highest and lowest burden months between 2015 and 2019, across Uganda

The map at the top represents the distribution of significant clusters of malaria risk across the 15 regions of the country during the highest risk month of June 2017. The map at the bottom represents a similar distribution but for the lowest risk month of February 2018.

High-High Clusters: The black and dark red areas represent the clusters of high-risk health facility catchments that are spatially located next to other high-risk catchments, with significant positive spatial autocorrelation at <=0.01 and <=0.05 levels of significance, respectively.

High-Low Outliers: These orange areas represent high-risk clusters that are significantly disparate from their surrounding low-risk catchments. These outliers have significant negative spatial autocorrelation.

Low-High Outliers: These blue areas represent the low-risk clusters that are significantly disparate from their surrounding high-risk catchments. These outliers also have significant negative spatial autocorrelation.

Low-Low Clusters: These green areas represent the clusters of low-risk health facility catchments that are spatially located next to other low-risk catchments, with significant positive spatial autocorrelation.

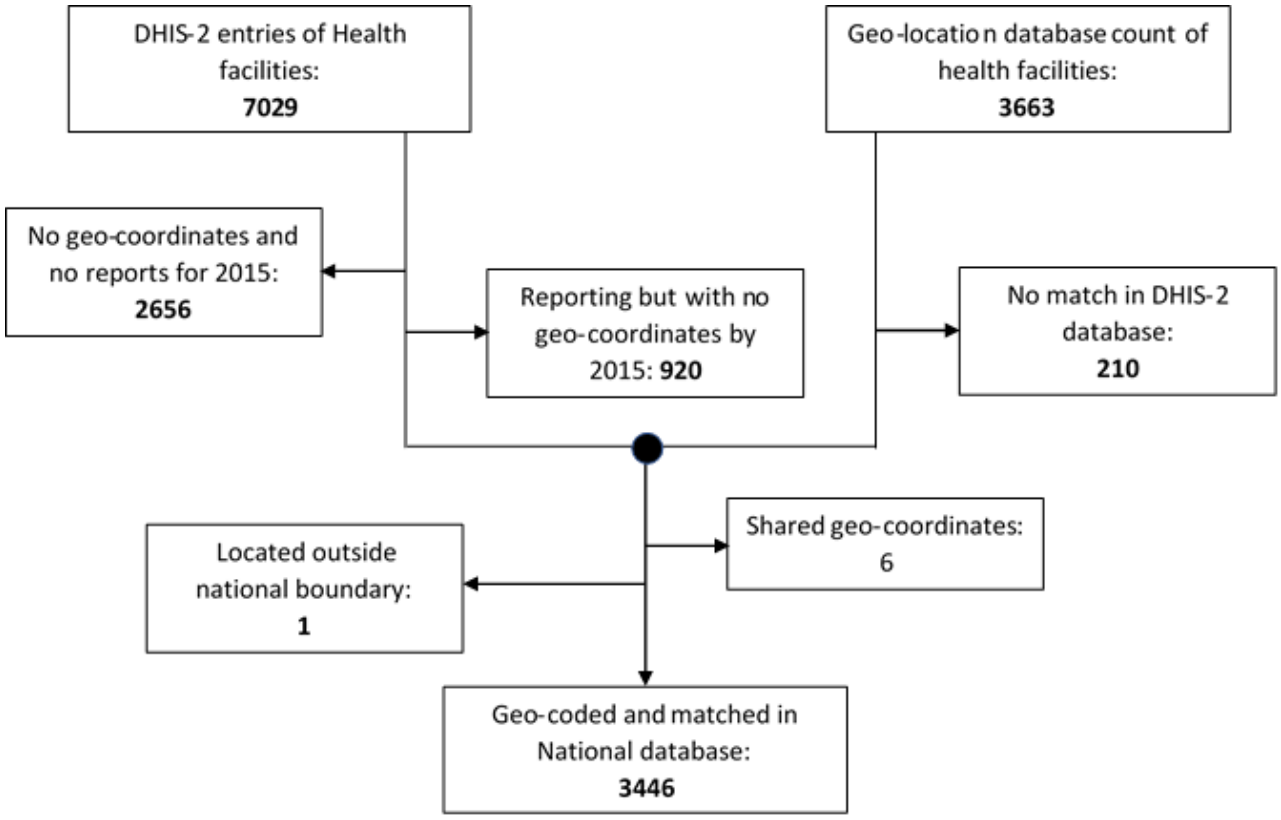
Not significant: The light grey areas represent the health facility catchments that did not show any significant spatial autocorrelation or clustering of either high, low or outlier distribution of risk of malaria. They are areas of highly random spatial distribution of risk of malaria.

6.1 Additional Information for Paper 4

6.1.1 Study health facilities' selection

Health facilities were selected for inclusion in this study primarily in the basis of presence of geo-ordinates from the publicly available database as well existence in the national routine HMIS data repository (DHIS-2) by 2015. The geo-location database contained nearly half the number of entries as the reported malaria cases DHIS-2 database given that 50.9% were without a match in the geo-location database (Fig. S1).

Fig S1. Flow-diagram of the recruitment process for the study health facilities



By 2015, there were at least 7029 health facilities (public and private) registered in the DHIS-2 database, in anticipation of them submitting monthly malaria reports. Of these, 2656 (37.8%) did not submit any reports through the year, though majority of these were private clinics that often do not comply with MoH reporting requirements (Fig. S1). A total of 3663 health facilities with associated geo-coordinates were identified by name and these comprised a health facility geolocation database. Comparing the two databases, identities of 3446, by facility name, were matched and the same constituted the study health facilities for this work.

6.1.2 Diagnostic confirmation of malaria and reporting

Whereas diagnostic testing rates varied across the country, overall proportions of the reported malaria cases confirmed by a diagnostic test increased nationally over time from 62.2 to 88.7% between 2015 and 2019 (Fig. S2). These increases were observed across 14 of the 15 endemicity regions of the country, with the three best

performing regions of Lango in Northern Uganda, Kigezi in South Western Uganda and Teso in Eastern Uganda, increasing from 70.9 to 98.1%, 58.8 to 97.4% and 45.8 to 98.1% respectively. Notably however, Kampala in central Uganda recorded declining performance, with the proportion of reported cases that were diagnostically confirmed reducing from 62.0 to 16.0% between 2015 and 2017. Notably however, majority of excluded facilities, were disproportionately concentrated in Kampala (the capital city and most urban district in the country) making it an outlier. Along with the improved diagnostic testing, national reporting rates defined as the proportion of the expected reports received within the DHIS-2 system per year recorded an 22.1% increase between 2015 and 2019 (Fig. S3).

Fig S2. Distribution of proportions of reported malaria cases that were test confirmed, by the 2018 MIS endemicity regions and nationally.

Region	2015*	2016	2017	2018	2019*
Acholi	69.1%	73.3%	74.0%	73.1%	88.0%
Ankole	46.5%	52.2%	85.5%	83.2%	92.3%
Bugisu	49.5%	57.2%	56.9%	72.1%	86.8%
Bukedi	47.4%	52.7%	60.0%	80.3%	88.7%
Bunyoro	73.3%	69.3%	73.0%	80.9%	87.2%
East Central - Busoga	64.2%	60.4%	63.6%	74.7%	86.4%
Kampala	62.0%	37.6%	17.9%	16.0%	21.1%
Karamoja	47.3%	47.4%	54.8%	67.7%	71.0%
Kigezi	60.0%	58.8%	77.3%	82.3%	97.4%
Lango	75.0%	70.9%	75.3%	83.6%	98.1%
North Buganda	72.8%	79.6%	81.3%	80.3%	90.2%
South Buganda	69.7%	68.1%	74.6%	79.8%	91.5%
Teso	52.0%	45.8%	50.4%	85.2%	98.1%
Tooro	60.2%	55.8%	72.3%	84.1%	92.9%
West Nile	65.6%	69.4%	78.5%	85.8%	88.6%
Nationally	62.2%	62.6%	70.4%	79.3%	88.7%

**Not complete calendar years, included July-December 2015 and January-September 2019*

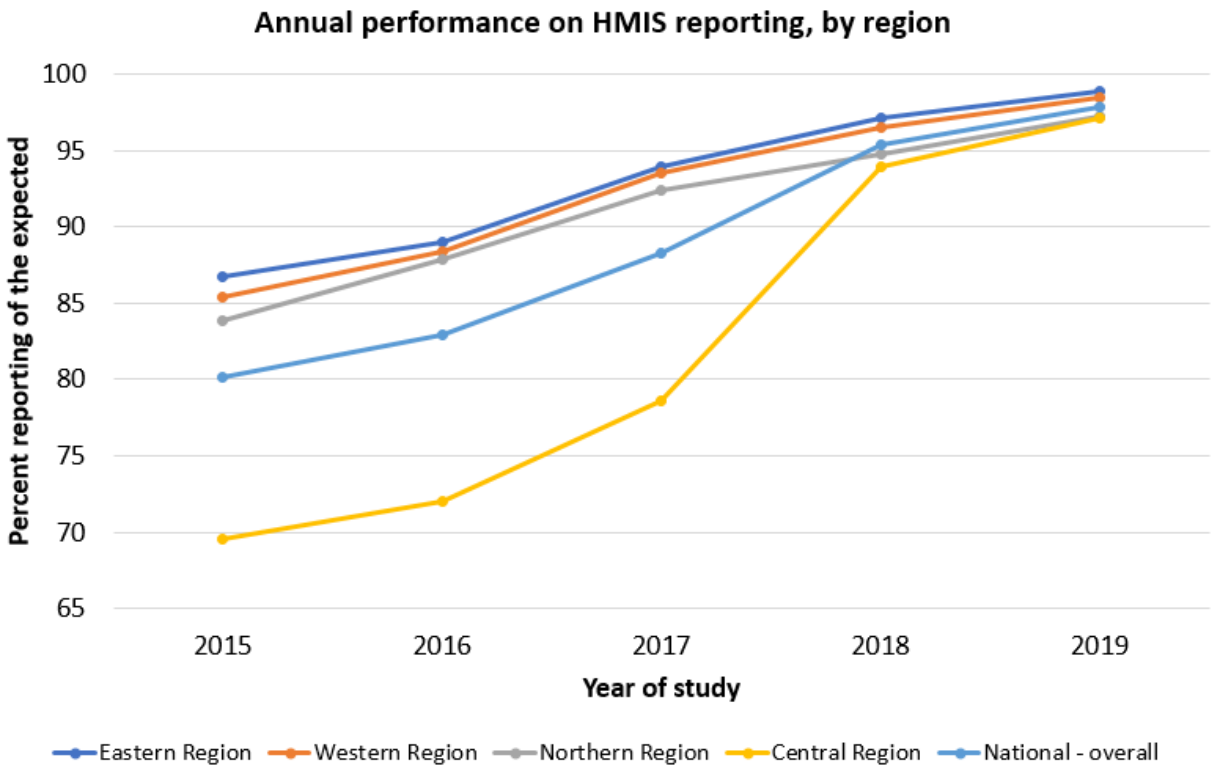
The red colours represent lowest performance of reporting test confirmed cases while the green colours, improving performance

6.1.3 Reporting completeness and timeliness

Part of the output from the DHIS-2 system is an assessment of reporting completeness and timeliness of report submission. Reports of national reporting were generated from the system for the study duration and these were available by MIS regions of 2014 during which the country was divided into five regions. Fig. S3 below shows the

trends in reporting rates, defined as the proportion of expected reports for a given year that were submitted from regions. Notably however, reporting timeliness was lower than eventual submission indicating an area of much needed improvement in HMIS reporting.

Fig S3. Proportion of expected HMIS reports that were submited, by the four regions and the overall national trend

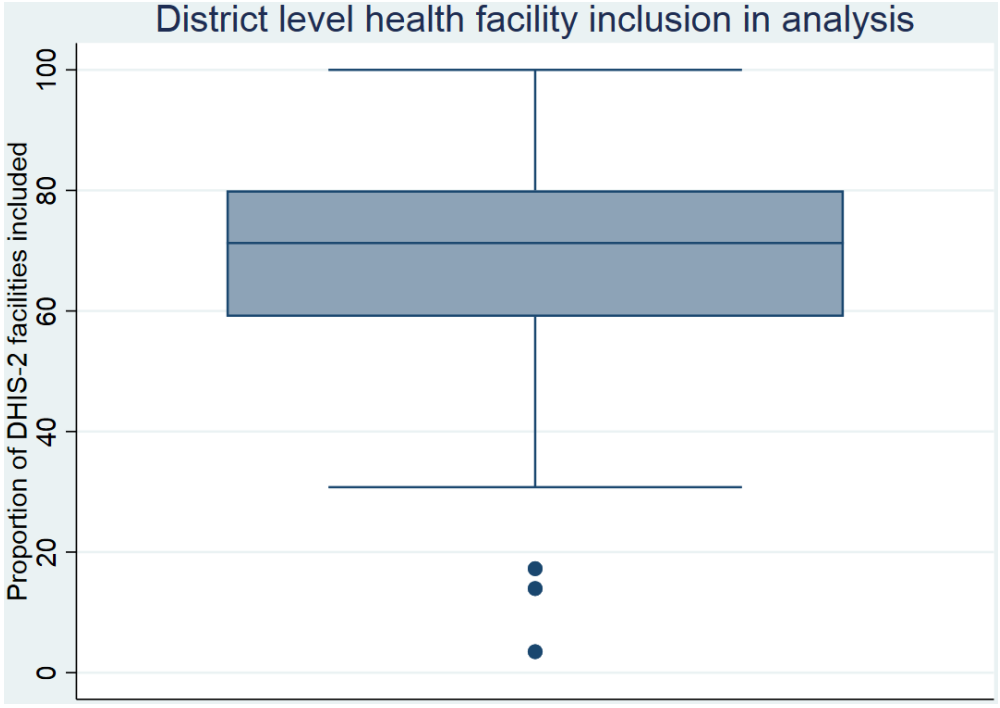


The central region which includes Kampala and the other highly urbanized areas of the country showed the lowest performance, consistent with findings from our analysis where urban areas with high proportions of PFP that report the least. However, the region showed considerable improvement over the years.

6.1.4 Impact of study design on results

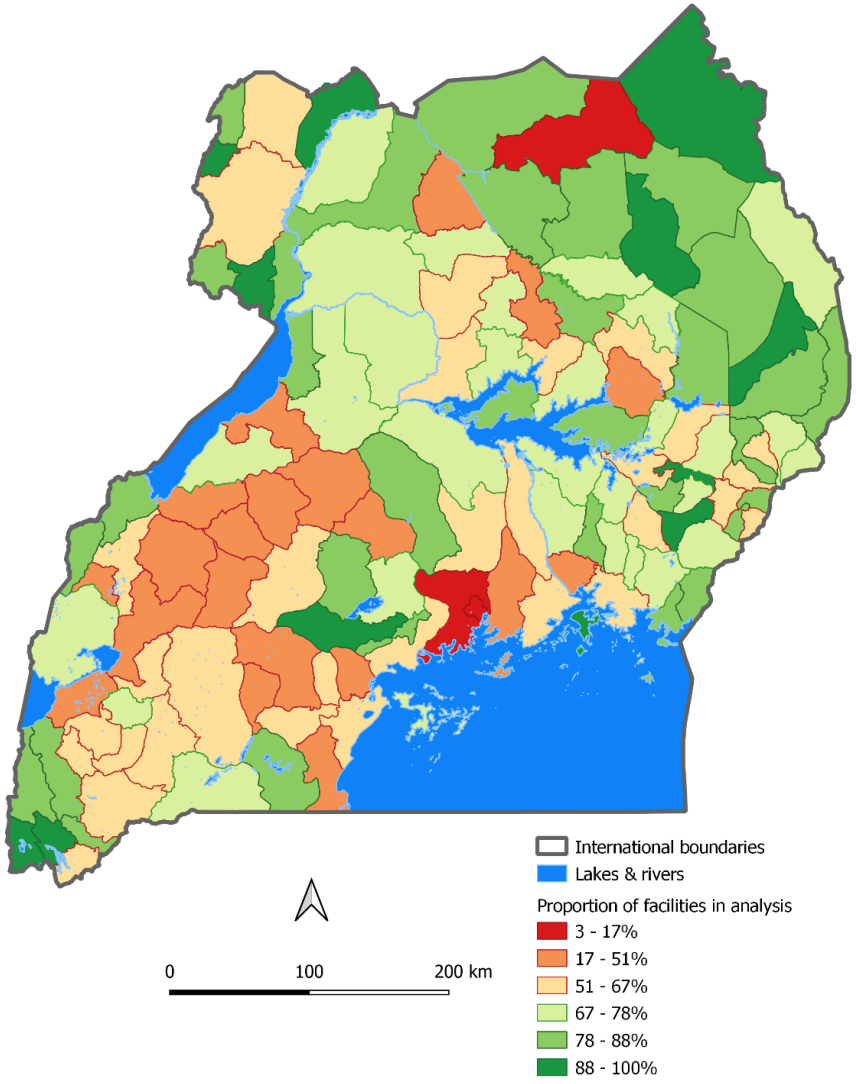
Whilst the study health facilities included in the analysis were only a proportion of the full list of health facilities registered in the national data repository of surveillance data (DHIS-2), these were a good general representation for the national distribution. This was supported by a general left skewed distribution with a median proportion of 71.3% (IQR: 59.0 – 80.0) (Fig. S4) and three districts (Kampala and Wakiso - the most urban settings of the country, plus Kitgum - a fairly rural district in the previously war ravaged Northern Uganda) presenting as outliers with low proportions of registered health facilities included in the analysis (Figs. S4 & S5).

Fig S4. Distribution of health facility inclusion in the analysis from amongst the list registered in the DHIS-2 database by 2015



This distribution suggests a fairly good representation of routine surveillance data from across the country with the exception of the three unique districts for which additional efforts would be required to improve their status.

Fig S5. Geographical representation of the proportion of DHIS-2 health facilities included in this study by district.



We observed here that the mid-western region of the country is another that is fairly under-represented. However, it's not clear why in spite of the fairly good overage of catchments across this region, there is still a large number of facilities unaccounted for. Interestingly however, the north-eastern region that had a sparse distribution of catchments than the rest of the country presented some the highest representation districts. Nevertheless, the population in this region leads a characteristically nomadic lifestyle with far-flung permanent settlements among which health facilities would be viable.

To assess the impact of exclusion of facilities from the study on the basis of no geo-location information, however, we examined the raw reported confirmed malaria cases by for each of the 15 endemicity regions. Here, we assessed confirmed malaria cases from excluded regions as a proportion of overall regional confirmed malaria cases reported over each calendar year of the study duration. Results showed that the proportion of confirmed

cases among non-geolocated facilities increased slightly between 2015 and 2019 with the highest increase in Tooro from 11.7 to 32.9%, followed by West Nile from 3.6 to 16.4% respectively (Table S1). Overall, the proportion of malaria cases reported from non-study health facilities moderately increased between 2015 and 2019 from 11.7% to 16.1%, which may indicate an increase in reporting, especially from private facilities with time.

Table S1. Comparison of observed raw (reported) confirmed malaria cases between study and non-study facilities by region, per calendar year

Region	2015*		2016		2017		2018		2019*	
	Included	Excluded (% Missed)	Included	Excluded (% Missed)	Included	Excluded (% Missed)	Included	Excluded (% Missed)	Included	Excluded (% Missed)
Acholi	469,343	89,535 (16.0%)	1,011,420	230,923 (18.6%)	648,221	158,629 (19.7%)	405,519	92,890 (18.6%)	1,028,892	222,463 (17.8%)
Ankole	223,367	37,015 (14.2%)	635,577	99,989 (13.6%)	690,194	111,088 (13.9%)	217,762	53,439 (19.7%)	201,695	51,337 (20.3%)
Bugisu	121,622	15,058 (11.0%)	299,722	35,663 (10.6%)	316,752	44,194 (12.2%)	199,340	35,895 (15.3%)	242,374	42,895 (15.0%)
Bukedi	147,273	11,500 (7.2%)	365,374	28,915 (7.3%)	413,057	35,863 (8.0%)	255,022	22,221 (8.0%)	237,194	21,658 (8.4%)
Bunyoro	190,002	27,645 (12.7%)	485,290	79,643 (14.1%)	496,471	70,328 (12.4%)	298,454	48,038 (13.9%)	460,368	67,734 (12.8%)
East Central Busoga	459,733	58,858 (11.3%)	1,315,082	159,976 (10.8%)	1,102,382	172,844 (13.6%)	936,740	147,221 (13.6%)	946,492	133,679 (12.4%)
Kampala [#]	39,321	63,481 (61.8%)	65,791	118,693 (64.3%)	44,852	107,817 (70.6%)	25,245	107,912 (81.0%)	29,196	96,983 (76.9%)
Karamoja	123,951	5,676 (4.4%)	228,299	11,836 (4.9%)	261,321	14,978 (5.4%)	309,448	14,427 (4.5%)	300,314	23,727 (7.3%)
Kigezi	63,218	4,790 (7.0%)	151,663	19,676 (11.5%)	160,423	21,471 (11.8%)	80,466	11,094 (12.1%)	94,855	7,795 (7.6%)
Lango	274,980	20,689 (7.0%)	649,071	57,531 (8.1%)	364,469	41,071 (10.1%)	245,398	34,690 (12.4%)	515,354	57,704 (10.1%)
North Buganda	359,740	37,244 (9.4%)	899,172	104,118 (10.4%)	1,075,636	126,180 (10.5%)	598,192	80,623 (11.9%)	806,892	104,170 (11.4%)
South Buganda	339,566	50,641 (13.0%)	857,379	132,875 (13.4%)	924,703	180,585 (16.3%)	446,711	124,387 (21.8%)	511,241	114,152 (18.3%)
Teso	250,834	27,289 (9.8%)	610,964	58,393 (8.7%)	406,337	53,759 (11.7%)	478,098	57,421 (10.7%)	495,778	57,500 (10.4%)
Tooro	256,001	33,934 (11.7%)	610,964	95,187 (13.5%)	613,590	84,107 (12.1%)	304,932	122,434 (28.7%)	379,662	186,336 (32.9%)
West Nile	479,283	17,873 (3.6%)	1,185,521	66,045 (5.3%)	1,367,089	132,152 (8.8%)	1,486,110	140,520 (8.6%)	1,445,541	283,897 (16.4%)
Overall	3,798,234	501,228 (11.7%)	9,371,289	1,299,463 (12.2%)	8,885,497	1,355,066 (13.2%)	6,287,437	1,093,212 (14.8%)	7,695,848	1,472,030 (16.1%)

***Incomplete calendar years included in study: July-December for 2015 and January – September for 2019*

[#] *Region comprised of one district*

Notably, the total predicted number of confirmed cases recorded during the nine months of 2019 included in the study period were higher than the total for all of 2018 by at least 1.4 million cases indicating a major increase in malaria burden and confirming 2018 as the lowest burden year between 2016 and 2019.

Population estimates: To determine incidence rates, total confirmed cases at either health facility catchment, district or regional resolution provided the numerator while total population estimates at each respective resolution provided the denominator. Population estimates were extracted from WorldPop gridded surfaces. Compared to Uganda Bureau of Statistics' (UBOS) national population projection for 2015 of approximately 35.5 million, this study's estimated Uganda's at 35.9 million in 2015 from WorldPop. The estimated total study population of 34.9 million therefore, accounted for 97.2% of the 2015 national population estimate. Similarly, whilst UBOS population projection for 2019 was 40.3 million, this study estimated 40.8 million. The study population estimate of 39.6 million during 2019 therefore, accounted for 97.1% of the national population estimate. Whilst the differences between UBOS and WorldPop estimates may be due to in model approaches, differences between national total and study population are attributable to populations located beyond our defined catchments and in locations where very few and sparse health facilities were geolocated.

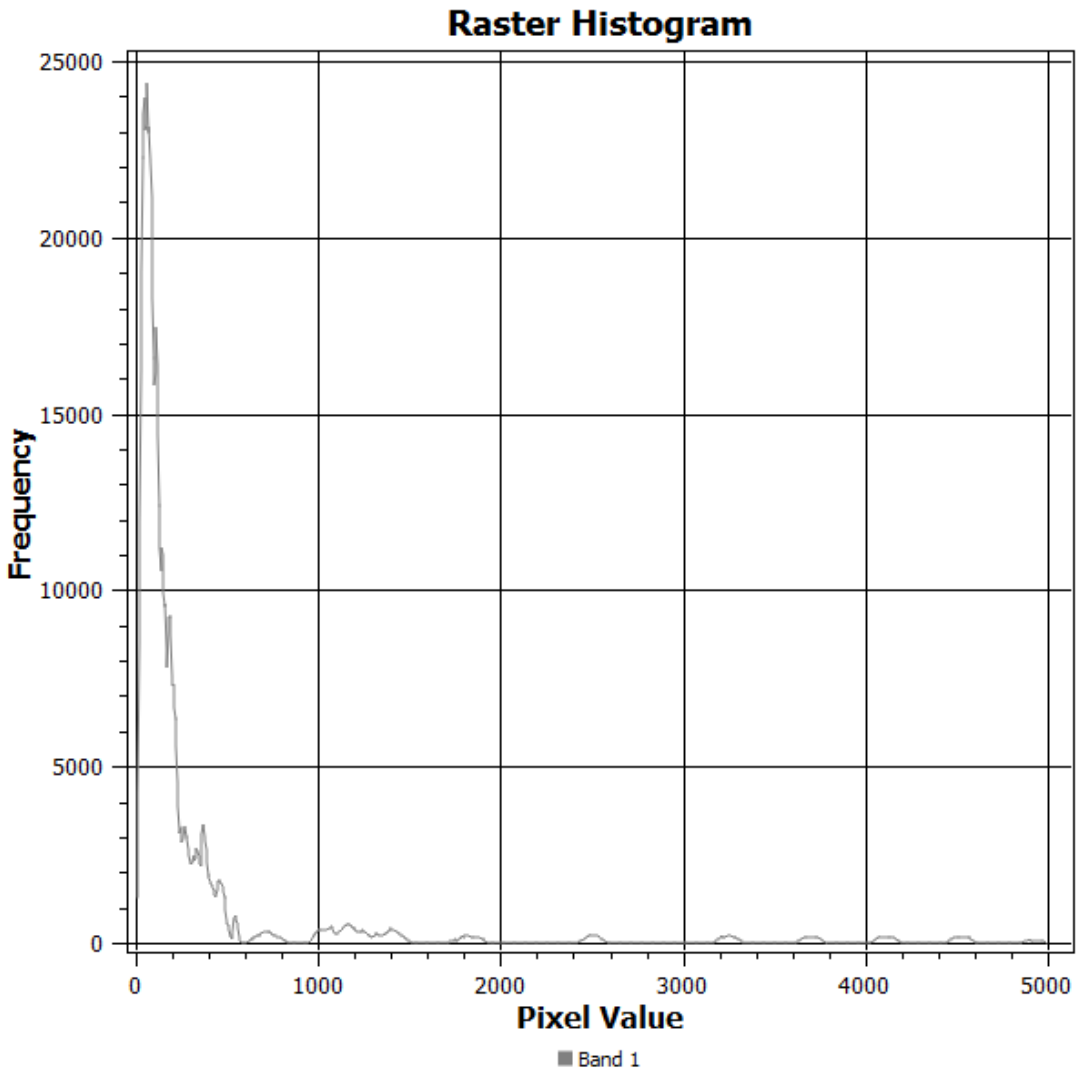
6.1.5 Cost distance surface

Generally in Uganda, geographical catchment for each level of health facility have been conceptualized as level II serving a parish, level III a sub-county, and level IV a group of sub-counties otherwise known as a health sub-district (MoH), among others, though it may not be the case in actual practice. Treatment seeking for malaria has been reported as influenced by multiple factors like: knowledge about malaria and its outcomes, severity of disease, reputation of a health facility and affordability of its services, available alternative remedies, as well as age of head of household [225-227]. However, proximity of a health facility may be one of the strongest influences on treatment seeking, often cited in the rampant use of private versus public health facilities [141]. We, therefore, defined health facility catchments under the assumption that people seek care for uncomplicated malaria from the most proximal health facility, using the AccessMod tool supported by the WHO [195].

AccessMod is a web-enabled spatial analysis tool that provides extended ArcView 3.x functionality. Among others, it is used for modelling catchment areas associated with geo-located sources of care as an estimate of physical accessibility, using travel time [195]. Here, first we generated a cost-distance surface of the entire country, at a 100x100 meter pixel resolution. For this, several geographical covariates, within the WGS 1984 UTM Zone 36S coordinate system under the Transverse Mercator projection, were included. Respective covariate classifications were first assigned an intuitive characteristic speed of travel across them, taking into consideration the most likely means of travel useable, to define an overall travel scenario. This scenario included most likely modes of travel across different surfaces such as walking and cycling, driving and riding, as well as using a canoe or boat across water surfaces. Along with these modes, average travel speeds across the respective surfaces or covariates were

estimated. The covariates included: 1) a digital elevation model (DEM) of the country that defines the elevation variability across the surface at a 100x100m resolution. This measure of slope per pixel in the DEM was used to penalize the speed of crossing a pixel, particularly for walking and bicycle means, making direction of travel important to define. With interest in access to the health facility in this study, direction of travel was chosen to be towards the health facility. 2) Road network across the country classified as Primary, secondary, tertiary, and other roads, were each considered to enable varied speeds of travel across them, ranging from 30 kilometres per hour (Km/h) on other roads such as feeder or country roads to 100 Km/h on primary roads such as highways. 3) Land use and land cover surface covariate, also at a 100x100 meter resolution, was defined in ten classifications of: tree cover, shrub, grassland, cropland, aquatic vegetation, sparse vegetation or lichens-Mosses, bare ground, built-up, open water, and no data areas. The predominant land cover type per pixel, was assigned a characteristic speed of travel, ranging from a low of one Km/h such as across open water to ten Km/h across bare ground. 4) Lakes and rivers. Whereas the first three covariates were classified to enable travel across them, lakes and rivers were considered primarily as barriers with limited capability to cross them and no likelihood of being residential areas for populations or locations for a health facility. Any health facility that would have its geo-coordinates within water was excluded. This did not include health facilities on islands of which there were several. 5) Another barrier considered in this cost-distance evaluation were swamps with limited likelihood of travel across them except if there was a certain type of road network through them, assumed to likely be a bridge-type crossing. These too were included in this process. Majority of country was within four hours (240 minutes) of travel time to the health facility (Fig. S6) however, there were some outliers, especially far into the lakes and fairly high travel times in parts of the country such as the North-Eastern areas, among others.

Fig S6. Distribution of travel time to the health facility across the country in minutes from country-wide raster surface generated



6.1.6 Model selection

To select the model covariates to include, time varied covariates including rainfall estimates, land surface temperature and vegetation amounts quantified as NDVI were evaluated for selection between the current monthly estimate and the mean of current monthly estimate and either one, two or three months' lags using akaike's information criteria values of a multi-variate regression model of crude incidence rates as dependent variable. While keeping all others constant, each covariate was varied to obtain its best quantity for inclusion in the model and the choice of covariate quantity was based on the lowest value of AIC between each covariate's varied values as summarized in Table S2 below.

Table S2. Akaike's information criteria values with corresponding value of covariate included

Covariate	Akaike's information criteria (AIC)		
	Rainfall	Land surface temperature	Vegetation (NDVI)
Current month's estimate	-478514.4	-479,047.50	-481,472.60
Mean of current & 1 month's lag	-479,561.30	-479,232.40	**-481,537.80**
Mean of current & 2 months' lag	-480,403.10	-479,743.80	-481,127.30
Mean of current & 3 months' lag	**-480,553.30**	**-480,553.30**	-480,553.30

Best choice covariate AIC indicated with bold value

From the final selection of covariates including: years of education for women of childbearing age, nighttime light emissivity, mean of current and three months' lags for rainfall and land surface temperature, and mean of current and one month's lag of NDVI, it was clear that all except years of education for women of childbearing age were significantly and positively associated with the outcome, while education was negatively but also significantly (p<0.001) associated as shown in Table S3 below.

