

Assessing the accuracy of the recording and reporting of malaria rapid diagnostic test results in four African countries: Methods and key results

Kim A. Lindblade

kalindblade@gmail.com

PMI Insights project/PATH

Arthur Mpimbaza

Child Health and Development Centre

Corine Ngufor

Centre de Recherche Entomologique de Cotonou

William Yavo

Institut National de Santé Publique

Sunday Atobatele

Sydani Group

Ese Akpiroroh

Sydani Group

Abibatou Konate-Toure

Institut National de Santé Publique

Idelphonse Ahogni

Centre de Recherche Entomologique de Cotonou

Nelson Ssewante

Child Health and Development Centre

Bosco Agaba

National Malaria Control Division

Augustin Kpemasse

Programme National de Lutte contre le Paludisme

Jacques Agnon

Programme National de Lutte Contre le Paludisme

Onyebuchi Okoro

National Malaria Elimination Programme

Godwin Ntadom

National Malaria Elimination Programme

Antoine Tanoh

Programme National de Lutte Contre le Paludisme

Cyriaque Affoukou

Programme National de Lutte contre le Paludisme

Jimmy Opigo

National Malaria Control Division

Shawna Cooper

Audere

John J. Aponte

PMI Insights project/PATH

Kevin Griffith

U.S. President's Malaria Initiative

Radina Soebiyanto

U.S. President's Malaria Initiative

Michael Humes

U.S. President's Malaria Initiative

Research Article

Keywords: malaria, rapid diagnostic tests, diagnosis, health management information systems, surveillance, data quality, case management, sub-Saharan Africa

Posted Date: May 14th, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-6577975/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Malaria Journal on July 1st, 2025. See the published version at <https://doi.org/10.1186/s12936-025-05459-7>.

Abstract

Background

Rapid diagnostic test (RDT) results are the foundation of both case management and malaria surveillance across sub-Saharan Africa. However, RDT results may be misrecorded in health facility registers, either unintentionally or deliberately, for example, to justify treatment based on clinical judgment. We conducted a multi-country evaluation to quantify the extent of RDT misrecording and identify factors associated with recording and reporting accuracy. This report summarizes the study's methods, key findings, and implications for improving malaria RDT data quality.

Methods

We conducted a mixed-methods observational study in Benin, Côte d'Ivoire, Nigeria, and Uganda in 2023. At participating health facilities, images of RDTs were captured at the point of care and later reviewed by a trained, independent panel. Agreement between the panel's interpretation and the result recorded in the register was assessed using weighted mean Cohen's kappa (κ). Secondary objectives included identifying factors associated with inaccurate recording, assessing the durability of RDT results after one month of storage, evaluating data accuracy in national health management information systems (HMIS), and measuring the observer effect on test positivity rates (TPR) reported to HMIS.

Results

We analyzed agreement on 102,091 RDT results. Agreement between register-recorded results and the external panel was high, ranging from $\kappa = 0.81$ (95% confidence interval [CI] 0.76, 0.85) in Nigeria to $\kappa = 0.88$ (95% CI 0.84, 0.92) in Benin. HCWs were more likely to misrecord results as positive (range: 5.0, 7.1%) than negative (range: 0.7, 3.5%). After the study began, TPRs reported from the HMIS declined more in study than control facilities in all countries except Benin. Decreases in study facilities ranged from 6.4 to 20.4 percentage points more than in control facilities, although only Côte d'Ivoire showed a statistically significant difference-in-difference (rate ratio: 0.80, 95% CI 0.76, 0.84).

Conclusions

Although HCWs generally recorded RDT results accurately, the disproportionately higher number misrecorded as positive raises concern about possible intentional misreporting. The observed TPR decrease after study onset in three of four countries suggests HCWs can be motivated to reduce misrecording and improve the quality of malaria case management and surveillance data.

BACKGROUND

The foundation of both case management and surveillance for malaria is the diagnostic confirmation of infection in persons with suspected malaria. The development of malaria rapid diagnostic tests (RDTs), which are small cassettes containing immunochromatographic strips that detect malaria-specific antigens in blood obtained from a finger-prick, has revolutionized testing for malaria. From 2010 to 2023, the annual number of suspected malaria cases across sub-Saharan Africa (SSA) receiving a diagnostic test increased more than five-fold from 54 million to 334 million [1]. This increase was almost entirely driven by the scale-up of malaria RDTs, which rose from 20 million instances in 2010 to 266 million in 2023. As a result, malaria treatment specificity has improved and routine surveillance data now rely primarily on parasitological confirmation rather than presumptive diagnosis.

The World Health Organization (WHO) recommends that all patients with suspected malaria undergo parasitological confirmation with either microscopy or RDT before starting antimalarial treatment. However, presumptive treatment may be initiated if diagnostics are not immediately available [2]. Patients with confirmed malaria should receive a full course of antimalarial treatment while those who test negative should not be given antimalarials. Instead, further evaluation is needed to identify the cause of illness and guide appropriate treatment. In SSA, however, prescription of antimalarials to those who test negative for malaria is widespread: a systematic review found that 22% (95% confidence interval [CI] 11, 34) of RDT-negative patients were still treated with antimalarials by healthcare workers (HCWs) [3].

While diagnostic tests are essential tools in clinical decision-making, they do not, on their own, determine a clinical diagnosis. HCWs may choose to rely on clinical judgment over RDT results due to known limitations of the tests: depending on the specific product and country context, RDTs can miss low density infections [4]; may fail to detect parasites lacking the histidine-rich protein 2 (HRP2) gene [5]; can remain positive for two weeks or longer following treatment [6]; and may not be designed to detect non-*Plasmodium falciparum* species that could be responsible for illness [7]. Beyond these technical limitations of RDTs, treatment decisions may also be shaped by external factors. Public demand for medication can pressure HCWs to prescribe antimalarials irrespective of test results, and the frequent lack of diagnostic tools and appropriate treatments for other febrile illnesses may further contribute to HCWs' treatment of RDT-negative patients with antimalarials [8, 9].

Several initiatives aimed at reducing presumptive treatment, promoting adherence to RDT results, and strengthening antimalarial supply chain integrity have placed considerable emphasis on aligning the number of positive malaria test results with the number of antimalarial treatments prescribed. The WHO Test, Treat & Track initiative that launched in 2012 emphasized the importance of testing before treating. In alignment with this approach, ministries of health (MOH) routinely assess the frequency of presumptive treatment and HCW adherence to negative test results through supportive supervision and end-user verification surveys [10, 12]. Similarly, the US President's Malaria Initiative recommends monitoring the rational use of antimalarial treatments by routinely comparing the number of test results to numbers of treatments administered [13].

An unintended consequence of these initiatives may have been the creation of incentives for HCWs to intentionally misrecord RDT results in health facility registers to account for the treatments provided. As a result, the reported number of confirmed malaria cases and the test positivity rate (TPR) may be artificially inflated. This misreporting can bias estimates of malaria burden and potentially lead to misguided programmatic or policy decisions.

Although multiple studies have found suboptimal alignment between RDT results, diagnoses and prescribed treatment, most evaluations have relied on patient reinterviews where adherence is assessed based on the result of a second RDT performed by trained research staff [14, 16]. These studies typically do not include an objective review of the original RDT cassette interpreted by the HCW at the point of care and are often limited by small sample sizes given the resources required to reinterview a large number of patients. To more accurately assess the extent to which RDT results recorded in health facility registers reflect an objective interpretation of the test outcome, we implemented the Malaria RDT Capture and Reporting Assessment (MaCRA) across public primary health care facilities in Benin, Côte d'Ivoire, Nigeria and Uganda. This paper outlines the common methodology used in the MaCRA study, summarizes key results and discusses their potential implications for malaria surveillance and case management in SSA.

METHODS

While this publication reports on the methods and key findings of the MaCRA study, more detailed statistical methods and complete results will be presented in future publications.

Study objectives

The primary objective of this study was to assess the accuracy of malaria RDT results recorded in public health facility outpatient and antenatal care registers and reported to the national health management information systems (HMIS) in Benin, Côte d'Ivoire, Nigeria, and Uganda, which are all based on the district health information system 2 (DHIS2, University of Oslo, Oslo, Norway) platform. The study explored several secondary objectives including: identification of characteristics of health facilities, HCWs, and patients associated with misrecording RDT results; measurement of the stability of RDT cassette results over time; comparison of RDT results reported from the DHIS2 with monthly summary forms and health facility registers; and measurement of the observer effect on TPRs reported from the DHIS2.

