Update of the response plan to *pfhrp2* gene deletions

Meeting report, 26 January 2023

World Health Organization

Update of the response plan to *pfhrp2* gene deletions

Meeting report, 26 January 2023

World Health Organization

Update of the response plan to *pfhrp2* gene deletions: meeting report, 26 January 2023

ISBN 978-92-4-009791-9 (electronic version)
ISBN 978-92-4-009792-6 (print version)

Suggested citation. Update of the response plan to *pfhrp2* gene deletions: meeting report, 26 January 2023. Geneva: World Health Organization; 2024. Licence: CC BY-NC-SA 3.0 IGO.

Cataloguing-in-Publication (CIP) data. CIP data are available at https://iris.who.int/.

Sales, rights and licensing. To purchase WHO publications, see https://www.who.int/publications/book-orders. To submit requests for commercial use and queries on rights and licensing, see https://www.who.int/copyright.

Contents

Abbreviations

CHAI	Clinton Health Access Initiative
CI	confidence interval
DHS	Demographic and Health Survey
HRP2	histidine-rich protein 2
LDH	lactate dehydrogenase
LOD	limit of detection
MOI	multiplicity of infection
PET	photo-induced electron transfer
RDT	rapid diagnostic test
uRDT	ultra-sensitive rapid diagnostic test
WHO	World Health Organization

1. Background

Rapid diagnostic tests (RDTs) are the most commonly used method for malaria diagnosis in endemic countries, especially in rural areas where microscopy is not available *(1)*. The most commonly used RDTs for the detection of *Plasmodium falciparum* detect histidine-rich protein 2 (HRP2), either alone or in combination with another antigen such as lactate dehydrogenase (LDH) or aldolase. LDH can either detect all *Plasmodium* species (pan-LDH) or be *P. falciparum*-specific (*Pf*-LDH) or *P. vivax*-specific (*Pv*-LDH). HRP2-detecting RDTs are more sensitive and heat-stable than non-HRP2 RDTs, and they are specific to *P. falciparum*. These RDTs have come to dominate the malaria RDT market.

However, *P. falciparum* parasites with deletions in the *pfhrp2* gene and paralogue *pfhrp3* gene have now been identified in all malaria-endemic regions *(2)*. These parasites do not produce HRP2 and/or HRP3 (a cross-reactive protein). Therefore, they are often undetectable by HRP2-based RDTs, resulting in false-negative RDT results. The World Health Organization (WHO) recommends that the diagnostic strategy be changed from RDTs that exclusively detect HRP2 when the prevalence of false-negative RDT results caused by *pfhrp2* gene deletions reaches or exceeds 5%. Due to the challenges of using microscopy for diagnosis, especially in rural areas, it is most common to switch to non-HRP2-based RDTs, mainly those that detect *Pf*-LDH, for diagnosis. When the prevalence of false-negative RDT results caused by *pfhrp2* gene deletions is over 5%, in many epidemiological settings, the benefit of switching RDTs will outweigh the lower sensitivity of the alternative *Pf*-LDH RDTs.

In 2016, WHO hosted a technical consultation related to the threat of *pfhrp2* gene deletions, which led to the development of an information note outlining when and how to investigate and report on *pfhrp2* deletions *(3)*. Subsequently, the WHO *Response plan to* pfhrp2 *deletions* was published in 2019 *(4)*.

The key objectives laid out in the response plan were to:

- define the frequency and distribution of diagnostically relevant mutations in circulating *P. falciparum* strains;
- provide concrete guidance to countries on malaria diagnosis and treatment in settings where such mutations are found to be frequent;
- identify gaps in knowledge on the emergence and spread of strains with *pfhrp2* and/or *pfhrp3* deletions and the actions required to develop new, accurate tests for malaria based on alternative antigens; and
- coordinate advocacy and communication with donors, policy-makers, test developers, research agencies, technical partners and disease control programmes to assist in planning.

In line with these objectives, the core activities that WHO has undertaken include:

- developing and updating the Malaria Threats Map;
- developing harmonized protocol templates to support surveillance;
- establishing an international network of laboratories to support molecular and serological analysis;
- providing support to countries relating to implementation of surveys/surveillance, when to switch diagnostics, what to switch to, and so on;

- advising commercial manufacturers on priorities and market forecasts; and
- engaging in advocacy and dissemination, including through Malaria Policy Advisory Group statements, videos, information to national malaria programmes and main funding agencies, such as the Global Fund to Fight AIDS, Tuberculosis and Malaria, and integration and expansion of *pfhrp2* gene deletion surveillance as part of malaria genomic surveillance.

Since the first technical consultation, many aspects of the epidemiology and understanding of *pfhrp2/3* gene deletions have evolved. There has been an increase in and spread of *pfhrp2/3* deletions; it is often a case of "where we look, we find". However, there is still a need to improve the quality and consistency of surveys on the prevalence of these deletions. Such surveys are usually decentralized, slow and resource-intensive, competing with other priorities. Studies that do not confirm deletions are sometimes not reported, as these findings are not seen as relevant and warranting dissemination. However, to understand the situation in a given area, it is necessary to report the lack of deletions as well as the presence of deletions. In addition, *pfhrp2/3* gene deletions often appear in very localized foci and their prevalence can be very heterogeneous throughout a country, further complicating decisions as to when and how to switch RDTs.

A very high prevalence of deletions has been reported in some countries in South America (e.g. Brazil, Peru) and Africa (e.g. Djibouti, Eritrea), triggering a change in RDT policy *(5)*. In light of the threat posed by these gene deletions and the post-treatment persistence of HRP2, which can cause false-positive results for clinical malaria, new RDTs have been developed since 2016 and are now commercially available, while others are in the development pipeline. The manufacturer Rapigen has developed three novel malaria RDTs with lines that can apparently detect low concentrations of *Pf*-LDH. These tests have shown promising results in the detection of pLDH antigens and have passed the independent laboratory evaluation for WHO prequalification. These tests are:

- BIOCREDIT Malaria Ag Pf (pLDH), which detects *Pf*-LDH antigens;
- BIOCREDIT Malaria Ag Pf (pLDH/HRP2), which detects *Pf*-LDH and HRP2 antigens on separate lines; and
- BIOCREDIT Malaria Ag Pf/Pv, which detects *Pf*-LDH and *Pv*-LDH antigens on separate lines.

In light of recent developments, the *Response plan to* pfhrp2 *deletions (4)* is to be updated and some fundamental aspects require review and expert consultation. These include:

- the validity of the threshold for switching to an alternative RDT;
- how to accelerate the response among programmes and manufacturers;
- lessons learned from settings where RDT policy has been changed;
- whether the decision to switch to an alternative RDT should be data-driven or pre-emptive; and
- what additional measures could be taken to accelerate the response.

A virtual meeting was held on 26 January 2023 to discuss these aspects for inclusion in the updated response plan. The agenda is provided in Annex 1 and the list of participants in Annex 2.

2. Presentations

2.1 Revisiting how the 5% threshold was determined

Presenter: Michelle Gatton

2.1.1 Knowledge of *pfhrp2/3* deletions available in 2016

At the time of the technical consultation in 2016, it was known that there were parasites with *pfhrp2* deletions in South America, in particular in Peru and other Amazonian regions of South America *(6)*. The first evidence of parasites with *pfhrp2* deletions impacting diagnosis of clinical malaria in Africa was found in Eritrea, where *pfhrp2*-deleted parasites dominated clinical cases in sampled locations. Mathematical modelling suggested that the continued use of HRP2-only RDTs in areas where *pfhrp2*-deleted parasites existed would rapidly select for this parasite subpopulation *(7)*. It was also understood that non-HRP2-detecting RDTs could yield positive results.

However, in 2016, most of the non-HRP2-detecting RDTs in use detected pan-LDH, which was not appropriate in areas where it was important to discriminate between *P. falciparum* and *P. vivax* malaria. Three RDTs with a separate *Pf*-LDH test band from two manufacturers had just progressed through WHO RDT product testing (round 6) *(8)*, but were not yet WHO-prequalified. The overall RDT positivity in round 6 product testing showed sensitivities of:

- 90.1% and 74.7% for HRP2 and *Pf*-LDH test bands, respectively, at 200 parasites/µL; and
- 99.8% and 98.3% for HRP2 and *Pf*-LDH test bands, respectively, at 2000 parasites/µL.

However, there were still many unknowns in 2016. It was not yet understood whether *pfhrp2*-deleted parasites in Eritrea were genetically related to those in Peru (i.e. whether the finding represented the spread of an existing population or a new deletion event), or whether these parasites were impacting clinical diagnosis elsewhere in Africa. The threshold parasite density at which patients seek diagnosis was not well understood. Another important consideration at the time was whether manufacturers would be able to quickly scale up production of LDH-detecting RDTs.