Table S3. Association between best fitted covariates and crude incidence rates from multi-variable regression

Covariate	Regression coefficient	95% Conf I	p-value
Education	-0.007338	(0.007578 - 0.007099)	<0.001
Nighttime light emissivity	0.000052	0.000042 - 0.000063	<0.001
Mean of current & 1 month's lag (NDVI)	0.000477	0.000453 - 0.000501	<0.001
Mean of current & 3 months' lag (Rainfall)	0.000128	0.000122 - 0.000135	<0.001
Mean of current & 3 months' lag (Temperature)	0.001654	0.001584 - 0.001725	<0.001

6.1.7 Bayesian model validation

The Bayesian model fit to generate the posterior estimates of malaria incidence accounted for four main explanatory factors including: education of women of child-bearing age, the same being the predominant primary care givers in homes; mean rainfall estimates over the current and three previous months; land surface temperature; vegetation amounts, and nighttime emissivity. In order to validate this model, we randomly selected 20% of catchments to be withheld from the posterior estimates and re-run the model to thereafter compare estimates generated from the model with 80% data to posterior estimates determined from the model with the full data.

Results showed that consistent with strong correlation between observed incidence rates and posterior estimates for all data as shown in Fig. S7 below, out of 990 randomly selected catchments, posterior estimates for 942 were within 95% credible interval of the full model prediction for the same. Moreover, there was high correlation between these estimates with Spearman's rho = 0.6988, P<0.001 also represented in Fig. S8 below.

Fig S7. Scatter plot of predicted against observed confirmed number of malaria cases from all 3446 catchments in the study

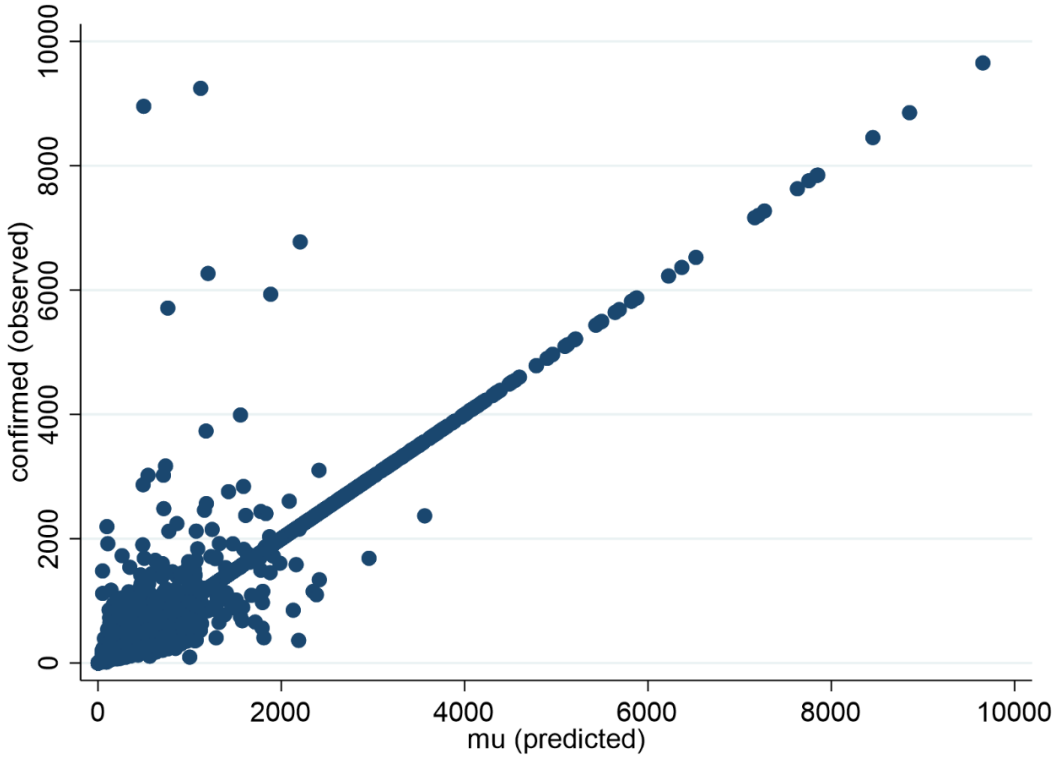
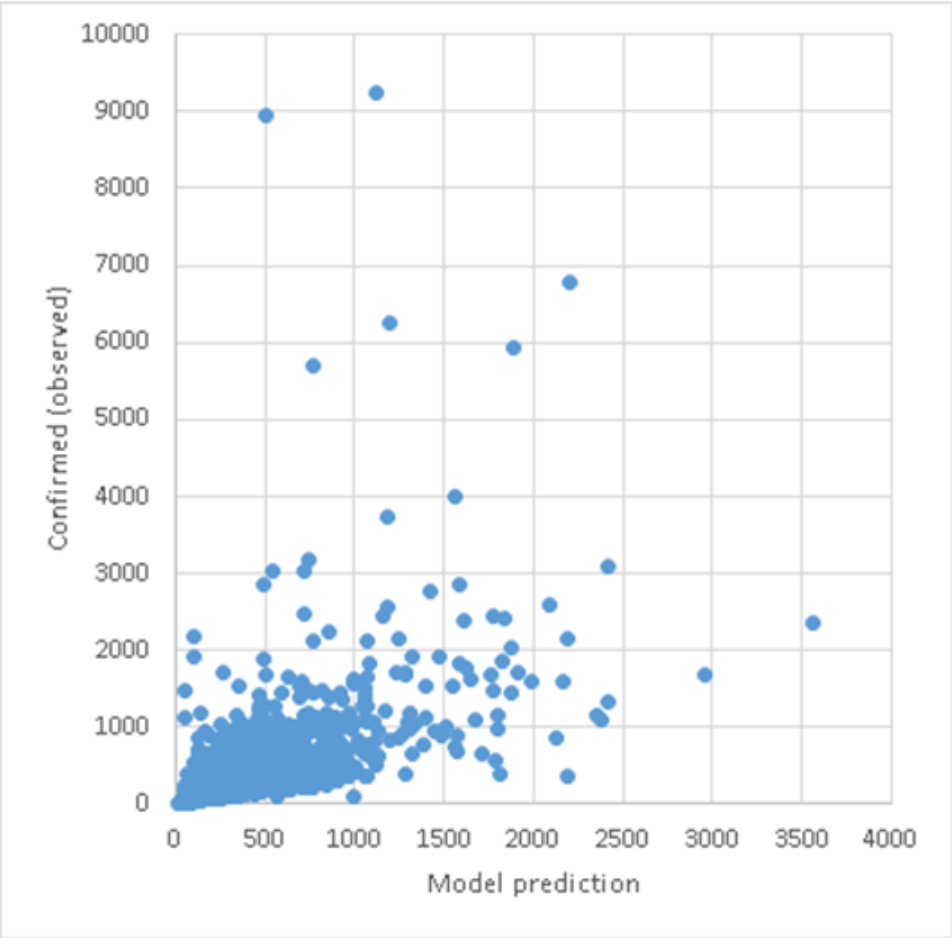


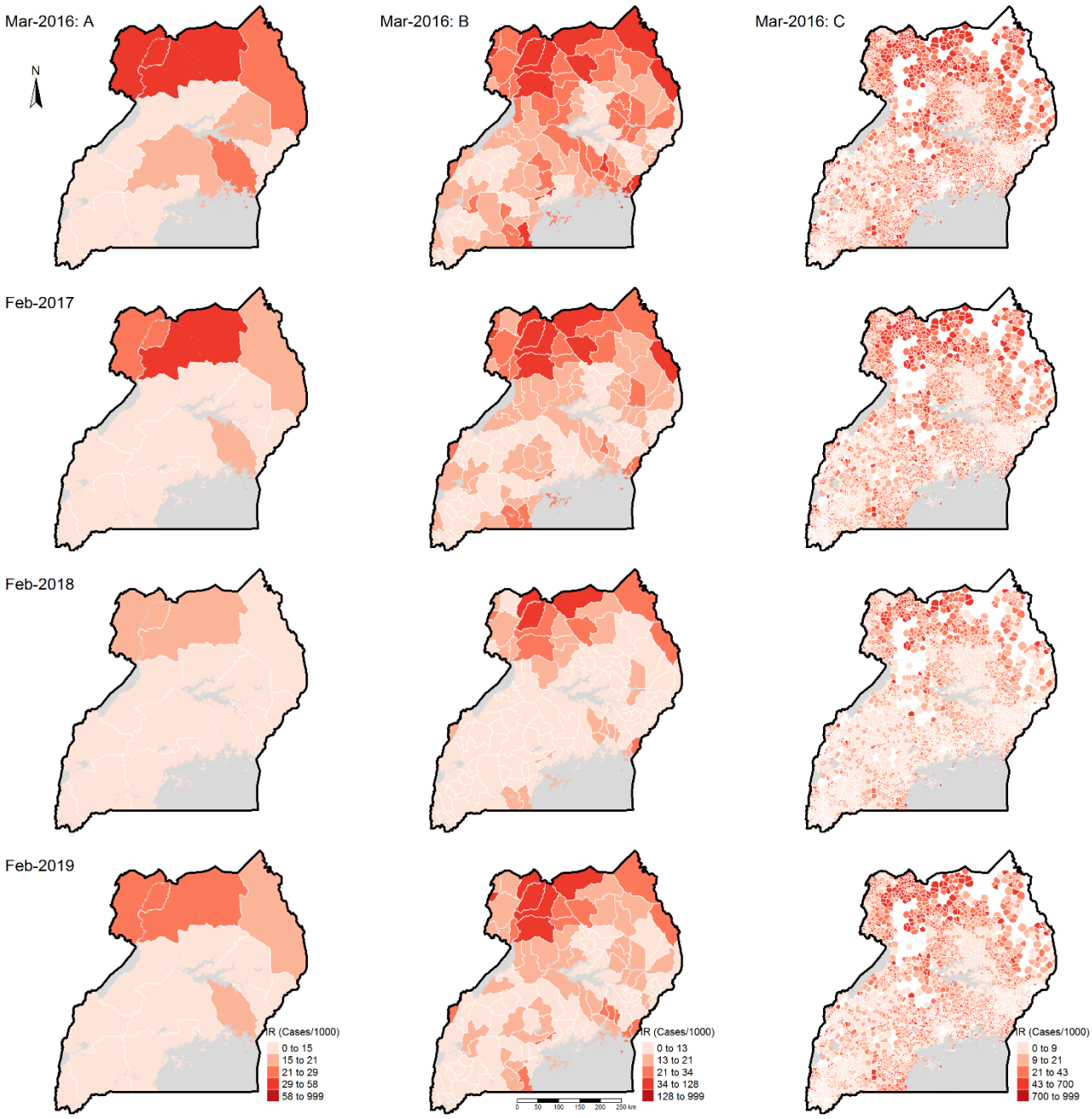
Fig S8. Scatter plot of model predicted against observed confirmed malaria cases for a random sample of 990 catchments for validation of Bayesian model



6.1.8 Distribution of catchment-level incidence rates

Consistent with the observed distribution of incidence rates during the high burden months of the study duration, for all the low burden months, a distinct distribution was observed. In the latter like the former, highest burden districts - potential drivers of their respective regional burden were identifiable. Similarly, highest burden catchments – potential drivers of district burden were also identifiable as shown in Fig. S9 below. Moreover, the pattern of reducing burden from 2016 through 2018, followed by a rebound in 2019 was also observable, particularly among regions and districts.

Fig S9. Spatial distribution of malaria incidence rates during low burden months of study duration



Maps in column A represent the distribution of incidence rates by MIS regions (15 regions) of the country; B represent the distribution of incidence rates by district (128 districts); while C represent the distribution of incidence rates by health facility catchments (3446 catchments in all).

Examination of the association between mean and standard deviation (SD) of monthly incidence rates, considering SD the coefficient of variation in these monthly incidence rates, showed that there was a strong linear association between the two, both at district and health facility catchment levels. This was shown using scatter plots of the coefficient of variation versus mean monthly incidence rates at each level as shown in Fig. S10 below as well as

the heat-map of district monthly incidence rates in Fig. S11 below. Moreover, increase in mean monthly incidence rates was associated with an increased coefficient of variation and therefore, variability.

Fig S10. Scatterplot for association between Standard deviation and Mean monthly incidence rates at district and catchment levels

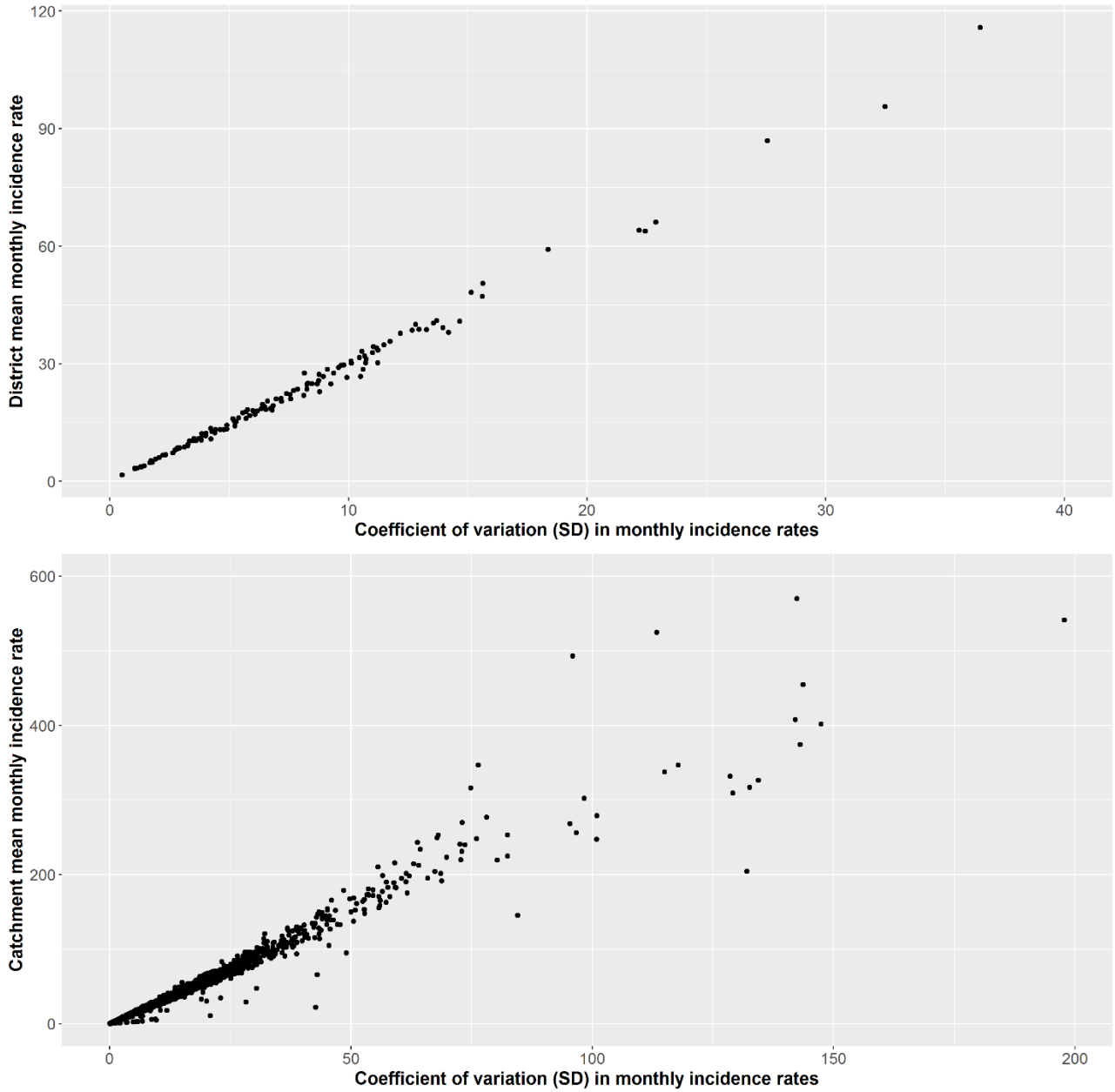
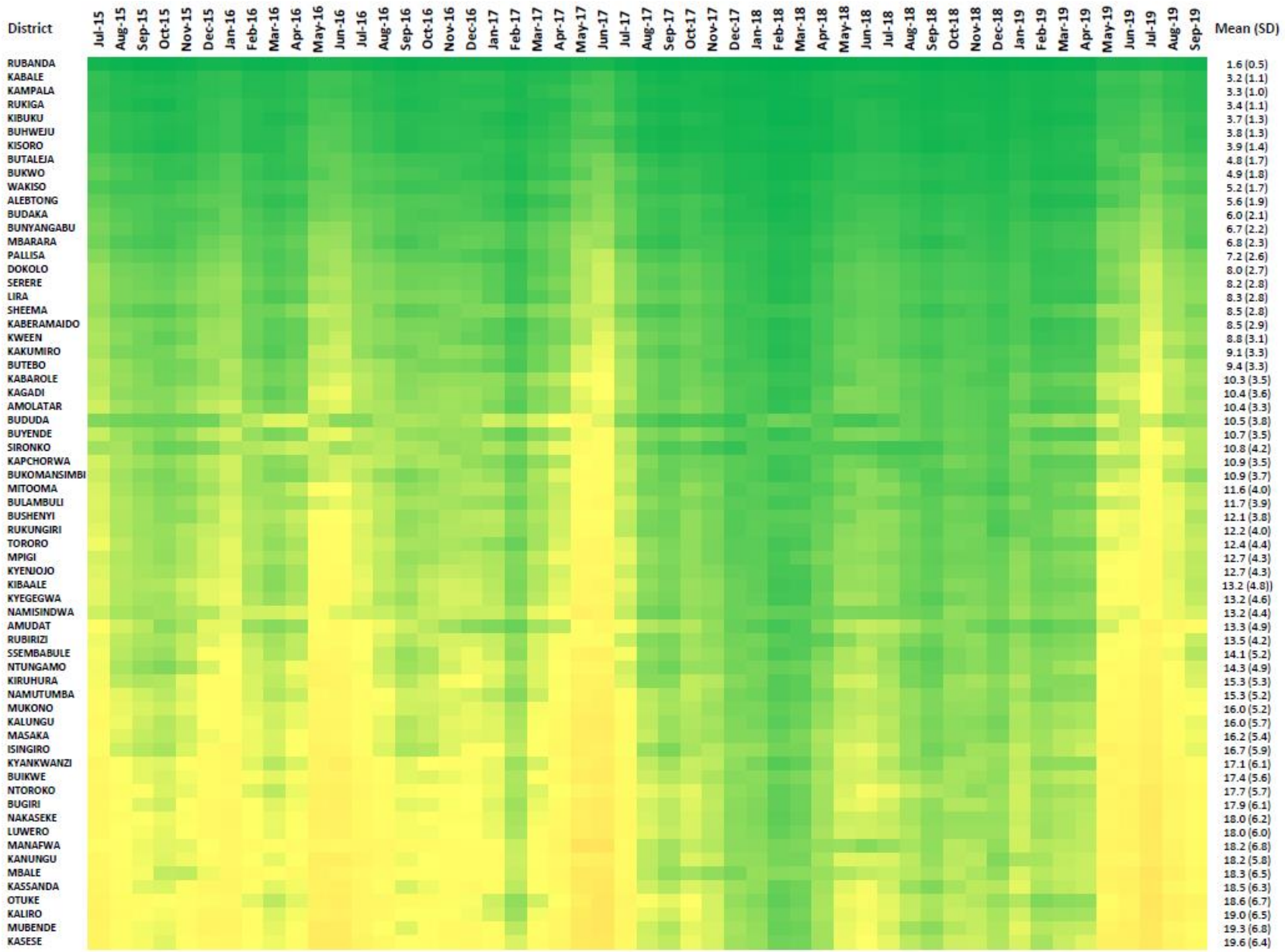
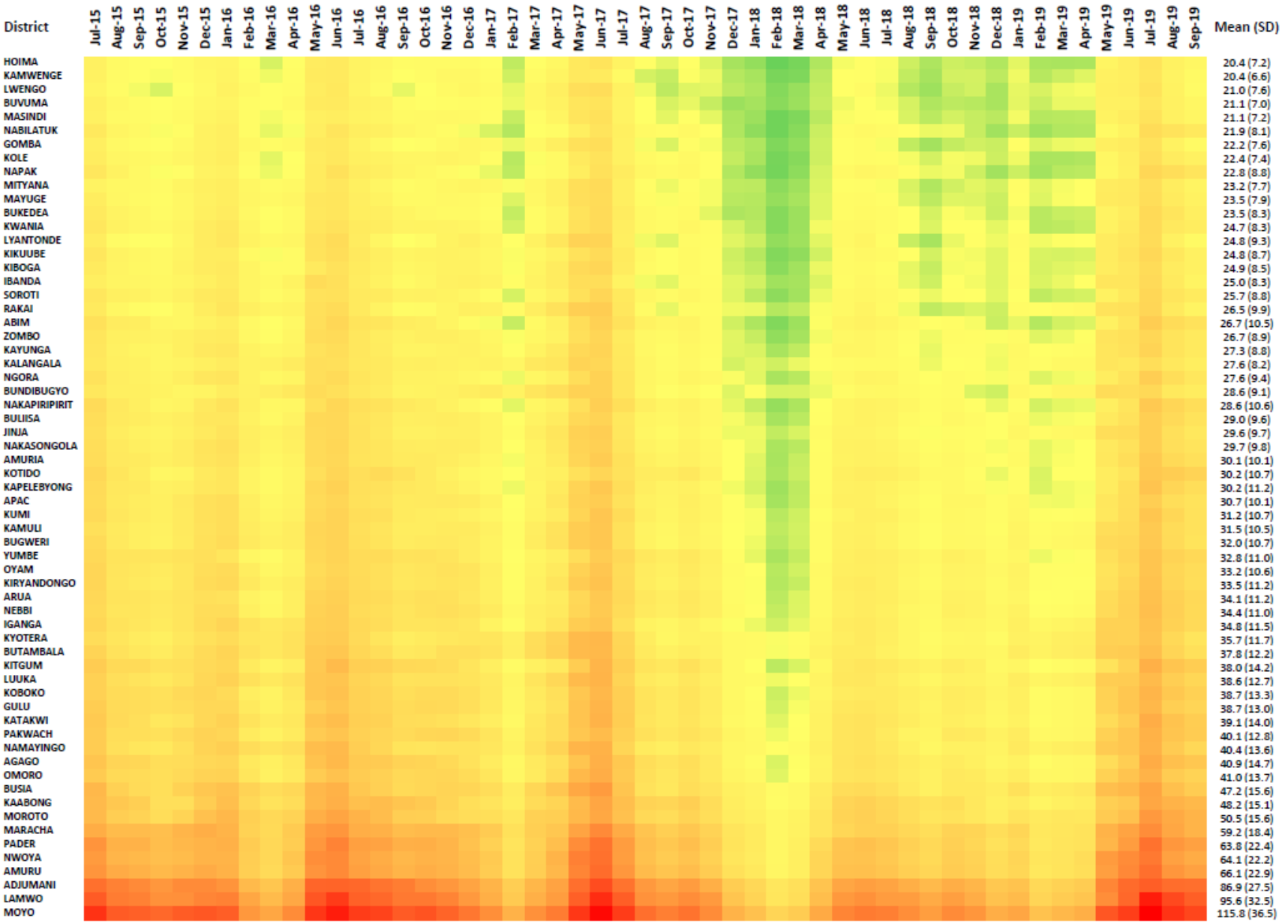


Fig S11. Distribution of malaria incidence rates per month for the 128 districts in Uganda between July 2015 and September 2019

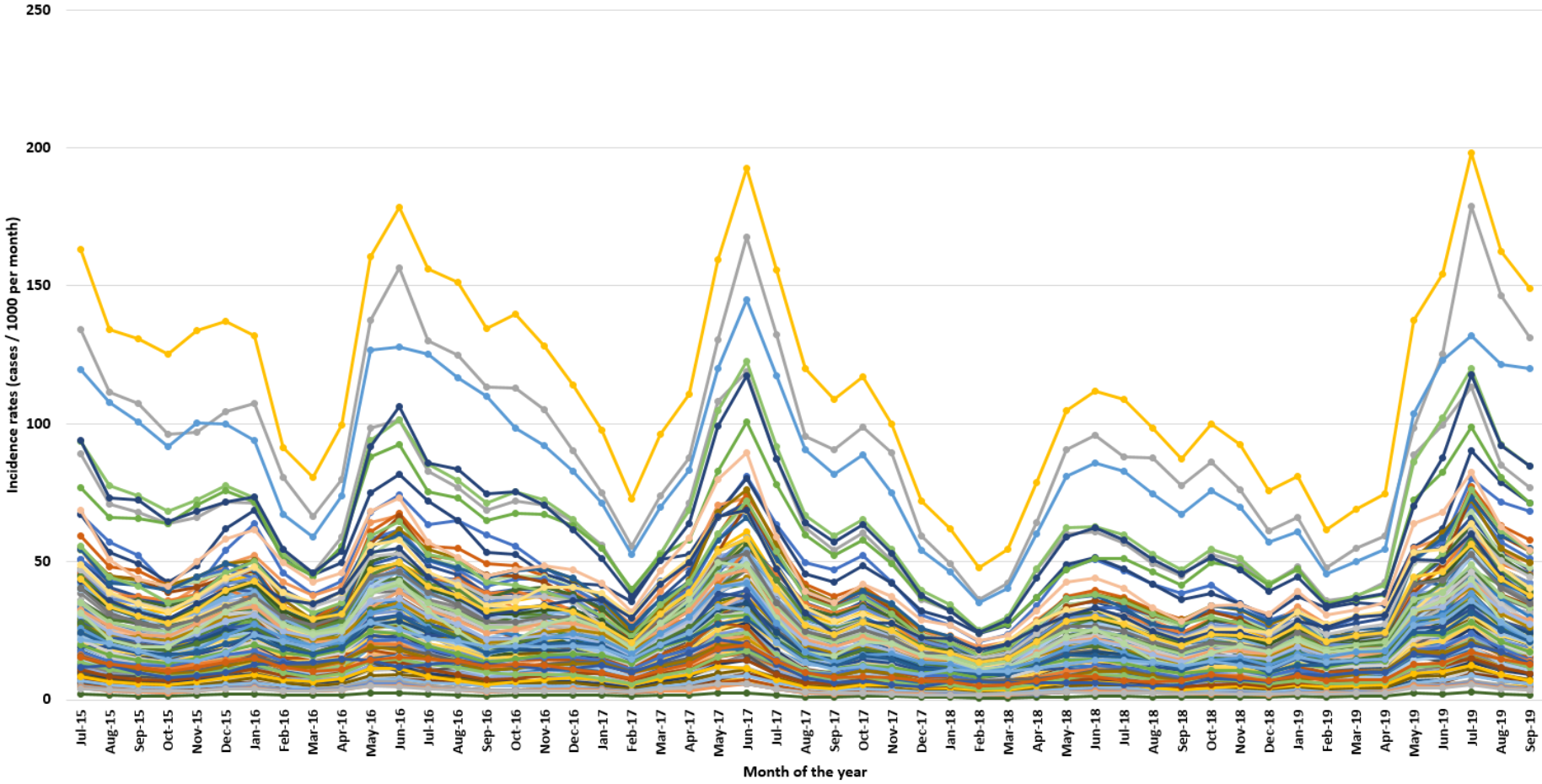



District
Jul-15 Aug-15 Sep-15 Oct-15 Nov-15 Dec-15 Jan-16 Feb-16 Mar-16 Apr-16 May-16 Jun-16 Jul-16 Aug-16 Sep-16 Oct-16 Nov-16 Dec-16 Jan-17 Feb-17 Mar-17 Apr-17 May-17 Jun-17 Jul-17 Aug-17 Sep-17 Oct-17 Nov-17 Dec-17 Jan-18 Feb-18 Mar-18 Apr-18 May-18 Jun-18 Jul-18 Aug-18 Sep-18 Oct-18 Nov-18 Dec-18 Jan-19 Feb-19 Mar-19 Apr-19 May-19 Jun-19 Jul-19 Aug-19 Sep-19
Mean (SD)
HOIMA 20.4 (7.2)
KAMWENGE 20.4 (6.6)
LWENGO 21.0 (7.6)
BUVUMA 21.1 (7.0)
MASINDI 21.1 (7.2)
NABILATUK 21.9 (8.1)
GOMBA 22.2 (7.6)
KOLE 22.4 (7.4)
NAPAK 22.8 (8.8)
MITYANA 23.2 (7.7)
MAYUGE 23.5 (7.9)
BUKEDEA 23.5 (8.3)
KWANIA 24.7 (8.3)
LYANTONDE 24.8 (9.3)
KIKUUBE 24.8 (8.7)
KIBOGA 24.9 (8.5)
IBANDA 25.0 (8.3)
SOROTI 25.7 (8.8)
RAKAI 26.5 (9.9)
ABIM 26.7 (10.5)
ZOMBO 26.7 (8.9)
KAYUNGA 27.3 (8.8)
KALANGALA 27.6 (8.2)
NGORA 27.6 (9.4)
BUNDIBUGYO 28.6 (9.1)
NAKAPIRIPIRIT 28.6 (10.6)
BULIISA 29.0 (9.6)
JINJA 29.6 (9.7)
NAKASONGOLA 29.7 (9.8)
AMURIA 30.1 (10.1)
KOTIDO 30.2 (10.7)
KAPELEBYONG 30.2 (11.2)
APAC 30.7 (10.1)
KUMI 31.2 (10.7)
KAMULI 31.5 (10.5)
BUGWERI 32.0 (10.7)
YUMBE 32.8 (11.0)
OYAM 33.2 (10.6)
KIRYANDONGO 33.5 (11.2)
ARUA 34.1 (11.2)
NEBBI 34.4 (11.0)
IGANGA 34.8 (11.5)
KYOTERA 35.7 (11.7)
BUTAMBALA 37.8 (12.2)
KITGUM 38.0 (14.2)
LUUKA 38.6 (12.7)
KOBOKO 38.7 (13.3)
GULU 38.7 (13.0)
KATAKWI 39.1 (14.0)
PAKWACH 40.1 (12.8)
NAMAYINGO 40.4 (13.6)
AGAGO 40.9 (14.7)
OMORO 41.0 (13.7)
BUSIA 47.2 (15.6)
KAABONG 48.2 (15.1)
MOROTO 50.5 (15.6)
MARACHA 59.2 (18.4)
PADER 63.8 (22.4)
NWOYA 64.1 (22.2)
AMURU 66.1 (22.9)
ADJUMANI 86.9 (27.5)
LAMWO 95.6 (32.5)
MOYO 115.8 (36.5)

6.1.9 Trends in monthly incidence rates

Whilst clear and fairly strong trends in malaria incidence rates were observed at the national level with seasonality, similar trends and seasonality were also reflected among regions and districts. The highest burden areas also sustained the highest levels of mean monthly incidence rates across the study duration both at regional and also as shown among districts in Fig. S12 below.

Fig S12. Trends in the mean monthly incidence rates by the 128 districts of Uganda as of 2018



Here, we observe two groups of districts with the highest incidence rates. The first comprised the three districts of Moyo, Lamwo, and Adjumani while the second group included Amuru, Pader, Nwoya, and Maracha. Consistent with seasonality observed at national level, all districts both how and high burden,

showed the same seasonality pattern with June-July as the highest incidence rates months and February-March being the lowest incidence rates months across the study duration.

6.1.10 Annual estimates of confirmed malaria cases

National estimates of annual total confirmed malaria cases in Uganda are obtainable through routine surveillance data. However, these are not fully utilized in global burden estimates of malaria for Uganda or other countries in the region. To evaluate how HMIS-based estimates compared with global estimates, annual estimates for each calendar year were compared with the most recent estimates reported in the WHO's world malaria report 2019. Given that global reports are provided in complete years, the respective calendar years for which data included in our study was not from the entire 12 months of the year, the comparisons were considered not applicable. These results are presented in Table S4 below.

Table S4. Estimated number of confirmed malaria cases from this study compared with estimates from the WHO estimates

Year / duration	Confirmed malaria cases in millions (95% CI)	Estimated cases from Malaria atlas project in millions (95% Conf. I)	WHO malaria report 2019 in millions (95% Conf. I)
Jul-Dec, 2015	4.759 (4.642 - 5.017)	N/A	N/A
Jan-Dec, 2016	10.151 (9.904 - 10.688)	10.876 (8.439 - 14.007)	12.070 (9.342 - 15.300)
Jan-Dec, 2017	9.439 (9.210 - 9.927)	11.096 (8.613 - 14.286)	13.863 (10.840 - 17.470)
Jan-Dec, 2018	6.527 (6.368 - 6.865)	N/A	12.357 (7.623 - 18.970)
Jan-Sep, 2019	7.951 (7.760 - 8.362)	N/A	N/A

CI – Credible interval

Conf. I – Confidence interval

Given that estimates from this study included only the health facilities that were geo-located across the country, we argue that this may explain the lower estimates of confirmed malaria cases per year from this study compared to WHO reported estimates as seen in Table S4. Notably however, the trend observed from this study with total annual confirmed cases reducing between 2016 and 2018 is not observable from WHO reported estimates [3] or with Malaria Atlas project estimates. From this study, 2016 registered the highest total confirmed number of cases and 2018 the lowest of the three years, while the global reports indicated 2017 as the highest with 2016 the lowest of the three years. This may be attributable to differences in the approaches used for the estimates reported between these sources. Both MAP and WHO estimates, that showed an increase in cases from 2016 to 2017, were

generated predominantly using survey-based data [3, 207] that may not capture nuances observed from routine data that showed a clear downward trend in malaria cases between 2016 and 2018 from this study.