Study design

We conducted a mixed-methods, prospective observational study in selected health facilities across Benin, Côte d'Ivoire, Nigeria, and Uganda in 2023. Baseline health facility surveys and assessments of HCWs characteristics, knowledge, attitudes, perceptions, behaviors and skill in performing RDTs were undertaken in study facilities. Trained research assistants used a digital, artificial intelligence (AI)-

powered RDT reader (HealthPulse, Audere, Seattle, WA USA) on a smartphone to take images of all RDTs performed in the study facilities and record basic patient data, test results, diagnosis and treatment prescribed. Patients' RDT results as written in health facility registers were compared to RDT results obtained from a trained, external panel who reviewed and interpreted images of the same RDTs. RDT results aggregated from health facility registers were compared to data downloaded from the national DHIS2 for the same facilities and time period. The change in TPR reported from the DHIS2 after the start of the study was compared between study facilities and matched controls.

Study sites

The countries selected for this evaluation were chosen purposively to reflect a range in baseline TPRs in countries supported by the US President's Malaria Initiative (PMI) (Fig. 1). In all four countries, malaria is a leading cause of morbidity and mortality in children: malaria incidence in 2022 ranged from 262 cases per 1000 population in Côte d'Ivoire to 383 per 1000 population in Benin while malaria mortality per 100,000 population ranged from 40 in Côte d'Ivoire to 87 in Nigeria [17]. In 2022, Côte d'Ivoire tested 50.7% of its suspected malaria cases by RDT, while Benin, Uganda and Nigeria used RDTs to test 79.3%, 81.4% and 84.5%, respectively, of their malaria cases.

Two regions in each country (i.e. departments [Benin], health regions [Côte d'Ivoire], states [Nigeria] or subregions [Uganda]) were purposively selected for the evaluation through discussions with the ministry of health (MOH) and PMI. Benin, Côte d'Ivoire and Nigeria restricted selection to the areas supported by PMI to ensure availability of RDTs and antimalarial medicines during the study and excluded any regions experiencing violence or insurgency, while in Uganda, no geographic restriction was applied as PMI only supported commodities for private facilities. For Benin, Côte d'Ivoire and Nigeria, one region in the north with seasonal transmission of malaria and one region in the south with more perennial transmission were selected. In Uganda, two high-burden subregions were chosen. Within each region, two sub-regional areas (i.e. health zones [Benin], health district [Côte d'Ivoire], local governmental area [Nigeria] or districts [Uganda]) were purposively selected for the evaluation. Benin conducted a further selection of communes within health zones.

Only basic public health facilities were included in this evaluation. In Benin, Côte d'Ivoire and Nigeria, only health facilities without trained laboratorians were eligible for inclusion while in Uganda, facilities with microscopy services were also eligible. No hospitals or referral facilities were included in the evaluation. Public health facilities were further eligible if they reported a minimum average of 50 RDTs per month and had reported data on RDTs to the DHIS2 for at least nine months out of 12 for each year over the previous two years. Health facilities were stratified using medians by patient volume (based on the number of outpatient visits) and average TPR for 2021–2022. Within each of the four strata in each sub-region, one health facility was randomly selected as a study facility for a total of four study facilities per stratum and 16 study facilities per country. One additional health facility was selected from each stratum in each sub-region to serve as a control to measure the observer impact on TPRs reported from the DHIS2; no activities were undertaken and there was no engagement with staff in control health facilities.

Data collection instruments

Computer-assisted personal interviews and observations were conducted by trained research assistants using data collection instruments designed using KoboToolbox (Kobo, Cambridge, MA USA). KoboToolbox was also used to record information from health facility registers and monthly summary forms. RDT images and matching patient information from health facility registers were collected by research assistants using the HealthPulse application that was customized for this study. All data collection instruments were implemented in English in Nigeria and Uganda and translated into French for Benin and Côte d'Ivoire using DeepL (Cologne, Germany). French translations were verified by the principal investigators in Benin and Côte d'Ivoire. Data collection tools were piloted in non-evaluation areas and subsequently refined to improve the clarity and flow of the questions.

Study procedures

Research assistants in each country underwent a three- to four-day training that included an orientation to the study objectives and research ethics, a review of standardized operating procedures for data collection, and guidance on obtaining informed consent from participants. They were also trained to administer RDTs in order to accurately observe and document how health care workers performed the tests.

Research assistants were present during the main working hours of each facility. Most countries deployed one research assistant per facility; however, Benin assigned two per facility, as most sites included both outpatient and maternal health clinics. In Nigeria, facilities operating 24 hours a day were supported by two shifts of research assistants.

Facility survey

A survey was conducted at baseline to record the operating characteristics of each study facility through an interview with the facility officer in-charge and direct observation. The survey included: a count of the HCWs working in the facility by position and qualifications; presence of registers, guidelines and job aids; availability of laboratory diagnostic tests; first- and second-line antimalarial medicines; stockouts of medicines and supplies over the previous three months (a stockout was described as at least seven consecutive days without the medicine or supply); and status of infrastructure such as electricity and internet connection.

Knowledge, attitudes, perceptions and behavior (KAPB) survey and RDT performance evaluation

At baseline, all HCWs in the study health facilities who had current or potential responsibility for testing, interpretation, or recording of RDT results or treatment of patients with malaria were administered a

knowledge, attitudes, perceptions and behavior (KAPB) survey. The survey gathered data on: HCWs training and experience; knowledge of malaria transmission and case management; attitudes towards RDTs; perceptions of the accuracy of RDTs; behaviors associated with prescription of antimalarial medicines; and subjective norms with respect to RDTs, malaria case management and surveillance practices.

Health care workers (HCWs) who routinely performed RDTs at the facility were observed conducting an RDT in the presence of a research assistant. Their performance was assessed using a standardized 19-point checklist derived from a validated tool, covering biosafety, procedural steps, and result interpretation [18]. HCWs who joined study facilities after the start of the study were also observed performing an RDT and completed the KAPB survey, following written informed consent.

Each HCW who participated in the KAPB survey was assigned a unique, random identification code that was recorded with their name in a notebook kept locked outside of working hours. This unique identification code was used in the HealthPulse application to link HCWs who performed or recorded RDT results with their responses to the KAPB survey and RDT performance evaluation. To preserve the anonymity of results, all notebooks were gathered and destroyed at the end of the study after data cleaning was completed.

Collection of RDT images, results and patient data

The HealthPulse digital RDT reader application was originally developed as a smartphone tool for HCWs to improve performance and interpretation of RDTs. The application has an image quality assurance (IQA) component that leverages computer vision and machine learning processes to assess the quality of images of RDTs, immediately flagging those that do not meet quality standards (such as those with blur or multiple RDTs) and prompting the user to retake the photo. The HealthPulse application includes an AI algorithm that interprets the RDT result from the image but this result was not provided to the research assistant, HCW, external panel or study investigators (except for SC) during the study period and was not used to inform the clinical management of the patient.

Eight RDTs products expected to be used in one or more of the four evaluation countries were identified prior to the start of the study by the MOH; one additional product was added during the course of the study (Supplementary Table 1). All RDT products detected *P. falciparum* using the *P. falciparum*-specific histidine-rich protein 2 (HRP2) antigen. One product had a second test line with a biomarker to detect *P. falciparum*-specific lactate dehydrogenase (PfLDH) and two products had a second test line that detected pan-*Plasmodium* lactate dehydrogenase (pLDH).

Research assistants captured images of all RDTs performed at the facility as soon as possible after interpretation by HCWs, without disrupting patient care. Patients were managed by HCWs in accordance with national case management guidelines. Research assistants were instructed not to interact directly with patients and to minimize their impact on routine facility operations. Each RDT was labeled on the

back with a unique preprinted barcode, with a matching barcode affixed next to the corresponding entry in the health facility register. Barcodes were scanned using the HealthPulse application prior to image capture or data entry, enabling linkage of RDT images to register data. Research assistants used the application to record the unique identification codes of both the HCW who performed the RDT and the HCW who documented the result in the register. In addition, they recorded the patient's age, sex, RDT result as documented in the register, clinical diagnosis, and whether an antimalarial or antibiotic was prescribed.

Research assistants were unable to edit existing records in the HealthPulse application. If corrections were necessary, they were instructed to rescan the barcode and re-enter the data. The application retained only the most recent image or data entry associated with each barcode. The HealthPulse application had identical functionality offline as it did online. RDT images and patient data were automatically uploaded to a secure server once an internet connection became available. In Uganda, research assistants working in high-volume facilities with limited connectivity addressed upload backlogs by enabling mobile hotspots or relocating devices to areas with better signal coverage.

Objective interpretation of RDT results

Images of RDTs from the HealthPulse application were sent to a trained and quality-controlled external panel of 20 RDT interpreters who reviewed randomly-assigned images independently and were blinded to any metadata or patient data. Panelists interpreted the RDT result for each image (positive, negative, invalid, or indeterminate) by determining the presence, absence or obscuring of the control and test lines and flagged problematic photo conditions such as blur, glare, darkness, skew, and presence of multiple RDTs in a single photo. Panelists also identified excess blood in the result window and noted when test lines were faint. Panelists were trained on both good and poor-quality photos and were required to pass a test that involved correctly interpreting and labeling a set of 50 images. Quality control of results involved two stages: an initial review of 100% of images and a second stage that reviewed a random selection of 30% of images. If a reviewer identified errors in interpretation or labeling anomalies at any stage, the images were sent back to the pool for a fresh interpretation.