It was decided that the most immediate need was to focus on clinical malaria and that it was important to adopt a pragmatic approach that could balance the need to investigate and take action to maintain effective diagnosis with the manufacturers' ability to supply the market. It was considered critical to avoid a situation in which countries started to distrust RDT results and disregard all HRP2-based RDT results. However, it was acknowledged that *Pf*-LDH-detecting RDTs did not perform as well as HRP2-detecting RDTs at lower densities.

With these considerations in mind, the following recommendation was presented to the Malaria Policy Advisory Committee (now the Malaria Policy Advisory Group) in 2016: A nationwide change to non-HRP2-based RDTs is recommended when the prevalence of patients carrying *pfhrp2*-deleted parasites causing false-negative RDT results meets or exceeds the lower 95% confidence interval (CI) for 5% prevalence. If the prevalence is less than 5%, the recommendation is to plan for change over a longer time frame, as it is anticipated that *pfhrp2/3*-deleted parasites will persist and spread. Acquiring additional surveillance data would help to prioritize the roll-out of non-HRP2-based RDTs.

2.1.2 The selection of 5% as the threshold

The threshold of 5% was selected because it is around this point that the public health impact and proportion of cases missed by less sensitive non-HRP2-based tests is likely to be lower than that associated with continued use of HRP2-based tests. In other words, it was considered to be the ideal pivot point when the false-negative rate of RDTs caused by *pfhrp2*-deleted parasites in patients equals the false-negative rate of RDTs due to the differential performance of *Pf*-LDH- and HRP2-based RDTs. There was no solid evidence to determine this pivot point for all scenarios, but the following assumptions were made: (i) most (75%) patients have higher parasite densities (~2000/µL) and a smaller proportion (25%) have lower densities (~200/µL); and (ii) HRP2 tests detect 1.5% more high-density infections and 15% more low-density infections. The average nets out at approximately 5%.

2.2 Summary of comparisons between HRP2-based and *Pf*-LDH-based RDTs from published literature: exploring the validity of the 5% threshold

Presenter: Rebecca Thomson

2.2.1 Methodology

A literature search was conducted to compare the sensitivity and specificity of HRP2-based RDTs and *Pf*-LDH-based RDTs in a range of malaria-endemic settings. The search terms "LDH", "HRP2" and "RDT performance" were used, and additional articles were identified by reviewing the citations in articles. Articles published from 2011 until September 2022 were included. Papers comparing HRP2 and pan-LDH RDTs only were excluded, along with those not directly comparing RDT results from the same set of samples.

2.2.2 Results

In total, 24 publications were identified that compared HRP2-based and *Pf*-LDH-based RDTs. Some of these publications compared more than two brands of RDTs or compared RDT results from more than one batch of samples. Consequently, these publications were broken down into 36 different comparisons. The majority of the papers directly compared two or more brands of RDTs or compared the test lines from combination RDTs that detected both HRP2 and *Pf*-LDH. Three publications were systematic reviews involving meta-analysis and presenting data for several RDTs. One review presented data separately by brand of RDT and therefore this information was included as separate studies; two reviews presented only amalgamated data. Data on study type, country, number of samples, sensitivity and specificity of each RDT along with information on the RDTs used (brand, code, antigens detected, etc.) were extracted.

The studies from which the samples were taken came from 21 different countries, including seven in East Africa or the Horn of Africa, six in West Africa, three in Asia and one in South America; the remaining four countries were not endemic for malaria and samples were taken from returning travellers. Most samples came from studies designed for different malaria research questions, for example malaria during pregnancy or antimalarial efficacy studies. About two thirds of the studies used microscopy as the reference standard, whereas the remaining studies used PCR, microscopy corrected by PCR, or both PCR and microscopy with results presented separately for each method. The patients enrolled in the studies included different cohorts, with some studies enrolling only children under 5, children under 15 or pregnant women. The number of *P. falciparum* samples in the studies ranged from nine to 655.

Among the studies that included an HRP2-based RDT and an RDT relying mainly on *Pf*-LDH, the average sensitivity of the HRP2-detecting RDT was 4.3% higher than that of the *Pf*-LDH-detecting RDT, whereas the specificity was 8.8% lower, with microscopy as the reference standard (Table 1). With PCR as the reference standard, HRP2-detecting RDTs had on average 12.9% higher sensitivity and 1.3% lower specificity than *Pf*-LDH-detecting RDTs. However, one study of 21 samples from Peru showed that sensitivity was 27.1% lower when using the HRP2-detecting RDTs than when using the *Pf*-LDH-detecting RDTs (71.5% vs 98.7%). This difference was mainly due to the very high prevalence of *pfhrp2* gene deletions (26%) in the setting. When this study was excluded from the analysis, HRP2-detecting RDTs or strips were found to be 5.9% more sensitive than *Pf*-LDH-detecting tests, with microscopy as the reference standard. The range of the difference spanned from 27% lower to 43% higher sensitivity of HRP2-detecting RDTs compared to *Pf*-LDH-detecting RDTs. However, the study in which the *Pf*-LDH RDTs were 27% more sensitive was the aforementioned study with the very high prevalence of *pfhrp2* gene deletions.

Table 1. Sensitivity and specificity of HRP2-detecting RDTs relative to Pf-LDH-detecting RDTs

Microscopy as the reference standard		PCR as the reference standard		Overall	
Sensitivity %	Specificity %	Sensitivity %	Specificity %	Sensitivity %	Specificity %
4.3	-8.8	12.9	1.3	8.0	-5.3
Excluding one study in Peru with a very high prevalence (26%) of *pfhrp2* gene deletions					
5.9		-		9.0	

Fourteen studies focused on the sensitivity of the HRP2 line only and the sensitivity of the *Pf*-LDH line only, in contrast to some of the other studies that included additional test lines (e.g. often an HRP2-based RDT includes a pan-LDH line). The sensitivity of the HRP2-based RDT or strip was 5.6% higher than that of the *Pf*-LDH-based strips, whereas the specificity was 10.0% lower, with microscopy as the reference standard. With PCR as the reference standard, HRP2-based RDTs or strips had a 13.3% higher sensitivity and 1.1% lower specificity than *Pf*-LDH-based RDTs. When comparing RDTs regardless of the reference standard used, the HRP2 strips were 9.3% more sensitive than the *Pf*-LDH strips.

The average sensitivities of HRP2-detecting and *Pf*-LDH-detecting RDTs were both quite high compared to microscopy: 94% and 88%, respectively. However, the sensitivities of both types of RDT were much lower compared to PCR: 72.9% for HRP2-detecting RDTs and 60.0% for *Pf*-LDH-detecting RDTs. PCR has greater sensitivity and a lower limit of detection (LOD) than microscopy; therefore, PCR can detect samples that are not detected by microscopy or either type of RDT. The difference in sensitivity between HRP2-detecting and *Pf*-LDH-detecting RDTs was generally greater when using PCR as the reference standard due to the increased sensitivity of the PCR method. The smaller difference in sensitivity with microscopy as the reference also resulted in a larger drop in the specificity of HRP2-based tests compared to *Pf*-LDH-based tests with PCR as the reference.

One study compared samples from asymptomatic and symptomatic individuals and compared the results to PCR, using the same RDTs for each cohort. The average difference in sensitivity was 14.5% among asymptomatic people, but only 3.5% among symptomatic people.

2.2.3 Main points

- Overall, HRP2-based RDTs showed better sensitivity than *Pf*-LDH-detecting RDTs, but with slightly lower specificity. The specificity of HRP2 test lines was generally lower than that of *Pf*-LDH lines, in part due to the persistence of HRP2 antigens following a recent *P. falciparum* infection.

- The difference in sensitivity between HRP2-detecting and *Pf*-LDH-detecting RDTs was generally greater when using PCR as the reference standard than when using microscopy, due to the higher sensitivity of PCR. The reference standard used when comparing data must be considered in making policy decisions. While it may be optimal for analytical purposes to compare the accuracy of RDTs using PCR, using microscopy as the reference standard better reflects the difference between the two types of RDTs in clinical performance, or for detecting the majority of clinically relevant infections.

- The difference in sensitivity was greater when using samples from subclinical or asymptomatic patients than when using samples from symptomatic patients. However, the results from symptomatic patients better represent a clinical setting, where the majority of RDTs are used.