National monthly incidence rates: Examining monthly estimates of national mean incidence rates, results showed that across the 51 months, June-July experienced the highest incidence rates for all the years, while the lowest estimates were observed variably but mostly during February-March as shown in Table S5. For 2015 where only 7 months were included in this study starting from July, however, the lowest estimate was observed during October.

Table S5. National, regional and health facility catchment highest and lowest estimated monthly incidence rates per study calendar year in Uganda

Annual duration	Peak monthly incidence rates		Lowest monthly incidence rates	
	month	IR (95% CI)	month	IR (95% CI)
Jul-Dec 2015	July	27.6 (27.0 - 29.1)	October	20.1 (20.6 – 21.7)
Jan-Dec 2016	June	32.1 (31.3 - 33.8)	March	17.2 (16.8 - 18.1)
Jan-Dec 2017	June	36.6 (35.7 - 38.5)	December	12.3 (12.0 - 13.0)
Jan-Dec 2018	June	18.9 (18.4 - 19.9)	February	8.9 (8.7 - 9.4)
Jan-Sep 2019	July	36.3 (35.4 - 38.1)	February	12.7 (12.3 - 13.3)

CI – Credible interval

IR – Incidence rate estimated

Further, the five highest risk regions both during the highest and lowest burden months being identified across the country as Acholi, West Nile, Karamoja, East Central – Busoga, and Teso. Within these regions, the highest burden districts were also identifiable, the highest four districts in each shown in Table S6 below between the two seasons.

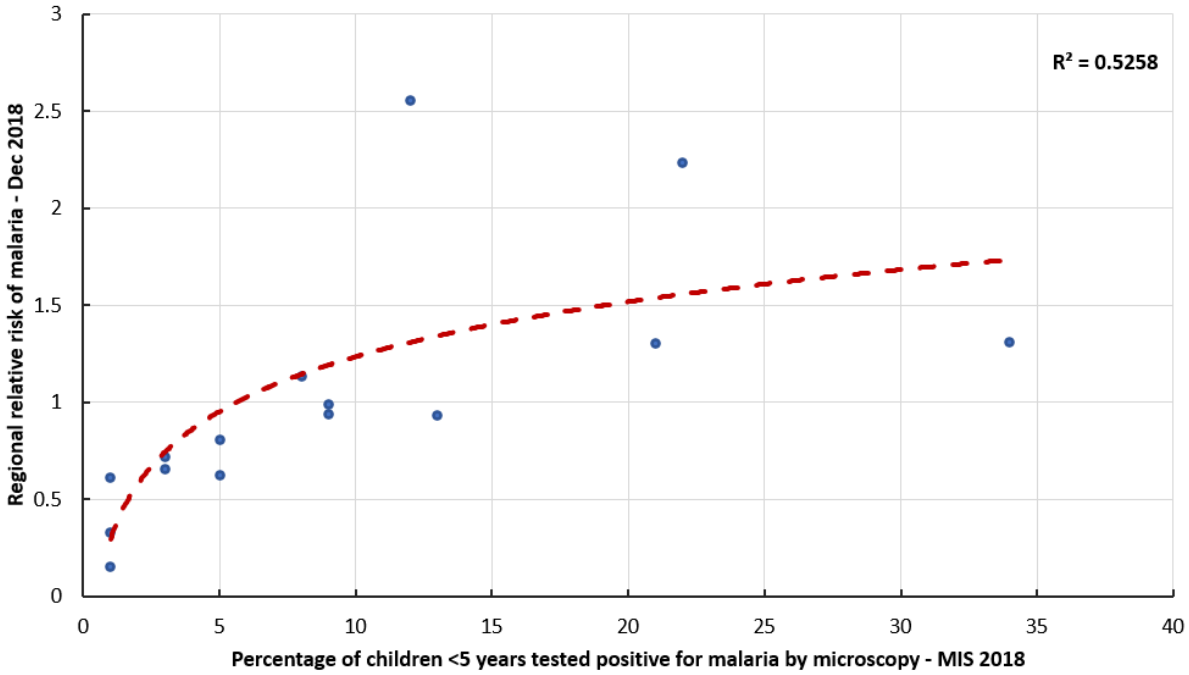
Table S6. The four highest burden districts within the five highest risk regions across the country during July-2017 and February-2018, the highest and lowest burden months of the study duration.

	July 2017 (the highest burden month)		February 2018 (the lowest burden month)	
Region	**District**	**IR (95% CI)**	**District**	**IR (95% CI)**
Acholi	Lamwo	167.6 (165.6 - 169.8)	Lamwo	36.4 (35.9 - 37.0)
	Amuru	122.5 (118.4 - 138.0)	Amuru	25.1 (24.2 - 28.2)
	Nwoya	118.8 (117.2 - 120.3)	Pader	24.3 (23.2 - 28.7)
	Pader	117.3 (111.5 - 118.2)	Nwoya	23.9 (23.5 - 24.3)
West Nile	Moyo	192.5 (189.9 - 195.1)	Moyo	48.0 (47.2 - 48.8)
	Adjumani	145.0 (143.5 - 146.4)	Adjumani	35.2 (34.8 - 35.7)
	Maracha	100.8 (99.3 - 102.4)	Maracha	24.4 (24.1 - 24.8)
	Koboko	68.5 (67.8 - 69.3)	Pakwach	17.2 (16.9 - 17.7)
Karamoja	Kaabong	81.1 (79.6 - 82.5)	Moroto	23.9 (23.4 - 24.4)
	Moroto	80.2 (78.7 - 81.7)	Kaabong	21.0 (20.6 - 21.4)
	Nakapiripirit	52.4 (51.5 - 53.3)	Kotido	12.6 (12.4 - 12.9)
	Abim	51.1 (50.1 -52.0)	Nakapiripirit	11.1 (10.9 - 11.3)
East Central - Busoga	Namayingo	73.1 (71.9 - 75.0)	Namayingo	19.1 (18.7 - 19.6)
	Luuka	71.3 (69.9 - 72.7)	Luuka	17.2 (16.8 - 17.6)
	Iganga	58.5 (57.6 - 59.5)	Iganga	15.1 (14.9 - 15.4)
	Bugweri	58.4 (57.5 - 59.4)	Jinja	14.1 (13.5 - 16.0)
Teso	Katakwi	72.0 (70.9 - 73.1)	Katakwi	16.2 (15.9 - 16.5)
	Kumi	58.6 (57.7 - 59.5)	Kumi	13.5 (13.3 - 13.7)
	Kapelebyong	57.5 (56.5 - 58.5)	Ngora	12.5 (12.3 - 12.7)
	Amuria	54.2 (53.4 - 55.1)	Amuria	12.4 (12.1 - 12.6)

To evaluate the estimated relative risk against other known estimates, a scatter plot of the regional prevalence of malaria estimated among children 0-59 months of age tested by microscopy within the 2018 Malaria Indicator

Survey [61], against the study estimated relative risk of malaria during December 2018, the month when the MIS survey was conducted was plotted. Fig. S13 below shows a positive relationship between these two estimates and indication of a positive association between the two.

Fig S13. Relationship between the 2018 MIS regional prevalence of malaria and estimated relative risk of malaria for December 2018



The blue points represent the (prevalence, risk) coordinates and the red dotted line, the fitted curve for the relationship. This observed relationship may provide some evidence of the important effect of age that is largely precluded from evaluations of risk among all populations based on data from children. In these cases, low transmission setting estimates of malaria burden may be under-estimated while being over-estimated in the high transmission settings.

6.1.11 Malaria risk distribution

Given variability of risk of malaria through the spatial hierarchy, risk at the two higher levels (region and district) was assessed and presented in Table S7 below. The results show regional risk of malaria relative to national average, during the lowest and highest incidence rate months at national level as independent columns as

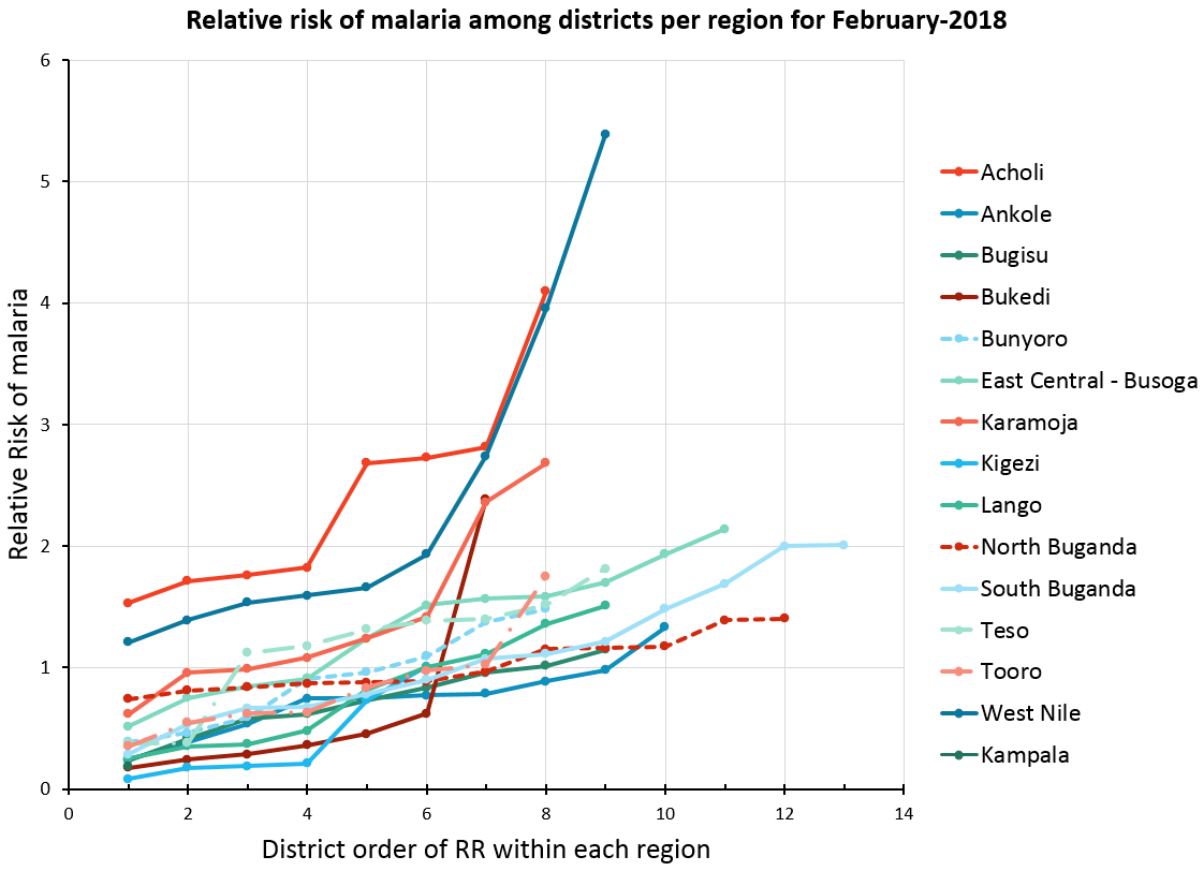
evaluated through national trends in incidence rates. Additionally, for each region, the range of district risk of malaria for the districts that comprise the respective region was also included

Table S7. Risk of malaria relative to national average, for the lowest and highest burden months between July-2015 and September-2019, by region and district.

	Lowest burden month (February-2018)		Highest burden month (June-2017)	
Region	Regional risk	District risk range	Regional risk	District risk range
Acholi	2.2	1.5 – 4.1	2.6	1.9 – 4.6
Ankole	0.8	0.2 – 1.3	0.6	0.2 – 1.1
Bugisu	0.8	0.2 – 1.1	0.7	0.3 – 1.0
Bukedi	0.7	1.2 – 2.4	0.8	0.2 – 2.4
Bunyoro	0.8	0.4 – 1.5	0.9	0.5 – 1.7
East Central - Busoga	1.3	0.5 – 2.1	1.3	0.5 – 2
Kampala	0.2	N/A	0.2	N/A
Karamoja	1.4	0.6 – 2.7	1.4	0.7 – 2.2
Kigezi	0.4	0.1 – 1.0	0.4	0.1 – 0.8
Lango	0.8	0.3 – 1.5	0.9	0.3 – 1.6
North Buganda	1	0.7 – 1.4	1	0.8 – 1.4
South Buganda	0.7	0.3 – 2.0	0.7	0.3 – 1.9
Teso	1.1	0.4 – 1.8	1.2	0.4 – 2.0
Tooro	0.9	0.4 – 1.8	0.8	0.3 – 1.4
West Nile	1.9	1.2 – 5.4	2	1.3 – 5.3

With 15 regions define across the country, we assessed risk distribution by region and the consequent distribution of risk across each region by district. For the low burden month of February-2018, the four highest risk regions in descending order were Acholi, West Nile, Karamoja, and East Central – Busoga, each with greater than one times the national average. Comparatively during the highest burden month of June-2017, the same regions maintained their position in rank of risk (Table S7). On the other hand, the four lowest risk regions during the lowest burden month in ascending order were Kampala, Kigezi, Bukedi, and South Buganda each at lower than national average risk. However, during the highest risk month, the four lowest risk regions changed order to include (in ascending order) Kampala, Kigezi, Ankole and Bugisu, implying that only Kampala and Kigezi maintained their lowest rank of risk. The distribution of risk across districts within each region were further explored using scatter plots for both the lowest and highest burden months as presented in Figs. S14 and S15 below.

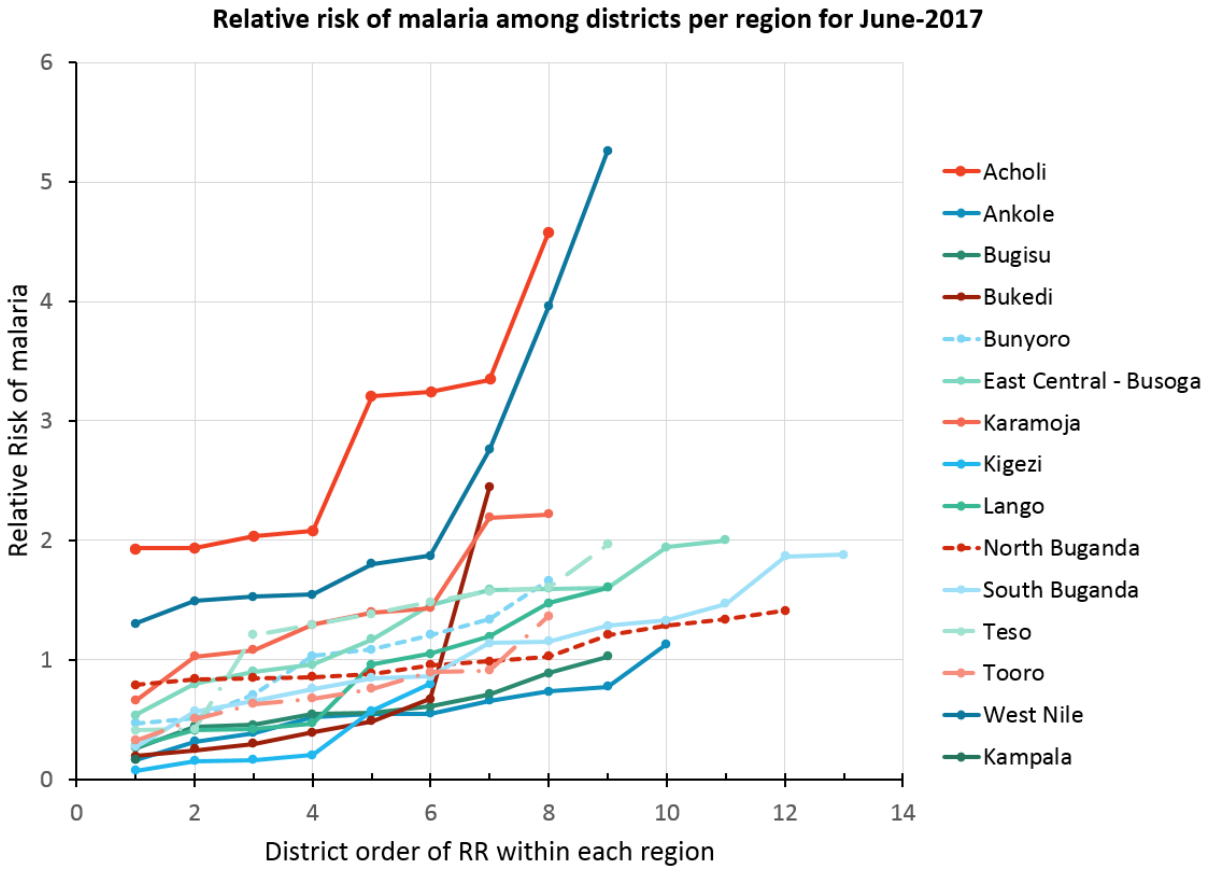
Fig S14



Results showed Acholi and West Nile as the regions with districts that are at the highest risk of malaria, will all their comprising districts at higher relative risk than national average regardless of season. However, while Kigezi region was one of the lowest burden regions with four districts having the lowest risk, the region also had two districts with notably higher risk of malaria than the rest in the region. This pattern is observable among all regions and could play a role in identification of higher priority districts per region over any given observation period.

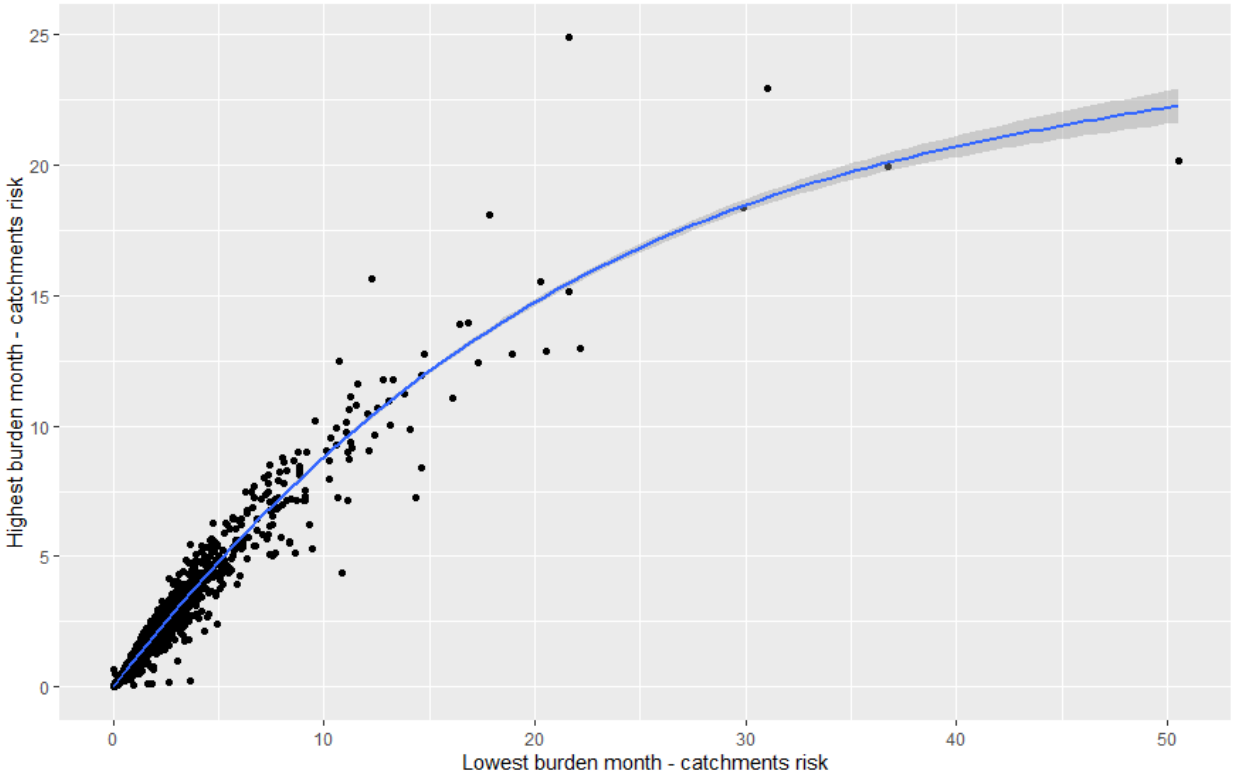
Notably, among districts within the two lowest risk regions of Kampala and Kigezi the mean relative risk was higher during the lowest burden month at 0.4 (95% Conf. I:0.0 – 0.7) than the highest at 0.3 (95% Conf. I:0.1 – 0.6), but with no significant difference (P=0.706). Similarly, for the middle-ranked risk regions, mean relative risk among their districts was higher during the lowest burden month at 0.9 (95% Conf. I:0.8 – 1.0) than during the highest at 0.8 (95% Conf. I:0.7 – 0.9) with no significant difference (P=0.717).

Fig S15.



Furthermore, examining the relationship between catchment-level risk during lowest and highest burden months, results showed that the catchments at highest risk were at disproportionately higher risk during lowest than highest burden seasons of the year, as shown in Fig. S16 below.

Fig S16. Relationship of catchment-level risk of malaria between Lowest and highest burden seasons

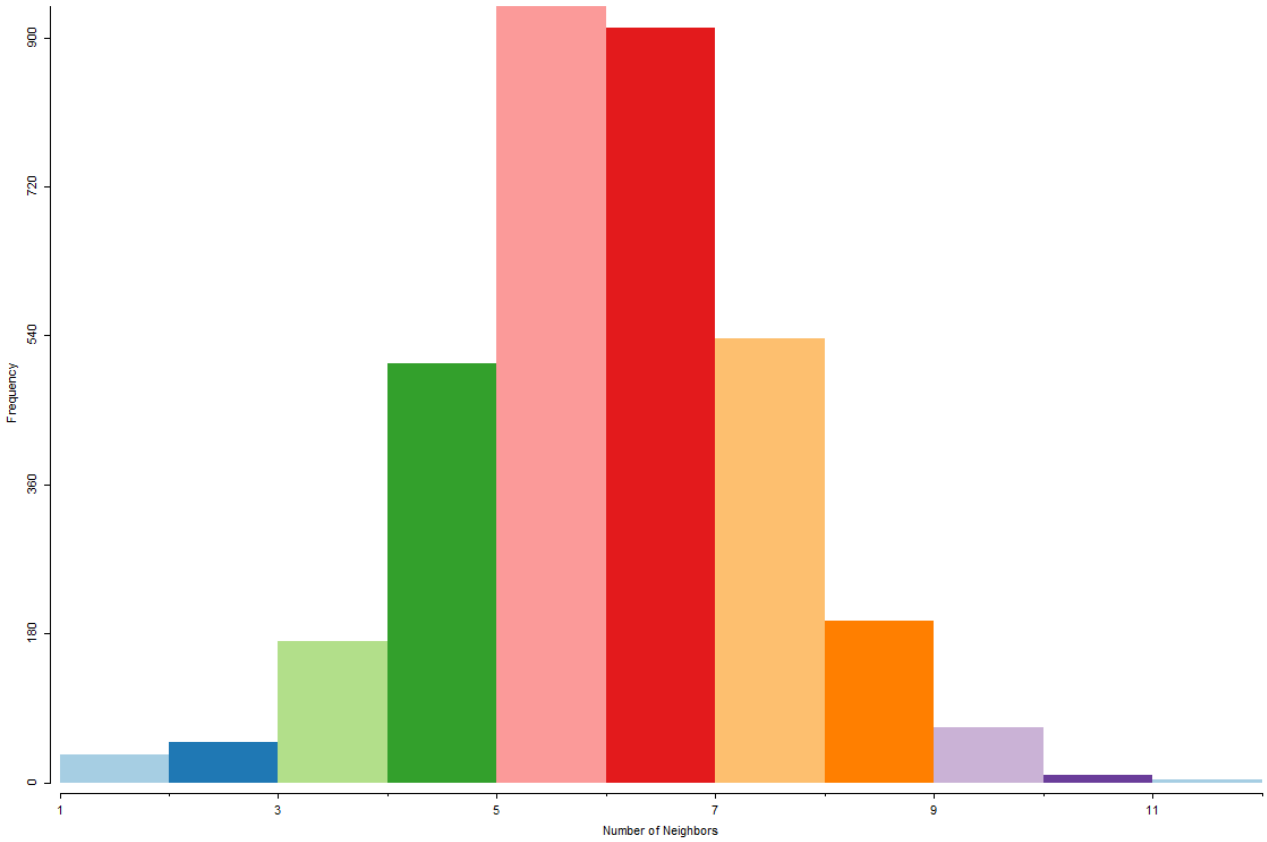


A scatter plot of health facility catchment risk of malaria during the lowest burden month (Feb-2018, x-axis) against the highest burden month (Jul-2017, y-axis) with a fitted smoother curve (blue) and 95% confidence band (grey). The plot indicates that a large majority of locations maintain similar levels of risk of malaria both between high and low burden seasons. However, the non-linear form of the relationship suggests that a few locations bear a much higher risk of malaria during their lower than higher burden seasons, and these could be identified through such small area assessments for further intervention.

6.1.12 Spatial autocorrelation of risk

Given the identifiable distribution of risk across the 15 regions of the country, we assessed for spatial autocorrelation of risk at district and catchment levels so as to test for spatial randomness also known as heterogeneity versus spatial clustering of risk of malaria at these scales. For these assessments, the global Moran's Index (Moran's I) test was performed in R. At district-level, this included all 128 districts as they were known by 2018. However, owing to the requirement of contiguity among neighbours for this analysis, isolated catchment areas without a neighbour with a shared border were excluded. Consequently, 27 catchments were excluded from this assessment leaving 3419 (99.2%) catchments. Catchment neighbourhood was approximately normally distributed as shown in Fig. S17 below, with the highest number of contiguous neighbours a single catchment had being 11 and the least being one.

Fig S17. Neighbourhood distribution among health facility catchment areas.



This analysis was conducted for both the lowest and highest burden months of the study duration and results are presented using Moran's scatter plots for the two durations in Figs. S18 and S19 for the district level and Figs. S20 and S21 at the catchment level below.

The test statistics for spatial autocorrelation (Moran's I) showed that both at the catchment and district levels, there was increased variability of clustering during the highest burden periods that were associated with increased relative risk. Moreover, we observed higher relative risk among lower transmission areas during lower than higher burden seasons. Together, these may provide some indication of disproportionately higher increase in burden among highest risk locations than increases among the lowest burden areas when malaria upsurges occur. This may be consistent with the notion of the 80:20 Pareto rule [228] indicating here that 20% of the population may bear 80% of the burden of malaria infections.

Fig S18. Moran's scatter plot for district-level risk of malaria during the lowest burden month of the study duration

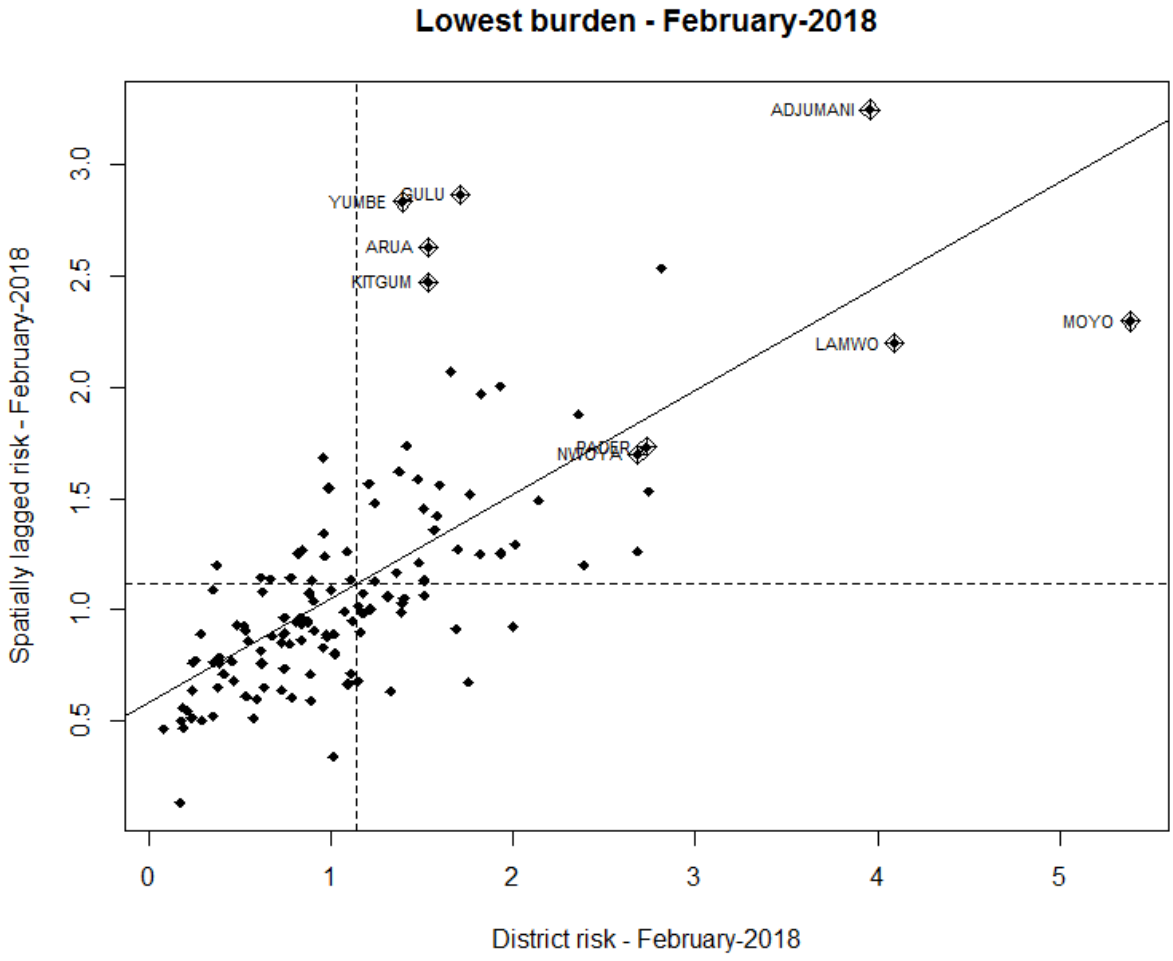


Fig S19. Moran's scatter plot for district-level risk of malaria during the highest burden month of the study duration