For the first two to four weeks of the study in each country, all images were reviewed by three panelists. After validation of the results indicated that there was near perfect agreement among panelists (Fleiss' kappa > 0.99 for each country, Supplementary Table 2), each RDT image was reviewed by a single panelist, randomly assigned. Panel interpretations of the RDT results were not made available to the study investigators (apart from SC) until the end of the study except for the limited initial validation.

Stability of RDT results

For the first three months of the study in Benin, Nigeria and Uganda, RDT cassettes were placed in paper envelopes and stored in cardboard boxes at ambient temperature and humidity in health facilities after their original interpretation. Research assistants rephotographed RDTs one week and one month after the initial interpretation. The barcode placed on the back of the cassette was scanned at each time

interval to allow for images to be matched over time. After the final image capture, RDT cassettes were disposed of according to national guidelines. All images were interpreted as positive, negative, invalid or indeterminate by the external panel.

Health facility registers, monthly summary forms and DHIS2 data

Research assistants summarized data from health facility registers on a daily basis, including the number and results of malaria RDT tests performed. Each month, research assistants digitized the aggregated malaria results included on the monthly summary form that was prepared for entry into the DHIS2. After the study was completed, data on the number of RDTs administered and positive RDT results were downloaded from the DHIS2 for the facilities and the time period covered by the study.

Investigation of the observer effect on test positivity rates

We assessed whether the initiation of the study and the presence of research assistants photographing RDTs influenced the TPR reported in DHIS2, indicating a potential observer effect. Specifically, we evaluated the presence of a Hawthorne effect, that is, whether HCWs altered their reporting or recording of RDT results due to awareness of being observed. The term "Hawthorne effect" originates from studies conducted at the Hawthorne Works of the Western Electric Company in Cicero, Illinois, during the 1920s and 1930s, which suggested that individuals modify their behavior when they know they are being studied [19, 20]. DHIS2 data on the number of positive RDT results and the total number of RDTs conducted were extracted for each study and control facility from January 2021 to December 2023. This period includes 30 months prior to study initiation and the 6-month study period.

Sample size

The sample size for the main objective was based on the level of precision for estimating Cohen's kappa (κ) for an individual HCW [21]. We assumed a range of κ scores between 0.7, 0.9 and calculated the number of RDTs per HCW needed for different levels of precision. The probability of a positive rating was estimated by the TPR. Assuming TPRs between 30, 70%, the maximum sample size needed to calculate 95% confidence intervals for κ with a width of no more than 0.2 was 236 RDTs per HCW. We estimated that individual HCWs would likely interpret between 40 and 80 RDTs each month, for a possible range of 320–480 observations per HCW over the course of the evaluation, which would be sufficient for a precise measurement of κ at both the HCW and health facility level. The number of study facilities included was fixed at 16 due to budget limitations.

Data management

The geographic coordinates of each study facility were recorded and used to calculate average *P. falciparum* parasite prevalence for children 2–10 years ($PfPR_{2-10}$) within 5 km of the facility using the malariaAtlas package in R (R Foundation for Statistical Computing, Vienna, Austria) [22].

Because no HCW ever recorded an invalid RDT result in the health facility register, and indeterminate outcomes were relevant only to external reviewers, records with invalid or indeterminate interpretations were excluded from analysis. These outcomes were rare, comprising fewer than 0.9% of all RDTs across the four countries. When HCWs and the external panel agreed on their interpretation of the RDT result, the outcome was classified as ‘positive agreement’ or ‘negative agreement.’ When a HCW recorded a positive RDT result in the register but the external panel interpreted a negative result from the image, the HCW result was classified as ‘misrecorded as positive.’ When a HCW recorded a negative RDT result but the external panel interpreted a positive result from the image, the HCW result was classified as ‘misrecorded as negative.’

HCWs who performed RDTs were classified as clinical (e.g. nurse, medical doctor, midwife, community health worker) or non-clinical staff (e.g. cleaner, guard, volunteer). Questions of perceptions or beliefs based on the Likert scale were dichotomized into agree (agree or strongly agree) or do not agree (neutral, disagree or strongly disagree). The 19-point RDT performance score was divided into terciles based on the distribution of scores in each country.

Statistical analysis

Statistical methods for the primary and secondary objectives are described briefly while more detailed statistical methods will be included in individual reports. Statistical analyses were conducted in R.

Accuracy of recorded RDT results

We calculated the interrater agreement between RDT results recorded by HCWs in the health facility register and RDT results interpreted by the external panel from images as a measure of accuracy. A bespoke R function based on the kappa2 function from the *irr* package was developed to summarize Cohen’s kappa scores by calculating a great mean and weighting by the inverse of the variance for each health facility [23, 24].

Factors associated with results misrecorded as positive or negative

Multinomial logistic regression was used to model the odds ratio (OR) of an RDT result being misrecorded as positive or negative based on characteristics of the facility, HCW and RDT, adjusted for health facility. 95% confidence intervals (95% CI) were calculated for all OR with robust standard errors adjusted for clustering of RDT results by HCW.

Stability of RDT results

RDTs that had an image taken one month (i.e. 25 to 35 days) after the RDT was administered were included in the analysis for this report; images categorized as ‘indeterminate’ were discarded as these reflected problems with the image rather than the RDT result. Simple descriptive statistics were used to calculate the proportion of initially positive and negative RDTs that retained their original result one month after the initial interpretation. Positive and negative predictive values (PPV and NPV, respectively)

of the one-month outcome for the original result were calculated and 95% CI were calculated using the Wilson score method for binomial proportions using the *binom* package in R [25].

Comparing numbers of positive RDTs by data source

The numbers of total and positive RDT results were compared between health facility registers and the DHIS2 in each country. Data reporting accuracy was calculated as the weighted absolute percentage error (WAPE)-based aggregate data reporting accuracy (WADRA). WAPE was calculated for each facility as the sum of the absolute values of the differences between the health facility register and the DHIS2 each month divided by the total number of observations in the register over the time period. WADRA was calculated as $100\% - \text{WAPE}$. The accuracy of each facility was categorized as high ($\text{WADRA} \geq 85\%$), moderate (70, 84%) or low ($< 70\%$) based on previously published thresholds [26]. The analysis included only months for which the study team was in place for the entire period.

Measuring the observer effect on test positivity rates

A difference-in-difference analysis was conducted with study and control health facilities, comparing the TPR from RDTs before and after the study was initiated and after it concluded. A negative binomial, mixed effects linear regression model was constructed using integrated nested Laplace approximation with health facility-month as the unit of analysis, the number of positive RDT results as the dependent variable and log-transformed counts of RDTs as the model offset [27]. Risk ratios (RR) with 95% credibility intervals (CI) were calculated.

Ethical issues

Written, informed consent was obtained from HCWs who participated in the KAPB survey and were observed administering an RDT. HCWs were informed that anonymous and unique identifiers would be used in the HealthPulse app to link information from the KAPB survey and RDT observation to RDTs performed or recorded by HCWs. No patients were consented by the study team as the RDT images and data from the health facility registers were recorded anonymously and obtained from secondary data sources, respectively. No personally identifiable information was obtained from patient records.

The PATH institutional review board approved the multi-country study protocol. In Benin, the Comité National d'Éthique pour la Recherche en Santé provided approval. In Côte d'Ivoire, the Comité National d'Éthique des Sciences de la Vie et de la Santé reviewed and approved the study. In Nigeria, approval was received from Oyo State Ministry of Health Research Ethics Committee, Sokoto State Health Research Ethics Committee and the National Health Research Ethics Committee of Nigeria. The Uganda National Council for Science and Technology and Vector Control Division-Research & Ethics Committee both reviewed and approved the study in Uganda.

RESULTS

The key results of the MaCRA study are presented in this report while more detailed results will be presented in separate publications.

Accuracy of recorded RDT results

Between June and December 2023, 110,843 RDTs were observed across the four countries; after matching images with patient records, 102,091 (92.1%) RDTs were available for analysis. TPRs calculated from RDT results recorded by HCWs in health facility registers were lowest in Nigeria (48.0%) and highest in Côte d'Ivoire (63.9%) and were higher than those calculated from the external panel by between 3 and 6 percentage-points (pct-pts; Table 1). The level of interrater agreement as measured by κ was high for each country and ranged from 0.81 (95% CI 0.76, 0.85) in Nigeria to 0.88 (95% CI 0.84, 0.92) in Benin. Among results not in agreement between HCWs and the external panel, HCWs were more likely to misrecord RDT results as positive than negative; the proportion of results misrecorded as positive ranged from 5.0% in Benin to 7.1% in Uganda while results misrecorded as negative ranged from 0.7% in Benin to 3.5% in Nigeria (Table 1).