2.3 Heterogeneity in the distribution of *pfhrp2/3* deletions in Ethiopia and current alternatives to RDTs that exclusively detect HRP2

Presenter: Fitsum Tadesse

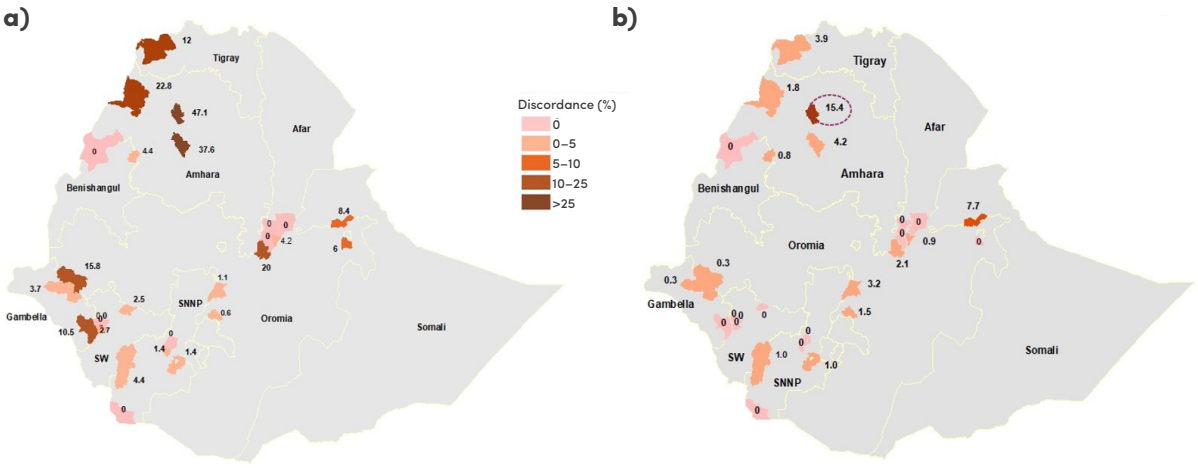
Although the majority of Ethiopia has very low to no malaria transmission, the western side of the country has high malaria transmission. *P. vivax* is endemic in Ethiopia, with 40% of malaria cases due to this species. Therefore, there is a critical need for diagnostics that can differentiate between *P. falciparum* and *P. vivax*. The African Center for *hrp2/3* Deletion Surveillance (ACHIDES) was set up in February 2021 to monitor *pfhrp2/3* deletions, with the two main objectives discussed below.

2.3.1 Objective 1: examining the prevalence of parasites with *pfhrp2/3 gene* deletions

A nationwide survey was conducted in 22 districts, enrolling 370 patients per district as per the WHO protocol. Microscopy was conducted on suspected malaria patients, and if positive, two RDTs were performed: SD Bioline Malaria Ag Pf/Pv (HRP2/*Pv*-LDH) and Rapigen BIOCREDIT Ag Pf (HRP2/*Pf*-LDH).

A high level of heterogeneity was detected in the levels of discordance between microscopy and the SD Bioline RDT in small geographical areas (Fig. 1). Some areas were 100% concordant, whereas areas in the north-west near the borders with Sudan and Eritrea had levels of discordance up to 47.1%. The Rapigen RDT was able to detect the vast majority of infections missed by the SD Bioline RDT, and this difference was seen across the country. In the district where the discrepancy between the microscopy and SD Bioline RDT results was 47.1%, the discrepancy between microscopy and the Rapigen RDT was only 15.4%. While this was an improvement on the discordance seen with the SD Bioline RDT, the 15.4% discordance with the Rapigen *P. falciparum*-only RDT may have been partly due to microscopists misclassifying *P. vivax* infections as *P. falciparum* infections (i.e. leading to microscopy *P. falciparum* positive and Rapigen *P. falciparum* RDT negative results); there is a high prevalence of *P. vivax* in this area of Ethiopia and species-specific identification by microscopy is imperfect.

Fig. 1. Discordance between microscopy and (a) existing/routine SD Bioline Malaria Ag Pf/Pv (HRP2/*Pv*-LDH) RDT and (b) Rapigen BIOCREDIT Ag Pf (HRP2/*Pf*-LDH) RDT



Source: Fitsum Tadesse, Armauer Hansen Research Institute, unpublished data, 26 January 2023.

2.3.2 Objective 2: examining the performance of candidate RDTs in Ethiopia in malaria-suspected febrile patients

Three hundred suspected malaria patients from six districts were tested with microscopy, the two RDTs mentioned above, and a third Rapigen RDT that detected *Pf*-LDH and *Pv*-LDH in order to detect both *P. vivax* and *P. falciparum* infection, along with digital PCR and MAGPIX antigen quantification. Table 2 shows the positivity rate for *P. falciparum* and *P. vivax* infection by microscopy in each district.

Table 2. Positivity rate for *P. falciparum* and *P. vivax* infection by microscopy by district in Ethiopia

District	Microscopy, % (n/N)	
	P. falciparum	*P. vivax*
Gonder	48.6 (151/311)	17.7 (55/311)
Metema	42.9 (133/310)	12.3 (38/310)
Gambela	37.5 (112/299)	2.0 (6/299)
Fentale	33.3 (101/303)	5.3 (16/303)
Arba Minch	30.6 (89/291)	19.9 (58/291)
Pawe	32.8 (103/314)	4.5 (14/314)
Total	**37.7 (689/1828)**	**10.2 (187/1828)**

The results showed that HRP2-based RDTs missed a significant proportion of infections. A high level of discordance was seen in the north-west of the country, in line with the results from the survey described in section 2.3.1 (Fig. 2). The Rapigen HRP2/*Pf*-LDH RDT detected the vast majority of *P. falciparum* infections, while the Rapigen Pf/Pv (*Pf*-LDH/*Pv*-LDH) RDT was also a good alternative test, as it was able to detect the majority of infections missed by the HRP2-only RDT.

Fig. 2. Discordance between microscopy and (a) existing/routine SD Bioline Malaria Ag Pf/Pv (HRP2/*Pv*-LDH) RDT, (b) Rapigen BIOCREDIT Ag Pf (HRP2/*Pf*-LDH) RDT and (c) Rapigen BIOCREDIT Ag Pf/Pv (*Pf*-LDH/*Pv*-LDH) RDT

Source: Fitsum Tadesse, Armauer Hansen Research Institute, unpublished data, 26 January 2023.

Among all PCR-positive *P. falciparum* samples, 17.7% (109/615) had *pfhrp2* gene deletions. Among the discordant samples, 60.4% (64/106) had *pfhrp2* gene deletions, while among the concordant samples, 4.5% (16/359) were shown to have *pfhrp2* gene deletions by PCR. Nearly 50% (304/615) of the PCR-positive samples had *pfhrp3* gene deletions. Table 3 shows the prevalence of gene deletions by district.

Table 3. Prevalence of pfhrp2/3 gene deletions by district

District	*pfhrp2* deletion, % (n/N)	*pfhrp3* deletion, % (n/N)
Gonder	26.2 (43/164)	61.6 (101/164)
Metema	16.1 (20/124)	62.9 (78/124)
Gambela	8.2 (7/85)	4.7 (4/85)
Fentale	23.3 (29/126)	75.4 (95/126)
Arba Minch	10.1 (8/79)	48.1 (38/79)
Pawe	8.6 (6/70)	12.9 (9/70)
Total	**17.7 (109/615)**	**49.4 (304/615)**

Table 4 shows the sensitivity and specificity of the three RDTs based on *pfhrp2* deletion status. The Rapigen BIOCREDIT Ag Pf (HRP2/*Pf*-LDH) RDT showed the best sensitivity to all strains of *P. falciparum*, regardless of *pfhrp2* deletion status; however, the specificity was compromised.

Table 4. Sensitivity and specificity of three RDTs based on molecular-confirmed presence of *pfhrp2/3* gene deletions

Category	Value	Excluding *P. vivax* species			Including *P. vivax* species	
		SD Bioline Pf/Pv	BIOCREDIT Pf/Pv	BIOCREDIT Pf/Pf	SD Bioline Pf/Pv	BIOCREDIT Pf/Pv
Any *Pf*-detected infections	Sensitivity	77.1 (75.1–79.2)	93.9 (92.7–95.0)	97.1 (96.3–97.9)	80.1 (78.3–82.0)	94.9 (93.9–95.9)
	Specificity	97.2 (96.4–98.0)	96.1 (95.1–97.0)	93.1 (91.9–94.3)	96.5 (95.7–97.4)	95.0 (94.0–96.0)
pfhrp2-positive infections	Sensitivity	89.6 (87.2–92.1)	97.4 (96.1–98.7)	99.4 (98.7–100.0)	88.2 (86.3–91.4)	97.1 (95.7–98.4)
	Specificity	83.5 (80.5–86.5)	85.1 (82.2–88.0)	77.4 (74.0–80.8)	85.0 (82.1–87.9)	85.0 (82.1–87.9)
pfhrp3-positive infections	Sensitivity	85.3 (81.7–88.9)	96.9 (95.2–98.7)	99.3 (98.5–100.2)	84.8 (81.1–88.4)	96.0 (94.1–98.0)
	Specificity	82.7 (78.8–86.5)	87.8 (84.5–91.2)	77.3 (73.1–81.6)	84.9 (81.3–88.6)	86.3 (82.8–89.8)
pfhrp2/3-positive infections	Sensitivity	92.0 (89.1–94.9)	96.6 (94.6–98.6)	99.4 (98.3–100.2)	91.1 (88.0–94.1)	95.9 (93.8–98.0)
	Specificity	82.6 (78.5–86.7)	88.2 (84.8–91.7)	78.3 (73.8–82.7)	83.8 (79.9–87.8)	86.8 (83.2–90.4)