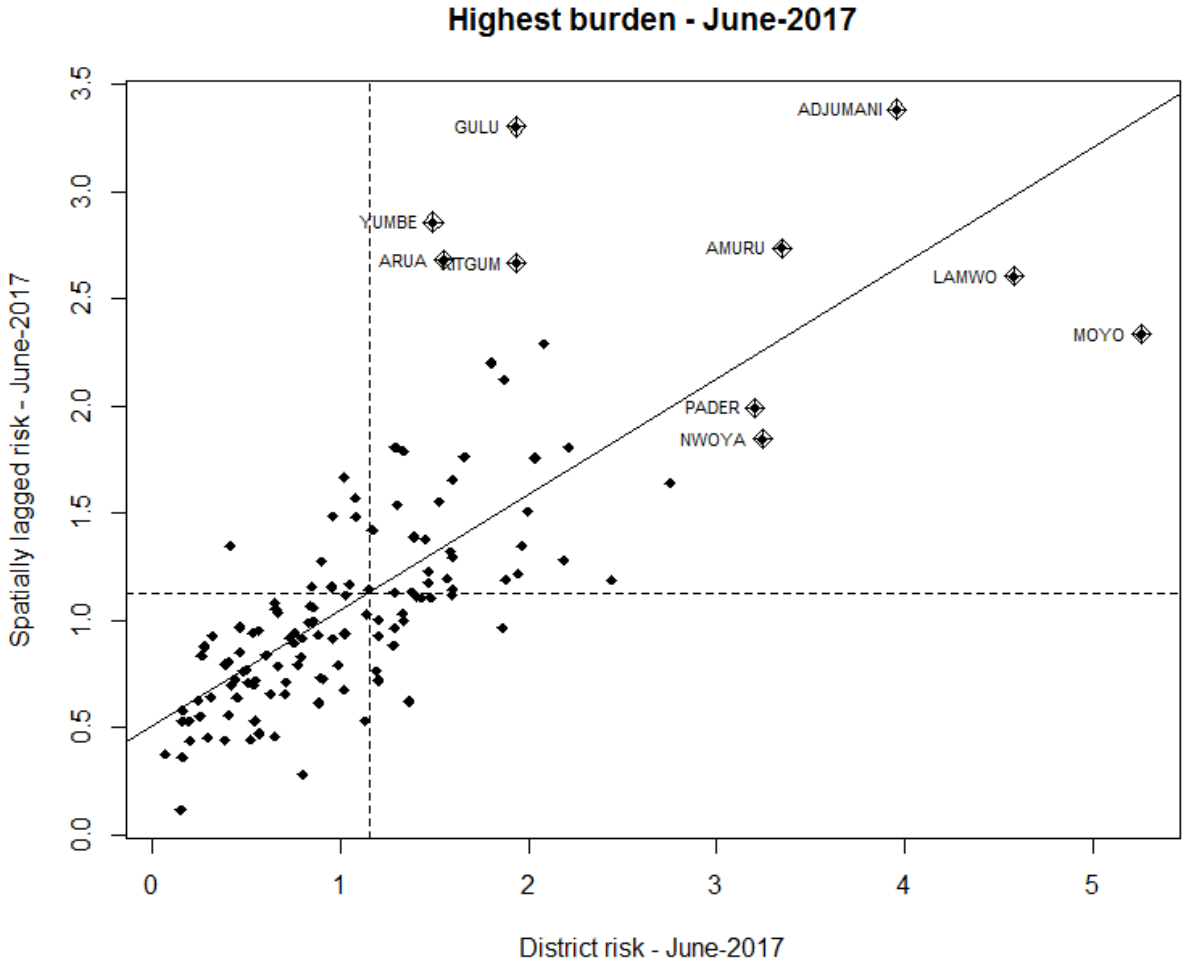


Fig S20. Moran's scatter plot for catchment-level risk of malaria during the lowest burden month of the study duration

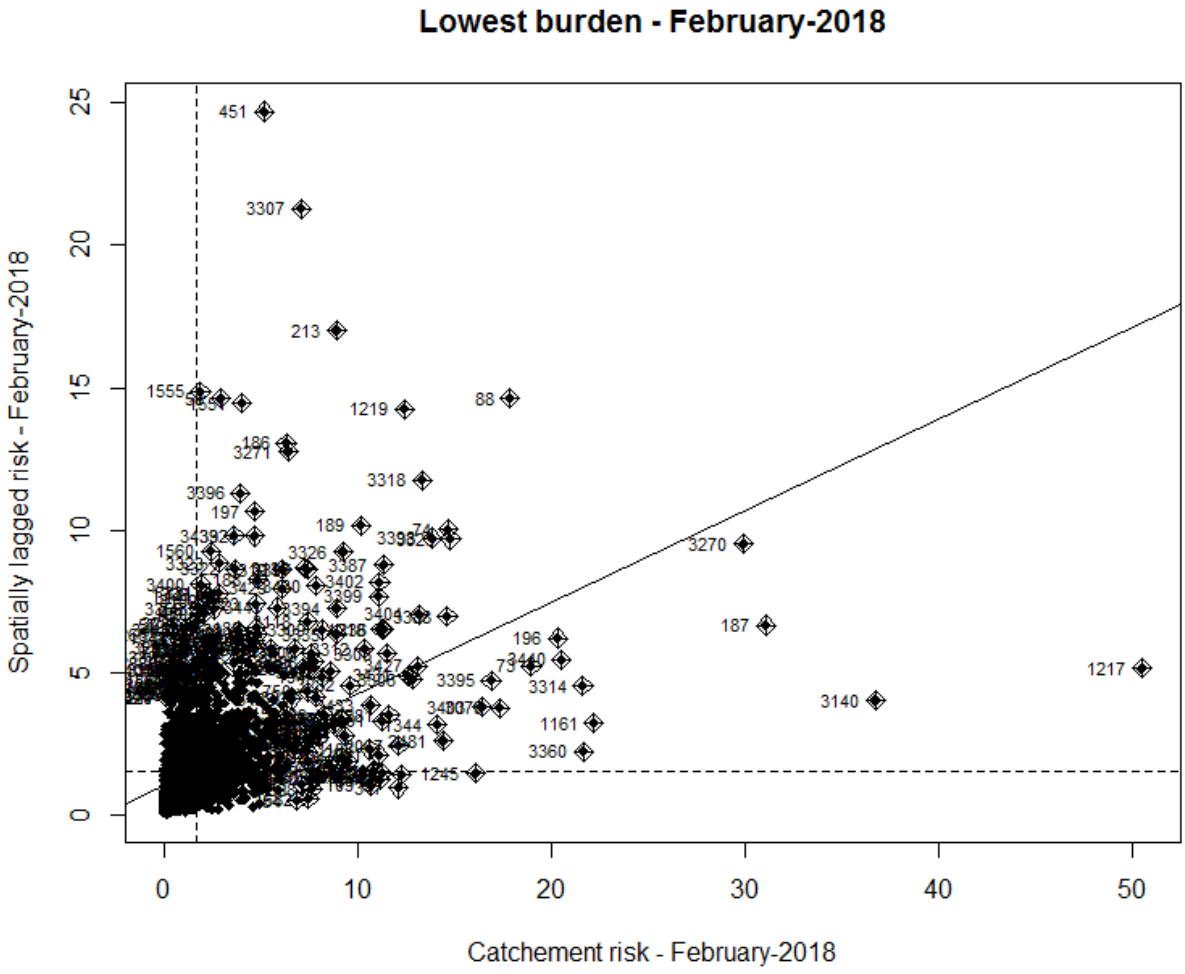
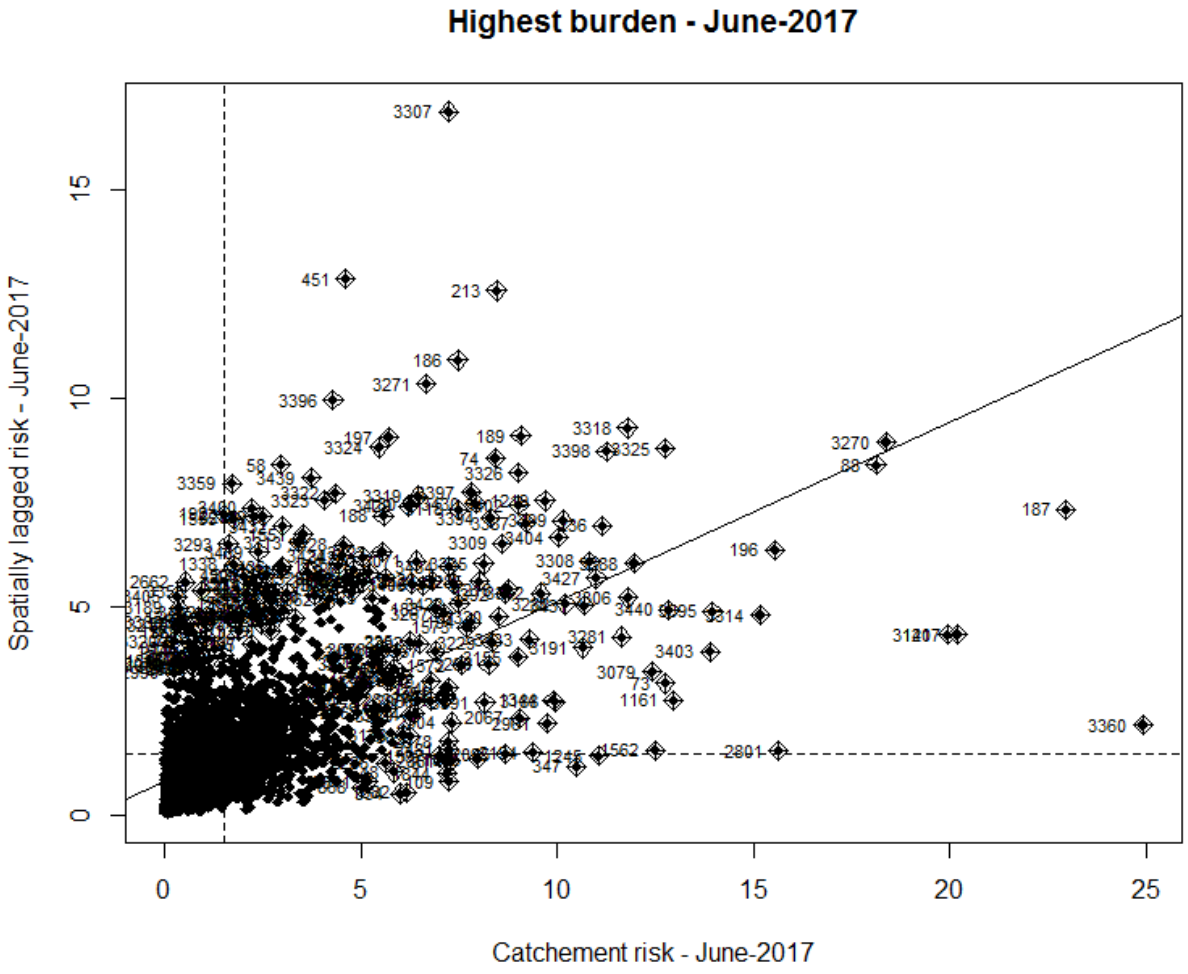
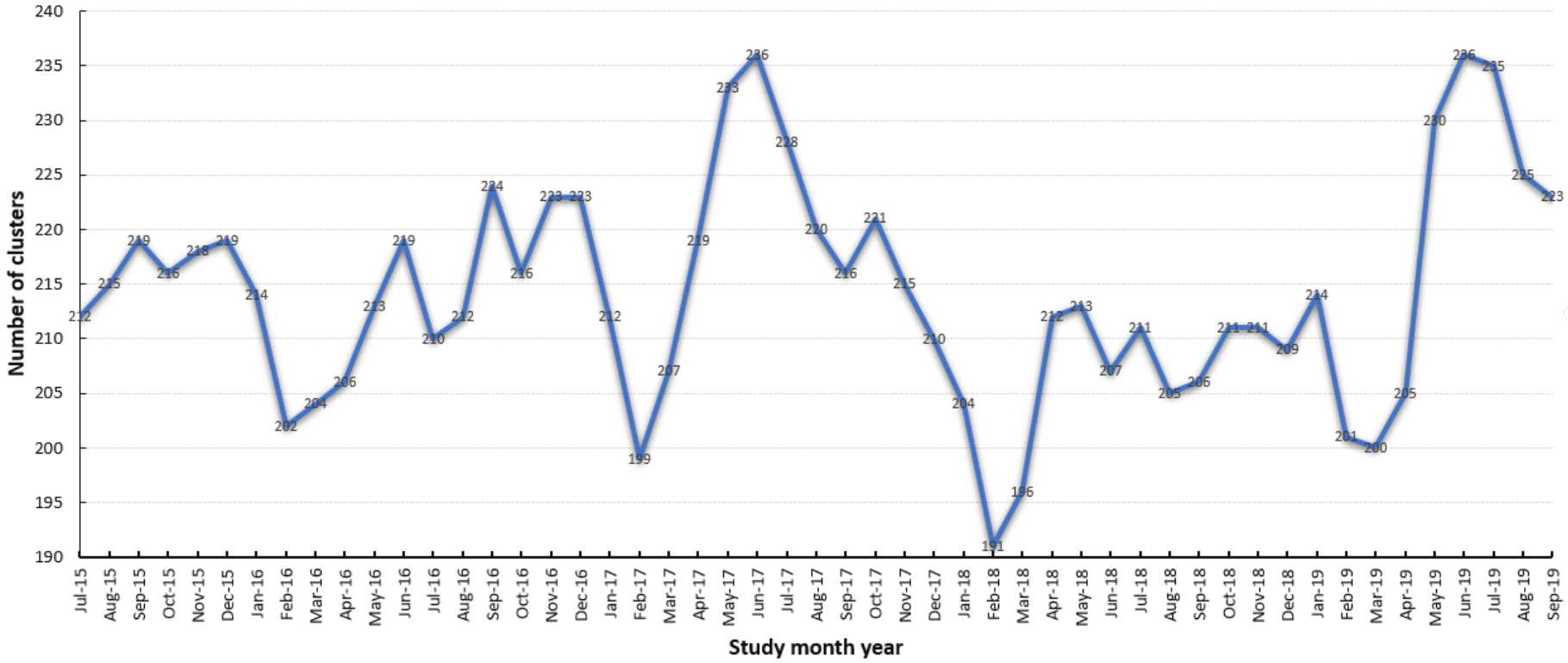


Fig S21. Moran's scatter plot for catchment-level risk of malaria during the highest burden month of the study duration



Whilst the location of HH clusters did not change much over time, the number of health facility catchments comprising the identified significant clusters of high-high risk of malaria per month varied over time, showing larger numbers of health facility catchments during high burden seasons that reduced during the low burden seasons (Fig. S22). The total number of health facility catchments was lowest during February 2018 (the lowest burden month) at 191 and highest during both June 2017 and 2019 at 236 health facility catchments.

Fig S22. Changes in number of catchments comprising monthly high-high clusters of malaria risk, identified using the Local Moran's I statistics



7 Summary, General discussion, conclusion, and Recommendation

7.1 Summary

This thesis investigated the utility of routine HMIS estimates of malaria incidence through intra-system indicator comparisons, evaluation of impacts of interventions, evaluation of bias, and identification of high burden locations. It provided substantial evidence supporting HMIS data as a viable source of reliable indicators of malaria trends and seasonality, incidence and/or burden, and ultimately risk. First, it showed agreement in trends between HMIS incidence and test positivity rates - the current common metric in use, providing evidence of its reliability for malaria trends evaluations. However, predictable change in incidence with distance of residence from health facility, absent for TPR, could be evidence of its enhanced utility, compared with TPR. Second, this research provided evidence on impact of control interventions on age distribution of malaria cases, indicating that estimation of malaria burden based on a single age group such as children under 5 years to infer burden across the full age spectrum, is likely to be a considerable source of bias. This may explain the widely acknowledged and increasingly reported differences, between national reports and global estimates. This, therefore, points to the need to strengthen routine surveillance and encourage use of this routine data for burden assessment, especially in the low resource setting countries. Third in relation to strengthened HMIS, the rare evidence of a strong linear relationship between enhanced HMIS incidence and the gold standard cohort incidence in this research setting, further highlights the viability of HMIS to generate reliable estimates of malaria burden in similar locations. Lastly, clear identification of seasonality, as well as areas of high incidence and high-risk in the perennial transmission setting of Uganda using HMIS data, is profound testament to the vital role that routine national HMIS data can play in providing low-cost indicators of malaria burden, and therefore, a vital resource for control efforts in similar settings.

7.2 Discussion of findings

Malaria remains a heavy global health challenge that is responsible for 405,000 deaths globally, a significantly high proportion of these (85%) being from the low income countries of sub-Saharan Africa as well as India [3]. Moreover, children under 5 years of age (67% of deaths in 2018) and pregnant women remain highly vulnerable population groups that suffer the most severe outcomes. There is a growing body of evidence that not only the global but also African burden reduced significantly over the first 15 of the past 20 years thanks to expanded control activity [8, 229-233] but this decline then reportedly has stalled over the past few year [6, 7, 71]. These intractable dynamics in malaria burden further emphasize the importance of routine surveillance and the understanding of the data thereof, for malaria control for the following reasons. First, our understanding of the common indicators of burden derived from routine data or their relationships with each other has been limited. Second, impacts of the scaled-up control intervention activities and/or the subsequent changes in burden or transmission levels have not been fully explored or adequately prospectively accounted for in estimating malaria

burden, with important ramifications. Third, the indicators of malaria burden from routine surveillance data have not been assessed for their representativeness or level of bias, as evaluated against their de facto gold-standard estimates. Fourth, routine surveillance estimates of burden for monitoring and impact assessment have been largely shunned, based on historical status of quality and instead, settled for less informative approaches. Responding to these issues may not be sufficiently achieved in this one thesis, however, investigations into these important subjects are vital for global public health, and especially for the low resource malaria endemic nations of the world.

Evaluating the relationship between various indicators of malaria burden derived from routine surveillance data is important in the interpretation of the burden. Test positivity rate (TPR), that is, the proportion of patients tested that have a positive result for malaria - has been very widely used for estimation of malaria burden in endemic settings, as recommended by WHO. However, other than for the low cost and ease of generating this indicator up to the lowest level of health facility, there was no defined association between TPR and incidence of malaria or other indicator [44, 45, 234]. TPR was therefore, a widely utilised indicator of burden, primarily based on convenience, rendering estimates less likely to be epidemiologically understandable. Here, I undertook to study the relationship between more common TPR and the less utilised but better understood indicator - malaria incidence rate, all derived from enhanced routine surveillance data. The temporal relationship between TPR and incidence is nonlinear and strongest when TPR values are lower than 50% [45] and this was also observed through closer trends during low transmission seasons. My original contribution to knowledge was revealing that when the spatial dimension was incorporated, the relationship was much more complex, owing to unparallel response to spatial dimensions by the two indicators. With seasonal variations in endemic settings, some of which often recording TPR's higher than 50%, coupled with the non-specific nature of symptoms indicative of suspected malaria [235], TPR has close dependence on extra-malarial factors, which may render it a less reliable indicator of malaria burden. With limited such dependences, higher sensitivity to spatial variations, and therefore environmental factors, incidence rate is a better measure of burden from routine surveillance data. Importantly, incidence rate was also shown to be significantly influenced by age, after accounting for transmission intensity.

When considering age, major research activities and routine malaria burden assessments are predominantly conducted in children, due to reported comparatively higher vulnerability to malaria and greater sensitivity to changes in transmission within this group [93]. Whilst subsequent projections of burden estimates in entire populations using these data are commonplace, particularly in sub-Saharan Africa, the impact of this approach on the true burden is less well understood. These burden estimates, however, have not gone without criticism [71]. Here, I set out to investigate the possible impacts of malaria control interventions on age-specific burden of malaria, using routine surveillance data from across diverse transmission settings. Interestingly, the risk of being positive for malaria progressively increased with effective control activities in place, while reducing among

children and vice versa. These findings are important in addressing the two vital malaria challenges including stalled declines in malaria burden both in *Plasmodium falciparum* and *vivax* endemic areas [236, 237] and the discrepancies between reported case estimates and data from individual high burden countries [71]. With effective interventions reducing the burden among population groups central to malaria burden estimation (children <5 years of age), progress may be duly registered. However, with the burden shifting to overlooked population groups (older children and adults) within the same locations, onward transmission is likely to be facilitated even among the vulnerable groups. This then sustains unintended outcomes such as high mortality rates in children under 5 years of age at 26.8% in 2017 [238]. It may also explain the discordancy between model estimates and routinely reported burden of malaria from among endemic countries [239]. My original contribution to knowledge is the revelation that with effective large-scale interventions continued, surveillance based on data from younger children may be misleading and consequently, interventions targeting these vulnerable groups may be sub-optimal. From evidence provided, however, routine surveillance data may provide a more reliable source of burden assessments.

Evaluating HMIS-based incidence rates against the gold-standard of incidence estimates from community cohort studies, was considered a vital step in understanding its representativeness of the true burden, within communities served by health facilities providing the HMIS data. By this, we could assess the level of bias in HMIS indicators, a commonly cited limitation of HMIS, relative to their benchmarks. Whilst HMIS has seen extensive quality improvement efforts associated with: improved neonatal clinical outcomes and decreased neonatal mortality in Uganda [240]; improved recording of deaths in Viet Nam [241]; and, improved HIV service delivery monitoring in Kenya [133], little has been done to evaluate HMIS indicators of burden for representativeness. Here, I set out to evaluate the relationship between HMIS and community cohort incidence rates, as well as assess bias in HMIS-based incidence due to factors associated with health facility data recording, to inform the representativeness by HMIS of the true burden. My original contribution to knowledge was the revelation of a strong linear and unbiased relationship between HMIS-based incidence and cohort incidence of malaria. This provided evidence of representativeness of HMIS-based incidence of the true burden of malaria, in these high and moderate Ugandan transmission settings.

For meaningful use of malaria control resources, the timely and continuous identification of high-burden or high-risk locations may be the most important functions of surveillance. Nevertheless, this capacity is yet to be exploited in the sub-Saharan African high transmission settings, where it is most needed and burden reduction been slowest [242]. Using well established geostatistical models to analyse incidence of confirmed malaria cases, high-burden locations were identified across the country. Whilst these approaches have been widely used and reported from multiple studies as favourable [96, 100, 101, 243], my original contribution to knowledge is the use of small area approaches to not only confirm regional stratification of malaria burden as estimated from malaria

indicator survey data [244], but concurrently define fine-resolution spatial-temporal distribution of this burden, within regions of interest across high transmission Uganda. More importantly, at this fine resolution of health facility catchments, control programs can both assess and implement control interventions, providing an effective surveillance-based approach to malaria control [102, 245]. In addition, my research revealed nation-wide strong seasonality of malaria incidence amidst a perennial transmission setting. With renewed efforts in targeting interventions for high impact [4, 160], this knowledge will facilitate improved use of surveillance data for identification of multi-scale high-risk locations, an important resource for the achievement of malaria eradication targets as we aim for elimination

7.3 Limitations

First, the limited number of sites included in the assessments performed under the first three objectives, may have limited wider generalizability of these findings, given the heterogeneity of transmission across the country. Moreover, whilst there were nearly five years of data available from community cohorts, in the assessment of the relationship between HMIS-based incidence and cohort incidence, I was limited to only three years of data, owing to available funding. This reduced, older dataset may have limited inferences amenable to the current status, however, overall applications of findings in this thesis, transcend this indicated limitation.

Second, given the temporal frequency used in all the studies in this research as is required for estimating incidence of malaria, some important metrics may have been excluded in this work. These may include, social economic status, community-initiated malaria control interventions, research associated control interventions, or other national health programs with secondary impacts on malaria transmission or burden over time. Though these may not have varied from month to month or by location, estimates of these were not accessible at either scale or duration considered in this study.

Third, the disproportionate under representation of private facilities, especially clinics and drug shops that are active care providers for uncomplicated malaria but are not actively providing reports, may have impacted findings in this research. Circumstantially, these facilities tend not to stay operational for very long for multiple reasons but either close shop or may change locations over time as observed in our study. However, during their operational time, be it limited, their routine reporting would nonetheless be beneficial for surveillance. Particularly, this status quo may disadvantage the urbanized communities where private facilities are the main care providers. Presently, therefore, routine HMIS estimates of burden in these locations remain unreliable.

Lastly, overarching influences from political and other governance associated factors were not at all accounted for in this research. Whilst these could have had varied impacts from one location to another, they were beyond the scope of this research and may have had unexplained impacts on finding in this thesis.

7.4 Future direction for HMIS-based risk mapping

This research has been foundational in its assessment of indicators of malaria burden from routinely reported health facility data, for spatial and/or temporal risk assessments. It builds on remarkable work from the malaria research community in using many other sources of data to assess risk within endemic settings, since inception of the Roll Back Malaria initiative. However, there is growing need to use timely data to inform optimal use of available effective control tools. This is especially through the global technical strategy for malaria 2016-2030 pillar of transforming surveillance into a core intervention, as well as the quest to address inconsistences in global reports. Moving forward, research is needed first, in the full and regular evaluation of the quality of routine surveillance data to facilitate assurances concerning the burden estimated from these data. Second, given findings from this study, implementation of near real-time estimates of burden from routine data accessible to policy makers as identified locations of highest burden or risk of malaria is not only necessary but possible. Such a system would support their decision making and guided allocation of treatment and disease control resources. Third, with these low-cost indicators of burden, the identification of seasonality of malaria even in perennial transmission settings will inform vector control programs, among others, on optimal timing of interventions especially IRS, which to date though most effective, is also most expensive and thereby limited. Further research on the effectiveness of surveillance supported timing of known control interventions is therefore imperative. Fourth, the increased understanding of indicators of burden and risk derived from routine data demands for additional studies on control interventions impacts using these data. Fifth, research on impacts of regular feedback from district health management offices to their respective reporting health facilities may foster increased capacity to use routine data and facilitate improved HMIS data quality and timeliness of reporting. Lastly, approaches to estimate the residual burden of malaria that is not reflected under current routine surveillance and is potentially responsible for on-ward transmission, remain ununderstood. Studies addressing this gap may facilitate improved understanding of surveillance and its utility to address these and other gaps that keep the burden of malaria high and thereby support other innovative methods to reduce it.

8 References

1. WHO, *World Malaria Report 2015*. 2015, World Health Organization: Geneva, Switzerland.
2. Gaye, O., *There should be a World Health Assembly resolution for malaria eradication.* Malar J, 2019. **18**(1): p. 352.
3. WHO, *World Malaria Report 2019* 2019, World Health Organization: Geneva, Switzerland.
4. WHO, *Global technical strategy for malaria 2016-2030*. 2015, World Health Organization: Geneva, Switzerland.
5. UNDP, *Sustainable development goal 3: Good Health and Well-being*. 2015, United Nations Development Programme: New York, USA.
6. WHO, *World Malaria Report 2017*. 2017, World Health Organization: Geneva, Switzerland.
7. Whitty, C.J.M. and E. Ansah, *Malaria control stalls in high incidence areas.* BMJ, 2019. **365**: p. l2216.
8. Bhatt, S., et al., *The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015.* Nature, 2015. **526**(7572): p. 207-11.
9. de Savigny, D. and F. Binka, *Monitoring future impact on malaria burden in sub-saharan Africa.* Am J Trop Med Hyg, 2004. **71**(2 Suppl): p. 224-31.
10. Gething, P.W., et al., *Information for decision making from imperfect national data: tracking major changes in health care use in Kenya using geostatistics.* BMC Med, 2007. **5**: p. 37.
11. Gething, P.W., et al., *Improving imperfect data from health management information systems in Africa using space-time geostatistics.* PLoS Med, 2006. **3**(6): p. e271.
12. Alegana, V.A., E.A. Okiro, and R.W. Snow, *Routine data for malaria morbidity estimation in Africa: challenges and prospects.* BMC Med, 2020. **18**(1): p. 121.
13. Bennett, A., et al., *A methodological framework for the improved use of routine health system data to evaluate national malaria control programs: evidence from Zambia.* Popul Health Metr, 2014. **12**(1): p. 30.
14. Kiberu, V.M., et al., *Strengthening district-based health reporting through the district health management information software system: the Ugandan experience.* BMC Med Inform Decis Mak, 2014. **14**: p. 40.
15. Hotchkiss, D.R., et al., *Evaluation of the Performance of Routine Information System Management (PRISM) framework: evidence from Uganda.* BMC Health Serv Res, 2010. **10**: p. 188.
16. WHO, *World Malaria Report 2005*. 2005, World Health Organization and UNICEF: Geneva, Switzerland.
17. Karunamoorthi, K., *Malaria vaccine: a future hope to curtail the global malaria burden.* Int J Prev Med, 2014. **5**(5): p. 529-38.
18. Lefevre, T., et al., *Non-genetic determinants of mosquito competence for malaria parasites.* PLoS Pathog, 2013. **9**(6): p. e1003365.

19. Meibalan, E. and M. Marti, *Biology of Malaria Transmission.* Cold Spring Harb Perspect Med, 2017. **7**(3).
20. Kelly-Hope, L.A., J. Hemingway, and F.E. McKenzie, *Environmental factors associated with the malaria vectors Anopheles gambiae and Anopheles funestus in Kenya.* Malar J, 2009. **8**: p. 268.
21. Zhou, G., et al., *Spatial relationship between adult malaria vector abundance and environmental factors in western Kenya highlands.* Am J Trop Med Hyg, 2007. **77**(1): p. 29-35.
22. Nkuo-Akenji, T., et al., *Environmental factors affecting malaria parasite prevalence in rural Bolifamba, South West Cameroon.* Afr J Health Sci, 2006. **13**(1-2): p. 40-6.
23. Castro, M.C., *Malaria Transmission and Prospects for Malaria Eradication: The Role of the Environment.* Cold Spring Harb Perspect Med, 2017. **7**(10).
24. World Health Organization. Department of Child, A.H., World Health Organization, & UNICEF, *Handbook IMCI: Integrated management of childhood illness*. 2005, World Health Organization.
25. Tangpukdee, N., et al., *Malaria diagnosis: a brief review.* Korean J Parasitol, 2009. **47**(2): p. 93-102.
26. Wellems, T.E. and C.V. Plowe, *Chloroquine-resistant malaria.* J Infect Dis, 2001. **184**(6): p. 770-6.
27. Kamya, M.R., et al., *Increasing antimalarial drug resistance in Uganda and revision of the national drug policy.* Trop Med Int Health, 2002. **7**(12): p. 1031-41.
28. WHO, *World Malaria Report 2009*. 2009, World Health Organization: Geneva, Switzerland.
29. WHO, *Guidelines for the treatment of malaria - 2nd edition*. 2010, World Health Organization: Geneva, Switzerland.
30. Murray, C.K., et al., *Rapid diagnostic testing for malaria.* Trop Med Int Health, 2003. **8**(10): p. 876-83.
31. Falade, C.O., et al., *Malaria Rapid Diagnostic Tests and Malaria Microscopy for Guiding Malaria Treatment of Uncomplicated Fevers in Nigeria and Prereferral Cases in 3 African Countries.* Clin Infect Dis, 2016. **63**(suppl 5): p. S290-S297.
32. WHO, *Guidelines for the treatment of malaria, 3rd Edition*. 2015, World Health Organization: Geneva.
33. Kyabayinze, D.J., et al., *Use of RDTs to improve malaria diagnosis and fever case management at primary health care facilities in Uganda.* Malar J, 2010. **9**: p. 200.
34. Kyabayinze, D.J., et al., *Parasite-based malaria diagnosis: are health systems in Uganda equipped enough to implement the policy?* BMC Public Health, 2012. **12**: p. 695.
35. MoH, *Uganda National Malaria Control Policy 2011*. 2011.
36. MoH, *Uganda Malaria Reduction Strategic Plan 2014-2020*. 2014, Uganda Ministry of Health: Kampala, Uganda.
37. MoH, *Annual health sector performance report 2018/2019*. 2019, Uganda Ministry of Health: Kampala.
38. MoH, *Uganda National Malaria Control Strategyic Plan 2010/11 - 2014/15*. 2011, Ministry of Health: Kampala, Uganda.

39. Berhane, A., et al., *Major Threat to Malaria Control Programs by Plasmodium falciparum Lacking Histidine-Rich Protein 2, Eritrea.* Emerg Infect Dis, 2018. **24**(3): p. 462-470.
40. Kozycki, C.T., et al., *False-negative malaria rapid diagnostic tests in Rwanda: impact of Plasmodium falciparum isolates lacking hrp2 and declining malaria transmission.* Malar J, 2017. **16**(1): p. 123.
41. Parr, J.B., et al., *Pfhrp2-Deleted Plasmodium falciparum Parasites in the Democratic Republic of the Congo: A National Cross-sectional Survey.* J Infect Dis, 2017. **216**(1): p. 36-44.
42. Gomez-Elipe, A., et al., *Forecasting malaria incidence based on monthly case reports and environmental factors in Karuzi, Burundi, 1997-2003.* Malaria Journal. **6**: p. 129.
43. Oduro, A.R., et al., *Health centre surveys as a potential tool for monitoring malaria epidemiology by area and over time.* PLoS One, 2011. **6**(11): p. e26305.
44. Jensen, T.P., et al., *Use of the slide positivity rate to estimate changes in malaria incidence in a cohort of Ugandan children.* Malaria Journal, 2009. **8**.
45. Boyce, R.M., et al., *Practical Implications of the Non-Linear Relationship between the Test Positivity Rate and Malaria Incidence.* PLoS One, 2016. **11**(3): p. e0152410.
46. Benelli, G. and J.C. Beier, *Current vector control challenges in the fight against malaria.* Acta Trop, 2017. **174**: p. 91-96.
47. Chen, I., et al., *Mass Drug Administration for Malaria: A Means to What end?* J Infect Dis, 2016. **214**(12): p. 1790-1792.
48. WHO, *Updated WHO Policy Recommendation. Intermittent Preventive Treatment of malaria in pregnancy using Sulfadoxine-Pyrimethamine (IPTp-SP)*. 2012, World Health Organization & Global Malaria Program: Geneva, Switzerland.
49. Carter, R. and K.N. Mendis, *Evolutionary and historical aspects of the burden of malaria.* Clin Microbiol Rev, 2002. **15**(4): p. 564-94.
50. WHO, *The Abuja declaration and the plan of action. Extract of: African Summit on Roll Back Malaria, 25 April 2000*, in *African Summit on Roll Back Malaria*. 2003, World Health Organization: Abuja. p. 11.
51. Okello, P.E., et al., *Variation in malaria transmission intensity in seven sites throughout Uganda.* American Journal of Tropical Medicine and Hygiene, 2006. **75**(2): p. 219-225.
52. Helinski, M.E., et al., *Entomological surveillance following a long-lasting insecticidal net universal coverage campaign in Midwestern Uganda.* Parasit Vectors, 2015. **8**: p. 458.
53. Okello, P.E., et al., *Variation in malaria transmission intensity in seven sites throughout Uganda.* Am J Trop Med Hyg, 2006. **75**(2): p. 219-25.
54. Okia, M., et al., *Insecticide resistance status of the malaria mosquitoes: Anopheles gambiae and Anopheles funestus in eastern and northern Uganda.* Malar J, 2018. **17**(1): p. 157.

