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, their strong predictive power of unbiased malaria burden when improved highlights the important role they 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]
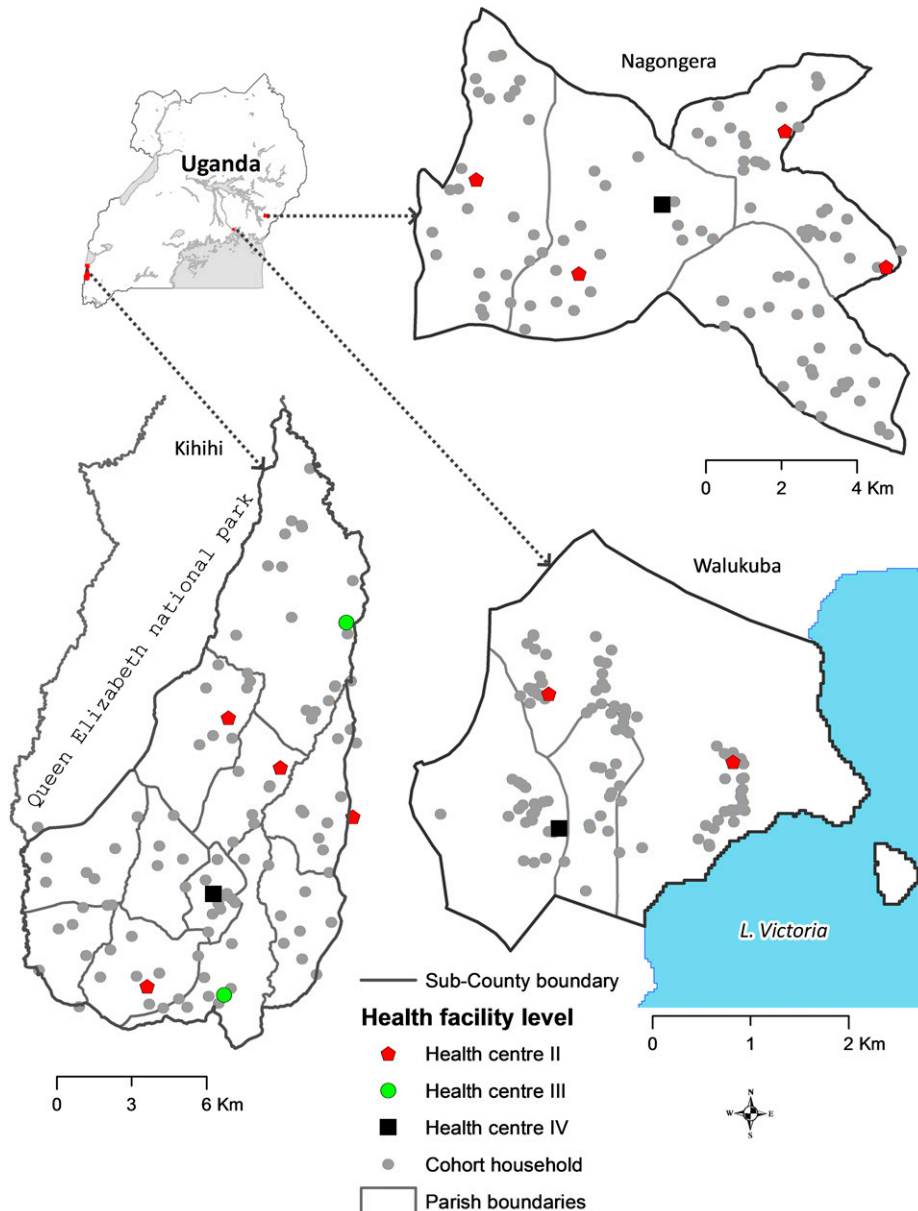

FIGURE 1.   Site map showing the sites and health facilities included. This figure appears in color at www.ajtmh.org.

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 $\approx$ 50,000 people providing inpatient, laboratory, and maternity services, whereas level III facilities serve $\approx$ 20,000 with inpatient and laboratory services, and level II $\approx$ 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 Tropical Applications of Meteorology using SATelline data and ground-based observations (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 (‡) + cases of laboratory-confirmed malaria (§).