2.3.3 Main findings

- The SD Bioline Malaria Ag Pf/Pv RDT performed best when it was used to detect parasites with no *pfhrp2/3* gene deletions compared to when it was used to detect infections with deletions of either one of these genes (92.0% compared to 89.6% and 85.3%, respectively).
- The Rapigen BIOCREDIT Ag Pf RDT showed the best performance, followed by the Rapigen BIOCREDIT Ag Pf/Pv RDT, in detecting both *pfhrp2/3*-deleted and non-deleted parasite populations, compared to the SD Bioline Malaria Ag Pf/Pv RDT. These results suggest that in clinical practice, there is no trade-off in sensitivity between HRP2 and *Pf*-LDH.
- The Rapigen BIOCREDIT Ag Pf/Pv RDT, which relies exclusively on *Pf*-LDH for detection of *P. falciparum*, showed a higher level of sensitivity and comparable specificity compared to the SD Bioline Malaria Ag Pf/Pv RDT among populations of parasites with both *P. falciparum* and *P. vivax* species.

2.4 Results from a *pfhrp2/3* gene deletion survey conducted in the United Republic of Tanzania

Presenter: Deus Ishengoma

2.4.1 Survey methodology

In 2021, a survey was conducted, as per the WHO protocol, to evaluate the prevalence of false-negative RDT results caused by *pfhrp2* gene deletions in the United Republic of Tanzania. The survey included 100 health facilities from 10 regions – from areas with high (two regions), moderate (two regions), low (three regions) and very low (three regions) malaria transmission. As per the WHO protocol, 37 febrile patients were enrolled from each health facility. Patients were tested with two RDTs: one of the RDTs

used by the national malaria control programme and a *Pf*-LDH-detecting RDT as a comparator. The national malaria control programme uses three RDTs, all detecting HRP2 and pan-LDH: First Response, SD Bioline and CareStart. The comparator RDT was the Rapigen BIOCREDIT Malaria Ag Pf RDT detecting *Pf*-LDH.

2.4.2 Survey results

A laboratory-based multiplex antigen assay was conducted on all selected dried blood spot samples, and all of those that were discordant or negative for *pfhrp2/3* in the antigen assay were tested with photo-induced electron transfer (PET)-PCR for the presence of *P. falciparum*. Of these samples, all that were found to be positive by PET-PCR underwent *pfhrp2/3* genotyping; they were confirmed to have deletions if both *msp1/2* genes were amplified in the *pfhrp2/3*-deleted samples.

Among all the national malaria control programme RDTs (n = 7863), there was a 46.7% positivity rate using the HRP2 line, a 33.6% positivity rate using the pan-LDH line and a 47.4% positivity rate when using both test lines to obtain the RDT result. The Rapigen BIOCREDIT *Pf*-LDH RDT showed a 40.2% positivity rate. There were 88 (2.3%) discordant samples found, with no clustering in any region.

Table 5 shows the performance of the HRP2-detecting RDTs used routinely by the programme and the Rapigen BIOCREDIT *Pf*-LDH-detecting RDT, using quantitative PCR as the reference standard. The Rapigen BIOCREDIT RDT showed higher sensitivity and positive predictive value, but slightly lower specificity than the HRP2-detecting RDTs. These results again suggest that there is no trade-off between HRP2 and *Pf*-LDH in terms of sensitivity.

Table 5. Performance of HRP2-detecting and Pf-LDH-detecting RDTs

Value	Programme HRP2-detecting RDTs	Rapigen BIOCREDIT *Pf*-LDH-detecting RDTs
Sensitivity	89.5%	96.7%
Specificity	78.3%	72.2%
Negative predictive value	80.8%	76.8%
Positive predictive value	88.3%	95.9%
Concordance with quantitative PCR	84.1%	84.4%
Concordance with pLDH RDT	90.3%	N/A

2.5 RDTs addressing *pfhrp2/3* gene deletions

Presenter: Gonzalo Domingo

2.5.1 Study overview

The performance of two of the new Rapigen tests was compared with that of the current "best-in-class" WHO-prequalified *P. falciparum* RDTs (targeting HRP2 and/or *Pf*-LDH), using the WHO international standards for *P. falciparum* and *P. vivax* antigens (National Institute for Biological Standards and Control (NIBSC) code 16/376 and NIBSC code 19/116, respectively). The two Rapigen tests evaluated were Rapigen BIOCREDIT Ag Pf (HRP2/*Pf*-LDH) and Rapigen BIOCREDIT Ag Pf/Pv (*Pf*-LDH/*Pv*-LDH).

The Rapigen Pf and Rapigen Pf/Pv RDTs showed about a 10-fold lower LOD on the *Pf*-LDH lines (3.91 units/mL and 1.95 units/mL, respectively), compared to the WHO-prequalified *P. falciparum* RDT (31.25 units/mL) (Fig. 3) in terms of analytical sensitivity. The HRP2 line on the Rapigen Pf test showed a slightly lower LOD than the WHO-prequalified test, but the difference was not as great as with the *Pf*-LDH lines.

Fig. 3. Performance of two Rapigen BIOCREDIT RDTs and two WHO-prequalified RDTs, using NIBSC code 16/376 antigens

NIBSC Pf Ag 16/376				RapiGEN Pf		RapiGEN Pf/Pv		WHO PQ Pf/Pv RDT		WHO PQ Pf RDT	
		Quansys measured, pg/mL		HRP2	PfLDH	PfLDH	PvLDH	HRP2	PvLDH	HRP2	PfLDH
Units/mL	Dil. Factor	HRP2	PfLDH	Fraction positive: number positive/number run							
62.50	32	4,353	11,661	5/5	5/5	5/5	0/5	5/5	0/5	40/40	40/40
31.25	64	1,873	4,760	5/5	5/5	5/5	0/5	5/5	0/5	39/40	29/40
15.63	128	1,001	2,702	5/5	5/5	5/5	0/5	40/40	0/40	36/40	0/40
7.81	256	566	1,610	38/40	38/40	5/5	0/5	31/40	0/40	0/40	0/40
3.91	512	295	902	40/40	40/40	40/40	0/40	0/40	0/40	0/5	0/5
1.95	1,024	146	433	0/40	0/40	28/40	0/40	0/5	0/5	0/5	0/5
0.98	2,048	76	205	0/5	0/5	0/40	0/40			0/5	0/5
0.49	4,096	35	94			0/5	0/5				

Test line
← HRP2
← PfLDH
← PvLDH

Source: Gonzalo Domingo, PATH, unpublished data, 26 January 2023.

The difference in LOD between the Rapigen and WHO-prequalified RDTs remained for those tests with a *Pf*-LDH line (Fig. 4) when comparing the RDTs using a *pfhrp2/3*-deleted strain (3BD5, cultured at PATH).

Fig. 4. Performance of two Rapigen BIOCREDIT RDTs and two WHO-prequalified RDTs, using 3BD5 cell culture (*pfhrp2* negative and *pfhrp3* negative)

3BD5 Culture Lot M:20200212AAR		RapiGEN Pf		RapiGEN Pf/Pv		WHO PQ Pf/Pv RDT		WHO PQ Pf RDT	
Quansys measured pg/mL		HRP2	PfLDH	PfLDH	PvLDH	HRP2	PvLDH	HRP2	PfLDH
HRP2	Pf-LDH	Fraction positive: number positive/number run (highlighted indicates majority of replicates were positive)							
1	20,518	0/5	5/5	5/5	0/5	0/5	0/5	0/40	40/40
1	6,026	0/5	5/5	5/5	0/5	0/5	0/5	0/40	36/40
1	3,143	0/5	5/5	5/5	0/5	0/5	0/5	0/40	0/40
< 0.68	1,389	0/40	38/40	5/5	0/5	0/5	0/5	0/5	0/5
< 0.68	750	0/40	29/40	40/40	0/40	0/5	0/5	0/5	0/5
< 0.68	406	0/40	0/40	40/40	0/40				
< 0.68	244			0/40	0/40				
< 0.68	142			0/5	0/5				

Test line
← HRP2
← PfLDH
← PvLDH

Source: Gonzalo Domingo, PATH, unpublished data, 26 January 2023.

The Rapigen Pf/Pv RDT also had a lower LOD on the *Pv*-LDH test line when tested against *P. vivax* antigens (Fig. 5).