55. Yeka, A., et al., *Malaria in Uganda: challenges to control on the long road to elimination: I. Epidemiology and current control efforts.* Acta Trop, 2012. **121**(3): p. 184-95.
56. Kassam, R., et al., *Narrative review of current context of malaria and management strategies in Uganda (Part I).* Acta Trop, 2015. **152**: p. 252-268.
57. MoH, M.C.P.-. *Uganda Malaria Control Strategy 2005/06 - 2009/10*. [cited 2016 20/Aug]; Available from: archiverbm.rollbackmalaria.org/countryaction/nsp/uganda.pdf.
58. Kiwuwa, M.S. and P. Mufubenga, *Use of antenatal care, maternity services, intermittent presumptive treatment and insecticide treated bed nets by pregnant women in Luwero district, Uganda.* Malar J, 2008. **7**: p. 44.
59. Macro, U.B.o.S.U.a.I., *Uganda Malaria Indicator Survey 2009*. 2010, UBOS and ICF Macro: Calverton, Maryland, USA.
60. International, U.B.o.S.U.a.I., *Uganda Malaria Indicator Survey 2014-15*. 2015, UBOS and ICF International: Kampala, Uganda, and Rockville, Maryland, USA.
61. (NMCD), U.N.M.C.D., U.B.o.S. (UBOS), and ICF, *Uganda Malaria Indicator Survey 2018-19*. 2020, NMCD, UBOS, and ICF: Kampala, Uganda, and Rockville, Maryland, USA.
62. WHO, *World Malaria Report 2015*. 2015: Geneva. p. 40.
63. WHO, *World Malaria Report 2016*. 2016: Geneva. p. 38-44.
64. Talisuna, A.O., et al., *The past, present and future use of epidemiological intelligence to plan malaria vector control and parasite prevention in Uganda.* Malar J, 2015. **14**: p. 158.
65. Muhindo, R., N. Joloba E., and D. Nakanjako, *Health Management Information System (HMIS): Whose Data is it Anyway? Contextual Challenges.* Review of Public Administration and Management, 2016. **4**(2): p. 190.
66. Shaikh, B.T. and F. Rabbani, *Health management information system: a tool to gauge patient satisfaction and quality of care.* East Mediterr Health J, 2005. **11**(1-2): p. 192-8.
67. PMI, *President's malaria initiative: Uganda, Malaria operational plan FY 2018*. 2018.
68. Sserwanga, A., et al., *Improved malaria case management through the implementation of a health facility-based sentinel site surveillance system in Uganda.* PLoS One, 2011. **6**(1): p. e16316.
69. Consortium, M., *USAID's Malaria Action Program for Districts*. 2019, Malaria Consortium: Kampala, Uganda.
70. Cibulskis, R.E., et al., *Worldwide incidence of malaria in 2009: estimates, time trends, and a critique of methods.* PLoS Medicine / Public Library of Science. **8**(12): p. e1001142.
71. WHO, *World Malaria Report 2018*. 2018, World Health Organization: Geneva, Switzerland.

72. NMCP, A. Associates, and I. Project, *An epidemiological profile of malaria and its control in Uganda.* 2013, Ministry of Health, Roll Back Malaria Partnership and Department for International Development: UK.
73. Talisuna, A.O., et al., *Intensity of malaria transmission, antimalarial-drug use and resistance in Uganda: what is the relationship between these three factors?* Trans R Soc Trop Med Hyg, 2002. **96**(3): p. 310-7.
74. Smith, D.L., et al., *Standardizing estimates of the Plasmodium falciparum parasite rate.* Malar J, 2007. **6**: p. 131.
75. Yeka, A., et al., *Factors associated with malaria parasitemia, anemia and serological responses in a spectrum of epidemiological settings in Uganda.* PLoS One, 2015. **10**(3): p. e0118901.
76. Noor, A.M., et al., *Malaria risk mapping for control in the republic of Sudan.* Am J Trop Med Hyg, 2012. **87**(6): p. 1012-21.
77. Bennett, A., et al., *Mapping malaria transmission intensity in Malawi, 2000-2010.* Am J Trop Med Hyg, 2013. **89**(5): p. 840-9.
78. Dalrymple, U., B. Mappin, and P.W. Gething, *Malaria mapping: understanding the global endemicity of falciparum and vivax malaria.* BMC Med, 2015. **13**: p. 140.
79. NMCP, KEMRI-Wellcome-Trust, and LSHTM, *Uganda: A profile of Malaria Control and Epidemiology.* 2018, Uganda Ministry of Health: Kampala, Uganda.
80. Ssempiira, J., et al., *The effect of case management and vector-control interventions on space-time patterns of malaria incidence in Uganda.* Malar J, 2018. **17**(1): p. 162.
81. WHO, *WHO Technical brief for countries preparing malaria funding requests for the Global Fund (2020-2022)*. 2020, World Health Organization: Geneva, Switzerland.
82. Guerra, C.A., et al., *Assembling a global database of malaria parasite prevalence for the Malaria Atlas Project.* Malar J, 2007. **6**: p. 17.
83. Hay, S.I. and R.W. Snow, *The malaria Atlas Project: developing global maps of malaria risk.* PLoS Med, 2006. **3**(12): p. e473.
84. Cibulskis, R.E., et al., *Estimating trends in the burden of malaria at country level.* Am J Trop Med Hyg, 2007. **77**(6 Suppl): p. 133-7.
85. Weiss, D.J., et al., *Re-examining environmental correlates of Plasmodium falciparum malaria endemicity: a data-intensive variable selection approach.* Malar J, 2015. **14**: p. 68.
86. Craig, M.H., et al., *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**: p. 44.
87. Howes, R.E., et al., *The global distribution of the Duffy blood group.* Nat Commun, 2011. **2**: p. 266.
88. Rishikesh, K. and K. Saravu, *Primaquine treatment and relapse in Plasmodium vivax malaria.* Pathog Glob Health, 2016. **110**(1): p. 1-8.

89. Dayananda, K.K., R.N. Achur, and D.C. Gowda, *Epidemiology, drug resistance, and pathophysiology of Plasmodium vivax malaria.* J Vector Borne Dis, 2018. **55**(1): p. 1-8.

90. Kleinschmidt, I., et al., *An empirical malaria distribution map for West Africa.* Trop Med Int Health, 2001. **6**(10): p. 779-86.

91. Rue, H., S. Martino, and N. Chopin, *Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplase approximations.* Journal of the Royal Statistical Society: Series B, 2009. **71**(2): p. 877-892.

92. Gosoniu, L., et al., *Bayesian modelling of geostatistical malaria risk data.* Geospat Health, 2006. **1**(1): p. 127-39.

93. Mabaso, M.L., et al., *Towards empirical description of malaria seasonality in southern Africa: the example of Zimbabwe.* Trop Med Int Health, 2005. **10**(9): p. 909-18.

94. Mabaso, M.L., et al., *Spatio-temporal analysis of the role of climate in inter-annual variation of malaria incidence in Zimbabwe.* Int J Health Geogr, 2006. **5**: p. 20.

95. Zacarias, O.P. and M. Andersson, *Spatial and temporal patterns of malaria incidence in Mozambique.* Malar J, 2011. **10**: p. 189.

96. Riedel, N., et al., *Geographical patterns and predictors of malaria risk in Zambia: Bayesian geostatistical modelling of the 2006 Zambia national malaria indicator survey (ZMIS).* Malar J, 2010. **9**: p. 37.

97. Stensgaard, A.S., et al., *Bayesian geostatistical modelling of malaria and lymphatic filariasis infections in Uganda: predictors of risk and geographical patterns of co-endemicity.* Malar J, 2011. **10**: p. 298.

98. Lorenz, C., et al., *Spatial and temporal epidemiology of malaria in extra-Amazonian regions of Brazil.* Malar J, 2015. **14**: p. 408.

99. Ssempiira, J., et al., *Interactions between climatic changes and intervention effects on malaria spatio-temporal dynamics in Uganda.* Parasite Epidemiol Control, 2018. **3**(3): p. e00070.

100. Zayeri, F., M. Salehi, and H. Pirhosseini, *Geographical mapping and Bayesian spatial modeling of malaria incidence in Sistan and Baluchistan province, Iran.* Asian Pac J Trop Med, 2011. **4**(12): p. 985-92.

101. Kazembe, L.N., *Spatial modelling and risk factors of malaria incidence in northern Malawi.* Acta Trop, 2007. **102**(2): p. 126-37.

102. Wickremasinghe, A.R., D.M. Gunawardena, and S.T. Mahawithanage, *Use of routinely collected past surveillance data in identifying and mapping high-risk areas in a malaria endemic area of Sri Lanka.* Southeast Asian J Trop Med Public Health, 2002. **33**(4): p. 678-84.

103. Nkurunziza, H., A. Gebhardt, and J. Pilz, *Bayesian modelling of the effect of climate on malaria in Burundi.* Malar J, 2010. **9**: p. 114.

104. Nkurunziza, H., A. Gebhardt, and J. Pilz, *Geo-additive modelling of malaria in Burundi.* Malar J, 2011. **10**: p. 234.

105. Kifle, M.M., et al., *Malaria Risk Stratification and Modeling the Effect of Rainfall on Malaria Incidence in Eritrea.* J Environ Public Health, 2019. **2019**: p. 7314129.
106. Group, R.M.I.E., *Evaluation of the Impact of Malaria Control Interventions on All-Cause Mortality in Children under Five Years of Age in Rwanda, 2000-2010*. 2015, USAID, President's Malaria Initiative & CDC: Kigali, Rwanda.
107. Group, M.M.I.E., *Evaluation of the Impact of Malaria Control Interventions on All-Cause Mortality in Children under Five Years of Age in Mozambique*. 2016, USAID, President's Malaria Initiative & CDC: Maputo, Mozambique.
108. Lowe, R., J. Chirombo, and A.M. Tompkins, *Relative importance of climatic, geographic and socio-economic determinants of malaria in Malawi.* Malar J, 2013. **12**: p. 416.
109. WHO, *World Malaria Report 2014*. 2014: Geneva. p. 174-175.
110. Carter, R., K.N. Mendis, and D. Roberts, *Spatial targeting of interventions against malaria.* Bull World Health Organ, 2000. **78**(12): p. 1401-11.
111. WHO. *Public Health Surveillance*. 2017 2017 [cited 2017 31 Jan.]; Available from: http://www.who.int/topics/public_health_surveillance/en/.
112. MoH, *The Health Management Information System - Health Unit Procedure Manual*. 2010, Minitry of Health - Uganda: Kampala, Uganda.
113. MoH, *National Technical Guidelines for Integrated Disease Surveillance and Response*. 2012, Ministry of Health - Uganda: Kampala, Uganda.
114. Kajungu, D., et al., *Cohort profile: The Iganga-Mayuge Health and Demographic Surveillance Site, Uganda (IMHDSS, Uganda).* Int J Epidemiol, 2020.
115. Alegana, V.A., et al., *Estimation of malaria incidence in northern Namibia in 2009 using Bayesian conditional-autoregressive spatial-temporal models.* Spat Spatiotemporal Epidemiol, 2013. **7**: p. 25-36.
116. Heywood, A. and D. Boone, *Guidelines for Data Management Standards in Routine Health Information Systems*. 2015, MEASURE Evaluation.
117. Gething, P.W., et al., *A long neglected world malaria map: Plasmodium vivax endemicity in 2010.* PLoS Negl Trop Dis, 2012. **6**(9): p. e1814.
118. Nguyen, H.V., et al., *Marked age-dependent prevalence of symptomatic and patent infections and complexity of distribution of human Plasmodium species in central Vietnam.* Am J Trop Med Hyg, 2012. **87**(6): p. 989-95.
119. Charles H. Hennekens, a.J.E.B., *Epidemiology in Medicine*, ed. S.L. Mayrent. 1987, USA: Little, Brown and Company.

120. Alemu, K., A. Worku, and Y. Berhane, *Malaria infection has spatial, temporal, and spatiotemporal heterogeneity in unstable malaria transmission areas in northwest Ethiopia.* PLoS One, 2013. **8**(11): p. e79966.

121. Krefis, A.C., et al., *Spatial analysis of land cover determinants of malaria incidence in the Ashanti Region, Ghana.* PLoS One, 2011. **6**(3): p. e17905.

122. Midekisa, A., et al., *Seasonal associations of climatic drivers and malaria in the highlands of Ethiopia.* Parasit Vectors, 2015. **8**: p. 339.

123. Musa, M.I., et al., *A climate distribution model of malaria transmission in Sudan.* Geospat Health, 2012. **7**(1): p. 27-36.

124. Bejon, P., et al., *A micro-epidemiological analysis of febrile malaria in Coastal Kenya showing hotspots within hotspots.* Elife, 2014. **3**: p. e02130.

125. Gething, P., et al., *A local space-time kriging approach applied to a national outpatient malaria dataset.* Comput Geosci, 2007. **33**(10): p. 1337-1350.

126. Ferrao, J.L., et al., *Spatio-temporal variation and socio-demographic characters of malaria in Chimoio municipality, Mozambique.* Malar J, 2016. **15**: p. 329.

127. Midekisa, A., G.B. Senay, and M.C. Wimberly, *Multisensor earth observations to characterize wetlands and malaria epidemiology in Ethiopia.* Water Resour Res, 2014. **50**(11): p. 8791-8806.

128. Wimberly, M.C., et al., *Spatial synchrony of malaria outbreaks in a highland region of Ethiopia.* Trop Med Int Health, 2012. **17**(10): p. 1192-201.

129. Alemu, K., et al., *Spatiotemporal clusters of malaria cases at village level, northwest Ethiopia.* Malar J, 2014. **13**: p. 223.

130. Ernst, K.C., et al., *Malaria hotspot areas in a highland Kenya site are consistent in epidemic and non-epidemic years and are associated with ecological factors.* Malar J, 2006. **5**: p. 78.

131. WHO, *Managerial process for National Health Development*, in *Guidelines principles for use in support of strategies for health for all by the year 2000*. 1981, World Health Organization: Geneva, Switzerland. p. 53.

132. Mutemwa, R.I., *HMIS and decision-making in Zambia: re-thinking information solutions for district health management in decentralized health systems.* Health Policy Plan, 2006. **21**(1): p. 40-52.

133. Lahuerta, M., et al., *Monitoring the transition to new antiretroviral treatment regimens through an enhanced data system in Kenya.* PLoS One, 2020. **15**(4): p. e0232104.

134. Kamuliwo, M., et al., *Spatial patterns and determinants of malaria infection during pregnancy in Zambia.* Trans R Soc Trop Med Hyg, 2015. **109**(8): p. 514-21.

135. Kagoya, H.R. and D. Kibuule, *Quality assurance of health management information systems in Kayunga district, Uganda.* African Evaluation Journal, 2018. **6**(2): p. a238.

136. MoH, *National Village Health Teams Assessment in Uganda, report 2015.* 2015.

137. Ndyomugyenyi, R., P. Magnussen, and S. Clarke, *Malaria treatment-seeking behaviour and drug prescription practices in an area of low transmission in Uganda: implications for prevention and control.* Trans R Soc Trop Med Hyg, 2007. **101**(3): p. 209-15.

138. Awor, P., et al., *Private sector drug shops in integrated community case management of malaria, pneumonia, and diarrhea in children in Uganda.* Am J Trop Med Hyg, 2012. **87**(5 Suppl): p. 92-6.

139. ACTwatch-Group, P., *ACTwatch Study Reference Document: Uganda Outlet Survey 2013.* Washington DC: PSI., 2014.

140. Littrell, M., et al., *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, 2011. **10**: p. 327.

141. Cohen, J., et al., *Determinants of malaria diagnostic uptake in the retail sector: qualitative analysis from focus groups in Uganda.* Malar J, 2015. **14**: p. 89.

142. Bawate, C., et al., *Factors affecting adherence to national malaria treatment guidelines in management of malaria among public healthcare workers in Kamuli District, Uganda.* Malar J, 2016. **15**: p. 112.

143. Cheng, Q., et al., *Plasmodium falciparum parasites lacking histidine-rich protein 2 and 3: a review and recommendations for accurate reporting.* Malar J, 2014. **13**: p. 283.

144. Johansson, E.W., *Beyond 'test and treat' - malaria diagnosis for improved pediatric fever management in sub-Saharan Africa.* Glob Health Action, 2016. **9**: p. 31744.

145. WHO. *Fact Sheet about Malaria*. 2016 Dec. 2016 [cited 2017 31 Jan.]; Available from: http://www.who.int/mediacentre/factsheets/fs094/en/.

146. WHO, *World Malaria Report 2016*. 2016: Geneva. p. 64.

147. Abeku, T.A., et al., *Spatial and temporal variations of malaria epidemic risk in Ethiopia: factors involved and implications.* Acta Trop, 2003. **87**(3): p. 331-40.

148. Bisanzio, D., et al., *Use of prospective hospital surveillance data to define spatiotemporal heterogeneity of malaria risk in coastal Kenya.* Malar J, 2015. **14**(1): p. 482.

149. Frank, C., et al., *Spatial heterogeneity of malaria in Ghana: a cross-sectional study on the association between urbanicity and the acquisition of immunity.* Malar J, 2016. **15**(1): p. 84.

150. Ndiath, M.M., et al., *Application of geographically-weighted regression analysis to assess risk factors for malaria hotspots in Keur Soce health and demographic surveillance site.* Malar J, 2015. **14**(1): p. 463.

151. Oesterholt, M.J., et al., *Spatial and temporal variation in malaria transmission in a low endemicity area in northern Tanzania.* Malar J, 2006. **5**: p. 98.

152. Siraj, A.S., et al., *Altitudinal changes in malaria incidence in highlands of Ethiopia and Colombia.* Science, 2014. **343**(6175): p. 1154-8.

153. Zacarias, O.P. and M. Andersson, *Mapping malaria incidence distribution that accounts for environmental factors in Maputo Province--Mozambique.* Malar J, 2010. **9**: p. 79.

154. Zacarias, O.P. and P. Majlender, *Comparison of infant malaria incidence in districts of Maputo province, Mozambique.* Malar J, 2011. **10**: p. 93.

155. Zhou, G., et al., *Association between climate variability and malaria epidemics in the East African highlands.* Proc Natl Acad Sci U S A, 2004. **101**(8): p. 2375-80.

156. Mlacha, Y.P., et al., *Fine scale mapping of malaria infection clusters by using routinely collected health facility data in urban Dar es Salaam, Tanzania.* Geospat Health, 2017. **12**(1): p. 494.

157. Adeola, A.M., et al., *Environmental factors and population at risk of malaria in Nkomazi municipality, South Africa.* Trop Med Int Health, 2016. **21**(5): p. 675-86.

158. Akachi, Y. and R. Atun, *Effect of investment in malaria control on child mortality in sub-Saharan Africa in 2002-2008.* PLoS One, 2011. **6**(6): p. e21309.

159. Pigott, D.M., et al., *Funding for malaria control 2006-2010: a comprehensive global assessment.* Malar J, 2012. **11**: p. 246.

160. WHO and RBM, *High burden to high impact: A targeted malaria response*. 2019, World Health Organization: Geneva, Switzerland.

161. Chandler, C.I., et al., *The impact of an intervention to introduce malaria rapid diagnostic tests on fever case management in a high transmission setting in Uganda: A mixed-methods cluster-randomized trial (PRIME).* PLoS One, 2017. **12**(3): p. e0170998.

162. Victora, C.G., et al., *Measuring impact in the Millennium Development Goal era and beyond: a new approach to large-scale effectiveness evaluations.* Lancet, 2011. **377**(9759): p. 85-95.

163. Kenea, O., et al., *Impact of combining indoor residual spraying and long-lasting insecticidal nets on Anopheles arabiensis in Ethiopia: results from a cluster randomized controlled trial.* Malar J, 2019. **18**(1): p. 182.

164. Staedke, S.G., et al., *The Impact of an Intervention to Improve Malaria Care in Public Health Centers on Health Indicators of Children in Tororo, Uganda (PRIME): A Cluster-Randomized Trial.* Am J Trop Med Hyg, 2016. **95**(2): p. 358-67.

165. Ashton, R.A., et al., *Use of Routine Health Information System Data to Evaluate Impact of Malaria Control Interventions in Zanzibar, Tanzania from 2000 to 2015.* EClinicalMedicine, 2019. **12**: p. 11-19.

166. Iyer, H.S., et al., *Impact of a district-wide health center strengthening intervention on healthcare utilization in rural Rwanda: Use of interrupted time series analysis.* PLoS One, 2017. **12**(8): p. e0182418.

167. MoH, *Annual Health Sector Performance Report - FY 2010/2011*. 2011, Ministry of Health - Uganda: Kampala, Uganda.

168. MoH, *Annual Health Sector Performance Report - FY 2011/2012*. 2012, Ministry of Health - Uganda: Kampala, Uganda.
169. MoH, *Annual Health Sector Performance Report - FY 2014/2015*. 2015, Ministry of Health - Uganda: Kampala, Uganda.
170. MoH, *National Health Facility Master List 2018*. 2018, Ministry of Health, Division of Health Information: Kampala, Uganda.
171. Consortium, M., *Sustaining Integrated Community Case Management to improve Child Health in Uganda.* 2013.
172. MoH, *The Uganda Hospital and Health Centre IV Census Survey 2014.* 2014.
173. Gladwin, J., R.A. Dixon, and T.D. Wilson, *Implementing a new health management information system in Uganda.* Health Policy Plan, 2003. **18**(2): p. 214-24.
174. WHO, *Diseases surveillance for malaria control: an operational manual.* Geneva: World Health Organization; 2012, 2012.
175. UMSP, *UMSP: Malaria Reference Centres Surveillance Report. January 2016*. 2016, Uganda Malaria Surveillance Project: Kampala, Uganda.
176. Talisuna, A., et al., *Malaria in Uganda: challenges to control on the long road to elimination. II. The path forward.* Acta Trop, 2012. **121**(3): p. 196-201.
177. Kamya, M.R., et al., *Malaria transmission, infection, and disease at three sites with varied transmission intensity in Uganda: implications for malaria control.* Am J Trop Med Hyg, 2015. **92**(5): p. 903-12.
178. Rek, J.C., et al., *Rapid improvements to rural Ugandan housing and their association with malaria from intense to reduced transmission: a cohort study.* Lancet Planet Health, 2018. **2**(2): p. e83-e94.
179. UBOS, *The National Population and Housing Census 2014 - Main Report*. 2016, Uganda Bureau of Statistics: Kampala, Uganda.
180. Atsumi, S., *More than 3,000 Congolese flee to Uganda to escape clashes*. 2012, UNHCR The UN Refugee Agency.
181. Maina, J., et al., *A spatial database of health facilities managed by the public health sector in sub Saharan Africa.* Sci Data, 2019. **6**(1): p. 134.
182. Kwiecien, R., A. Kopp-Schneider, and M. Blettner, *Concordance analysis: part 16 of a series on evaluation of scientific publications.* Dtsch Arztebl Int, 2011. **108**(30): p. 515-21.
183. Kigozi, R., et al., *Assessing temporal associations between environmental factors and malaria morbidity at varying transmission settings in Uganda.* Malar J, 2016. **15**(1): p. 511.
184. Vliegenthart-Jongbloed, K., et al., *Severity of imported malaria: protective effect of taking malaria chemoprophylaxis.* Malar J, 2013. **12**: p. 265.

185. Nisingizwe, M.P., et al., *Toward utilization of data for program management and evaluation: quality assessment of five years of health management information system data in Rwanda.* Glob Health Action, 2014. **7**: p. 25829.

186. Wagenaar, B.H., et al., *Effects of a health information system data quality intervention on concordance in Mozambique: time-series analyses from 2009-2012.* Popul Health Metr, 2015. **13**: p. 9.

187. Maokola, W., et al., *Enhancing the routine health information system in rural southern Tanzania: successes, challenges and lessons learned.* Trop Med Int Health, 2011. **16**(6): p. 721-30.

188. Mate, K.S., et al., *Challenges for routine health system data management in a large public programme to prevent mother-to-child HIV transmission in South Africa.* PLoS One, 2009. **4**(5): p. e5483.

189. Ploubidis, G.B., et al., *Social determinants of tuberculosis in Europe: a prospective ecological study.* Eur Respir J, 2012. **40**(4): p. 925-30.

190. P., C., *Representing spatial dependence and spatial discontinuity in ecological epidemiology: a scale mixture approach.* Stoch Environ Res Risk Assess, 2016. **31**: p. 291-304.

191. Okami, S. and N. Kohtake, *Spatiotemporal Modeling for Fine-Scale Maps of Regional Malaria Endemicity and Its Implications for Transitional Complexities in a Routine Surveillance Network in Western Cambodia.* Front Public Health, 2017. **5**: p. 262.

192. Rodriguez-Rodriguez, D., et al., *Mapping routine malaria incidence at village level for targeted control in Papua New Guinea.* Geospat Health, 2019. **14**(2).

193. Ver Hoef, J.M., et al., *Spatial autoregressive models for statistical inference from ecological data.* Ecological Society of America, 2017. **88**(1): p. 36-59.

194. Graetz, N., et al., *Mapping local variation in educational attainment across Africa.* Nature, 2018. **555**(7694): p. 48-53.

195. Ray, N. and S. Ebener, *AccessMod 3.0: computing geographic coverage and accessibility to health care services using anisotropic movement of patients.* Int J Health Geogr, 2008. **7**: p. 63.

196. Alegana, V.A., et al., *Spatial modelling of healthcare utilisation for treatment of fever in Namibia.* Int J Health Geogr, 2012. **11**: p. 6.

197. Aoun, N., H. Matsuda, and M. Sekiyama, *Geographical accessibility to healthcare and malnutrition in Rwanda.* Soc Sci Med, 2015. **130**: p. 135-45.

198. Ouma, P.O., et al., *Access to emergency hospital care provided by the public sector in sub-Saharan Africa in 2015: a geocoded inventory and spatial analysis.* Lancet Glob Health, 2018. **6**(3): p. e342-e350.

199. Lee, D., *A comparison of conditional autoregressive models used in Bayesian disease mapping.* Spat Spatiotemporal Epidemiol, 2011. **2**(2): p. 79-89.

200. Riebler, A., et al., *An intuitive Bayesian spatial model for disease mapping that accounts for scaling.* Stat Methods Med Res, 2016. **25**(4): p. 1145-65.

201. Goovaerts, P. and S. Gebreab, *How does Poisson kriging compare to the popular BYM model for mapping disease risks?* Int J Health Geogr, 2008. **7**: p. 6.

202. Semakula, M., F.I. Niragire, and C. Faes, *Bayesian spatio-temporal modeling of malaria risk in Rwanda.* PLoS One, 2020. **15**(9): p. e0238504.

203. Team, R.C., *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria., 2018.

204. Congdon, P., *Bayesian statistical modelling*. 2nd ed. Wiley series in probability and statistics. 2007, Chichester, England ; Hoboken, NJ: John Wiley & Sons. 304.

205. Morris, M., et al., *Bayesian hierarchical spatial models: Implementing the Besag York Mollie model in stan.* Spat Spatiotemporal Epidemiol, 2019. **31**: p. 100301.

206. Bernardinelli, L., D. Clayton, and C. Montomoli, *Bayesian estimates of disease maps: how important are priors?* Stat Med, 1995. **14**(21-22): p. 2411-31.

207. Gething, P.W., et al., *A new world malaria map: Plasmodium falciparum endemicity in 2010.* Malar J, 2011. **10**: p. 378.

208. Nihei, N., et al., *Analysis of malaria endemic areas on the Indochina Peninsula using remote sensing.* Jpn J Infect Dis, 2002. **55**(5): p. 160-6.

209. Cohen, J.M., et al., *Local topographic wetness indices predict household malaria risk better than land-use and land-cover in the western Kenya highlands.* Malar J, 2010. **9**: p. 328.

210. J, P., et al., *Analysing Spatial-Temporal Clustering of Meningococcal Meningitis Outbreaks in Niger Reveals Opportunities for Improved Disease Control.* PLoS Negl Trop Dis, 2012. **6**(3): p. e1577.

211. Kigozi, R., et al., *Indoor residual spraying of insecticide and malaria morbidity in a high transmission intensity area of Uganda.* PLoS One, 2012. **7**(8): p. e42857.

212. Bodker, R., et al., *Relationship between altitude and intensity of malaria transmission in the Usambara Mountains, Tanzania.* J Med Entomol, 2003. **40**(5): p. 706-17.

213. Steinhardt, L.C., et al., *The effect of indoor residual spraying on malaria and anemia in a high-transmission area of northern Uganda.* Am J Trop Med Hyg, 2013. **88**(5): p. 855-61.

214. Tugume, A., et al., *Effects and factors associated with indoor residual spraying with Actellic 300 CS on malaria morbidity in Lira District, Northern Uganda.* Malar J, 2019. **18**(1): p. 44.

215. Hay, S.I., et al., *Urbanization, malaria transmission and disease burden in Africa.* Nat Rev Microbiol, 2005. **3**(1): p. 81-90.

216. Kabaria, C.W., et al., *The impact of urbanization and population density on childhood Plasmodium falciparum parasite prevalence rates in Africa.* Malar J, 2017. **16**(1): p. 49.

217. Kigozi, S.P., et al., *Rapid shifts in the age-specific burden of malaria following successful control interventions in four regions of Uganda.* Malar J, 2020. **19**(1): p. 128.

218. Mouchet, J., et al., *Evolution of malaria in Africa for the past 40 years: impact of climatic and human factors.* J Am Mosq Control Assoc, 1998. **14**(2): p. 121-30.
219. Simple, O., et al., *Influence of Climatic Factors on Malaria Epidemic in Gulu District, Northern Uganda: A 10-Year Retrospective Study.* Malar Res Treat, 2018. **2018**: p. 5482136.
220. Tukei, B.B., A. Beke, and H. Lamadrid-Figueroa, *Assessing the effect of indoor residual spraying (IRS) on malaria morbidity in Northern Uganda: a before and after study.* Malar J, 2017. **16**(1): p. 4.
221. Keiser, J., et al., *Urbanization in sub-saharan Africa and implication for malaria control.* Am J Trop Med Hyg, 2004. **71**(2 Suppl): p. 118-27.
222. Hamel, M.J., et al., *A reversal in reductions of child mortality in western Kenya, 2003-2009.* Am J Trop Med Hyg, 2011. **85**(4): p. 597-605.
223. Bousema, T., et al., *The Impact of Hotspot-Targeted Interventions on Malaria Transmission in Rachuonyo South District in the Western Kenyan Highlands: A Cluster-Randomized Controlled Trial.* PLoS Med, 2016. **13**(4): p. e1001993.
224. Tatem, A.J., *Innovation to impact in spatial epidemiology.* BMC Med, 2018. **16**(1): p. 209.
225. Amuge, B., et al., *Health-seeking behavior for malaria among child and adult headed households in Rakai district, Uganda.* Afr Health Sci, 2004. **4**(2): p. 119-24.
226. Menon, M.P., J.D. Njau, and D.A. McFarland, *Cost and Predictors of Care-Seeking Behaviors Among Caregivers of Febrile Children-Uganda, 2009.* Am J Trop Med Hyg, 2016. **94**(4): p. 932-7.
227. Tumwesigire, S. and S. Watson, *Health seeking behavior by families of children suspected to have malaria in Kabale: Uganda.* Afr Health Sci, 2002. **2**(3): p. 94-8.
228. Cooper, L., et al., *Pareto rules for malaria super-spreaders and super-spreading.* Nat Commun, 2019. **10**(1): p. 3939.
229. Aregawi, M., et al., *Effect of anti-malarial interventions on trends of malaria cases, hospital admissions and deaths, 2005-2015, Ghana.* Malar J, 2017. **16**(1): p. 177.
230. Karema, C., et al., *Trends in malaria cases, hospital admissions and deaths following scale-up of anti-malarial interventions, 2000-2010, Rwanda.* Malar J, 2012. **11**: p. 236.
231. Eckert, E., et al., *Impact Evaluation of Malaria Control Interventions on Morbidity and All-Cause Child Mortality in Rwanda, 2000-2010.* Am J Trop Med Hyg, 2017. **97**(3_Suppl): p. 99-110.
232. Ceesay, S.J., et al., *Changes in malaria indices between 1999 and 2007 in The Gambia: a retrospective analysis.* Lancet, 2008. **372**(9649): p. 1545-54.
233. Teklehaimanot, H.D., et al., *Malaria in Sao Tome and principe: on the brink of elimination after three years of effective antimalarial measures.* Am J Trop Med Hyg, 2009. **80**(1): p. 133-40.
234. WHO, *World Malaria Report 2008*. 2008, World Health Organization: Geneva, Switzerland.
235. Taylor, S.M., et al., *Does this patient have malaria?* JAMA, 2010. **304**(18): p. 2048-56.