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
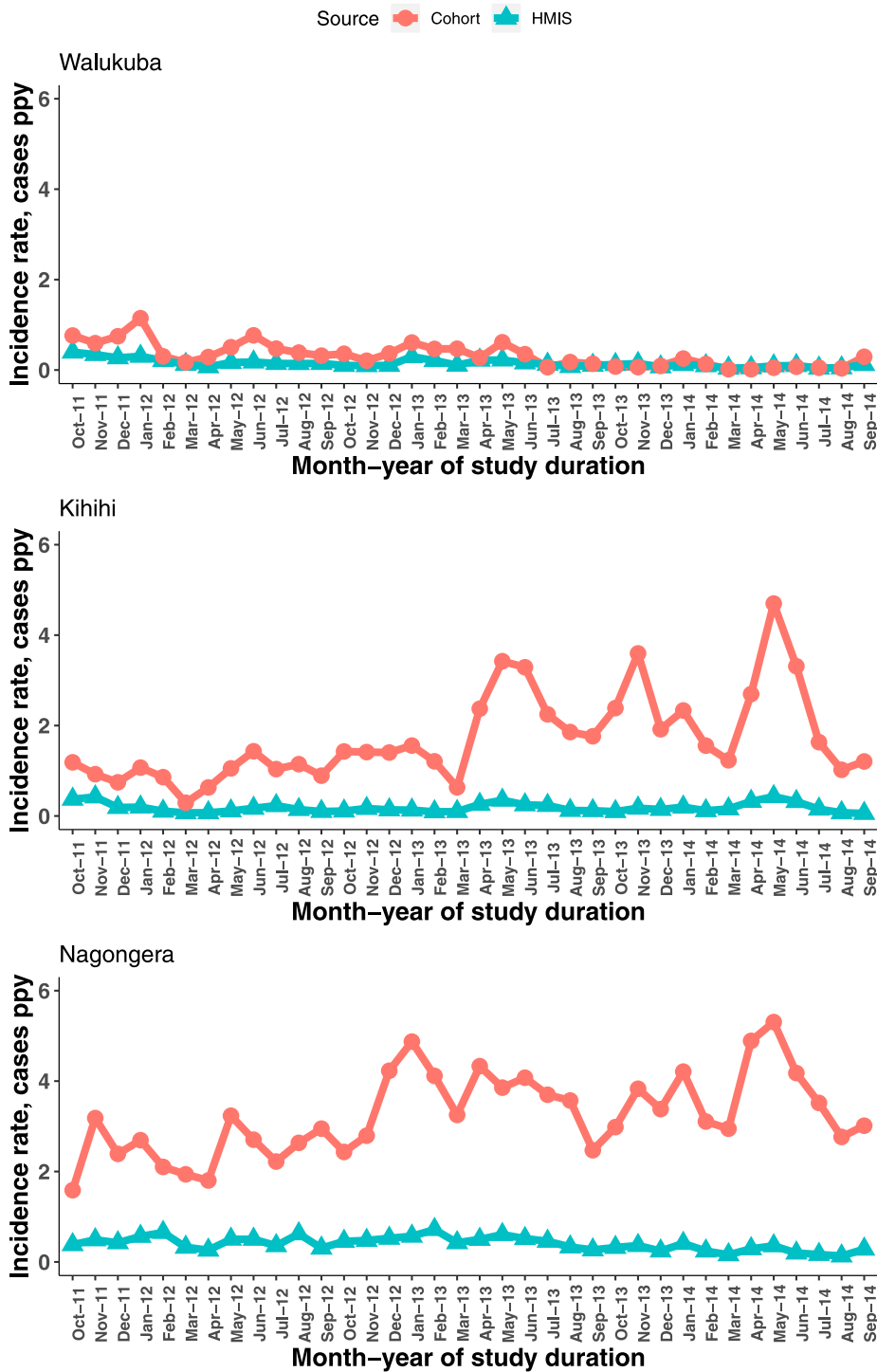
FIGURE 2. Trends in raw incidence rates per month. This figure appears in color at www.ajtmh.org.
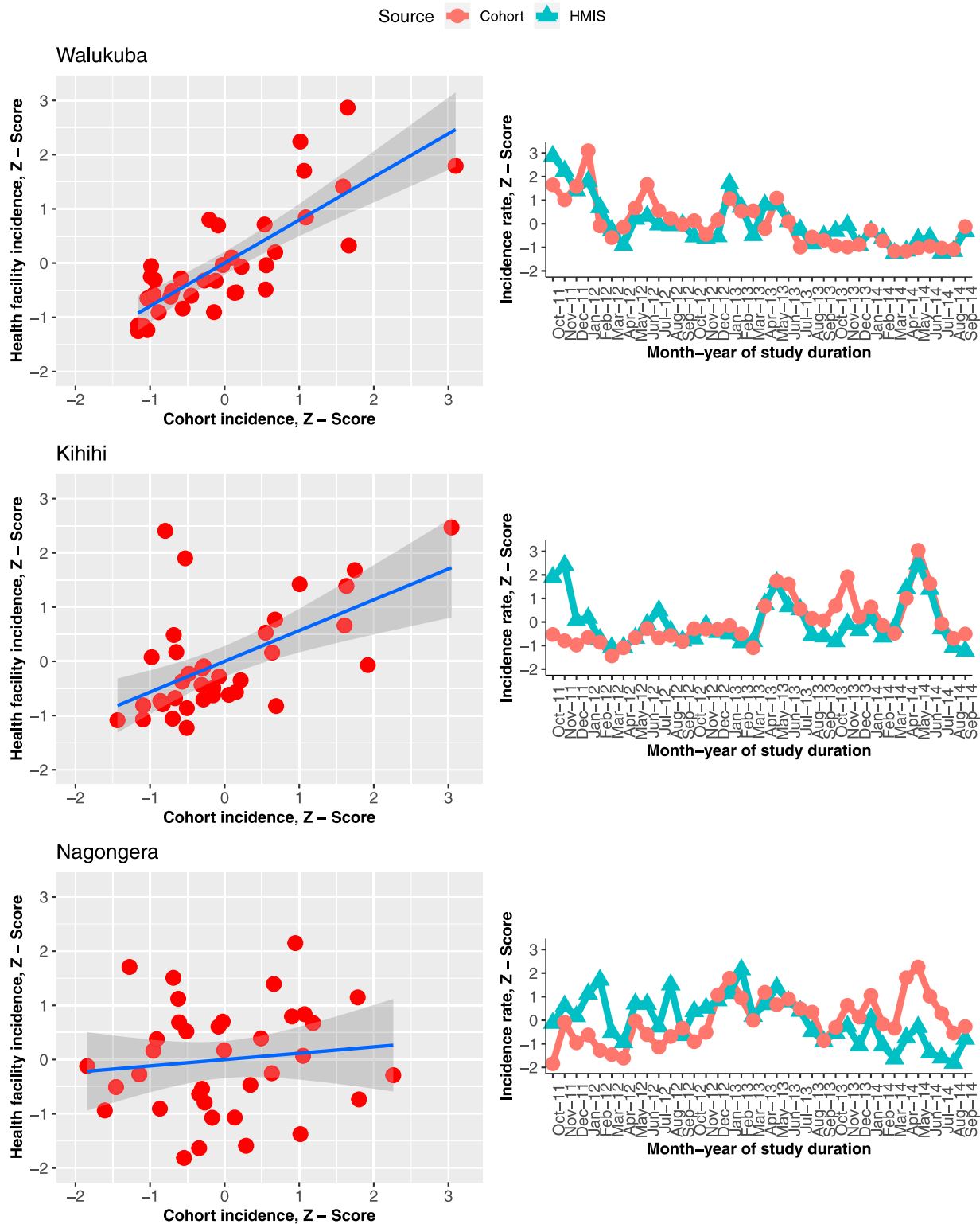
FIGURE 3.   Scatterplots and trends in health management information systems and cohort incidence rates on respective mean-standardized scales (z-scores). This figure appears in color at www.ajtmh.org.