Fig. 5. Performance of Rapigen *Pf/Pv* RDT and WHO-prequalified RDT on *P. vivax* antigens

NIBSC Pv antigen NIBSC# 19/116		RapiGEN Pf/Pv	WHO PQ Pf/Pv RDT
Units	Q-Plex measured pg/mL	PvLDH test line	PvLDH test line
IU/mL	PvLDH	Fraction positive	
4,000	> ULQ	5/5	5/5
400	99,553	5/5	5/5
200	45,480	5/5	5/5
100	23,196	5/5	5/5
50	13,506	5/5	5/5
25	4,122	5/5	3/5
15	2,534	5/5	0/5
10	1,807	5/5	0/5
5	927	5/5	0/5
2.5	484	3/5	0/5
1	191	2/5	

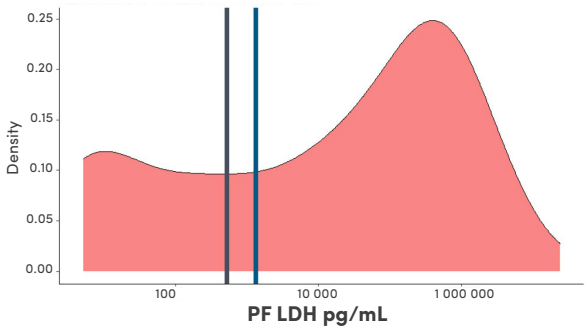
Source: Gonzalo Domingo, PATH, unpublished data, 26 January 2023.

2.5.2 Performance of Rapigen RDTs using clinical samples

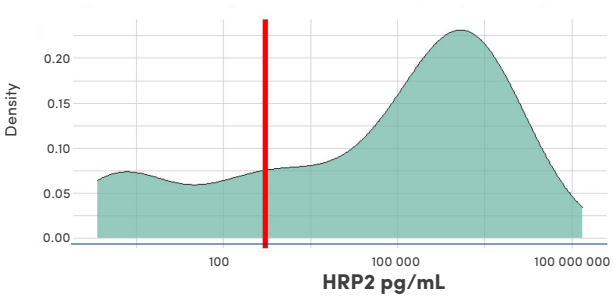
The Rapigen tests were then tested against clinical samples from a study led by Makhtar Niang, Institut Pasteur de Dakar, Senegal. In this study, a total of 220 febrile patients were recruited – 154 who were positive for malaria by PCR and 66 who were negative. Of these, the antigen concentration was obtained for 200 samples; 133 were positive for antigens (two non-*P. falciparum*) and 67 were antigen-negative. One sample was suspected of having *pfhrp2/3* gene deletions. Fig. 6 shows the antigen concentration distribution in patients with malaria.

Fig. 6. *Pf*-LDH and HRP2 antigen concentration distribution in patients presenting with malaria symptoms

Density plot of PCR+ Senegal Pf LDH results
Hypothetical LODs of the LDH lines in two Rapigen RDTs (HRP2/Pf-LDH RDT and Pf-LDH/ Pv-LDH RDT)



Density plot of PCR+ Senegal HRP2 results
Hypothetical LOD of the HRP2 line in the HRP2/Pf-LDH RDT



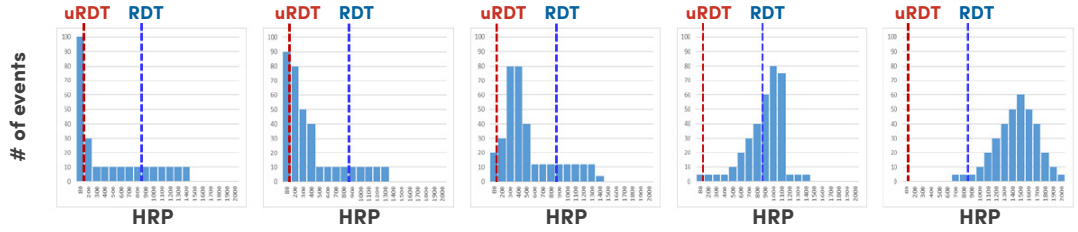
Note: This is a logarithmic scale, so the proportion of samples lying above the LOD is significantly higher than the proportion below the LOD for the RDT, and the scale of the x-axis for the HRP2 antigen is two orders of magnitude greater than for the LDH antigen. Therefore, the average concentration of HRP2 in the samples was significantly higher than that of the *Pf*-LDH antigen.

Source: Gonzalo Domingo, PATH, unpublished data, 26 January 2023.

The analysis of results for the Rapigen RDTs used at the point of care is ongoing. The following results were tested against venous blood for which the antigen concentration had been determined. Then, based on previous data on the LODs for the Rapigen tests, the performance was estimated. The sensitivity of the programme RDT used in Senegal (SD Bioline Ag Pf (05FK50)) showed 71.4% sensitivity, compared to PCR using capillary blood, while microscopy showed a lower performance of 53.4%. The two Rapigen tests also showed higher sensitivity than microscopy, with the RDT with HRP2 and *Pf*-LDH lines at 77.7% and the *Pf*-LDH-only test (which included a *Pv*-LDH line) at 71.2% sensitivity. These early evaluation data from Senegal support the evidence that in settings with no *pfhrp2/3* deletions, the HRP2 line confers additional clinical sensitivity over the LDH line, with enhanced analytical sensitivity.

Fig. 7 depicts a simplistic model showing the relationship between the distribution in antigen concentration and analytical sensitivity (lower LOD for the antigen) on clinical sensitivity and prevalence estimates for two RDTs, including an ultra-sensitive RDT (uRDT) with improved analytical sensitivity over the other RDT. While all settings have a 5% prevalence of *P. falciparum* in the model, the underlying distribution of parasite density and consequently antigen (in this case HRP2) concentration in each setting varies, leading to varied performance of the RDTs in each setting. The performance and malaria prevalence estimates of the uRDT with the improved analytical sensitivity are more resilient to underlying changes in parasite density and fluctuations in antigen concentration, compared to the RDT with the poorer analytical sensitivity.

Fig. 7. Performance of ultra-sensitive and normal RDTs at 5% *P. falciparum* prevalence at varying parasite densities (by PCR) represented by the distribution of HRP2 concentration



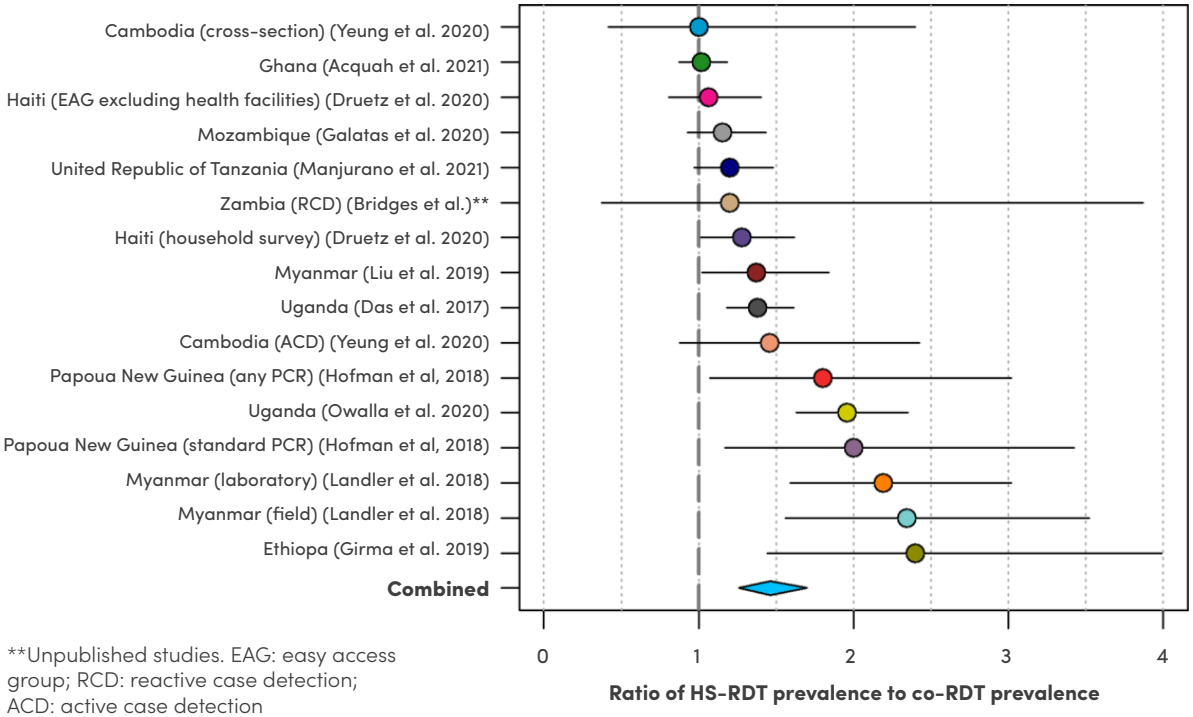
Prevalence	5%	5%	5%	5%	5%
% detected by uRDT	43	74	94	99	100
uRDT prevalence	2.1%	3.7%	4.7%	4.9%	5.0%
% detected by RDT	17	14	18	66	97
RDT prevalence	0.9%	0.7%	0.9%	3.3%	4.9%

Note: The red line indicates the analytical sensitivity of the uRDT and the blue line that of the standard RDT. The table indicates the percentage of PCR positives detected by the RDT under each scenario and the resulting *P. falciparum* prevalence estimates.