Table 1

Number of rapid diagnostic tests observed in the study and measurement of interrater agreement between results recorded in facility registers and an external panel by country, 2023

Country	No. RDTs observed and matched with patient data	Positive RDTs recorded by HCWs n (TPR %)	Positive RDTs recorded by the external panel n (TPR %)	Cohen's kappa (95% CI)	Results misrecorded as positive n (%)	Results misrecorded as negative n (%)
Benin	35,702	19,468 (54.5)	17,909 (50.2)	0.88 (0.84, 0.92)	1797 (5.0)	238 (0.7)
Côte d'Ivoire	11,129	7114 (63.9)	6662 (59.9)	0.83 (0.77, 0.88)	650 (5.8)	198 (1.8)
Nigeria	18,255	8764 (48.0)	8302 (45.5)	0.81 (0.76, 0.85)	1101 (6.0)	639 (3.5)
Uganda	37,005	22,860 (61.8)	20,902 (56.5)	0.82 (0.79, 0.84)	2617 (7.1)	659 (1.8)
CI: confidence interval; HCWs: healthcare workers; N: number; RDTs: rapid diagnostic tests; TPR: test positivity rate.						

Factors associated with misrecording RDT results

There were no factors consistently and statistically associated with misrecording results as positive across all countries (Table 2). However, patients aged 15 years and older were consistently more likely to have results misrecorded as positive than patients 0, 4 years old in Benin (OR 2.76, 95% CI 2.38, 3.20), Côte d'Ivoire (OR 1.94, 95% CI 1.65, 2.28) and Uganda (OR 1.42, 95% CI 1.25, 1.62) although not in Nigeria (OR 1.10, 95% CI 0.89, 1.34). In Benin and Côte d'Ivoire but not the other two countries, children 5, 14 years old were also more likely than children 0, 4 years old to have their RDT result misrecorded as positive (Benin: OR 1.49, 95% CI 1.25, 1.77; Côte d'Ivoire: OR 1.38, 95% CI 1.10, 1.74).

Table 2
Factors associated with rapid diagnostic test results misrecorded as positive or negative in a multinomial logistic regression model adjusted for health facility, 2023

Variable	Benin n = 35,706 OR (95% CI)	Cote d'Ivoire n = 11,129 OR (95% CI)	Nigeria n = 18,255 OR (95% CI)	Uganda n = 37,020 OR (95% CI)
Misrecorded as positive				
Patient sex	1.12 (0.97, 1.30)	1.09 (0.91, 1.31)	1.11 (1.01, 1.23)	1.03 (0.95, 1.13)
Female	1.0	1.0	1.0	1.0
Male				
Patient age (years)	1.0	1.0	1.0	1.0
0–4	1.49 (1.25, 1.77)	1.38 (1.10, 1.74)	0.98 (0.81, 1.20)	1.01 (0.89, 1.14)
5–14	2.76 (2.38, 3.20)	1.94 (1.65, 2.28)	1.10 (0.89, 1.34)	1.42 (1.25, 1.62)
15 +				
HCW sex	1.17 (0.91, 1.51)	1.29 (0.83, 2.01)	0.71 (0.45, 1.13)	0.96 (0.85, 1.08)
Female	1.0	1.0	1.0	1.0
Male				
HCW age (years)	1.0	1.0	1.0	1.0
< 30	0.76 (0.59, 0.98)	2.06 (1.16, 3.66)	0.87 (0.55, 1.38)	0.74 (0.58, 0.95)
30–39	0.47 (0.36, 0.62)	2.08 (1.21, 3.55)	0.87 (0.63, 1.21)	0.78 (0.62, 0.99)
40 +				
HCW occupation category	0.88 (0.71, 1.08)	1.08 (0.64, 1.84)	0.76 (0.55, 1.05)	1.02 (0.75, 1.40)
Clinical staff	1.0	1.0	1.0	1.0
Non-clinical staff				
HCW highest qualification achieved	1.0	1.0	1.0	1.0
Secondary school or below	0.89 (0.69, 1.14)	1.43 (0.87, 2.36)	1.20 (0.91, 1.57)	0.89 (0.70, 1.13)
University				
HCW years of experience	1.0	1.0	1.0	1.0
0 – 4	0.78 (0.64, 0.96)	1.18 (0.78, 1.79)	0.87 (0.69, 1.11)	0.72 (0.57, 0.89)
CI: confidence interval; HCW: healthcare worker; OR: odds ratio; RDTs: rapid diagnostic tests				

Variable	Benin n = 35,706 OR (95% CI)	Cote d'Ivoire n = 11,129 OR (95% CI)	Nigeria n = 18,255 OR (95% CI)	Uganda n = 37,020 OR (95% CI)
Misrecorded as positive				
5 +				
HCW frequency of administering RDT	0.55 (0.28, 1.09)	0.93 (0.66, 1.33)	1.03 (0.69, 1.53)	0.49 (0.30, 0.80)
Very often				
Once in a while to often	0.46 (0.21, 1.03)	1.49 (0.68, 3.26)	0.26 (0.11, 0.62)	0.51 (0.30, 0.84)
Never	1.0	1.0	1.0	1.0
HCW considers RDTs easy to perform	1.08 (0.77, 1.51)	0.81 (0.42, 1.57)	0.53 (0.38, 0.74)	1.18 (0.79, 1.76)
Agree	1.0	1.0	1.0	1.0
Do not agree				
HCW has sufficient time to perform RDTs	1.40 (1.00, 1.96)	1.28 (0.57, 2.91)	0.53 (0.38, 0.74)	1.18 (0.90, 1.55)
Agree	1.0	1.0	1.0	1.0
Do not agree				
HCW could provide antimalarial to the patient even if RDT negative	1.02 (0.65, 1.60)	2.03 (1.05, 3.96)	0.68 (0.45, 1.03)	1.02 (0.91, 1.15)
Yes	1.0	1.0	1.0	1.0
No				
HCW received training in RDTs	1.21 (0.87, 1.70)	0.65 (0.29, 1.44)	0.95 (0.67, 1.36)	1.02 (0.85, 1.22)
Yes	1.0	1.0	1.0	1.0
No				
Supervisor observed HCW perform an RDT during past visit	1.22 (0.86, 1.72)	0.41 (0.27, 0.62)	2.80 (0.81, 9.70)	1.11 (0.88, 1.39)
Yes	1.0	1.0	1.0	1.0
No				
RDT proficiency tercile	1.0	1.0	1.0	1.0
Low	1.31 (0.86, 2.01)	4.09 (1.13, 14.90)	1.34 (0.98, 1.84)	0.94 (0.64, 1.39)
Middle				
CI: confidence interval; HCW: healthcare worker; OR: odds ratio; RDTs: rapid diagnostic tests				

Variable	Benin n = 35,706 OR (95% CI)	Cote d' Ivoire n = 11,129 OR (95% CI)	Nigeria n = 18,255 OR (95% CI)	Uganda n = 37,020 OR (95% CI)
Misrecorded as positive				
High	1.56 (1.06, 2.30)	3.66 (0.97, 13.82)	0.85 (0.64, 1.14)	0.83 (0.58, 1.20)
Misrecorded as negative				
Patient sex	0.94 (0.70, 1.25)	0.93 (0.74, 1.17)	1.00 (0.82, 1.22)	1.08 (0.92, 1.27)
Female	1.0	1.0	1.0	1.0
Male				
Patient age (years)	1.0	1.0	1.0	1.0
0–4	0.73 (0.50, 1.07)	1.04 (0.69, 1.57)	0.94 (0.74, 1.21)	1.09 (0.87, 1.37)
5–14	1.81 (1.28, 2.56)	1.96 (1.52, 2.54)	1.17 (0.94, 1.44)	1.73 (1.36, 2.22)
15+				
HCW sex	1.15 (0.71, 1.89)	1.00 (0.74, 1.34)	1.03 (0.75, 1.40)	0.94 (0.80, 1.10)
Female	1.0	1.0	1.0	1.0
Male				
HCW age (years)	1.0	1.0	1.0	1.0
< 30	0.88 (0.49, 1.58)	0.87 (0.55, 1.37)	0.98 (0.79, 1.23)	1.08 (0.87, 1.35)
30–39	1.24 (0.64, 2.38)	0.77 (0.49, 1.20)	1.33 (0.93, 1.89)	1.12 (0.90, 1.40)
40 +				
HCW occupation category	1.05 (0.72, 1.55)	1.17 (0.74, 1.85)	0.86 (0.59, 1.25)	1.29 (0.92, 1.81)
Clinical staff	1.0	1.0	1.0	1.0
Non-clinical staff				
HCW highest qualification achieved	1.0	1.0	1.0	1.0
Secondary school or below	0.73 (0.31, 1.75)	0.80 (0.47, 1.36)	0.94 (0.62, 1.42)	1.34 (0.98, 1.83)
University				
HCW years of experience	1.0	1.0	1.0	1.0
0–4				
CI: confidence interval; HCW: healthcare worker; OR: odds ratio; RDTs: rapid diagnostic tests				