236. Battle, K.E., et al., *Mapping the global endemicity and clinical burden of Plasmodium vivax, 2000-17: a spatial and temporal modelling study.* Lancet, 2019. **394**(10195): p. 332-343.
237. Mbacham, W.F., et al., *Current Situation of Malaria in Africa.* Methods Mol Biol, 2019. **2013**: p. 29-44.
238. MoH, *Annual Health Sector Performance Report - FY 2016/2017*. 2017, Ministry of Health - Uganda: Kampala, Uganda.
239. Kamau, A., et al., *A systematic review of changing malaria disease burden in sub-Saharan Africa since 2000: comparing model predictions and empirical observations.* BMC Med, 2020. **18**(1): p. 94.
240. Achan, J., et al., *Improving the quality of neonatal data capture and clinical care at a tertiary-care hospital in Uganda through enhanced surveillance, training and mentorship.* Paediatr Int Child Health, 2020. **40**(2): p. 92-104.
241. Hong, T.T., et al., *Completeness and reliability of mortality data in Viet Nam: Implications for the national routine health management information system.* PLoS One, 2018. **13**(1): p. e0190755.
242. Cibulskis, R.E., et al., *Malaria: Global progress 2000 - 2015 and future challenges.* Infect Dis Poverty, 2016. **5**(1): p. 61.
243. Giardina, F., J. Franke, and P. Vounatsou, *Geostatistical modelling of the malaria risk in Mozambique: effect of the spatial resolution when using remotely-sensed imagery.* Geospat Health, 2015. **10**(2): p. 333.
244. Ssempiira, J., et al., *Geostatistical modelling of malaria indicator survey data to assess the effects of interventions on the geographical distribution of malaria prevalence in children less than 5 years in Uganda.* PLoS One, 2017. **12**(4): p. e0174948.
245. Wallis, D., et al., *Elevated serum anti-flagellin antibodies implicate subclinical bowel inflammation in ankylosing spondylitis: an observational study.* Arthritis Res Ther, 2013. **15**(5): p. R166.

9 Appendices

9.1 Appendix 1: Summary of literature on HMIS use in geo-spatial assessments for malaria

Results from literature review: Following the review of literature to evaluate the use of routine health facility data for mapping for malaria control, various categories of commonly used malaria data were identified, and these are presented in table 1 below.

Table 1. *Defined categories of subjects covered in malaria mapping for control and number of literature articles found, covering Jan-1990 through Feb-2016*

Category		Number of articles found
Malaria Mapping with:	HMIS	39
	Active case detection (cohort & cross-sectional)	52
	Aggregated & national (MIS/DHS) surveys	42
	LLINs	19
	Modelling + import/export	18
	Entomology and Mosquito density	201
Malaria studies non-spatial		13
GIS methodology		12
Severe malaria		8
Malaria Mortality		11
Reviews		7
Fever		1
Abstract and Full text of article not found		1
Global and continental malaria		14
Other		226

Table 2. *Breakdown of the literature reviewed for an evaluation of use of HMIS data in risk mapping for malaria in Sub-Saharan Africa*

Author	Year published	Location / study site	Aim	Data	Age of participants
Mabaso et al.	2005	Zimbabwe	Describe relationship between seasonality and incidence of malaria using environmental factors to define spatial variation in seasonality and derive seasonality concentration index	Monthly malaria cases at district level	Under 5 years
Mabaso et al.	2006	Zimbabwe	Describe year to year variation in climatic risk factors to support development of malaria early warning systems and determine areas prone to climate-driven epidemics, using Bayesian spatio-temporal analysis	Monthly malaria cases at district level	Under 5 years
Zacarius et al.	2011	Mozambique	Study influence of climatic explanatory variables on malaria incidence level	Monthly malaria cases at district level (2007 & 2008)	0 - 4 years
Krefis et al.	2011	Ashanti in Ghana	Investigate association between malaria incidence and classes of land cover that could influence vector abundance and human population density	Children visiting selected clinics over 2007-2008 (18months)	Under 15 years
Alegana et al.	2013	Northern Namibia	Predict malaria incidence at second administrative levels and Adjust public health utilization rates to estimate catchment population by a novel approach	Jan-Dec 2009	All ages
Alemu et al.	2013	North-western Ethiopia	Detect purely spatial, temporal, and space-time clusters at district levels	2003-2012 district aggregates	All ages
Bejon et al.	2014	Kilifi district - Coastal Kenya	Determine temporal and spatial scales of case clustering to inform targeting in malaria control	9 years, not specified	Under 5 years
Abeku et al.	2003	Ethiopia	Describe spatial and temporal variations in malaria epidemic risk in Ethiopia and examine factors involved in relation to implications for early warning and interpretation of geographical risk models	Data from 50 out of 59 sectors over Sep-1986 through Aug-1993	All ages
Alemu et al.	2014	North-western Ethiopia	Describe spatial and temporal patterns of malaria transmission and identify drivers of the spatio-temporal patterns in high-altitude villages of northwest Ethiopia with very low transmission intensity	Malaria case data in prospective survey at 4 clinics in the district for Aug-2012 to May-2013	All ages

Author	Year published	Location / study site	Aim	Data	Age of participants
Bennett et al.	2014	Zambia	Evaluate association between ITN program intensity and malaria incidence using a dose-response ecological analysis	District monthly data for 2009-2011	All ages
Bisanzio et al.	2015	Coastal Kenya	Calculate fraction of fevers due to malaria and describe the space-time pattern of malaria occurrence. Then identify areas of high non-malaria fever illness and assess HMIS use to capture short- and long-term effects of LLIN distribution	Oct-2012 to Mar-2015	All ages
Ernst et al.	2006	Western Kenya	Assess spatial patterns of malaria incidence and spatial distribution of ecological risk factors of malaria in an epidemic prone highland area	Data from one clinic for 2001-2004	All ages
Kamuliwo et al.	2015	Zambia	Study the burden of malaria in pregnancy, its spatial distribution and risk factors in the country.	District monthly data for 2009-2014, And IPTp data from malaria indicator survey of 2012	All pregnant women
Frank et al.	2016	Ghana	Assess the association between urbanicity and malaria and how this may influence development of immunity on a micro-epidemiological level	Data for 2012 from prospective survey	Under 15 years
Kazembe et al.	2007	Northern Malawi	Investigate spatial distribution of malaria using incidence data reported through HMIS	Jan-2002 to Dec-2003	Under 5 years
Lowe et al.	2013	Malawi	Investigate the spatial & inter-annual variations in malaria morbidity & determine how much, if any, of the inter-annual variability is due to climatic variability relative to other non-climatic factors	District level aggregates for Jul-2004 to Jun-2011	All ages
Midekisa et al.	2014	Amhara region - Ethiopia	Quantify the spatial distribution of herbaceous wetlands and test their association with malaria transmission at regional scale in Amhara region of Ethiopia	Outpatient cases for 2007-2009	All ages
Midekisa et al.	2015	Highlands in Amhara region - Ethiopia	Assess the effect of satellite derived climate variables summarised over different seasons of the year	District aggregates for 2001-2009	All ages
Zhou G. et al.	2004	7 east African highlands sites in Kenya, Uganda, and Ethiopia	Investigate the association between climatic variability and number of monthly malaria outpatients over 10-20 years in highlands where epidemics have been reported.	1978-1998	All ages for 6 sites and <15 years for one site

Author	Year published	Location / study site	Aim	Data	Age of participants
Nkurunziza H. et al.	2010	Burundi	Assess the climatic factors that are highly associated with monthly malaria incidence in Burundi using two models generalised linear and generalised additive mixed models	Data from all over the country for 1996 to 2007. Nearest neighbour method was used to fill in the 5% missing data	All ages
Zacarius O. et al.	2010	Mozambique	Identify important predictors and generate a malaria distribution map of Maputo using spatial statistical analysis of malaria incidence	District summaries from NMCP for 8 administrative districts with 92 health facilities for 2001-2002	All ages
Nkurunziza H. et al.	2011	Burundi	Use a geo-additive model to understand dependence of malaria cases on spatial effects and climatic covariates (rain, temperature, and humidity) in Burundi	12years (1996-2007) monthly data by province	All ages
Zacarius P. et al.	2011	Maputo province in Mozambique	Conduct spatial statistical analysis of malaria incidence to identify important predictor variables and generate a malaria distribution map of Maputo province	District summaries for 1999 to 2008	All ages
Wimberly M.C. et al.	2012	Ethiopian highlands	Assess consistency of indicators of temporal variability in malaria risk Test for presence of spatial and temporal patterns and their synchrony with variations in malaria cases.	Monthly district summaries for 2001 to 2009	All ages
Musa M. et al.	2012	Sudan	Create suitability maps through extension of Boolean logic called fuzzy logic that has no clear outcome	State level aggregates for 2004-2010 from 15 states	All ages
Siraj A.S. et al.	2014	Ethiopia and Columbia	Find evidence of changing spatial distribution of malaria with varying temperature of 10 years in highland regions of Ethiopia and Columbia	159 administrative units' summaries for 1993-2005 in Ethiopia, and 124 municipalities' summaries for 1990-2005	All ages
Oesterholt M. et al.	2006	Msitu village in Tanzania	Identify high-risk areas where interventions can be focused, by identification of micro-environmental factors influencing malaria risk	Clinic data for 2004	All ages
Ndiath et at.	2015	Ndoffane district in Senegal	Investigates malaria hotspots by using geographically weighted regression.	Malaria data for June-Dec 2013 in prospective survey	All ages
Gething et al.	2007	Kenya	Examine the effect on prediction accuracy of extension of spatial-only to space-time prediction approach, and replacement of stationary space-	District summaries	All ages

Author	Year published	Location / study site	Aim	Data	Age of participants
			time fandom function with a locally varying space-time random function.		

Table 3. *Summary of Results from literature that utilised routine HMIS data for malaria risk assessment*

Author	Case definition	Response variable used	Explanatory variables considered	Analysis and modelling approach
Mabaso et al. (2005)	Confirmed + Presumptive	Incidence per capita	Rainfall, Min temperature, Mean temperature, Max temperature, Vapour pressure, NDVI All lagged two months	Poisson for predictor selection, Mapping seasonality concentration index Account for spatial and temporal correlation using random effects Regression parameters estimation using Markov Chain Monte Carlo algorithm
Mabaso et al. (2006)	Confirmed + Presumptive	Incidence rate per 1000 persons	Rainfall, Min temperature, Mean temperature, Max temperature, Vapour pressure, NDVI	Bayesian negative binomial for selection of predictors Seasonality indexing approaches for year to year variations in data Account for spatial and temporal correlation using random effects Regression parameters estimation using Markov Chain Monte Carlo algorithm
Zacarius et al. (2011)	Confirmed + Presumptive	Observed cases, and Expected cases being district population each year	Rainfall, Min temperature, Mean Temperature, Max Temperature, Humidity	Bayesian hierarchical models in the presence of temporal and spatial correlations. Account for spatial correlation using conditional autoregressive approach, and temporal correlation by random walk. Consider lag between consecutive months
Krefis et al. (2011)	Confirmed	Village-level Incidence per year per 1000	NDVI with Vegetation classifications of: Banana, cocoa, palm, orange, swamp area, water, deforested area, road, built- up areas, and Population density	Poisson regression for predictor selection, then Spearman's rank correlation for cross correlation between predictors. Poisson regression for sensitivity analysis with highly significant predictors included.
Alegana et al. (2013)	Confirmed + Presumptive adjusted for test positivity rate	Incidence as cases/population at risk	Annual mean enhanced vegetation index, Monthly precipitation, Temperature suitability index, Proportion of urban population. All resampled to 1x1Km spatial resolution	Poisson regression for predictor selection. Significant predictors were fed into a Bayesian spatio-temporal zero-inflated conditional autoregressive model using integrated nested Laplace approximation
Alemu et al. (2013)	Not specific	Incidence for census tract polygons as cases/estimated mid-year population i.e. population at risk	None	Poisson model in Sat Scan using temporal, spatial and space-time scan statistics. A circular window was used for spatial scan and a cylindrical window for space-time scans, as well as for purely time scans and maximum likelihood indicated the most likely cluster verified by a p-value. Scans identified both observed and expected observations inside a window.

Author	Case definition	Response variable used	Explanatory variables considered	Analysis and modelling approach
Bejon et al. (2014)	Cases: Confirmed, Controls: febrile with negative results	Malaria positive fraction as number of cases divided by febrile cases	LLIN use for within-identified hotspots	Bernoulli model in Sat Scan with full data Re-analysed with children within identified hotspots. Circular moving window centred at each homestead was used
Abeku et al. (2003)	Confirmed	Epidemic status as log(cases) for sector and month. An epidemic was when log(cases) exceeded historical expected value by one standard deviation for at least 3 consecutive months and low incidence if it was less by the same amount and duration.	Abnormal weather conditions defined for sector and month as high, if value exceeds the historical expected by one standard deviation and vice versa for: Rainfall, Min temperature, and Max temperature	Using Chi-square test with 1 degree of freedom. Testing weather presence of abnormal weather in 3months preceding onset of epidemic or abnormally low incidence differed from expected
Zacarius O. et al. (2010)	Confirmed + Presumptive	Number of cases	Rain-precipitation, Temperature, Humidity, Vegetation, Stationary water pools, Human vector interaction	A Poisson distribution was assumed for observed case counts. A conditional autoregressive model was used with inverse Gamma distribution assumed for the priors.
Bennett et al. (2014)	Confirmed + Presumptive	Number of malaria cases, and Incidence as number of cases per 1000 population summarised as annual parasite index	ITN coverage, Standardised treatment seeking rates, Standardised percentage population 2hours from a facility, Standardised monthly reporting rates, Standardised testing rates, Standardised anomalies of: Enhanced vegetation index, Max and Max temperature and Rainfall	Poisson and negative binomial models to assess association between ITN coverage per district and response variables. Bayesian framework using integrated nested Laplace approximation. Model fit was compared using deviance information criteria
Bisanzio et al. (2015)	Confirmed	Probability of a febrile case being positive for malaria	Population size, Development category (more/less), Rainfall, Mosquito abundance based on select villages, Gender, Distance to shoreline, Presence of rice fields, and Development	Clusters of high and low febrile illness prevalence identified using Getis' local statistic. Structured additive regression models used to quantify contribution of patient demographics, village environmental characteristics, and seasonality to probability of a febrile case being malaria. Linear predictors and interaction terms were included in the model, as well as spatially correlated and unstructured random effects.

Author	Case definition	Response variable used	Explanatory variables considered	Analysis and modelling approach
Ernst et al. (2006)	Cases: Confirmed Controls: Others	Age adjusted incidence for hexagonal sub-units of area	Elevation, Distance to swamp	Risk ratio between highest and lowest incidence sub-units, Sat Scan used to identify clusters with a Poisson model, GEE used in individual and household analyses adjusting for correlation. Explanatory variables were added one at a time and log-likelihood test used to choose them
Kamuliwo et al. (2015)	Confirmed + Presumptive	Number of cases of malaria in pregnancy	Water bodies, roads/railroads, LLIN coverage	Maps for each year developed using cluster and outlier analysis and compared. Local Moran's Index, z-score & corresponding p-value, as well as the outlier type were generated
Kazembe et al. (2007)	Confirmed + Presumptive	Standardised incidence ratio	Total annual precipitation, total annual evapotranspiration, mean min annual temperature, mean max annual temperature, soil water holding capacity, altitude	Collinearity assessed between all possible covariate pairs and where found, the least biologically viable covariate dropped. Poisson regression used to select candidate covariates. Non-linear relation between malaria incidence and continuous covariates assessed using scatter plots. SIR plotted to investigate spatial variation in risk, and smoothed estimates of SIR produced using conditional autoregressive approach
Lowe et al. (2013)	Confirmed + Presumptive	Age stratified counts of cases per month	Socio-economic factors, Population density, urban dwelling proportion, Proportion of health facilities, ITN distribution, Housing, Sanitation and literacy levels, Precipitation, Temperature, Altitude	Maximal fixed effects model in a negative binomial generalised linear model framework was used to select explanatory variables Lags and polynomial terms were also included. Stepwise model selection based on Akaike's information criteria and removing non-significant interaction terms
Midekisa et al. (2014)	Confirmed + Presumptive	Incidence as a proportion of number of cases to total population per sub region	Rainfall estimate, Land surface temperature, Enhanced vegetation index, Evapotranspiration, Percent of herbaceous wetlands	Percent of herbaceous wetlands included to account for spatial variability. Natural log of total population included to account for spatial variability in the population. A conceptual model of cascading seasonal effects implemented
Zhou G. et al. (2004)	Not specific	From residuals of number monthly malaria cases (less effects of autocorrelation and seasonality), an epidemic measure was derived at a threshold of the average over past 5years + 2 standard deviations	Temperature, Rainfall	t-test to compare average monthly min, max temperature, and rainfall between two seasons of 1978-1988 and 1989-1998. A two-step approach: 1) Climatic variability playing no role and using forward stepwise regression. 2) Accounting for predicted effects of auto-regression and seasonality in monthly malaria cases then performed stepwise multiple regression analysis.

Author	Case definition	Response variable used	Explanatory variables considered	Analysis and modelling approach
Nkurunziza H. et al. (2010)	Not specific	Incidence as a proportion of number of cases to total population per province	Monthly precipitation, Monthly average maximum temperature, Monthly average maximum humidity, Monthly average minimum humidity	GLM: Metrics selected using Akaike's information criteria by stepwise algorithm. Regression coefficients estimated using Markov Chain Monte Carlo simulation, GAM: Metrics selected using Akaike's information criteria by simultaneous selection of variables. Interaction terms considered small and omitted
Nkurunziza H. et al. (2011)	Not specific	Incidence as Number of cases/Total provincial population	Monthly cumulative precipitation, Monthly average max & min temperature, Monthly max & min humidity	Generalised additive mixed model used for decision on explanatory variables Full analysis implemented using Geo-additive model incorporating spatial effects Spatial effect in two parts 1) structured or correlated and 2) structured random effect
Zacarius P. et al. (2011)	Presumptive + confirmed	None defined	Monthly average maximum temperature Monthly average rainfall	Poisson model to choose explanatory variables. Hierarchical model with 3 levels 1) using a simple Poisson model 2) Including environmental covariates and 3) including an intercept term
Wimberly M.C. et al. (2012)	Presumptive + confirmed	Number of cases	Proportion of outpatients with malaria as total malaria cases/ total outpatient visits, confirmed malaria cases, Proportion of confirmed cases that were P. falciparum	Likelihood ration tests used to determine significant seasonal and inter-annual effects. Spatial autocorrelation was quantified using Moran's I statistic for each year.
		Log of total number of cases	District, Year, Month	A mixed effects model was fitted with district, year and month as random effects. Random effects plotted to determine time of greatest deviation from global mean. Spline correlogram used to measure spatially lagged correlation between logs of cases time series at varied spatial scales.
Musa M. et al. (2012)	Not specific	Malaria case rate as Number of cases in a state/total state population	Interpolated quantities of: Rainfall, Temperature, Humidity from observation stations values	Two maps were generated 1) Case rate by state using thin plate spline interpolation and 2) predictive map using a suitability climate model. The two maps were compared using a condition that if 2/3 of area is within 1/3 difference, then maps are fairly similar

Author	Case definition	Response variable used	Explanatory variables considered	Analysis and modelling approach
Siraj A.S. et al. (2014)	Not specific	Number of cases	Temperature, Season, Altitude	Negative binomial linear models were used to assess explanatory variables, Multiple models with different numbers of covariates and their interaction terms were evaluated using Akaike's information criteria
Oesterholt M. et al. (2006)	Confirmed by microscopy	Incidence	Rainfall, Temperature, Season (3months), Distance to river, Housing condition, Age	Logistic regression, t-test, and single-way ANOVA, among many others

9.2 Appendix 2a: School of Medicine Research Ethics Committee Initial Approval

MAKERERE UNIVERSITY

P.O. Box 7072 Kampala, Uganda
E-mail: rresearch9@gmail.com

Phone: 256 414 533541
Fax: 256 414 541036/0414 532204

COLLEGE OF HEALTH SCIENCES
SCHOOL OF MEDICINE

RESEARCH ETHICS COMMITTEE

September 20, 2017

Mr. Simon Peter Kigozi
IDRC

Category of Review
[X] Initial review
[ ] Continuing review
[ ] Amendment
[ ] Termination of study
[ ] SAEs

Dear Mr. Kigozi,

Re: Approval of proposal #REC REF 2017-119

"Evaluating routine HMIS estimates of malaria incidence compared to cohort incidence for malaria mapping in Uganda" Version 1.1

Thank you for submitting an application for approval of the above – referenced **proposal**. The committee reviewed it and granted approval for one year, effective September 27th, 2017. Approval will expire on September 26th, 2018.

Continuing Review
In order to continue work on this study (including data analysis) beyond the expiration date, the School of Medicine Research and Ethics Committee must reapprove the protocol after conducting a substantive, meaningful, continuing review. This means that you must submit a continuing report form as a request for continuing review. To best avoid a lapse, you should submit the request six (6) to eight (8) weeks before the lapse date. Please use the forms supplied by our office.

Amendments
During the approval period, if you propose any change to the protocol such as its funding source, recruiting materials, or consent documents, you must seek School of Medicine Research and Ethics Committee approval before implementing it.
Please summarize the proposed change and the rationale for it in a letter to the School of Medicine Research and Ethics Committee. In addition, submit three (3) copies of an updated version of your original protocol application- one showing all proposed changes in bold or 'track changes,' and the other without bold or track changes.

Reporting
Other events which must be reported promptly in writing to the School of Medicine Research and Ethics Committee include: Suspension or termination of the protocol by you or the grantor. Unexpected problems involving risk to participants or others

Adverse events, including unanticipated or anticipated but severe physical harm to participants.

Do not hesitate to contact us if you have any questions. Thank you for your cooperation and commitment to the protection of human subjects in research.

Final approval is to be granted by Uganda National Council for Science and Technology.

Documents approved for use along with protocol:

- English and translated informed consent forms
- Data collection tool

Yours sincerely,

Assoc. Prof Joan Kalyango
Vice Chairperson School of Medicine Research and Ethics Committee

9.3 Appendix 2b: School of Medicine Research Ethics Committee Renewal Approval

MAKERERE UNIVERSITY

P.O. Box 7072 Kampala, Uganda
E-mail: rresearch9@gmail.com

Phone: 256 414 533541
Fax: 256 414 541036/0414 532204

COLLEGE OF HEALTH SCIENCES
SCHOOL OF MEDICINE

RESEARCH ETHICS COMMITTEE

September 07, 2018

Mr. Simon Peter Kigozi
IDRC

Category of review
[ ] Initial review
[X] Continuing review
[ ] Amendment
[ ] Termination of study
[ ] SAEs

Dear Mr. Kigozi,

RE: REC REF No. 2017-119

Title: "Evaluating routine HMIS estimates of malaria incidence compared to cohort incidence for malaria mapping in Uganda" Version 1.1"

Renewal Approval Date: 07th September, 2018
Effective Date: 20th September, 2018
Project Expiration Date: 19th September, 2019

On behalf of the committee, I write to inform you that the proposed extension has been approved.

The Makerere University School of Medicine Research and Ethics Committee (REC) initially reviewed and approved the above-referenced protocol on September 20th, 2017. Previous approval of this protocol expires on September 19th, 2018

The review by the committee has found that your renewal is consistent with the continued protection of the rights and welfare of human subjects which protection of human subjects is a partnership between the Research and Ethics Committee (REC) and the investigators. We look forward to working with you as we both fulfill our responsibilities.

Renewals: REC approval is valid until the expiration date given above. If you are continuing your project, you must submit an Application for renewal at least six (6) to eight (8) weeks before the lapse date. If the project is completed, please submit an application for permanent closure.

1

Amendments: The REC must review any changes in the project, prior to initiation of the change. Please submit an Application for Amendments to have your changes reviewed and summarize the proposed change and the rationale for it in a letter to the School of Medicine Research and Ethics Committee. If changes are made at the time of renewal, please include an Application for Amendments with the renewal application.
In addition, submit three (3) copies of an updated version of your original protocol application- one showing all proposed changes in bold or 'track changes,' and the other without bold or track changes.

Adverse Events: If issues should arise during the conduct of the research, such as unanticipated problems, severe adverse events or any other problem that may increase the risk to the human subjects, notify the REC Chairman promptly. The forms are available to report these issues.

Please use the REC REF number listed above on any forms submitted which relate to this project/study.

Good luck in your research. If we can be of further assistance, please contact us at (+256) 0414 -533541 or via email at research@chs.mak.ac.ug/rresearch9@gmail.com. Thank you for your cooperation.

Yours Sincerely,

[redacted]

Assoc. Prof. Ponsiano Ocama
Chairperson School of Medicine Research and Ethics Committee

9.4 Appendix 2c: School of Medicine Research Ethics Committee Amendment Approval

MAKERERE UNIVERSITY

P.O. Box 7072 Kampala, Uganda
E-mail: rresearch9@gmail.

Phone: 256 414 533541
Fax: 256 414 541036/0414 532204

COLLEGE OF HEALTH SCIENCES
SCHOOL OF MEDICINE

RESEARCH ETHICS COMMITTEE

April 23, 2020

Mr. Simon Peter Kigozi
IDRC

Category of review
[ ] Initial review
[ ] Continuing review
[X] Amendment
[ ] Termination of study
[ ] SAEs

Dear Mr. Kigozi,

Re: REC REF 2017-119

Title: "Evaluating routine HMIS estimates of malaria incidence compared to cohort incidence for malaria mapping in Uganda.

Your proposal entitled "**Evaluating routine HMIS estimates of malaria incidence compared to cohort incidence for malaria mapping in Uganda"** was initially reviewed and approved by the School of Medicine Research and Ethics committee on September 27th, 2017.

On 29th January 2020, you requested for permission to make some modifications in the study protocol; for objective 5, to add the national HMIS surveillance data that has been made more readily accessible due to the introduction of the district health information system – version 2 (DHIS-2). This is because considerable improvements have been made in national disease surveillance through the support of the on-line system known as DHIS-2. This improved accessibility as well as improved data quality at national level, will provide a better framework for assessment of the burden of disease in the community through risk mapping, to inform policy makers towards more effective control.

The committee considered these changes on April 23, 2020. On behalf of the committee, I am glad to inform you that these changes have been approved. You may now proceed with the study. Please forward regular reports on your study to the committee.

Yours sincerely,

[redacted]

Assoc. Prof. Ponsiano Ocama
Chairperson School of Medicine Research & Ethics Committee

MAKERERE UNIVERSITY
SCHOOL OF MEDICINE
APPROVED
★ 23 APR 2020 ★
RESEARCH & ETHICS COMMITTEE
P.O.BOX 7072, KAMPALA

9.5 Appendix 3: Uganda National Council for Science & Technology Initial Approval

Uganda National Council for Science and Technology

(Established by Act of Parliament of the Republic of Uganda)

Our Ref: SS 4455 **2nd January 2018**

Mr. Simon Peter Kigozi
IDRC-Uganda
KAMPALA

Re: Research Approval: **Evaluation of Routine HMIS estimates of incidence compared with cohort incidence for malaria mapping in Uganda**

I am pleased to inform you that on **20/12/2017**, the Uganda National Council for Science and Technology (UNCST) approved the above referenced research project. The Approval of the research project is for the period of **20/12/2017** to **20/12/2020**.

Your research registration number with the UNCST is **SS 4455**. Please, cite this number in all your future correspondences with UNCST in respect of the above research project.
As Principal Investigator of the research project, you are responsible for fulfilling the following requirements of approval:

7. All co-investigators must be kept informed of the status of the research.
8. Changes, amendments, and addenda to the research protocol or the consent form (where applicable) must be submitted to the designated Research Ethics Committee (REC) or Lead Agency for re-review and approval **prior** to the activation of the changes. UNCST must be notified of the approved changes within five working days.
9. For clinical trials, all serious adverse events must be reported promptly to the designated local IRC for review with copies to the National Drug Authority.
10. Unanticipated problems involving risks to research subjects/participants or other must be reported promptly to the UNCST. New information that becomes available which could change the risk/benefit ratio must be submitted promptly for UNCST review.
11. Only approved study procedures are to be implemented. The UNCST may conduct impromptu audits of all study records.
12. An annual progress report and approval letter of continuation from the REC must be submitted electronically to UNCST. Failure to do so may result in termination of the research project.

Below is a list of documents approved with this application:

	Document Title	Language	Version	Version Date
1	Research proposal	English	1.1	11 Aug 2017
2	Questionnaires	English	1.0	28 Feb 2017
3	Information Sheet and Consent Form	English	1.1	11 Aug 2017

Yours sincerely,

[redacted]
Beth Mutumba
For: Executive Secretary
UGANDA NATIONAL COUNCIL FOR SCIENCE AND TECHNOLOGY

cc The Chair, School of Medicine Research Ethics Committee

LOCATION/CORRESPONDENCE

Plot 6 Kimera Road, Ntinda
P. O. Box 6884
KAMPALA, UGANDA

COMMUNICATION

TEL: (256) 414 705500
FAX: (256) 414-234579
EMAIL: info@uncst.go.ug
WEBSITE: http://www.uncst.go.ug

9.6 Appendix 4a: London School of Hygiene & Tropical Medicine Research Ethics Committee Initial approval

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
United Kingdom
Switchboard: +44 (0)20 7636 8636

www.lshtm.ac.uk

Observational / Interventions Research Ethics Committee

Mr Simon Kigozi
LSHTM

19 October 2017

Dear Mr Simon Kigozi

Study Title: Evaluating routine HMIS estimates of malaria incidence compared to cohort incidence for malaria mapping in Uganda

LSHTM Ethics Ref: 13902

Thank you for your application for the above research project which has now been considered by the Observational Committee via Chair's Action.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

Conditions of the favourable opinion

Approval is dependent on local ethical approval having been received, where relevant.

Approved documents

The final list of documents reviewed and approved is as follows:

Document Type	File Name	Date	Version
Investigator CV	MosesKamya biosketch Malaria 14Feb2015	14/02/2015	1
Investigator CV	CV_Joaniter Immaculate Nankabirwa_21 Sep 2016	21/09/2016	1
Investigator CV	CV Sarah Staedke	12/11/2016	1
Information Sheet	APPENDIX C - English Information sheet	28/02/2017	1.0
Protocol / Proposal	Protocol_Simon P Kigozi	24/03/2017	1
Protocol / Proposal	APEENDIC D-1 - English VHT Reporting Survey Questionnaire for District Health Service	24/03/2017	1
Protocol / Proposal	APEENDIC D-2 - English VHT Reporting Survey Questionnaire for Health facility	24/03/2017	1
Investigator CV	cv Kigozi Simon Peter	24/03/2017	0
Investigator CV	CV_Rachel Pullan	05/04/2017	1
Investigator CV	Dorsey NIH Biosketch	05/04/2017	1
Protocol / Proposal	Protocol_Simon P Kigozi_v1pt1 with edit highlights	11/08/2017	1.1
Protocol / Proposal	Protocol_Simon P Kigozi_v1pt1	11/08/2017	1.1
Information Sheet	APPENDIX C - English Information sheet & Consent form_v1pt1	11/08/2017	1.1
Covering Letter	Cover letter - SPkigozi	11/08/2017	1.1
Information Sheet	APPENDIX C - English Information sheet & Consent form_v1pt1-with edits	11/08/2017	1.1

After ethical review

The Chief Investigator (CI) or delegate is responsible for informing the ethics committee of any subsequent changes to the application. These must be submitted to the committee for review using an Amendment form. Amendments must not be initiated before receipt of written favourable opinion from the committee.