Source: Gonzalo Domingo, PATH, unpublished data, 26 January 2023.

The results from the simplistic model have been replicated in clinical settings where uRDT performance has been compared to that of standard RDTs. There is consistent improvement in analytical sensitivity to HRP2 with the highly sensitive RDT (NxTek™ Eliminate Malaria Pf), but the level of improvement varies based on the underlying parasite density distribution (Fig. 8) *(9)*. The weighted mean estimated ratio was 1.46 (95% CI: 1.26–1.70).

Fig. 8. Ratio of highly sensitive RDT prevalence to conventional RDT prevalence in different studies and overall



**Unpublished studies. EAG: easy access group; RCD: reactive case detection; ACD: active case detection

Source: Figure taken from *(9)*

2.5.3 Summary of comparison

In sum, new tests with more sensitive *Pf*-LDH and *Pv*-LDH lines are becoming available, which should lead to improvement in clinical performance. The extent of this improvement will depend on parasite density distributions, but the new tests will be more resilient to variation in parasite density distributions. While the performance of RDTs detecting *Pf*-LDH is particularly relevant in areas with high *pfhrp2/3* gene deletions, these data suggest that the detection of *P. falciparum* HRP2 is still important to maintain clinical sensitivity, especially in settings where there are no *pfhrp2/3* deletions. For areas with *P. falciparum* and *P. vivax*, tests with a combination HRP2/*Pf*-LDH line (i.e. both antigens on the same line) and a *Pv*-LDH line will be useful.

The Rapigen tests presented above are already approved by the Expert Review Panel for Diagnostics, while new tests with HRP2/LDH combination lines are in development and anticipated in 2023–2024.

2.6 Modelling the impact of *pfhrp2* deletions on the performance of current and next-generation RDTs: implications for optimal time to switch strategy

Presenter: Hannah Slater

2.6.1 Modelling the sensitivity of RDTs at different prevalence levels of *pfhrp2* gene deletions

Question: Assuming that the HRP2 and *Pf*-LDH concentrations and RDT LODs in a given population are known, is it possible to simulate different levels of *pfhrp2* deletion to estimate the deletion threshold at which an RDT containing a *Pf*-LDH line is needed?

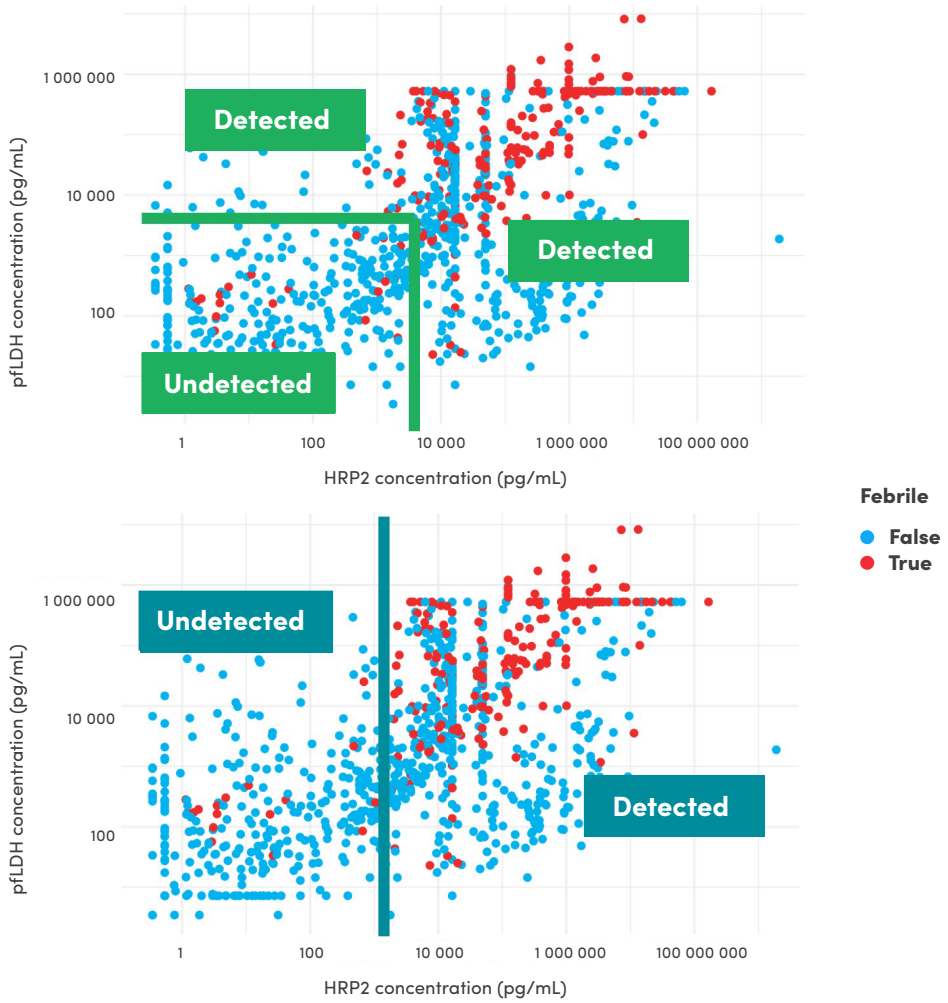
To estimate this, the following pieces of information are needed:

- the HRP2 and *Pf*-LDH concentrations in clinical (and asymptomatic) populations;
- the LODs of current and "next-generation" RDTs (i.e. the Rapigen BIOCREDIT RDTs); and
- the multiplicity of infection (MOI) and thus the probability that a *pfhrp2*-deleted infection is a monoinfection and will be missed by an HRP2-only RDT.

The MOI needs to be considered when estimating the number of infections that will be missed due to *pfhrp2* gene deletions, because a polyclonal infection with a negative and positive *pfhrp2* strain will usually result in a positive RDT result, masking the *pfhrp2* deletion. There is some evidence that the MOI is higher in areas with higher malaria transmission.

PATH has assembled a database of HRP2 and *Pf*-LDH concentrations from clinical and asymptomatic populations. It has also benchmarked several best-in-class and next-generation RDTs according to analytical estimates of the concentration at which the RDTs have a 90% probability of sensitivity, or the LOD for each test. Fig. 9 (a) shows the LOD of an HRP2-only RDT, indicating that any samples lying to the left of the vertical line would not be detected by the RDT and any samples to the right would be detected. A test with an HRP2 and a *Pf*-LDH line (or a dual line) would have an LOD for each antigen and would result in an LOD curve, as shown in Fig. 9 (b).

Fig. 9. *P. falciparum* infections that would be detected using (a) an HRP2-only RDT and (b) an RDT containing HRP2 and *Pf*-LDH lines



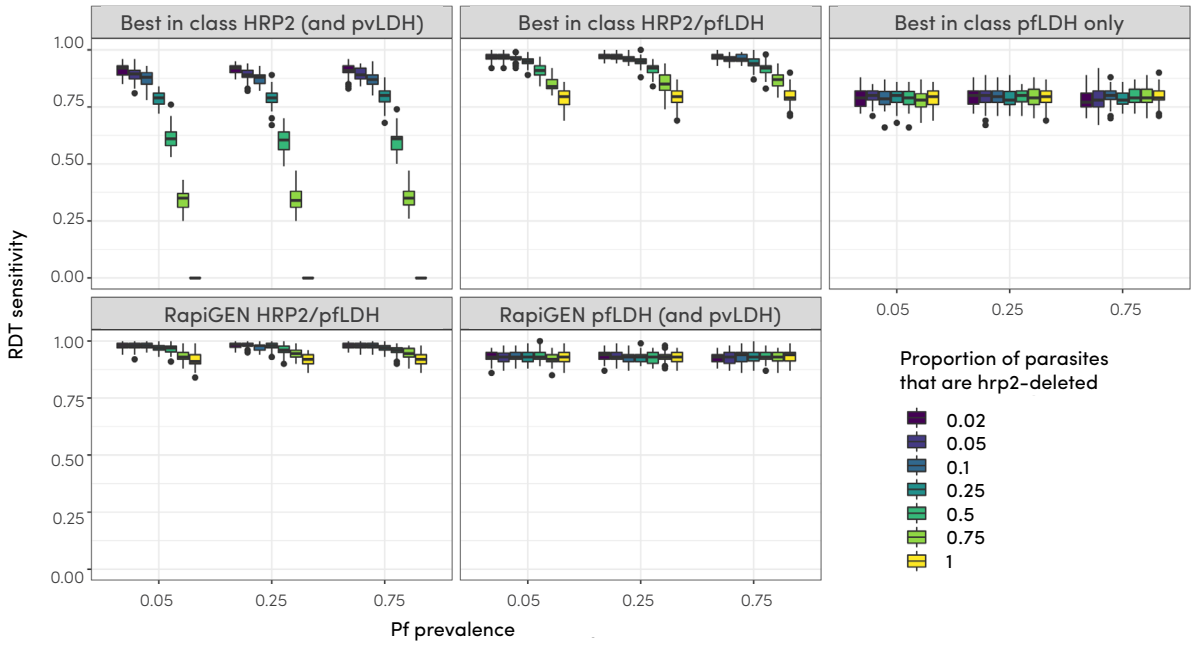
Source: Hanna Slater, PATH, unpublished data, 26 January 2023.