Variable	Benin n = 35,706 OR (95% CI)	Cote d'Ivoire n = 11,129 OR (95% CI)	Nigeria n = 18,255 OR (95% CI)	Uganda n = 37,020 OR (95% CI)
Misrecorded as positive				
5 +	1.09 (0.49, 2.42)	0.76 (0.57, 1.03)	1.22 (0.98, 1.53)	1.20 (0.95, 1.51)
HCW frequency of administering RDT	0.50 (0.19, 1.33)	1.03 (0.61, 1.75)	1.56 (0.95, 2.56)	0.46 (0.08, 2.72)
Very often	0.07 (0.01, 0.58)	0.69 (0.31, 1.54)	1.57 (0.75, 3.27)	0.34 (0.06, 2.06)
Once in a while to often				
Never	1.0	1.0	1.0	1.0
HCW considers RDTs easy to perform	1.29 (0.72, 2.31)	0.87 (0.44, 1.72)	0.91 (0.59, 1.41)	0.56 (0.45, 0.69)
Agree	1.0	1.0	1.0	1.0
Do not agree				
HCW sufficient time to perform RDTs	1.08 (0.70, 1.67)	0.77 (0.56, 1.07)	0.61 (0.33, 1.12)	1.09 (0.77, 1.54)
Agree	1.0	1.0	1.0	1.0
Do not agree				
HCW could provide antimalarial to patient with negative RDT	1.60 (0.89, 2.83)	1.01 (0.64, 1.58)	0.89 (0.60, 1.33)	0.95 (0.76, 1.17)
Yes	1.0	1.0	1.0	1.0
No				
Received training in RDTs	1.06 (0.71, 1.57)	0.59 (0.31, 1.13)	0.87 (0.69, 1.08)	0.99 (0.83, 1.19)
Yes	1.0	1.0	1.0	1.0
No				
Supervisor observed HCW perform an RDT during past visit	0.93 (0.65, 1.33)	1.13 (0.85, 1.50)	1.73 (0.72, 4.15)	0.76 (0.62, 0.94)
Yes	1.0	1.0	1.0	1.0
No				
RDT proficiency tercile	1.0	1.0	1.0	1.0
Low	1.68 (0.74, 3.85)	0.67 (0.26, 1.70)	1.11 (0.77, 1.58)	1.24 (0.96, 1.61)
Middle				
CI: confidence interval; HCW: healthcare worker; OR: odds ratio; RDTs: rapid diagnostic tests				

Variable	Benin n = 35,706 OR (95% CI)	Cote d'Ivoire n = 11,129 OR (95% CI)	Nigeria n = 18,255 OR (95% CI)	Uganda n = 37,020 OR (95% CI)
	Misrecorded as positive			
High	1.01 (0.62, 1.65)	0.50 (0.19, 1.28)	0.88 (0.57, 1.37)	0.89 (0.65, 1.22)
CI: confidence interval; HCW: healthcare worker; OR: odds ratio; RDTs: rapid diagnostic tests				

In Benin, older HCWs and those with more years of experience were less likely to misrecord RDT results as positive (Table 2). In Côte d'Ivoire, however, older HCWs were more likely to misrecord RDT results as positive. Additionally, HCWs who answered 'Yes' to the statement, "Might you provide an antimalarial medicine to a patient even if their RDT returns a negative result?" were more likely to misrecord RDT results as positive. HCWs whose supervisor had observed them performing an RDT were less likely to misrecord results as positive. In Nigeria, female patients were more likely to have their RDT results misrecorded as positive while HCWs that considered RDTs easy to perform were less likely to misrecord results as positive. In Uganda, older HCWs, those with more years of experience and those who frequently administered RDTs were less likely to misrecord a result as positive. HCWs with RDT proficiency scores in the top two terciles were conversely more likely to misrecord an RDT as positive in Benin as those in the lowest, but there was no statistical association between RDT proficiency scores and misrecording results as positive in any other country.

Few factors were statistically associated with misrecording results as negative (Table 2). Patients 15 years and older were more likely than patients aged 0, 4 years to have results misrecorded as negative in Benin (OR 1.81, 95% CI 1.28, 2.56), Côte d'Ivoire (OR 1.96, 95% CI 1.52, 2.54) and Uganda (OR 1.73, 95% CI 1.36, 2.22) but not Nigeria. In no country were patients 5, 14 years old more likely than younger children to have results misrecorded as negative.

Stability of RDT cassette results over time

There were 44,605 stored RDT cassettes that were re-imaged one month after they were administered. The external panel determined the result for each image and we found that 42,408 (95.1%) retained their original interpretation. At the time of administration, there were 25,129 (56.3%) positive, 19,420 (43.5%) negative, and 56 (0.1%) invalid results. At the end of one month, there were 24,864 (55.7%) positive, 19,733 (44.2%) negative and 8 (0.0%) invalid results. Over the one-month period, we observed that 1228 (4.9%) positive results changed to negative, 921 (4.7%) negative results changed to positive, 42 (75.0%) invalid results converted to positive and 6 (10.7%) invalid results converted to negative. The PPV of a positive outcome at one month was 96.1% (95% CI 95.9, 96.4), the NPV was 93.7% (95% CI 93.4, 94.1) and the equivalent for an invalid test was 100% (95% CI 67.5, 100).

Comparing numbers and results of RDTs between the DHIS2 and health facility registers

In Benin, nearly all health facilities were categorized as having a high level of reporting accuracy for their DHIS2 data, defined as a WADRA greater than 85% for both positive and total RDT results (Fig. 2). No facility in Benin was classified as having low reporting accuracy, defined as a WADRA below 70%. In contrast, between 19% and 38% of facilities in Nigeria and Uganda were classified as having low reporting accuracy, depending on the outcome assessed. Côte d'Ivoire had few facilities with low reporting accuracy but a substantial proportion of facilities with only moderate accuracy.

Measuring the observer effect on test positivity rates

Comparing the periods before and after the start of the study, TPRs reported by the DHIS2 declined more in study facilities than in control facilities across all countries (Table 3). However, the difference was only statistically significant in Côte d'Ivoire. In Côte d'Ivoire, TPRs in study facilities decreased by 20.4 pct-pts more than in control facilities and modeled results estimated an RR of 0.80 (95% credibility interval 0.76, 0.84). While Nigeria and Uganda also showed larger reductions in TPRs in study facilities compared to control facilities (10.2 pct-pts and 6.4 pct-pts, respectively), the credibility intervals for these estimates included the null, indicating that the differences were not statistically significant. Benin exhibited the smallest difference, with a 1.3 pct-pt greater decline in TPRs in study facilities compared to control facilities.

Table 3

Changes in observed test positivity rates before and after the MaCRA study was initiated in study and control facilities and modeled difference-in-difference results, January 2021, December 2023

Country	TPR before study start		TPR after study start		Difference of differences	Modeled results
	Study facilities	Control facilities	Study facilities	Control facilities	Pct-pts	RR (95% CI)
Benin	70.1	69.8	49.1	50.1	1.3	0.98 (0.87, 1.11)
Côte d'Ivoire	85.3	86.1	68.2	89.3	20.4	0.80 (0.76, 0.84)
Nigeria	72.7	69.0	58.5	64.8	10.2	0.90 (0.79, 1.04)
Uganda	67.9	66.2	58.4	63.1	6.4	0.92 (0.83, 1.00)
CI: credibility interval; Pct-pts: percentage points; RR: risk ratio; TPR: test positivity rate.						

DISCUSSION

We compared individual RDT results recorded by HCWs in health facility registers with those independently interpreted by an external panel reviewing images of the same RDTs and found high levels of agreement. However, a notable proportion of RDTs was misrecorded, either as positive (5.0, 7.1%) or negative (0.7, 3.5%), suggesting that a substantial proportion of RDT results may not have been accurately reported to the DHIS2. RDT results misrecorded as positive were consistently more frequent than those misrecorded as negative, suggesting that such discrepancies did not occur at random. These errors may occur when HCWs decide to treat an RDT-negative patient for malaria and intentionally misrecord a positive result to justify treatment.

We found only patient age to be consistently associated with results misrecorded as positive in three out of the four countries. While the potential consequences of a missed infection might be more severe in young children, we found the odds of negative results being misrecorded as positive to be higher in patients aged 15 years and older than in the youngest children. Several qualitative studies have pointed to the pressure that patients may place on HCWs to prescribe antimalarials despite negative RDT results [9,28,29]. Other factors identified in the literature, such as professional cadre, level of training and perceptions of RDTs, were not found to be associated with misrecording negative results as positive in this study [30–33].