The CI or delegate is also required to notify the ethics committee of any protocol violations and/or Suspected Unexpected Serious Adverse Reactions (SUSARs) which occur during the project by submitting a Serious Adverse Event form.

An annual report should be submitted to the committee using an Annual Report form on the anniversary of the approval of the study during the lifetime of the study.

At the end of the study, the CI or delegate must notify the committee using the End of Study form.

All aforementioned forms are available on the ethics online applications website and can only be submitted to the committee via the website at: http://leo.lshtm.ac.uk.

Further information is available at: www.lshtm.ac.uk/ethics.

Yours sincerely,

[redacted]

Professor John DH Porter
Chair

ethics@lshtm.ac.uk
http://www.lshtm.ac.uk/ethics/

Improving health worldwide

9.7 Appendix 4b: London School of Hygiene & Tropical Medicine Research Ethics Committee Renewal approval

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
United Kingdom
Switchboard: +44 (0)20 7636 8636

www.lshtm.ac.uk

Observational / Interventions Research Ethics Committee

Mr Simon Kigozi
LSHTM

18/04/2019

Dear Simon,

Project Title: Evaluating routine HMIS estimates of malaria incidence compared to cohort incidence for malaria mapping in Uganda

Project ID: 13902

Thank you for your annual report application for the continuation of your research dated 12/04/2019 23:34 , which has now been considered by the Chair on behalf of the Ethics Committee.

Confirmation of ethical opinion

This application is approved by the committee for a further year.

Conditions of the favourable opinion

Approval is dependent on local ethical approval having been received, where relevant.

After ethical review

Any changes to the application must be submitted to the committee via an Amendment form.

The CI or delegate is also required to notify the ethics committee of any protocol violations and/or Suspected Unexpected Serious Adverse Reaction (SUSARs) which occur during the project by submitting a SUSAR and Protocol Violation form.

An annual report should be submitted to the committee using an Annual Report form on the anniversary of the approval of the study during the lifetime of the study.

At the end of the study, the CI or delegate must notify the committee using an End of Study form.

All aforementioned forms are available on the ethics online applications website and can only be submitted to the committee via the website at http://leo.lshtm.ac.uk.

Additional information is available at: www.lshtm.ac.uk/ethics.

Yours sincerely,

[redacted]

Professor John DH Porter
Chair

ethics@lshtm.ac.uk
http://www.lshtm.ac.uk/ethics/

Improving health worldwide

9.8 Appendix 4c: London School of Hygiene & Tropical Medicine Research Ethics Committee Amendment approval

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
United Kingdom
Switchboard: +44 (0)20 7636 8636

www.lshtm.ac.uk

Research Ethics Committee

Mr Simon Kigozi
11 May 2020

Dear Simon,

Study Title: Evaluating routine HMIS estimates of malaria incidence compared to cohort incidence for malaria mapping in Uganda

LSHTM Ethics ref: 13902 - 1

Thank you for submitting your amendment for the above research project.

Your amendment has been assessed by the Research Governance & Integrity Office and has been approved as a non-substantial change. The amendment does not require further ethical approval from the observational ethics committee.

List of documents reviewed:

Document Type	File Name	Date	Version
Other	Protocol_Simon P Kigozi_v1pt2_Clean	29/01/2020	1.2
Other	Protocol_Simon P Kigozi_v1pt2_tracked	29/01/2020	1.2
Other	DHIS-2 Permission LETTER - MOH	29/01/2020	1.2
Other	SOM-REC Amendment Approval_Simon P Kigozi	29/01/2020	1.2

Any subsequent changes to the application must be submitted to the Committee via an Amendment form on the ethics online applications website: http://leo.lshtm.ac.uk .

Best of luck with your project.

Yours sincerely,

[redacted]

Rebecca Carter

Research Governance Coordinator

Ethics@lshtm.ac.uk
http://www.lshtm.ac.uk/ethics/

Improving health worldwide

9.9 Appendix 5: Ministry of Health Letter of Permission for National HMIS Data Access

Telephone: General Lines: 256 – 417 – 712260
Permanent Secretary's Office: 256 – 417 – 712221
Toll Free 0800100066

E-mail: ps@health.go.ug
Website: www.health.go.ug

IN ANY CORRESPONDE ON
THIS SUBJECT PLEASE QUOTE NO. **ADM.386/01**

Ministry of Health
P. O. Box 7272,
Plot 6, Lourdel Road,
Wandegeya
KAMPALA
UGANDA

24th February, 2020

Mr. Simon P. Kigozi
PhD Student
Department of Disease Control,
London School of Hygiene and Tropical Medicine

RE: PERMISSION TO USE DHIS2 OPD DATA (2014-2019) FOR YOUR PHD RESEARCH

Reference is made to your letter dated 12th February 2020 requesting for permission to use DHIS2 OPD data (2014-2019) for your PHD research.

The Ministry of Health hereby permits you to use outpatient malaria data from DHIS2 to conduct his research study titled "*Evaluate routine HMIS estimates of malaria incidence compared to cohort incidence for malaria mapping in Uganda*".

The Ministry of Health has provided you with access to variables including total attendance, total malaria cases, and total confirmed malaria cases on a monthly basis from all reporting health facilities over the duration of 2014 – 2019.

[redacted]
Dr. Olaro Charles

MINISTRY OF HEALTH
25 FEB 2020
P. O. BOX 7272, KAMPALA

FOR: DIRECTOR GENERAL HEALTH SERVICES

CC: Permanent Secretary

CC: CHS/Policy and Planning Department

9.10 Appendix 6: Evaluating sources of data on population estimates

The population estimates utilised in this research were obtained using Afripop data however, these estimates were evaluated against national estimates provided by the national bureau of statistics (UBOS), at district level. Afripop estimates were the preferred option given two specific properties:

- They were readily available and/or accessible,
- They allow for varied geographical scales such as the health facility catchment areas unlike the national census estimates, details of which are restricted to administrative scales including districts, sub-Counties, and parishes.

Whilst parishes are the lowest administrative units for which census population estimated summaries may be available, finer and non-administrative geographical scales such as health facility catchments in the assessment of incidence of malaria for this study, were not aligned with the population data in these administrative scales. This, therefore, influenced the choice of Afripop data with the possible flexibility.

Nevertheless, estimates from Afripop were evaluated against national census projections for the years between 2015 and 2019 at district level. To do this, scatterplots were evaluated stratified by the 15 regions of the country (Figure 1 below) and by year (Figure 2 below). In addition, the two population estimates were evaluated by examining the root mean square error in a bivariate regression between Afripop and UBOS district estimates, both among regions (Table 1) and years (Table 2).

Fig 1. Relationship between district population estimates from Afripop and UBOS between 2015 and 2019, by region.

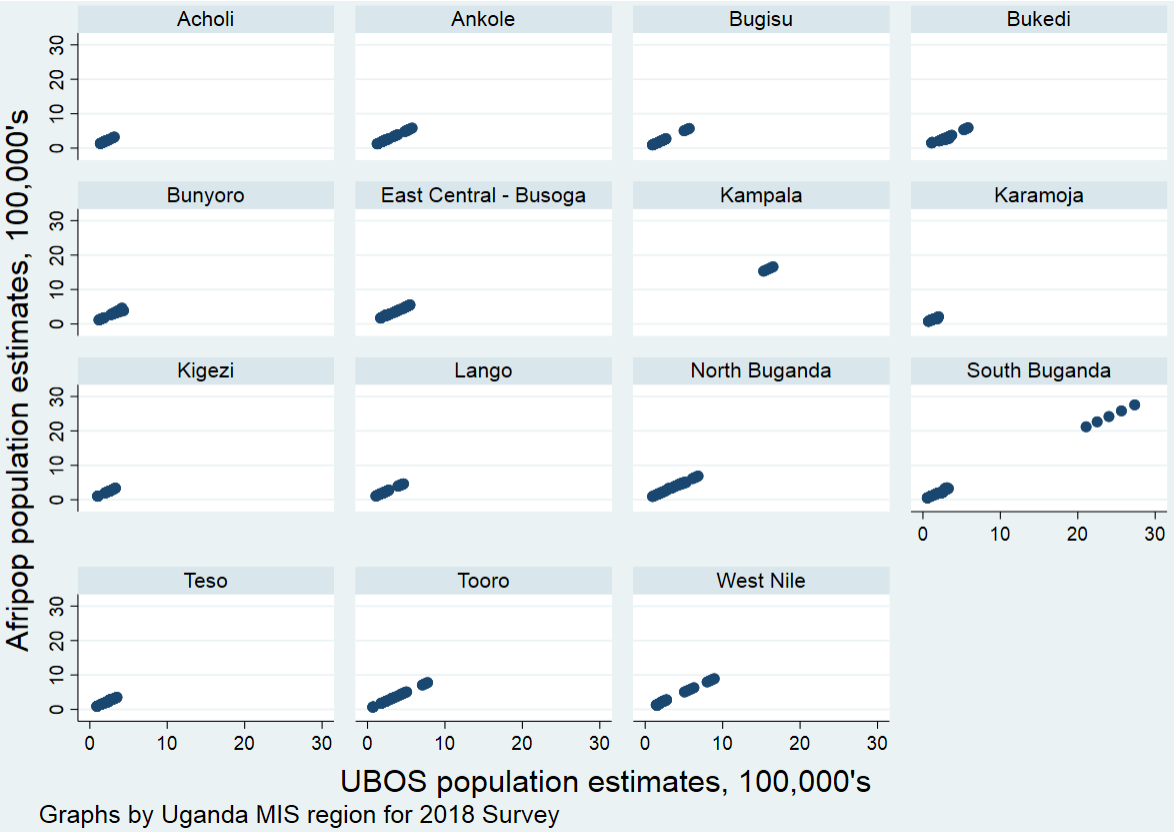
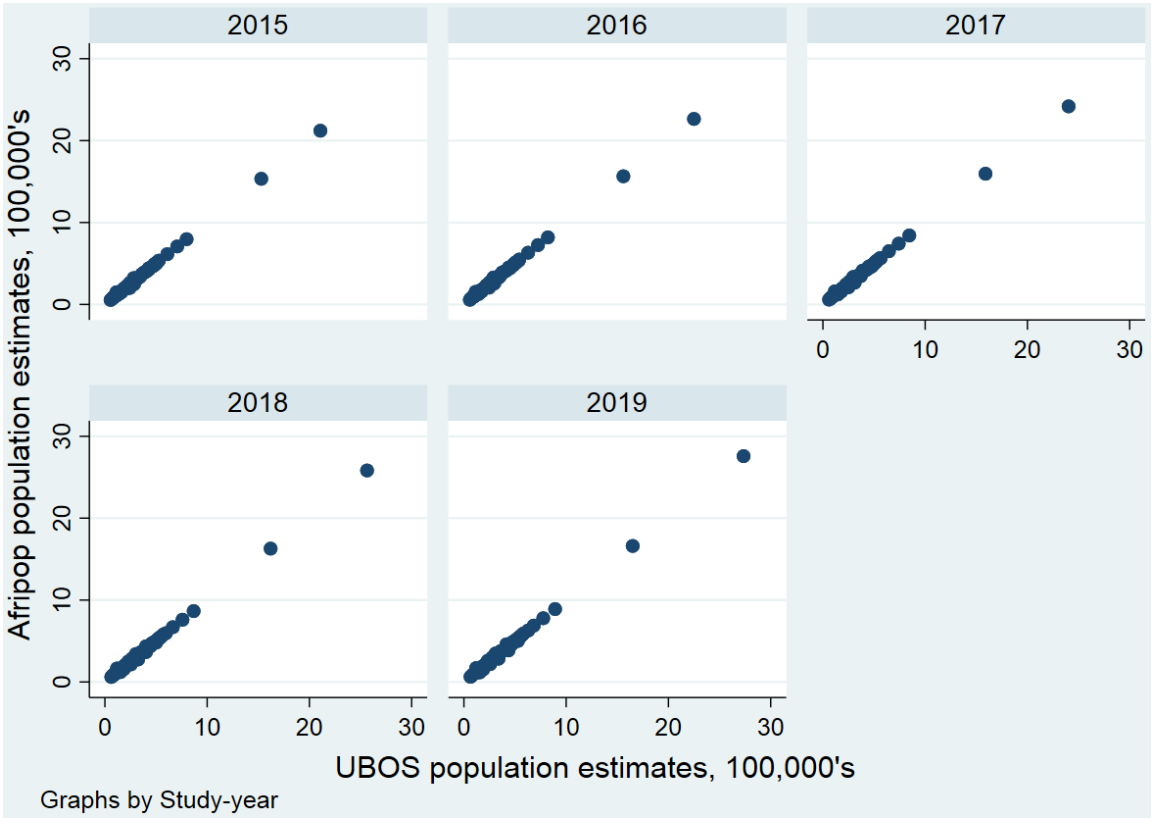


Fig 2. Relationship between district population estimates from Afripop and UBOS, by year.



Both the evaluation by region and year showed a strong correlation between these two sources of population data estimates.

Moreover, assessment of the differences between these Afripop estimates (predicted) versus the UBOS estimates (observed) using the root mean squared error from a bi-variate regression between the two, showed very minimal differences between these estimates.

Table 1. Root mean square deviation between district population estimates from UBOS and Afripop among regions.

Region	RMSE
Acholi	3,209.5
Ankole	2,182.3
Bugisu	2,420.5
Bukedi	23,578
Bunyoro	16,215
East Central - Busoga	6,375
Kampala	864.8
Karamoja	8305
Kigezi	3,345.4
Lango	3,564.1
North Buganda	8,239.3
South Buganda	15,681
Teso	8,933.5
Tooro	3,125.9
West Nile	9,220.4

Results here showed that the differences between regional population estimates from Afripop and UBOS were highest among districts in Bukedi, Bunyoro, and South Buganda regions and lowest in Kampala district/region.

Table 2. Root mean square deviation between district population estimates from UBOS and Afripop per year

Year	RMSE
2015	7,993.4
2016	8,783.6
2017	9,908.3
2018	11,324
2019	13,037

Notably over the years, differences between population estimates from the two approaches increased from year to year, being lowest in 2015 and highest in 2019. Notably, these differences were smaller during the year closest to the most recent census survey of 2014 and greatest further from that time.

9.11 Appendix 7: Response to reviewers in Paper 1

AJTMH-18-0901 – Response to reviewers

Malaria burden through routine reporting: relationships between incidence and test positivity rates

Reviewer #1:

Comments to the Author

The analysis presented here makes use of a rich dataset and provide findings which are of interest to all those concerned with malaria control planning. Which metric to use to generate plans for control remains a perennial question. The results are interesting but are presented for a research audience. Some additional context or clarification to put the outcomes in an operational context may improve the readership and uptake of the findings.

Specific comments

a) Major

a. It is mentioned several times that the three sites used are different transmission settings, without explicit mention of which site corresponds to what type of setting. Perhaps it could be labelled in Figure 1 or in the study setting methods which site is considered highest, medium, lowest, etc transmission perhaps based on the previous prevalence measures mentioned? This would help with the interpretation of the main findings.

Response: **We have revised Figure 1 to include the historical annual entomological inoculation rates of the three sites in the footnotes. In addition, we've indicated these details in the section on study settings.**

b. Line 114 – was the gridded population surface used also Worldpop? Outputs from Worldpop lack the granularity needed at this resolution. If enumeration data were available, can a brief explanation be provided for why that was not used to determine the population spatial pattern? Worldpop totals could be used if household population counts did not accompany the enumeration data.

Response: **We made several attempts to get access to population data from the national census bureau but without much success and hence the decision to use the Worldpop that is accessible. We have clarified in the text that this was due to inaccessible population estimates for our study area.**

The enumeration data from our study was an enumeration of household without a full human census in those households. This was helpful in estimating number of children (study population) based on the average household size.

c. The introduction begins with mention of the importance of these types of metrics for control and intervention planning, but the conclusions provided lack clarity for operational implementation. For example, 'confirmed malaria case rate that is sensitive to changes both in time and space, provides a better indicator of the burden of malaria on the health facility catchment, as estimated from the health facility than test positivity rate.' It may not be apparent to a program manager what burden in the health facility catchment, as estimated from the health facility means.

Response: **We have clarified further in the conclusion the operational impact of our findings on these indicators for implementation purposes.**

b) Minor

a. In abstract line 28-29 – specify if the pairwise comparisons were done by month, village or both

Response: **We have revised the abstract to clarify that pairwise comparisons were done first by month and then by village**

b. It makes sense that re-attendance episodes were excluded, but the mention in line 110 that CMCR is confirmed primary cases per 1000

Response: **We have revised the statement as suggested to "confirmed primary malaria cases per 1000"**

c. Figure 2- define abbreviations used. What is VOR?

Response: **We revised Figure 2 to include a full description of VOR as "village of residence" that was initially omitted**

Reviewer #2:

Comments to the Author

Manuscript Title: Malaria burden through routine reporting: Relationship between incidence and test positivity rate

Authors: Kigozi et al.

Overall Comments:

This article presents work comparing two metrics used as part of routine malaria surveillance, specifically focusing on the temporal and spatial dynamics at 3 hospitals in Uganda. Comparing the more operational test positivity rate with the ideal metric of malaria incidence rate is of interest to ensure that any decisions made based on the available data can be done in a robust manner. The strength of this work lies in the large number of data points available for analysis. However, I have some comments outlined below that would greatly improve this work.

Major Comments:

- The objectives of the work, including their significance is not very clear and lacks focus. The stated goal is to compare the two metrics, but ultimately the regression analysis adjusts TPR with the incidence suggesting the objective is to assess factors associated with TPR? A clearer focus should help clarify the main aims of the work, but also identify the most appropriate analysis.

Response: **The objective of this work was indicated as "to evaluate the relationship between IR and TPR both on the time and space dimensions to aid further understanding of the representativeness of HMIS estimates of burden" in the introduction.**

The regression analysis was used given it's the most well understood approach to evaluating relationships between an outcome (IR) and an exposure (TPR) where there are other factors known or presumed to play a role in the relationship being evaluated.

The regression approach was found to have been used in previous studies that formed the basis for our need to study this relationship further, especially by including the dimension of space.

- If the objective is to assess the bias between the metrics including spatial and temporal trends, the metrics should be on the same scale. This can be done by converting the metrics using established associations (e.g. publications by the MAP group) or standardizing the two metrics and work with their z-score. This will impact the scatter plots (and enable the line of concordance to be plotted), the Bland-Altman plots, etc.

Response: **Given that neither indicator is a gold standard measure of the other, the objective was not to assess bias between them**

o With the two metrics on a standard scale, the authors should consider using the difference between the metrics as the main outcome for the regression analysis to assess factors associated with any bias between metrics.

Response: **While IR and TPR are both indicators of burden of malaria, they are distinguished by the approach of generating them. As a result, we adjusted the IR scale to that of TPR, which is a proportion with limits of 0 to 1. No other valuable approach of standardizing these indicators to the same scale was found to our knowledge.**

o If once on the standardised scale, the Bland-Altman plots show similar patterns, there is a strong trend in Kihihi suggesting non-normal differences in this site. This should be acknowledged and assessed.

Response: **We have clarified the approach used in determining which sites fulfilled the Bland Altman criteria and this was fully assessed.**

- Based on the data presented, the authors are correct in that the association between the two metrics is non-linear. Instead of relying on lowess models with assumptions of localized weighting etc., I suggest that the authors explore non-linear forms of regression (e.g. log-linear) which can account for the observed association while adjusting for covariates

Response: **We have included quadratic prediction plots where Lowess was initially assumed sufficient and this has particularly improved our plots.**

However, regarding the final regression model, the multi-level mixed effect poisson model was chosen based on it being non-linear regression approach for count data. This model sufficiently accounts for observed associations while adjusting for covariates, in addition to addressing multi-level variability in the parameters and in our case the two levels being month and village.

- The authors have justified why they have only considered hospital attendees <11 years of age from an epidemiological perspective, but less so for the clinical perspective. Instead of ignoring data on the adults, if the data is available, can all age results be assessed, adjusted for age? E.g. Are the trends similar when considering all ages the same as <11? <5? Are the trends for adults only different? This is a very interesting question and operationally relevant for surveillance programs in being able to accurately interpret their data.

Response: **We have clarified further that trends of TPR were not different between <5 and 5-<11. Whereas the data on the adults is available, for this work we sought and received approval to conduct this study among study participants specified as children under 11 years with several other considerations made.**

Minor Comments:

- Why are you using the term CMCR instead of the more commonly used incidence rate (doing so would improve clarity of this work)?

Response: **We have revised this to incidence rate (IR) for the benefit indicated**

- Are there other smaller level facilities or community health workers within the study areas that may absorb some of the suspected malaria cases? If so, would this potentially bias the results?

Response: **We have acknowledged among study limitations that there are other lower level facilities that absorb some of the suspected malaria cases and this would impact our results**

- What was the pre-determined threshold for agreement using the Bland-Altman plots?

Response: **Other than the formal 95% confidence band, we had no pre-determined threshold of agreement in place and we have acknowledged this in the text**

- The temporal unit being assessed is not always clear with both monthly and annual scales being considered at different points. Please include the unit when discussing the temporal results.

Response: **We have included the units for further clarification**

- The authors showed that both metrics are able to pick up seasonality, but one interesting question that could be assessed with this data is whether the degree of bias changes between the high and low transmission seasons. This could be done with a simple interaction term.

Response: **Given that none of the metrics was a gold standard of the other, we did not consider bias however, we intend to compare this incidence rate to incidence from cohorts at which point we'll be in position to fully evaluate bias**

- LL146-154 is this referring to the total health care seeking population? Those suspected of malaria or those testing positive for malaria?

Response: **We have clarified these further in the text**

- LL153: removing data with missing variables will impact the precision as there are fewer data points available (e.g. make confidence bands wider). Do the authors mean accuracy (assuming non-differential missingness)?

Response: **We have revised this to "no considerable impact on our results" in the text**

- LL157: the authors state "suggesting agreement between them" when referring to plotting monthly TPR and CMCR. A timeseries plot doesn't test for agreement. Do they mean "show similar trends"?

Response: **We have revised this appropriately to indicate similarity of trends rather than agreement in the text**

- What is the case definition used by facilities for malaria? Who is supposed to be tested? (Also, note the potential for detecting opportunistic malaria infections.)

Response: **We have clarified this further in the outcome measures section**

- Figure 2: currently 'missing' is not qualified. Can the key missing variable (e.g. age, village of residence) be added to avoid confusion? Were all records complete with no missing diagnosis?

Response: **We have revised Figure 2 accordingly to qualify the 'missing'. In addition and importantly, there were no records with missing diagnosis**

- Figure 5: The figures suggest the mean annual figures are presented whereas the legend says 4-year mean. Please clarify and be consistent in labeling. Also, the data is available for 4.75 years. Why was this not all used and restricted only to the years with complete monthly data?

Response: **We have revised these to include the full 4.75 years**

- Figure 5: The color scales are all different which biases the visual interpretation when comparing between TPR and CMCR as well as across sites. The best approach would be to plot the standardized units so the maps are all directly comparable.

Response: **We have standardized the units in Figure 5 and a revised version is now included**

- Figure 7: Is a linear fit the best choice? Have any other forms been tested or residual diagnostics done?
  - Fit for Walukuba extends beyond data.
- There are no confidence bands on any of the linear fits presented in plots.

Response: **We have revised these plots to predicted quadratic plots that better explain the relationship; corrected the plot in Walukuba; and, included confidence bands in all three as suggested.**

Reviewer #3:

Comments to the Author

Overall Comments:

This paper comparing test positivity rate and incidence rate I settings in Uganda is much improved. I however have a couple of minor comments that should still be addressed before publication.

Minor Comments:

- Impact of lower-level facilities on results? Would you expect the trend with distance to be impacted if people further away prefer the lower level facility for something as routine as malaria? It would be helpful to include this in the discussion section. I'm guessing the linear trend with distance is due to cases seeking care in the more local lower-level facilities.

Response: **Previous unpublished data that we looked at showed that lower level facilities have very similar patterns of access with the higher-level facilities included in our study. All of them see patients as routine as uncomplicated malaria, from near (majority) and farther away (fewer) in a very similar way. I would, therefore, expect the crisscross swap of patients to minimize the potential effect of lower-level facilities on the results in our study. We have therefore, indicated this position in discussion as suggested.**

- Table 1: the denominators for testing rates and malaria diagnosis used for the proportions presented are not clear. Can you include a statement similar to the one for the gender category?

Response: **We have revised Table 1 to include clearer descriptions of sub-sections and the proportions presented for further clarity as suggested.**

9.12 Appendix 8: Response to reviewer comments in Paper 2

MALJ-D-19-00682 – Response to reviewers

Rapid shifts in the age-specific burden of malaria following successful control interventions in four regions of Uganda

Reviewer #1:

Overall comment

This is an interesting manuscript presenting results from a 10-year analysis of the age shift of malaria infections after vector control interventions in Uganda. Authors clearly present the methods and results and appropriately acknowledge the limitations of the study. There is also a good balance between the information presented in the main text and the information in the supplemental materials.

Response: **We appreciate the reviewer's overall comments on this work.**

Main comments

However, it is surprising to see how one of the main results of the study was neglected in the discussion: the fact that most malaria cases were diagnosed in women and that the percentage of confirmed malaria cases in women increased significantly in the group over 15 years old. Even when there was not an observable trend in the gender across intervention periods, data and figures showing these results could be included in the supplementary document and the authors should comment about them in the Discussion. Authors should analyze these results, present potential explanations for this, and discuss its implications for a vulnerable population group such as women in childbearing age.

Response: **We appreciate the reviewer's comments raised as concerns a key result that should have been further emphasized. To address these comments, we analyzed the results on most malaria cases being diagnosed in women as advised by the reviewer and included our findings in the results as indicated in line numbers 245-253 (in the tracked changes version)**

Further, figures 7 – 10 have also been included together with Table 2 in the additional file to document result details of this analysis and/or findings.

Findings showed a consistently higher proportion of females across all intervention periods in all four sites, as well as a consistent shift in age distribution of confirmed malaria cases between males and females. However,

there was a disproportionately larger increase in the proportion of older males than females among confirmed malaria cases across intervention periods.

A summary discussion of these results has been included as indicated in line numbers 303 – 313 (in the tracked changes version) in the discussion section.

Reviewer #2:

Overall summary

This paper presents findings from a study assessing the impact of control interventions on the age distribution of malaria cases, using malaria surveillance data from 4 sites in Uganda. This is generally a well written and presented paper, to a topic of importance to malaria epidemiology and impact evaluations of standard control measures. Below are issues the authors should address to improve the paper.

Response: **Thank you so much for your opinion of work from this study. We appreciate your notice of the importance of findings from this study**

Compulsory Revisions

Introduction

* Lines 37-43: this section should be qualified by the fact that children <5 contribute the most to malaria cases (burden) and are thus the focus of control measures and studies, in areas of stable malaria transmission in sub-Saharan Africa.

Response: **Revised as proposed by adding the text "and are thus the focus of control measures and studies, in areas of stable malaria transmission in sub-Saharan Africa"**

* Line 52: can the authors please specify 'confirmed malaria cases' if this is the case.

Response: **Revised as proposed by indicating the reported trends in 'confirmed malaria cases' from the reference provided**

Methods

* Line 134-135: can you please confirm the outcome was only 'confirmed malaria cases'. What does 'conveniently defined' mean?

Response: **Revised as proposed by precisely indicating 'confirmed malaria cases' and further elaborating on the fact that the three age categories were conveniently defined**

* Were any analyses of trends in confirmed malaria case incidence, by age, done, using OPD attendance as an offset to account for treatment seeking? This would have been a very nice check on the results and help interpret the multinomial model results.

Response: **Thank you for raising this important issue. However, OPD attendance was not considered as an appropriate proxy for the population at risk to estimate confirmed malaria case incidence. Whereas a good estimate of incidence would have been desirable to better understand these findings, the same was not possible and we acknowledged this among the limitations of the study but would consider or recommend this for future investigations among characterized populations. Nevertheless, we believe that the quality and scope of the longitudinal data within the present confines of the study, enables viable insights towards improved surveillance and control of malaria.**

* How were data pooled across MRCs? Were the cases from the MRCs treated as a random effect, as they should be, in the multinomial models?

Response: **Given differing transmission settings of the four sites, data were not pooled across MRCs but rather each site was analysed independently using the same approach for each. As such, site-specific multinomial models were fit, and results interpreted comparatively rather than as a pooled analysis.**

* How was time dealt with in the multinomial models, within intervention periods?

Response: **Given that the intervention period (defined by time) was the main exposure of interest, no further adjustment for time were deemed necessary. Further adjustment for time was expected to lead to controlling for the effect of 'time given intervention' that we were investigating as well as overfitting of the models.**

* How was suspected malaria (or TPR) dealt with in the final multinomial model? Was it included somehow as a covariate?

Response: **We did not expect any ecological association between being suspected for malaria and age category of confirmed malaria cases. However, we assumed an effect of diagnostic test used and age of confirmed cases and despite the predominance of microscopy (followed by gradual but gentle increase in RDT use at all sites), this was accounted for in the multinomial models at each site. Moreover, given the lack of variation in suspected malaria by age-category over time, it was assumed unnecessary to include this as a covariate in the multinomial models.**

* Was any attempt made to account for other potential confounding factors on the primary outcome, including MRC reporting completeness, rainfall patterns, seasonality, diagnostic stockouts, varying intervention coverages over time and across sites, treatment seeking, and access to treatment? I know these are difficult variables to ascertain, but there are a lot of publicly available sources for spatial surfaces (over time) for some of these, and proxies can be used for others. If not, these potential confounders (or influencing factors) should be discussed as a major limitation to the study in the discussion section.

Response: **We appreciate the insight raised in this comment and we have acknowledged in the limitations that we were unable to account for other potential influencing factors for which data were not available in our study. However, we have accounted for some of the factors listed in this comment. First, varying intervention coverage over time was accounted for by defining intervention periods based on the major large-scale control interventions programmatically implemented with high coverage as well as broad community effects expected among the pockets that may not have received these interventions. However, we agree that there could have been possible variation in intervention coverage especially when control interventions were more targeted, for instance during baseline. Nevertheless, we observed similar effects of the large-scale interventions across all sites, implying that the differences between sites at baseline would have no significant impact on the main findings in our study.**

Second, by virtue of being MRCs that are continuously monitored, we have very high reporting completeness rates with for instance, we reported that only 0.3% were missing a record of age implying over 99% completeness that hardly varied between sites or over time.

Third, the spatial scope of the reference sites' catchment areas was unknown. As such, evaluation of spatial surfaces with an unclear spatial scope could have introduced inadvertent bias in our evaluation. We have indicated this by stating in the limitations that without a well characterized population for each site, we were unable to effectively evaluate these additional factors and/or their effects.

* For the final model for marginal predictions, can the authors please somehow well the model fit the data (i.e. model fits), as these models can yield spurious results when models fir the data poorly.

Response: **Thank you for this very thoughtful comment.**

It is no clear to us what associations you would consider spurious. However, we have used two approaches to address this

First, in order to implement a goodness of fit for these multinomial models used for marginal predictions, we would need to know the expected age-distribution of the confirmed cases across intervention periods among the categories used to define the multinomial outcome. In this case we posit that if there were no external influences on the age distribution of confirmed cases, the closest distributions that confirmed cases may be expected to follow would be either the suspected malaria cases or the general attendance age-distributions. Both crude and adjusted results from this showed that these age distributions followed different patterns across intervention periods. In the interest of this comment therefore, we used the null hypothesis that the age-distribution of confirmed malaria cases would follow the same distribution as suspected malaria cases or as overall patient attendance. We evaluated both of these but consistent with our findings, there was strong evidence to reject the Null hypothesis in both cases.

In view of this, we strongly believe that whereas a possible effect of the limitations of multinomial models as implied through this review comment may not be ruled out, there is no basis to assume that it is the main driver of the findings from this study according to model goodness-of-fit evaluations.

Second, we conducted model evaluations, comparing the model with and without our defined intervention periods and used the likelihood ratio test to determine whether including intervention periods improved the model. In all cases, the model without intervention periods as a covariate was found to be nested within the final model (one with the intervention period as covariate) which was an improvement of the fit, based on Akaike's information criteria. This is now shown in the table 2 in the additional file and revised the text in results on page 15. Furthermore, we evaluated the relationship between model predicted and crude proportions using scatter plots as indicated in Figure 7 in the additional file. Results showed that the multinomial models fit the data very well.