To generate a model simulating the sensitivity of RDTs at different levels of *pfhrp2* gene deletion, a database of HRP2 and *Pf*-LDH paired PCR- or microscopy-positive clinical samples was used. For a given parasite prevalence and *pfhrp2* deletion proportion, the following steps were followed:

1. Randomly sample 100 individuals with replacement from the data.
2. Assign each individual a number of parasite strains from a random draw from a Poisson distribution, with the mean determined by the model parasite prevalence.
3. Given the *pfhrp2* deletion proportion in the population, use a random draw from a binomial distribution to estimate the probability that all the strains are *pfhrp2*-deleted (where the number of trials is determined for each individual's MOI in step 2).
4. If all strains are *pfhrp2*-deleted, set that individual's HRP2 concentration to zero.
5. Determine the proportion of the 100 individuals that would be correctly identified as positive using the LOD estimates.
6. Repeat steps 1–5 50 times with new random draws from the data.
7. Repeat steps 1–6 21 times to account for all combinations of three different prevalence levels (5%, 25%, 75%) and seven different levels of *pfhrp2* deletion (2%, 5%, 10%, 25%, 50%, 75%, 100%).

This approach attempts to estimate the sensitivity of different RDTs at different levels of *pfhrp2* deletion and transmission intensity. Results from the models are shown in Fig. 10. Panels (a), (b) and (c) show the results from current best-in-class RDTs. Here, when the proportion of parasites with *pfhrp2* deletion is fairly small (< 10%), the HRP2-only test (a) vastly outperforms the *Pf*-LDH-only RDT (c), but the combined test (b) is better than both of them. However, at a *pfhrp2* deletion prevalence of around 25%, the *Pf*-LDH-only test is predicted to perform equally as well as the best-in-class HRP2 test; at a *pfhrp2* deletion prevalence greater than 25%, the *Pf*-LDH-only test should outperform the best-in-class HRP2 test. At all levels of *pfhrp2* deletion, the combined HRP2/*Pf*-LDH test (b) outperforms both of the single-antigen best-in-class RDTs (a) and (c).

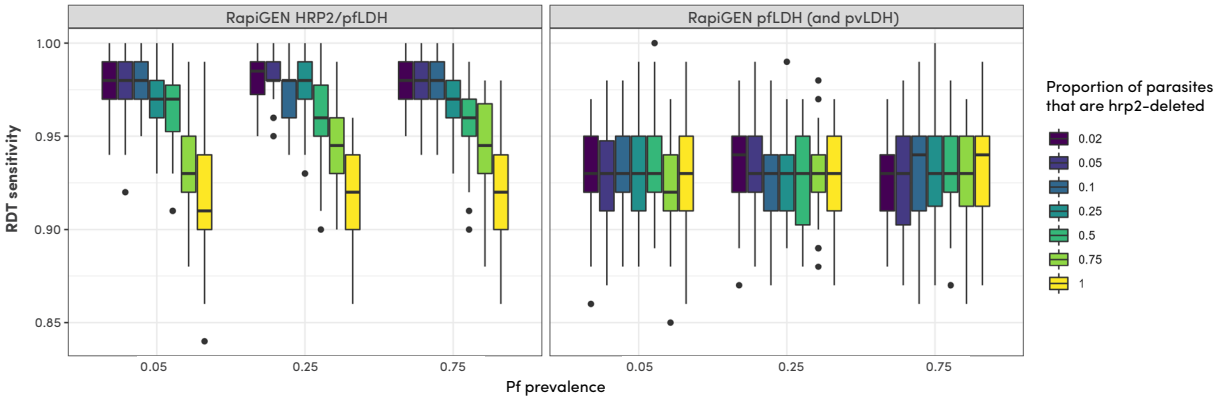
Fig. 10. Performance of three current best-in-class and two next-generation Rapigen BIOCREDIT RDTs against wild-type and *pfhrp2*-deleted parasites



Source: Hannah Slater, PATH, unpublished data, 26 January 2023.

Looking at the next-generation Rapigen BIOCREDIT RDTs (d) and (e), the Rapigen *Pf*-LDH-only test (e) shows a large improvement over the previous best-in-class *Pf*-LDH-only test (c). The Rapigen combined test (d) is still predicted to have higher sensitivity than the Rapigen *Pf*-LDH-only test (e). Fig. 11 shows the performance of the Rapigen tests in greater detail. The Rapigen combined test (a) is still predicted to have higher sensitivity than the Rapigen *Pf*-LDH-only test (b) until *pfhrp2* deletion reaches 100% saturation.

Fig. 11. Performance of two next-generation Rapigen tests against wild-type and *pfhrp2*-deleted parasites (zoom in from Fig. 10 (d) and (e) above)

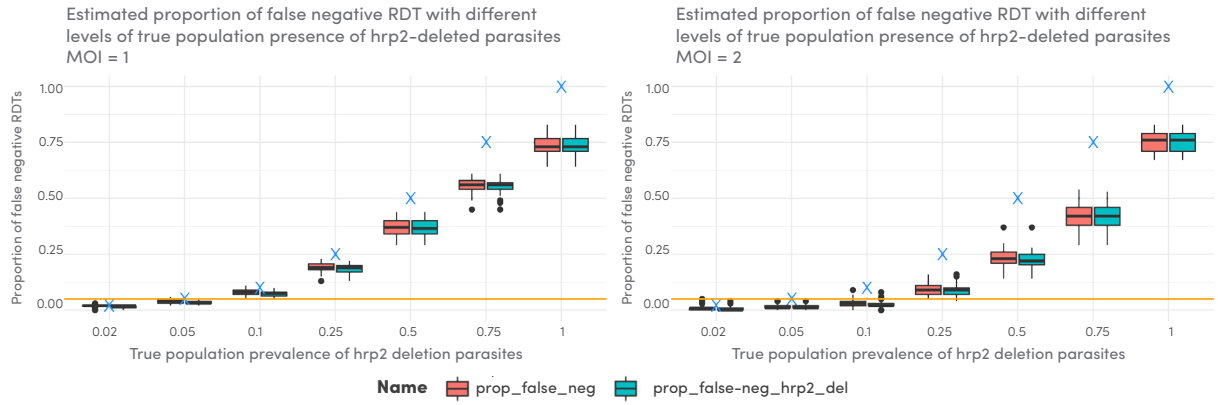


Note: It is important to remember that the models above are based on the true proportion of *pfhrp2* gene deletions in a given parasite population. This contrasts with the proportion of *pfhrp2* deletions leading to false-negative RDT results, which is the measure used by WHO for switching to a non-HRP2-based diagnostic method.

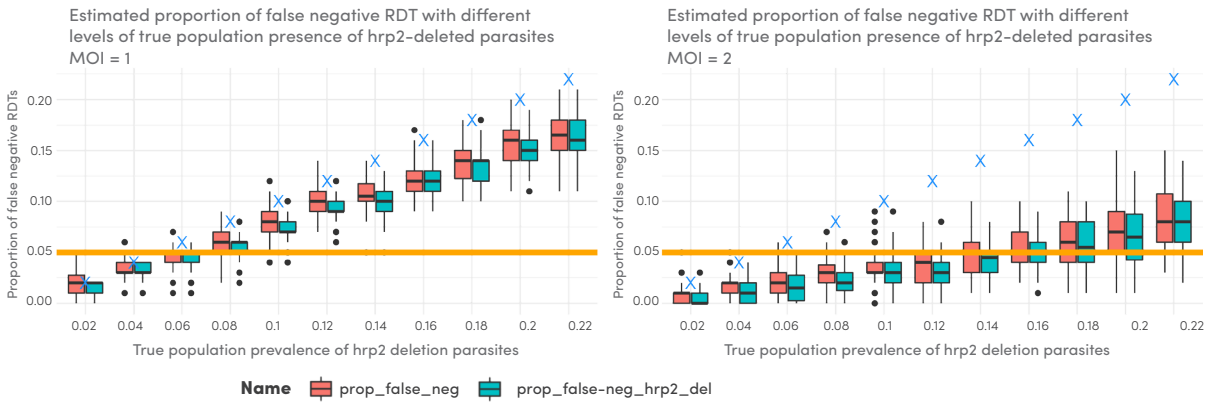
Source: Hannah Slater, PATH, unpublished data, 26 January 2023.