Results misrecorded as negative may arise from a failure to wait long enough for an RDT to fully develop before interpreting the result, or vision limitations that make faint lines difficult to see. HCWs may also decide to misrecord a positive result as negative if they are concerned that a positive result has arisen because of a prior malaria infection and want to avoid potential overtreatment [6]. However, not providing antimalarials to those who are symptomatic and test positive for malaria by RDT risks development of severe disease or death and HCWs should be counseled on appropriate practices.

In addition to errors resulting from the misrecording of individual RDT results, additional inaccuracies may occur during the processes of data aggregation, summarization, and manual entry into DHIS2. Using health facility registers as the source documents, we found that the accuracy of total and positive RDT numbers reported to the DHIS2 was high in Benin, lower in Côte d'Ivoire, and very low in a substantial proportion of study facilities in Nigeria and Uganda. Implementing a review process to verify monthly summary forms and ensure the accuracy of data aggregation at the health facility level may help address data concordance issues in Nigeria and Uganda.

In Côte d'Ivoire, the start of the study was associated with a large decline in TPR in study facilities while matched control facilities showed a small increase. Nigeria and Uganda also demonstrated substantial decreases in TPR in study facilities compared to control facilities, whereas Benin experienced only a minimal decline. These findings have two key implications. First, the high level of agreement observed in Côte d'Ivoire, Nigeria and Uganda between HCW RDT results recorded in health facility registers and those interpreted by an external panel may be an overestimate. Second, the results suggest that increased attention to RDT results can influence HCW behavior, improving the accuracy of recorded data. Supporting this conclusion, Benin, which had the highest agreement in RDT results and did not exhibit a

decline in TPR after the study began, had already implemented a national program for monthly RDT validation. Since January 2023, six months before the study's initiation. Benin began comparing health facility registers to RDT cassettes through monthly validation meetings [34]. A future publication will provide evidence of a nationwide reduction in TPR coinciding with the start of this validation program, further supporting the potential to influence HCW behavior and improve data accuracy. Our findings demonstrating the stability of RDT results on cassettes kept for one month provide evidence that stored cassettes can be used to validate health facility registers.

This study had several limitations. First, the reference standard used to assess the accuracy of recorded RDT results was based on photographs of the RDTs. As cameras may not capture all visual details observable to the human eye, this reference standard was imperfect [35]. However, given the large number of RDTs assessed, we do not expect this limitation to have meaningfully affected the overall findings. Second, HCW characteristics were linked to the individual who recorded the RDT result in the facility register as we considered that person to have made the final decision on the RDT result to be recorded. In some cases, however, the decision as to which RDT result to record may have been made by a different HCW, introducing potential non-differential misclassification and attenuating associations between HCW characteristics and misrecording. Finally, the study was limited in duration, number of facilities, and geographic scope, and thus may not be fully representative of the broader health system context in each country.

CONCLUSIONS

RDT results recorded in health facility registers demonstrated a high level of agreement with objective panel reviews of RDT images in Benin, Côte d'Ivoire, Nigeria, and Uganda. However, inaccuracies in recording results, particularly negative outcomes, remain a concern. Our study may have underestimated the true extent of misrecording, particularly in Côte d'Ivoire.

The observed change in TPR following the start of the study in Côte d'Ivoire, and to a lesser extent in Nigeria and Uganda, likely due to an observer effect generated by the presence of research assistants who were photographing all RDTs, indicates that HCWs can be motivated to improve the accuracy of malaria data. The experience in Benin, where monthly RDT validation meetings have been held at health zone level nation-wide since January 2023, highlights a promising approach to enhancing the accuracy of RDT results. Prospective testing of this strategy and evaluation of its acceptability to HCWs, feasibility of implementation and duration of effect will provide national malaria programs with evidence-based strategies to improve and sustain accuracy of routine malaria surveillance data.

Abbreviations

AI	Artificial intelligence
CI	Confidence interval or credibility interval
DHIS2	District health information system 2
HCW	Healthcare worker
HMIS	Health management information system
HRP2	Histidine-rich protein 2
IQA	Image quality assurance
IRB	Institutional review board
KAPB	Knowledge, attitudes, perceptions and behaviors
MOH	Ministry of health
N	Number
OPD	Outpatient department
OR	Odds ratio
PCR	Polymerase chain reaction
Pct-pts	Percentage-points
Pf	<i>Plasmodium falciparum</i>
pLDH	Parasite lactate dehydrogenase
PMI	President's Malaria Initiative
RDT	Rapid diagnostic test
RR	Risk ratio
SE	Standard error
SSA	Sub-Saharan Africa
TPR	Test positivity rate
WAPE	Weighted average percentage error
WHO	World Health Organization

Declarations

ACKNOWLEDGEMENTS

We gratefully acknowledge the healthcare workers who gave generously of their time to participate in this evaluation. Many staff members of the organizations implementing the study in country made important contributions: Manfred Accrombessi, Corneille Hueha (CREC, Benin); M. Anatole Mian, Orphée Kangah-Kouakou, Valérie Bedia-Tanoh (INSP, Côte d'Ivoire); Hilary Okagbue, Evelyn Orya, Shiva Gab-Deedam, Olufisayo Bademosi, (Sydani Group, Nigeria); Anne Katahoire, Jane Frances Namuganga, Jenipher Musoke (CHDC, Uganda). We are grateful to the national malaria programs in each country for their support. We received excellent research support from Saadjo Sow (PMI Insights). Megan Littrell and Kim Vu provided overall direction and financial support, respectively, to the PMI Insights project. Several PMI Insights staff supported important elements of the study: Samantha Herrera, Annie Arnzen and Maia Cullen. Aysu Uygur (Bill & Melinda Gates Foundation) is greatly appreciated for her contributions during the design of the evaluation. Many PMI staff in country provided introductions, country-specific context and additional insights that facilitated the study design, implementation and conclusions: Raoul Oloukoi, Virgile Gnanguenon (Benin); Pascal Zinzindohoue, Patricia Yepassis-Zembrou, Blaise Kouadio, Melaine Tape, Yao Stephane, Christie Billingsley (Côte d'Ivoire); Jules Mihigo, Cassandra Elagbaje, Veronica Momoh, Valerie Bampoe (Nigeria); Edgar Agba, Grace Appiah, Patrick Condo (Uganda). Emily Hilton and Natalie Galles (PATH) and Sarah Burnett (PMI) supported data analysis. Sasha Frade and Sam Smedinghoff (Audere, Johannesburg, South Africa) supported development and refinement of the HealthPulse application and designed dashboards used during the study. We acknowledge the contributions of Audere's AI data creation and labeling teams at the Centre for HIV-AIDS Prevention Studies (South Africa) and Indivillage (India). We appreciate the engagement of Erick Tshikamba (PMI) and Julia Mwesigwa (PATH) in the initial protocol harmonization meeting.

FUNDING

This evaluation was co-funded by PMI Insights and the Bill & Melinda Gates Foundation. PMI Insights was the global operational research and program evaluation project of the U.S. President's Malaria Initiative (PMI). Funding for this evaluation is made possible by the generous support of the American people through the United States Agency for International Development (USAID). The contents are the responsibility of the authors and do not necessarily reflect the views of USAID or the United States Government.

AUTHOR INFORMATION

Authors and Affiliations

PMI Insights Project/PATH, Geneva, Switzerland

John J. Aponte

Kim A. Lindblade

Child Health and Development Centre, Makerere University, Kampala, Uganda

Arthur Mpimbaza

Nelson Ssewante

Centre de Recherche Entomologique de Cotonou, Cotonou, Benin

Idelphonse Ahogni

Corine Ngufor

Institut National de Santé Publique, Abidjan, Côte d'Ivoire

Abibatou Konate-Toure

William Yavo

Sydani Group, Abuja, Nigeria

Ese Akpiroroh

Sunday Atobatele

Audere, Seattle, WA USA

Shawna Cooper

U.S. President's Malaria Initiative, United States Agency for International Development, Washington, DC USA

Kevin Griffith

Michael Humes

Radina Soebiyanto

National Malaria Control Division, Ministry of Health, Kampala, Uganda

Bosco Agaba

Jimmy Opigo

Programme National de Lutte contre le Paludisme, Cotonou, Benin

Cyriaque Affoukou

Augustin Kpemasse

Programme National de Lutte contre le Paludisme, Abidjan, Côte d'Ivoire

Jacques Agnon

Antoine Tanoh

National Malaria Elimination Programme, Abuja, Nigeria

Godwin Ntadom

Onyebuchi Okoro

Contributions

KL, MH and KG conceived and designed the evaluation. AK, AKT, AM, AT, BA, CA, CN, EA, GN, IA, JJA, JO, NS, OO, SA, and WY oversaw data collection activities. SC oversaw the development of the HealthPulse application used in the study. KL drafted the manuscript. KL, JJA, MH and RS analyzed the data. All authors critically reviewed the manuscript. All authors read and approved the final manuscript.