Results

* Lines 169-171: what was the lowest rate of microscopy testing during the period of observation?

Response: **Revised as suggested to include the lowest overall testing rate.**

* Lines 223-233: Are there statistical test results that should be presented, in addition to the 95% CIs, for the observed declines and increases in age groups over time periods?

Response: **Give that the reported proportions are marginal estimates from the final multinomial regression models whose results are reported in Table 2 with statistical test results included, we did not consider it necessary to further include statistical test results here.**

Discussion

* TPR is known to vary widely by treatment seeking patterns, over time and space, even within a given transmission level. What is the authors interpretation of the reported trends in TPR given this known bias?

Response: **We appreciate the concerns raised in this comment and agree that TPR varies over time and may be influenced by treatment seeking patterns. However, findings from our previously published work that included three MRCs from differing transmission settings showed no distinguishable patterns of TPR over space, regardless of endemicity but rather clear variations over time. That being said, evaluations in this study were predominantly temporal given unknown facility catchments' spatial scope.**

Concerning the influence of treatment seeking on TPR, we recognized that this study is limited to a passive surveillance context where accuracy of transmission-level through this indicator may only be limited to the population that seek care from the public health sector rather than that of a definite population at risk. We further acknowledged in the limitations that we were unable to evaluate changes in incidence in which case we'd have been able to characterize the population with an improved understanding and/or reduced impact of changes in care seeking within a given site.

Thus far, TPR is considered an indicator for changes in transmission as recommended by the WHO, especially with regards to temporal trends. Given that this study was investigating the effect of control interventions (expected to act by reducing transmission) over time, changes in TPR over time were not considered a source of bias in this study.

* How would the scale-up of iCCM have changed treatment seeking, and the main findings of the results for TPR trends and the age distributions of confirmed malaria cases once it was scaled?

Response: **The scale-up of iCCM may be expected to reduce patient attendance of children under 5 years (which is the target population of this approach) at the health facility within the site where it was implemented. If this effect were to be strong, the age distribution of suspected and malaria negative cases would have shifted upwards, consistent with confirmed cases. However, there was no**

evidence from our results that this was the case and this is compounded by the fact that no findings have been published from this iCCM activity. Furthermore, the reported implementation of iCCM was limited to one site and for a limited duration and despite this, the main findings from this study were consistent across all sites suggesting that the expected effect of iCCM on treatment seeking did not have a distinguishable effect on the findings in this study. Nevertheless, we acknowledged in the limitations that this intervention may have had some additional favourable effect to our findings. Concerning TPR, no particular change is expected in TPR trends as a result of iCCM scale-up, given that it may not disproportionately affect the number of malaria than non-malaria febrile cases.

* Were any sensitivity analyses conducted on the time periods being defined by intervention roll-out in the sites? What happened if the scale-up of LLINs, or iCCM, was expected to take longer to be fully scaled and adopted?

Response: **We appreciate this very insightful comment. However, the study was not powered to conduct either per-protocol or intention-to-treat analyses with the interventions included and/or evaluated here. These interventions were conducted programmatically, and we had no information concerning variations of when or how long implementation would take.**

9.13 Appendix 9: Response to reviewer comments in Paper 3

AJTMH-19-0950 – Response to reviewers

Practical implications of a relationship between Health Management Information System and community cohort-based malaria incidence rates

Reviewer #1:

Comments to the Author

General comments

This manuscript capitalises on a wealth of data from collaborative research sites to compare malaria incidence estimates from passive cohort and routine HMIS datasets. Considering the current tension between advocacy for increased use of surveillance data in decision making and lingering concerns over quality of HMIS data, this paper is a useful addition to the literature. My suggested revisions below mainly relate to sections requiring additional clarification or detail.

Response: **We appreciate the reviewer's opinion of this work and the suggestions for clarity provided.**

Major revisions

1. Please clarify if the HMIS data were restricted to the same age range as the cohort. In the abstract (line 36) you indicate that the focus was on children in HMIS data and in the methods (line 101) the HMIS population is stated as 6 months to 11 years. However, the rest of the paper (including the definition on line 141) suggests that HMIS data included all ages.

Response: **As the reviewer suggested, we have revised the text as follows:**

> 'The primary outcome in this study was the monthly malaria incidence rate derived from health facility HMIS OPD incident malaria cases data **for children between 6 months and less than 11 years of age**, …' *(Methods, pg. 8, lines 144-145)*

We also revised the text within the statistical analysis section as follows:

> 'We explored the relationship between HMIS and cohort incidence **among children between 6 months and less than 11 years of age** on a monthly timescale …' *(Methods, pg. 11, lines 218-219)*

2. There is some ambiguity about whether cohort incidence malaria cases were symptomatic or not. Line 136 explains when treatment was provided (symptoms & positive slide), but an explicit statement on the incidence malaria cases (infection only, or symptomatic infection) is needed earlier in the methods (I see there is a definition on line 159, but this needs to be stated sooner).

Response: **As the reviewer suggested, we have clarified the text as follows:**

'**By** these **symptomatic diagnostically confirmed infections**, incident malaria cases were identified, …' *(Methods, pg. 8, lines 140-141)*

3. Lines 144-146. I have some concerns about the assumptions used to estimate confirmed malaria cases in the absence of diagnostic testing at several of the facilities. Do you have any evidence to suggest that test positivity would be the same year-round at each of these sites? Do you expect the populations attending HCIV and HCII/III to be similar in terms of their risk of malaria, and therefore the proportion of fevers which are attributable to malaria? Any additional data or sensitivity analysis that can be provided to bolster your chosen approach to dealing with lack of diagnostic testing at lower level health facilities would be useful, and would strengthen the final conclusions you make about the comparison of HMIS data and cohort incidence. For example, did you explore comparing the fever incidence (clinical malaria diagnosis) between cohort and HMIS – this could give you an idea of whether HMIS is underestimating incidence due to lower health-seeking behaviour, or if the issue is with correct identification of malaria once patients get to the facility.

Response: **As far as each site is concerned, the only available reliable indication of the temporal level of transmission was the indicator from the national reference centres hosted within each site. We did not assume that test positivity would be the same all year round, however, we assumed that test positivity per month within each site would be very close to the respective reference centre estimate. As the reviewer suggested, we explored the relationship between monthly incident HMIS clinical malaria cases and monthly incident fever cases in the cohort and have included the results in Figure 1 in the additional file. We have also revised the text to indicate that these results provide further evidence supportive of the chosen approach to correct for diagnostic testing as follows:**

'This approach was also supported by a linear relationship between cohort fever incidence and HMIS clinical malaria incidence suggesting case identification at facilities as a major factor (Figure 1 in the additional file).' *(Methods, pg. 8-9, lines 152-154)*

Minor revisions

1. Abstract lines 44-47. You state that HMIS data has a strong predictive power in lower transmission settings, but then contradict this by closing the abstract with the statement that the findings have "important implications for surveillance in low resource, high burden countries". Do you mean that the implication is that HMIS is not valuable to high burden settings and there needs to be effort to use other methods / heavily invest in HMIS strengthening? Or do your findings are actually have relevance to all settings?

Response: **These findings have relevance in all settings, more so when surveillance is strengthened, for instance with improved testing of suspected cases and recording coupled with prompt reporting. We have revised the text to reflect this stating as follows:**

'HMIS still requires improvements, however its strong predictive power of unbiased burden **when improved**, highlights the important role it could play as a cost-effective tool for monitoring trends and estimating impact of control interventions. This has important implications for **malaria control** in low resource, high burden countries.' *(Abstract, pg. 3, lines 44-47)*

2. Lines 59-61. This statement bothers me a little and doesn't fully represent some of the progress made in this area. There have been efforts in the last few years to increasingly use HMIS data in impact evaluation, supported by substantial increases in access to confirmatory diagnosis in many settings following introduction of RDTs. While these studies are still outnumbered in comparison to those using DHS/MIS, the developments in this area should be briefly mentioned/referenced.

Response: **We agree that we may have undersold recent progress and as such now acknowledge the progress made and included reference to that effect. We also revised the text to emphasize the under-utilization of HMIS data in spite of these improvements as follows:**

'Whilst there are extensive HMIS improvements through standardized data formats as well as quality assessment tools, among others, HMIS data are **still underutilized** to provide rigorous…' *(Background, pg. 4, lines 59-61)*

3. Lines 66-70. A brief mention of which of these systems/initiatives are multi-disease and which are malaria-specific would be useful here. My expectation is that both DHIS2 and HMIS cover all priority diseases, but the sentinel sites are focussed on malaria?

Response: **We agree with the reviewer that DHIS-2 and HMIS cover all priority diseases but sentinel surveillance is focused on malaria in this case. We have revised the text to clarify this difference as follows:**

'A national HMIS was introduced in Uganda in 1997 to enable **priority** disease**s** surveillance at national levels … **Specific to malaria,** the HMIS was additionally supplemented by routine sentinel surveillance …' *(Background pg. 4, lines 67-69)*

4. Line 71. Which data do you mean by "these data"? HMIS, DHIS2, or sentinel site data?

Response: **As the reviewer suggested, we have clarified in the text as follows**

'These **sentinel site** data have been used to evaluate impact…' *(Background, pg. 5, line 72)*

5. Line 83-84. I would argue that a comparison of HMIS and cohort incidence will tell you more about representativeness of HMIS estimates of malaria burden compared to the true population burden, not about the quality of HMIS data (which you'd find out about by doing records review, consultation observations etc.)

Response: **As the reviewer suggested, we agree that this comparison informs representativeness of HMIS estimates of the true population burden. We have revised the text to reflect this as follows:**

'… providing important insight into the utility and representativeness of HMIS estimates of malaria burden compared to the true population burden …' *(Background, pg. 5, lines 84-85)*

6. Lines 109-111. Did you consider cross-referencing the OPD register with any laboratory registers? Depending on patient flow, patients with confirmed malaria may be recorded in the lab register but only listed as 'suspected' in OPD registers. There is an interesting paper by Okello et al. that looks at the issues around resolving case count estimates between different registers present in health facilities.

Response: **We appreciate this important point raised; however, we did not consider other sources of data from the health facility. That being said, we do not expect it to impact our results and if at, only minimally for the following reasons.**

a. **Level II facilities (majority in this study) do not have laboratory facilities and therefore have a single point of care for uncomplicated malaria case management. As such, the OPD register captures most, if not all the data in these facilities.**
b. **Whereas this may have affected level III facilities, there was only two in the study with extremely limited diagnostic testing done during the study period, which limits potential impact from this level.**
c. **Concerning level IV health facilities, this is not expected to have impacted the true case counts from this level because these were national reference centres with rigorous monitoring for data capture and quality**

7. Line 115. What is meant by "some secondary data"? Other covariates that were considered for inclusion in the models? Or additional case data?

Response: **We have clarified in the text that it was**

> **'… additional HMIS data'** *(Methods, pg. 7, line 116)*

8. Line 116. I would recommend a brief statement (or reference) to explain the services provided by each health centre level in Uganda and approximate population size served.

Response: **We have revised the text to include population served and services offered at each level of facility as follows:**

> **'Level IV facilities serve ≈ 50,000 people providing in-patient, laboratory, and maternity services while level III's serve ≈20,000 with in-patient and laboratory services, and level II ≈5,000 with basically outpatient and community outreach services**.' *(Methods, pg. 7, line 118-120)*

9. I found the description of correcting for non-residence on lines 146-149 a little difficult to follow and may need some rewording.

Response: **We have reworded the text here to further clarify the description as follows:**

> 'Level II and III health facilities **had very low testing rates and predominantly diagnosed malaria presumptively. Assuming that risk of malaria for children between 6 months and less than 11 years seen at each reference facility was the same as for similar age children seen at the respective lower level**

facilities each month, **total** monthly presumptive **malaria** cases from **site lower level** facilities …' *(Methods, pg. 8, line 148-151)*

10. Line 152. For clarity, I would suggest re-emphasising here that the estimated population at risk per month was age-restricted rather than the whole population.

Response: **We have revised the text as suggested indicating the age restriction as follows:**

'To generate monthly HMIS incidence rates, the site-level sum of incident cases of malaria **among children between 6 months and less than 11 years of age** after … was divided by the site estimated monthly population **of children under 11 years of age** at risk **of malaria**.' *(Methods, pg. 9, line 158-162)*

11. Line 163. Could you provide some more information about how you age-standardised cohort incidence? Did you check to make sure that the age groups captured by HMIS were comparable to the overall population structure? In some settings, younger children and boys are more likely to be taken to health facilities than older children and girls.

Response: **We have provided a description of how we age-standardized cohort incidence rates in the additional file (Section E). We have also revised the text in the manuscript to include a summary of this process as follows:**

'Consequently, we age-standardized incidence estimates **using six age categories defined between 6 month and < 11 years, based on the initial recruitment age distribution in these categories as the standard (as explained in section E of the additional file). Initial recruitment** into the cohort**s was** conducted primarily during August and September, 2011 for each site.' *(Methods, pg. 9-10, lines 173-276)*

Concerning comparison of age groups in HMIS with the overall structure, we did not find a reliable source of the overall age-structure comparable to our participants age range. Instead, we assumed that the same age structure within the HMIS would be maintained over the study duration.

12. Line 166. How did you account for community transmission? I would also suggest that you split this into a new sub-section here, as you start to discuss model-building.

Response: **We have revised the text to indicate how we account for community transmission "using cohort incidence" and re-defined this subsection as "Regression model" in the text.** *(Methods, pg. 10, line 176-177)*

13. Line 169. The last section of this sentence doesn't quite make sense.

Response: **We have revised to the text to further clarify the statement as follows:**

'… while health facility characteristics **were estimated using health facility** performance **in recording patients' diagnoses**.' *(Methods, pg. 10, lines 181-182)*

14. Line 170. Facility accessibility is surely a function of more than just rainfall. For example, whether roads are tarmac or unsurfaced, local topography etc. Other studies have used or generated travel time surfaces to take a

more nuanced look at health facility accessibility that includes hills, proximity to and quality of roads. Did you consider any of these approaches?

Response: **We did not take the other suggested approaches because we believe that the choice of rainfall as a proxy in an agriculturally dependent society covers the fundamental aspects of accessibility, not to mention seasonality. With regard to accessibility, rain seasons tend to draw families into cultivation activities as highest priority and therefore dropping or delaying any other prospective activities, given the risk of missing the season. Moreover, at the scale considered in this study where for instance there is no functional public transport, a composite metric of accessibility that incorporates factors such as road network, land use and topography among others, may not provide sufficient temporal or spatially variability. In addition, the suggested composite metric of accessibility is most suitable for the definition of health-facility-most-likely catchment areas, which was not the intention of using accessibility in this study. In this study, we perceive accessibility as an influence on the possibility of malaria cases being recorded in the health facility registry.**

15. Line 178. I was surprised that while you have a temporally variable indicator for accessibility (rainfall), you seem to just use a single average for facility availability. Do facilities closed more often around certain times of year (religious holidays, farming seasons etc)? Given that you had access to the OPD registers, are you able to actually determine the number of days that each facility was open each month (based on dates when no-one was registered at OPD)?

Response: **The temporal unit for evaluating effects of change in all covariates by site, including facility availability, was 'month' and not a single estimate.**

As regards number of days that a facility was open, we examined the OPD registers of each health facility to find which days had no patient records for each month. Here, like for accessibility, we generated a monthly site-level proportion of the days of each month that facilities within that site were open. This proportion was determined using a numerator = (average number of days site facilities were open in a month) and the calendar month's known number of days as denominator. Other than Christmas time when the majority would be closed on Christmas day and boxing day as well as new years' day, no other religious holidays or particular patterns were observed. Of note also, it is not uncommon for health workers to be involved in district or research organised trainings and/or meetings, besides personal engagements. For instance, level II health facilities are often run by a single health worker and if they are not available for whatever reason, the facility would be closed. However, unavailability is not unique to level II facilities. We have revised the text to further clarify how site-level facility availability was defined, as follows:

> 'To account for health facility availability, measured as ease of care availability at the facility, we generated the average proportion of days per month that health facilit**ies within each site were** open to see patients. **The average proportion was defined as mean number of days a site's facilities were open in a month, divided by the respective calendar month's total number of days.'** *(Methods, pg. 10-11, lines 195-198)*

16. Line 190. I don't follow the justification for using the reciprocal of the proportion of OPD patients missing a diagnosis, rather than just using the % missing a final diagnosis.

Response: **The choice of taking the reciprocal of the proportion missing a diagnosis was so that as a measure of performance, it can be interpreted intuitively as higher values corresponding to higher performance and lower values corresponding to lower performance. We have revised the text to further clarify this as follows:**

> 'The reciprocal of the proportion was derived so as to **enable intuitive** interpretation of its trend **as performance (high values correspond to high performance and vice-versa)**.' *(Methods, pg. 11, line 208-209)*

17. Lines 233-235. The term "participant recruitment" is a little confusing, particularly coming after a sentence about clinic visits. Does this mean proportion of people eligible for inclusion in the cohort who consented? Or the proportion of all cohort members recruited from each site?

Response: **We have revised the text to clarify that this is the proportion all cohort participants from all the three sites as follows:**

> 'Whereas Kihihi **had the highest number of participants recruited overall (**36%) and Walukuba **the lowest** at 31%, …' *(Results, pg. 13, line 254-255)*

18. Line 241. I suggest reemphasising here "symptomatic incident cases".

Response: **We have revised the text to clarify as follows:**

> 'Across the study duration, a total of 4,884, 12,058 and 18,960 **symptomatic test and residence corrected** incident malaria cases (Table 1) were **generated** among participants in Walukuba, Kihihi, and Nagongera respectively' *(Results, pg. 14, line 262-264)*

19. Line 251. You state the number of incident cases that were recorded, but then refer to these as age-standardized, suggesting that these are not the raw number of cases. In this case, the use of "recorded" is a little misleading, and perhaps the raw number of cases could be reported, then the age-standardised count.

Response: **The numbers reported were the raw incident cases recorded. Incidence rates were derived from these and were age-standardized and now we have revised the text to clarify this as follows:**

> 'From these, site mean monthly incidence rates were derived **and age standardized**.' *(Results, pg. 14, lines 272-273)*

20. Table 3. I suggest avoiding using the abbreviation "CI" to mean two different things in the table, and just write "cohort incidence".

Response: **We have revised the table as suggested**

21. Line 328. Do you have any information on what type of private facilities these are - private not for profit hospitals, pharmacies, clinics, drug shops? Since you mention a general preference for private sector in Uganda,

is it feasible that a substantial proportion of malaria cases in these locations are being addressed by these private facilities and additional ones outside your defined study area?

Response: **We have revised the text to clarify the type of private facilities included, as follows:**

> 'Previous studies have suggested that the majority of care seeking is conducted in private health facilities in Uganda **including private-for-profit hospitals/clinics, pharmacies and drug shops**' *(Discussion, pg. 18, lines 347-349)*

Notably, private-not-for-profit facilities are considered public in Uganda.

We acknowledged that people would have sought treatment outside of our study catchment area, but this could not be fully quantified within our study. However, we do not expect a substantial number of malaria cases to be addressed outside the defined study area in the case of uncomplicated malaria. Moreover, for two of the three sites where we had permission to visit private facilities and look at their data for a limited duration, we found one in each site and only one of these had accessible records.

Also, we expected the crossover of patients from one location to another to be highest in the peri-urban site, given greater availability of options within towns. We examined records from a mid-level facility (level III) in a town close to our study site but found extremely few patient visits from villages recognised as located within our study site. This provided some evidence that a sufficiently high number of cases were being captured within the expected catchment area.

22. Lines 341-344. Do you think that the overall workload (all cause OPD) at the facility could influence HMIS completeness and quality? Do the HCIV have additional staff to support data management and reporting? Do you have any data on the proportion of staff positions at each facility which were empty during the study period?

Response: **We believe that overall workload influences completeness and quality of HMIS and considered including this as metric in this study, however, having assessed data retrospectively, there were no records of staff availability over our study duration. HCIV being referral facilities for the lower level counterparts, have more resources including human resources such health information assistants (HIA). These HIA, however, are not involved in the direct management of patients but rather management and handling of data. This handling involves receipt of data reports from the HCIV where they are based and the lower level health facilities under its supervision, as well as the forward submission of reports to the district authorities. We have revised the text and indicated the factors that could potentially influence the observed heterogeneity across sites concerning HMIS recording completeness as follows:**

> **'We believe these effects may be due to variations in factors such as resource availability and staffing or workload, but these were not evaluated in this study**.' *(Discussion, pg. 19, lines 372-373)*

23. Lines 361-363. I would be surprised if the population preferences for self-medication and herbal medicines changed year-to-year, although it is feasible that these preferences could differ between sites. I'm not sure the reference cited here justified the statement about explaining year-to-year changes.

Response: **We believe that choice of alternatives can vary year-to-year. One study in Kenya [1] reported self-treatment to be dependent on an interaction of affordability, acceptability and availability all of which can vary with time. They report that affordability depends on seasonality of illness and income, transport costs and unofficial payment. On the other hand, provider patient relationships and distrust in quality of care, among others, were identified as influential to acceptability. Lastly, facility operating hours, drug and staff shortages were reported to influence availability. As such, the three main facets that are thought to directly but interactively influence choice of alternatives for care, are in turn directly influenced by factors that can change from year to year.**

Whereas we have revised the text for clarity on references *(Discussion, pg. 19 & 20, lines 391-392)*
we believe that preferences are not homogeneous within sites and due to these known drivers of change and would thus argue that they do explain a great deal of year-to-year changes as earlier indicated.

24. Line 388. Do you feel that the findings from this study are generalisable to the rest of Uganda, given that you captured three different epidemiological settings?

Response: **Given the diversity of epidemiological settings, we feel quietly confident that these findings are generalizable to most parts of Uganda, with some exceptions (such as areas with nomadic lifestyles). However, a larger number of sites may provide further insights unattainable with just three. We have indicated the generalizability of these findings by revising the text in the discussion as follows:**

> '… the diversity of settings and transmission provides important contributions of benefit to surveillance **and considerably generalizable findings in Uganda**.' *(Discussion, pg. 21, lines 419-421)*

25. I would be interested to read in the discussion about any costs associated with malaria diagnosis & treatment at the cohort clinics versus at the government health centres and if this could influence treatment-seeking behaviours. Presumably the cohort clinics provided free treatment and subsidised travel or were close enough to villages that there was no travel cost. Do patients have to pay any consultation fee at the health centres, even if treatment is free?

Response: **We have revised the discussion to include a summary paragraph indicating the difference in cost of treatment for HMIS versus cohort participants, as follows:**

> **'In addition, while treatment was free of monitory cost at public health facilities, regular patient visits to the health facility still costed them in the form of transport cost and long waiting times. Within the cohorts however, transport was reimbursed for every clinic visit made and waiting times minimized due**

to the dedicated clinics. This status quo may have limited potential HMIS clinic visits and thereby contributed to HMIS underestimating cohort incidence estimates.' *(Discussion, pg. 18, lines 356-360)*

We argue that this would have reduced and/or delayed potential clinic visits within the HMIS population and thereby contribute to HMIS underestimating cohort incidence estimates

Reviewer #2:

Comments to the Author

Overall, this is manuscript presents a well-structured study with a clear objective and interesting results. The figures are very well done, and the additional files were helpful to the reader to understand the analysis decisions better. Below I offer few comments/suggestions mostly with regard to clarity of the results and conclusions presented.

Response: **We truly appreciate your opinion of this study and suggestions provided to further improve it.**

Main Comments

- The authors have shown that there are similar trends and good concordance between the HMIS data and cohort data and go further to suggest that HMIS may be a reasonable estimate the malaria burden in low to moderate risk areas. The limitation is, of course, that this may only work in a highly effective health system or an area where the health system has been augmented to provide better care due to ongoing research, and that this assumption may not hold true where health systems are sub-optimal.

The manuscript would be stronger if this were to be mentioned somewhere. Specifically, as the authors state, this could be used in low-resource settings, but many of those countries also have weak health systems.

Response: **We agree with the suggestion made and have further emphasized the role of effective HMIS in better estimating malaria burden with revised text as follows:**

> **'Overall, whilst Nagongera had the highest recording completeness, it had the lowest availability and least accessibility making it the lowest HMIS performance site of the three.'** *(Methods, pg. 11, lines 214-215)*

Moreover, we've now revised the text in the discussion as follows:

> **'This suggests that these findings strongly depend on improved surveillance systems and can be reliable in all transmission settings.**' *(Discussion, pg. 18, lines 354-355)*

We've also revised the text in the conclusion as follows:

> 'These findings have important implications for malaria risk assessment in low resource settings that bear the majority of the burden of malaria**, given improved information systems**.' *(Conclusion, pg. 21, lines 432-433)*

- I wondered about the use of rainfall as a proxy for access to the health facilities. Often access is considered a patient accessing care. I wonder if there might be any data from national surveys on treatment-seeking and access, which could be presented in comparison to using this indicator? It seems that rainfall may have impeded physical access, but access may be more complex and contextual, so although not possible to capture, in this analysis, it would help to speak to this complexity.

Response: **We have compared rainfall and patient attendance trends by site as indicated in Figure 2 in the additional file. These trends show a general tendency for rainfall to be at peaks when attendance is at troughs for many months at all sites. We have also revised the text to indicate this as follows:**

> **'… and extracted as site monthly mean estimates. On examining monthly trends in rainfall and attendance, we observed a general pattern of peaks in rainfall corresponding to troughs in attendance for the same month and vice-versa, suggesting associations between rainfall and attendance and supporting its use as proxy for accessibility (Figure 2 in the additional file).'** *(Methods, pg. 10, lines 184-188)*

We've also clarified that rainfall may predominantly impede physical access in the revised text as follows:

> 'It was assumed that the higher the **mean** rainfall received per month, the less **physically** accessible the health facility for the population that month.' *(Methods, pg. 10, lines 189-191)*

Notably however, action of rain on physical access is the most important aspect of accessibility for this study, due to interest in the role accessibility plays in malaria cases being recorded in health facility registers.

In addition, we have revised the text among limitations stating that the estimates used for health facility performance, including accessibility, availability, and recording performance are but proxy measures as is, as follows:

> **'Sixth, health facility availability, accessibility and recording performance are more complex than this study proposed to estimate them. This could have masked any potentially observable associations otherwise not found**.' *(Discussion, pg. 20, lines 410-412)*

However, this study was not in position to fully define or quantify them with our available data.

- Use of days open as a proxy for availability was okay, but as mentioned access to stock, data would have made this stronger as a clinic may have been open but could not test or treat for malaria. This is definitely a limitation to this indicator, although what you have is better than nothing. As availability may be more complicated than just the number of days open, so this should be considered when considering these results.

Response: **We agree with the suggestion made here and have revised the text to include a limitation concerning data where we indicate that health facility performance factors are more complicated than our study proposed to estimate them, which may have masked potentially observable associations. The text revision is as follows:**

> **'Sixth, health facility availability, accessibility and recording performance are more complex than this study proposed to estimate them. This could have masked any potentially observable associations otherwise not found**.' *(Discussion, pg. 20, lines 410-412)*

Minor points:

1. Lines 242-245 in the results section should be moved to the methods section. “These were corrected for 1) non-testing among presumptively diagnosed cases, and 2) non-residence among those whose village of residence was missing. Also, cases from villages unknown within site boundaries were excluded.”

Response: **We have revised the text as suggested by removing this section from the results section and leaving it in the methods section only. We have also clarified the results as follows:**

> ‘Across the study duration, a total of 4,884, 12,058 and 18,960 **symptomatic test and residence corrected** incident malaria cases’ *(Results, pg. 13, lines 262-263)*

We’ve also revised the text in the methods as follows:

> ‘… villages of residence that were located or known within the study sites. **Notably, cases from villages unknown within site boundaries were excluded**.’ *(Methods, pg. 9, lines 157-158)*

2. Comments on Table 1.

a. It would have been helpful to include percentages in Table 1 to understand the proportion of patients who are tested and positive.

Response: **We have revised the table to include percentages for clarity as suggested**

b. Also, the right three columns for Table 1 could be presented better.

i. First, a line to delineate the adjusted results from the other would have been helpful.

ii. It is not clear if the adjusted numbers are for all health facilities, only HCII/HCIII.

Response: **We have revised the table for clarity as suggested.**

References

1. Chuma, J., V. Okungu, and C. Molyneux, *Barriers to prompt and effective malaria treatment among the poorest population in Kenya.* Malar J, 2010. **9**: p. 144.

9.14 Appendix 10: INLA code for Spatial-autoregressive model in chapter 6

```
# Set working directory
setwd("F:/Analysis_INLA")
#Load libraries
#install.packages("spdep")
library(spdep)
#install.packages("INLA", repos=c(getOption("repos"), INLA="https://inla.r-inla-
download.org/R/stable"), dep=TRUE)
library(INLA)
#install.packages("maptools")
library(rgeos)
library(maptools)
#install.packages('spDataLarge', repos='https://nowosad.github.io/drat/', type='source')
library(spDataLarge)
install.packages("bigmemory")
library(bigmemory)
#*****************************************Creating neigbourhood
Graph**********************************
#Load County Shapefile
shape<-readShapePoly("./hfaccess3hrs_Catchment",IDvar="idadj")
##plot(units)
#Create adjacency matrix
temp <- poly2nb(shape,queen=TRUE, row.names=shape$idadj)
#Convert the adjacency matrix into a file in the INLA format and save
nb2INLA("uganda.adj", temp )
#Visualize the graph and get summary
#g<-inla.read.graph("uganda.adj")
#summary(g)
#****************************************** Load Data
*********************************
#Read in the data
#data1<-read.csv("bugs_maindata1.csv")
library(haven)
```

```
data1 <- read_dta("bugs_maindata1.dta")
attach(data1)

#************* Introduce columns  for space and time ************************
data2<-cbind(data1,reg0=data1$idadj, reg1=data1$idadj, reg2=data1$idadj)
data<-cbind(data2,time0=data2$t, time1=data2$t, time2=data2$t)
# Display what the data set looks like with the new space and time variables included
head(data)
#Besag model with random spatial effect (i.e. BYM model) and structured (rw1) +unstructured temporal effects
hyper.besag <-list(prec=list(prior="loggamma", params=c(.1, .1)))
#hyper.besag <-list(prec=list(prior="loggamma", params=c(.5, .0005)))
hyper.iid<-list(prec=list(prior="loggamma", params=c(.001, .001)))
formula<-confirmed ~ 1 + rainfall + land_surface temperature + Night-time-light + education-level-for-women + vegetation amounts + f(reg0,model="besag",graph="uganda.adj",hyper=hyper.besag)
+    f(reg1, model="iid", hyper=hyper.iid)+
  f(time0,model="rw1", constr = TRUE, scale.model = TRUE,hyper = list(prec = list(prior = "pc.prec", param = c(1,0.01)))) +
  f(time1, model = "iid", constr = TRUE) + f(time2,model="iid", group=reg1,control.group=list(model="iid"),constr=TRUE)

starting.value <- inla(formula, family = "Binomial", Ntrials=pop, data = data,
                       control.compute = list(cpo = T, dic = T), control.inla = list(diagonal = 100, strategy = "gaussian", int.strategy = "eb"),
                     control.predictor = list(compute=TRUE),
                     control.family=list(link="logit"),
                      control.fixed = list(prec.intercept = 0.001),
                      verbose = TRUE)
plot (starting.value$summary.fitted.values$mean,pfpr)
pfpr.final<- inla(formula,family = "Binomial", Ntrials=pop, data = data,
                control.inla = list(strategy = "simplified.laplace"),
                control.predictor=list(compute=TRUE),
                control.family=list(link="logit"),
                control.mode = list(result = starting.value, restart = TRUE),
```

```
            verbose=TRUE)
summary(pfpr.final)
plot (pfpr.final$summary.fitted.values$mean,pfpr)

####### Get fitted values ################################################
## Predicted incidence value
predictedmean<-pfpr.final$summary.fitted.values$mean
# Write prediction to csv file
write.table(predictedmean,"pfpr_fitted.csv",row.names=T,sep=",")
## Predicted credible interval upper bound for incidence
predictedci97.5<-pfpr.final$summary.fitted.values$`0.975quant`
write.table(predictedci97.5,"pfpr_fitted_u.csv",row.names=T,sep=",")
## Predicted credible interval lower bound for incidence
predictedci2.5<-pfpr.final$summary.fitted.values$`0.025quant`
write.table(predictedci2.5,"pfpr_fitted_l.csv",row.names=T,sep=",")
```