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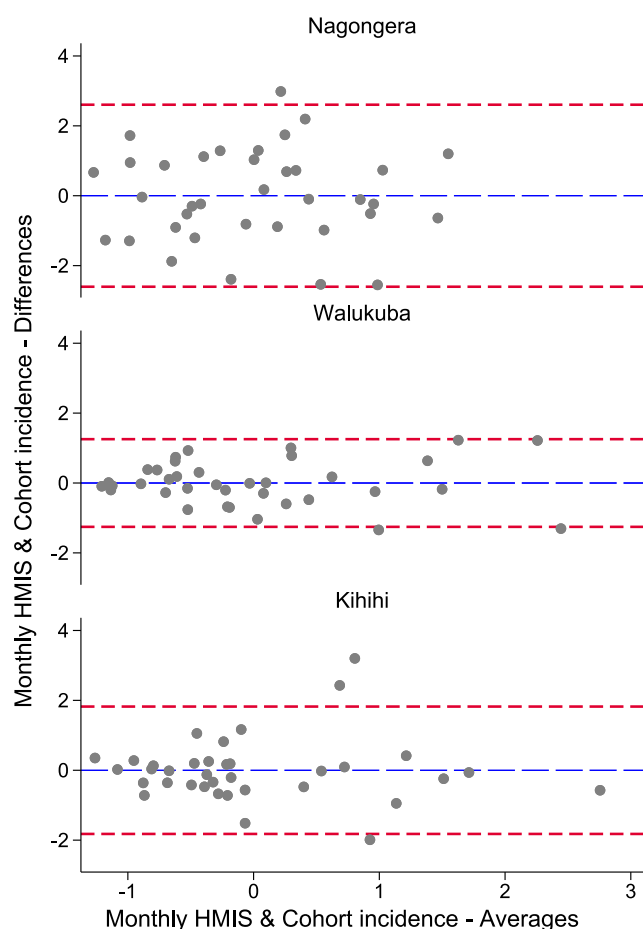
FIGURE 4. Concordance between health management information system incidence and cohort incidence rates using Bland–Altman diagrams by site, based on incidence z-scores. This figure appears in color at www.ajtmh.org.

(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

Site	Covariate	Year	Unadjusted			Adjusted		
			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.

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: UBOS and ICF International.
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*. Kampala, Uganda: 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.