Corresponding author

Correspondence to: Kim A. Lindblade, Rue du Grand-Pré, 1202 Geneva, Switzerland ;
kalindblade@gmail.com.

Ethical approval and consent to participate

Ethical approval was obtained from: the Comité National d’Ethique pour la Recherche en Santé (Benin); the Comité National d’Ethique des Sciences de la Vie et de la Santé (Côte d’Ivoire); the National Health Research Ethics Committee (Nigeria); the Oyo State Health Research Ethics Committee (Nigeria); the Sokoto State Health Research Ethics Committee (Nigeria); the Uganda National Council on Science and Technology (Uganda); the Vector Control Division Research and Ethics Committee (Uganda); and the WGC IRB in the USA. All participants provided written, informed consent to participate.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analyzed during the current evaluation can be provided by the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

References

1. WHO. World Malaria Report 2024. Geneva, Switzerland: World Health Organization; 2024.
2. WHO. WHO Guidelines for Malaria (30 November 2024) [Internet]. Geneva, Switzerland: World Health Organization; 2024 [cited 2024 Dec 17]. Available from: <https://app.magicapp.org/#/guideline/9683>
3. Kabaghe AN, Visser BJ, Spijker R, Phiri KS, Grobusch MP, van Vugt M. Health workers' compliance to rapid diagnostic tests (RDTs) to guide malaria treatment: a systematic review and meta-analysis. *Malar J*. 2016;15:163.
4. Wu L, van den Hoogen LL, Slater H, Walker PGT, Ghani AC, Drakeley CJ, et al. Comparison of diagnostics for the detection of asymptomatic *Plasmodium falciparum* infections to inform control and elimination strategies. *Nature*. 2015;528:S86–93.
5. Mekonen B, Dugassa S, Feleke SM, Dufera B, Gidisa B, Adamu A, et al. Widespread pfhrp2/3 deletions and HRP2-based false-negative results in southern Ethiopia. *Malar J*. 2024;23:108.
6. Dalrymple U, Arambepola R, Gething PW, Cameron E. How long do rapid diagnostic tests remain positive after anti-malarial treatment? *Malar J*. 2018;17:228.
7. Sendor R, Mitchell CL, Chacky F, Mohamed A, Mhamilawa LE, Molteni F, et al. Similar Prevalence of *Plasmodium falciparum* and Non-*P. falciparum* Malaria Infections among Schoolchildren, Tanzania. *Emerg Infect Dis*. 2023;29:1143.
8. Baltzell K, Kortz TB, Scarr E, Blair A, Mguntha A, Bandawe G, et al. “Not all fevers are malaria”: a mixed methods study of non-malarial fever management in rural southern Malawi. *Rural Remote Health* [Internet]. 2019 [cited 2024 Nov 5];19. Available from: <https://www.rrh.org.au/journal/article/4818/>
9. Altaras R, Nuwa A, Agaba B, Streat E, Tibenderana JK, Strachan CE. Why do health workers give anti-malarials to patients with negative rapid test results? A qualitative study at rural health facilities in western Uganda. *Malar J*. 2016;15:23.
10. Koko D, Arouna D, Bernard Y-M, Ba T, Mostel J, Abdou Y, et al. How Outreach Training and Supportive Supervision (OTSS) Affect Health Facility Readiness and Health-Care Worker Competency to Prevent and Treat Malaria in Niger: A Secondary Analysis of OTSS Data. *Am J Trop Med Hyg*. 2024;110:50–5.
11. USAID Global Health Supply Chain Program. EUV Survey Question Guide [Internet]. Washington, DC: USAID Global Health Supply Chain Program; 2020. Available from: <https://www.ghsupplychain.org/euv-survey-question-guide>
12. Agbemaflle EE, Addo-Lartey A, Odikro MA, Frimpong JA, Kubio C, Ameme DK, et al. Adherence to the test, treat and track strategy for malaria control among prescribers, Mfantseman Municipality, Central Region, Ghana. *PLOS ONE*. 2023;18:e0279712.
13. USAID DELIVER Project. Guidelines for Managing the Malaria Supply Chain: A Companion to the Logistics Handbook [Internet]. Arlington, VA: USAID DELIVER Project; 2011. Available from: https://ghsupplychain.org/sites/default/files/2019-07/GuidManaMalariaSC_0.pdf
14. Davlantes E, Camara A, Guilavogui T, Fofana A, Balde M, Diallo T, et al. Quality of Malaria Case Management and Reporting at Public Health Facilities in Six Health Districts in Guinea, 2018. *Am J*

Trop Med Hyg. 2019;101:148–56.

15. Zurovac D, Githinji S, Memusi D, Kigen S, Machini B, Muturi A, et al. Major Improvements in the Quality of Malaria Case-Management under the “Test and Treat” Policy in Kenya. PLoS ONE. 2014;9:e92782.
16. Plucinski MM, Ferreira M, Ferreira CM, Burns J, Gaparayi P, João L, et al. Evaluating malaria case management at public health facilities in two provinces in Angola. Malar J. 2017;16:186.
17. WHO. World Malaria Report 2023. Geneva, Switzerland: World Health Organization; 2023.
18. Boyce MR, Menya D, Turner EL, Laktabai J, Prudhomme-O’Meara W. Evaluation of malaria rapid diagnostic test (RDT) use by community health workers: a longitudinal study in western Kenya. Malar J. 2018;17:206.
19. Wickström G, Bendix T. The “Hawthorne effect” – what did the original Hawthorne studies actually show? Scand J Work Environ Health. 2000;26:363–7.
20. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. J Clin Epidemiol. 2014;67:267.
21. Shoukri MM, Asyali MH, Donner A. Sample size requirements for the design of reliability study: review and new results. Stat Methods Med Res. 2004;13:251–71.
22. Pfeffer DA, Lucas TCD, May D, Harris J, Rozier J, Twohig KA, et al. malariaAtlas: an R interface to global malariometric data hosted by the Malaria Atlas Project. Malar J. 2018;17:352.
23. Sun S. Meta-analysis of Cohen’s kappa. Health Serv Outcomes Res Methodol. 2011;11:145–63.
24. Gamer M, Lemon J, Fellows I, Singh P. Package “irr” [Internet]. Hamburg, Germany: University of Hamburg; [cited 2025 Apr 4]. Available from: <https://cran.r-project.org/web/packages/irr/irr.pdf>
25. Dorai-Raj S. Package “binom” [Internet]. 2022. Available from: <https://cran.r-project.org/web/packages/binom/binom.pdf>
26. Das S, Feltrer AR, Rutagwera M-RI, Lungu C, Malama P, Monde M, et al. Documenting Trends in Malaria Data Reporting Accuracy Using Routine Data Quality Audits in Zambia, 2015–2022. Am J Trop Med Hyg [Internet]. 2024 [cited 2025 Jan 24];1. Available from: <https://www.ajtmh.org/view/journals/tpmd/aop/article-10.4269-ajtmh.24-0429/article-10.4269-ajtmh.24-0429.xml>
27. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J R Stat Soc Ser B Stat Methodol. 2009;71:319–92.
28. Boadu NY, Amuasi J, Ansong D, Einsiedel E, Menon D, Yanow SK. Challenges with implementing malaria rapid diagnostic tests at primary care facilities in a Ghanaian district: a qualitative study. Malar J. 2016;15:126.
29. Diggle E, Asgary R, Gore-Langton G, Nahashon E, Mungai J, Harrison R, et al. Perceptions of malaria and acceptance of rapid diagnostic tests and related treatment practises among community members and health care providers in Greater Garissa, North Eastern Province, Kenya. Malar J. 2014;13:502.

30. Obi IF, Sabitu K, Olorukooba A, Adebowale AS, Usman R, Nwokoro U, et al. Health workers' perception of malaria rapid diagnostic test and factors influencing compliance with test results in Ebonyi state, Nigeria. *PLOS ONE*. 2019;14:e0223869.
31. Akinyode AO, Ajayi IO, Ibrahim MS, Akinyemi JO, Ajumobi OO. Practice of antimalarial prescription to patients with negative rapid test results and associated factors among health workers in Oyo State, Nigeria. *Pan Afr Med J [Internet]*. 2018 [cited 2024 Dec 17];30. Available from: <https://www.panafrican-med-journal.com//content/article/30/229/full>
32. Burchett HED, Leurent B, Baiden F, Baltzell K, Björkman A, Bruxvoort K, et al. Improving prescribing practices with rapid diagnostic tests (RDTs): synthesis of 10 studies to explore reasons for variation in malaria RDT uptake and adherence. *BMJ Open*. 2017;7:e012973.
33. Na'uzo AM, Tukur D, Sufiyan MB, Stephen AA, Ajayi I, Bamgboye E, et al. Adherence to malaria rapid diagnostic test result among healthcare workers in Sokoto metropolis, Nigeria. *Malar J*. 2020;19:2.
34. Ngufor, C, Ahogni, I. Documentation of the monthly validation of malaria rapid diagnostic test (RDTs) results in Benin. Cotonou, Benin; 2025 Apr.
35. Visser T, Ramachandra S, Pothin E, Jacobs J, Cunningham J, Menach AL, et al. A comparative evaluation of mobile medical APPS (MMAS) for reading and interpreting malaria rapid diagnostic tests. *Malar J*. 2021;20:39.

Figures

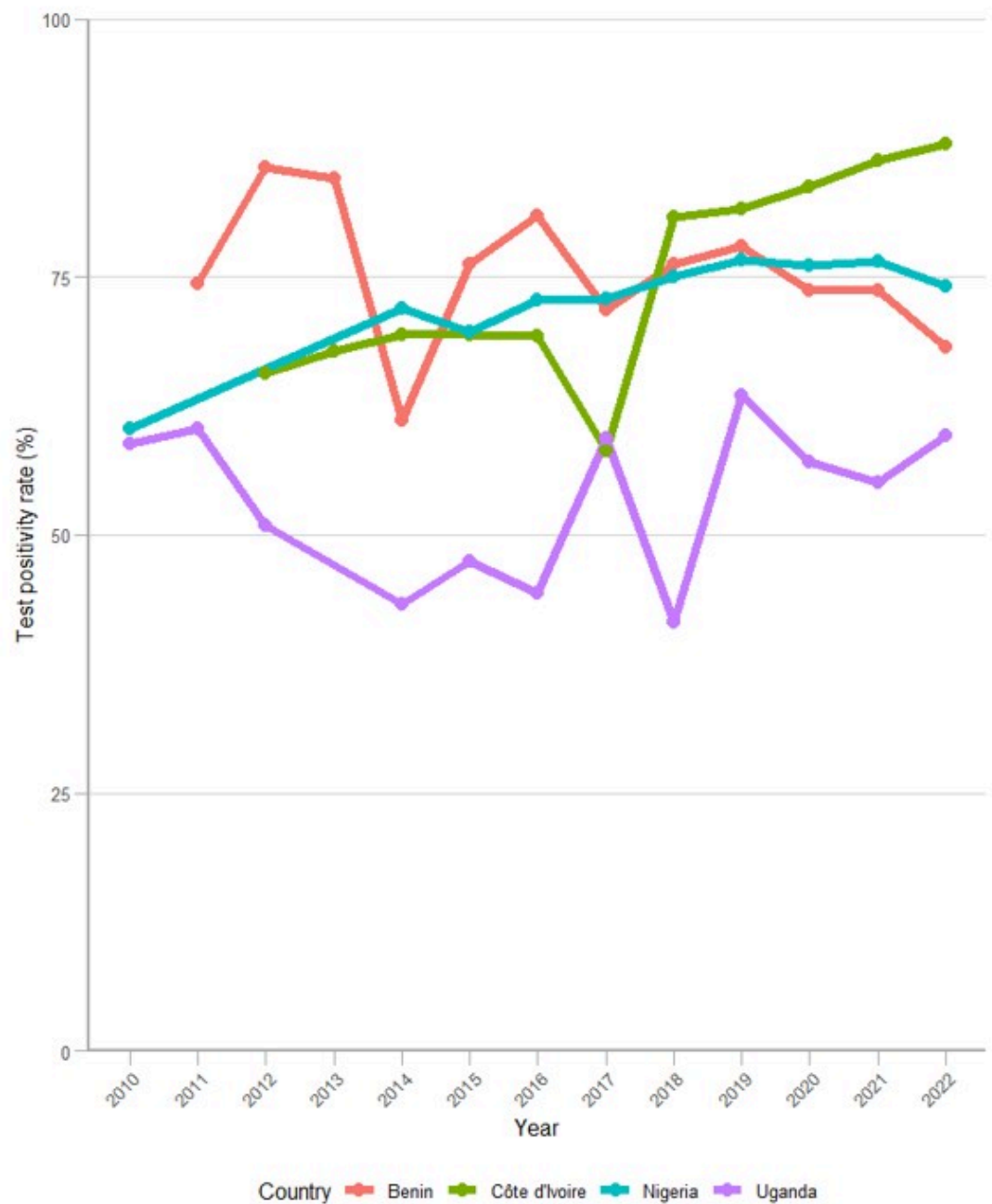


Figure 1

National test positivity rates from rapid diagnostic tests in the four study countries, 2010-2022 [17]

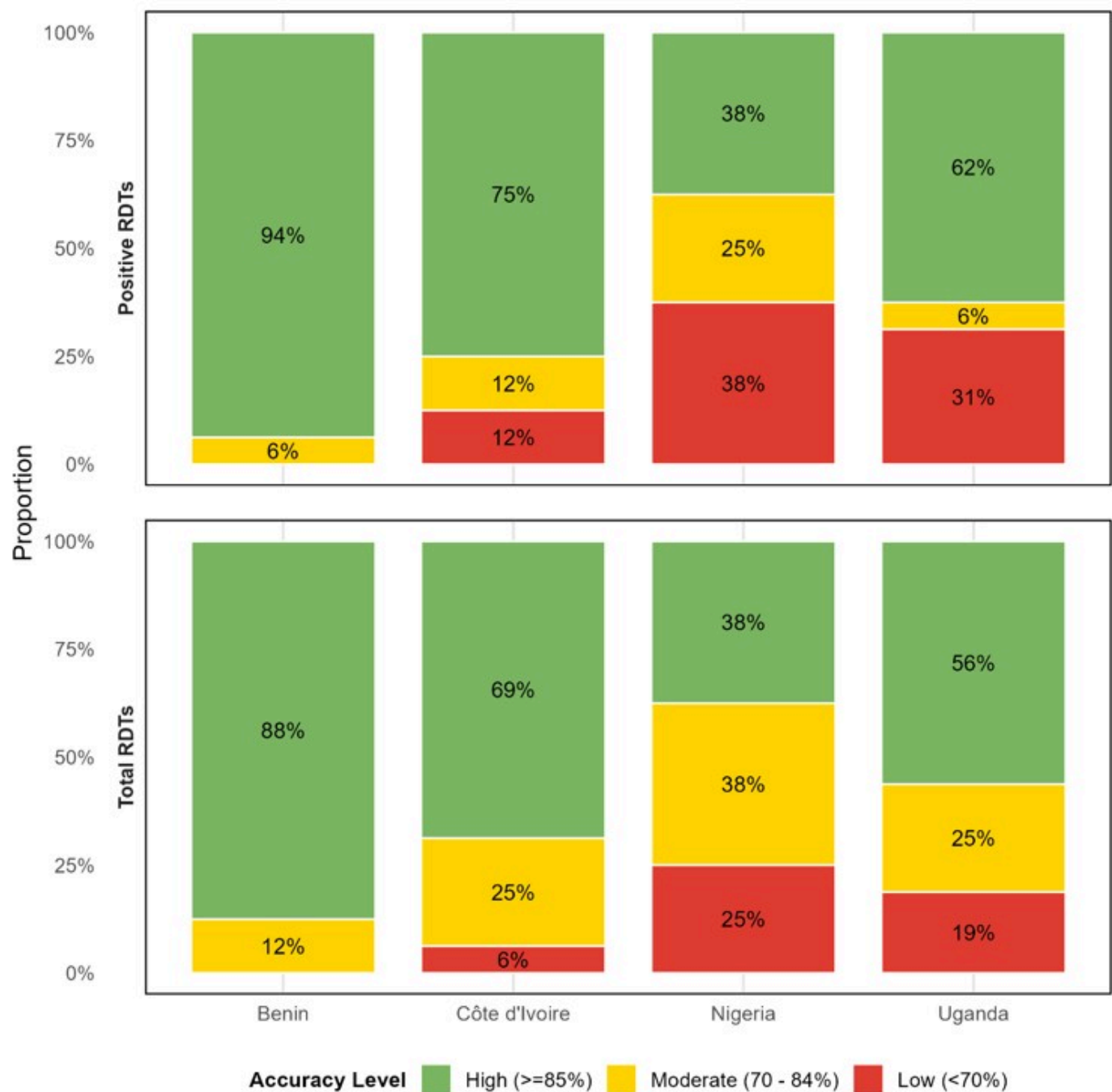


Figure 2

Proportion of health facilities with high, moderate or low levels of WAPE-based aggregate data reporting accuracy by indicator and country, 2023

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTables.docx](#)