Using the simulation approach outlined above, the proportion of RDT results that would be false-negative for a given population-level proportion of *pfhrp2*-deleted parasites can also be estimated. This relationship is shown in Fig. 12 (a) and (b) for two scenarios (MOI = 1 and MOI = 2). Fig. 12 (c) and (d) zoom in on the same data around the 5% level. Here, we see that a 5% level of false-negative RDT results corresponds to approximately a 6–8% true population prevalence of *pfhrp2*-deleted parasites; for an MOI of 2, this value is around 14–16%.

Fig. 12. Estimated proportion of false-negative RDT results with different levels of *pfhrp2* deletions in the population, at different MOIs



Source: Hannah Slater, PATH, unpublished data, 26 January 2023.



Source: Hannah Slater, PATH, unpublished data, 26 January 2023.

The models above indicate that even in the face of increasing *pfhrp2* deletions, there is still much value in having an HRP2 line on new tests. The next-generation HRP2/*Pf*-LDH RDT is predicted to maintain its sensitivity, even when *pfhrp2* deletion prevalence is high.

2.6.2 Implications of the analysis for the threshold at which countries should switch RDTs

- Countries that are using a best-in-class HRP2-only RDT will need to switch before the *pfhrp2* deletion prevalence in the parasite population reaches 25%, and likely considerably before this point to prevent additional spread and to account for the time needed for procurement and policy change.
- A 5% rate of false-negative RDT results roughly translates to around a 15% population-level *pfhrp2* deletion prevalence (given a mean MOI of 2).
- Switching to a *Pf*-LDH-only RDT does not appear to be advisable. If the country switches to a best-in-class *Pf*-LDH-only RDT, the model predicts that there will be a large reduction in sensitivity unless *pfhrp2* deletion prevalence reaches > 25%. If the country is able to procure a next-generation RDT, the combined HRP2/*Pf*-LDH RDT is predicted to be more sensitive than the *Pf*-LDH-only version.

Moving forward with this model, there is a need to increase the number of clinical samples to make sure that it is representative of symptomatic populations. In addition, there is a need to validate the benchmarking LODs against real-world data (upcoming from Senegal and Ethiopia studies).

2.7 Statistical considerations for updated *pfhrp2/3* deletion response plan

Presenter: Robert Verity

2.7.1 Statistical power

An important consideration when designing a study is the statistical power, defined as the probability of correctly rejecting the null hypothesis. Typically, a study aims for at least 80% power to ensure a low probability of missing interesting findings. A critical factor that influences power is the sample size, with larger sample sizes leading to higher power. This information can be presented using a sample size formula, or in a

sample size table that presents the minimum number of samples needed to achieve the target power. Once this information is in hand, it should be used alongside other information, such as financial, logistical and ethical considerations, to produce a study design that is both feasible and has good statistical properties. Statistical arguments alone are not enough to make a good study, but they should not be ignored completely.

An important feature of power and sample size arguments is that they only make sense in the context of a known statistical analysis plan. A sample size of 370 may be sufficient to reach 80% power in one statistical analysis plan, but insufficient in another. The original WHO protocol for measuring the prevalence of false-negative RDT results caused by *pfhrp2/3* gene deletions advised a sample size of 370 *P. falciparum* samples (10 clusters of 37) to reach the desired power. However, under the original statistical analysis plan, it has since been determined that the power is very low (less than 50%). This statistical analysis plan established a 95% CI around the estimated prevalence, concluding that prevalence is greater than 5% whenever the lower bound of the CI is above 5%.

2.7.2 Updated statistical analysis plan

For the updated statistical analysis plan, a Bayesian statistical approach has been developed, which has some advantages over the original method. Not only does it achieve higher power, but it also enables the integration of information from the many *pfhrp2/3* studies conducted to date in the form of priors. This method is made available through an interactive web app (https://shiny.dide.ic.ac.uk/DRpower-app/) that can be used in both the design and analysis phases of a study to ensure statistically valid results. The details and assumptions of the method are described extensively in the associated R package (https://mrc-ide.github.io/DRpower/).

Table 6 shows the updated sample sizes required to achieve at least 80% power using the DRpower method. This analysis assumes an intra-cluster correlation of 0.05 based on historical studies and a true prevalence of 10% at the population level. Minimum sample sizes are presented for 5–10 clusters. For fewer than five clusters, the sample size would exceed 2000 samples per cluster, which is considered to be prohibitively large and so values are not shown. The total sample size decreases as the number of clusters increases because intra-cluster correlation is less of an issue. This is one of many reasons why it is advisable to focus efforts on recruiting more clusters, rather than having larger sample sizes per cluster.

The values in Table 6 represent the number of confirmed malaria cases that make it into the final analysis. When considering the number of suspected malaria cases to enrol, these values should be increased to account for the positive fraction. For example, if 40% of suspected cases tend to come back as confirmed cases in a given cluster, then the sample size in Table 6 should be divided by 0.4 to convert it to the number of suspected cases. Similarly, values should be buffered to account for dropout.

Table 6 indicates that if 10 clusters can be recruited, then a minimum of 30 samples are needed per cluster. Note that this is fewer than the 37 samples in the original master protocol. The reason for the smaller sample size is that a more powerful analytical approach is being used, with a more pessimistic assumption of 10% prevalence at the population level compared to 8% in the original protocol.

Table 6. Sample sizes required under the DRpower statistical analysis plan based on the number of clusters

Number of clusters	Sample size per cluster	Total sample size
5	496	2480
6	113	678
7	68	476
8	51	408
9	37	333
10	30	300

2.8 Risk-based approach to transitioning to alternative RDTs

Presenter: Oliver Watson

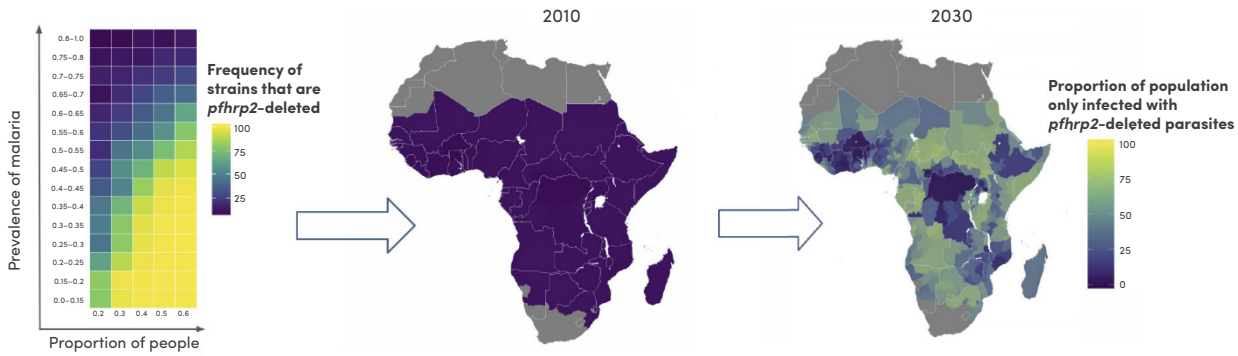
Modelling work has been performed to try to understand how susceptible regions are to the spread and increase of *pfhrp2* deletions in an area and what this may mean in terms of the timeline for transition to alternative RDTs.

2.8.1 Previous approaches to modelling *pfhrp2/3* risk

The original model was conducted in 2015–2016, when there were very few data available to parameterize key aspects of the model. New data can now be used to produce more up-to-date estimates of risk and timelines. The previous approach used a mathematical model of malaria transmission to estimate how RDT use would create selective pressure for *pfhrp2*-deleted parasites. Also included was an estimate of cross-reactivity with HRP3 epitopes, which was estimated at the time to be 25%. Using this model, the key determinants of *pfhrp2* deletions were identified as treatment-seeking rates and malaria prevalence.

Malaria prevalence and treatment-seeking rates from 2010 were incorporated into the model framework to generate a map of how deletions may be expected to increase because of HRP2-based RDTs. A flat 6% prevalence of *pfhrp2* deletions was used for all malaria-endemic countries. Fig. 13 shows the predicted proportions of the population infected with only *pfhrp2*-deleted parasites in 2010 and in 2030.

Fig. 13. Estimated proportion of population infected with only *pfhrp2*-deleted parasites in 2010 and 2030 using the model

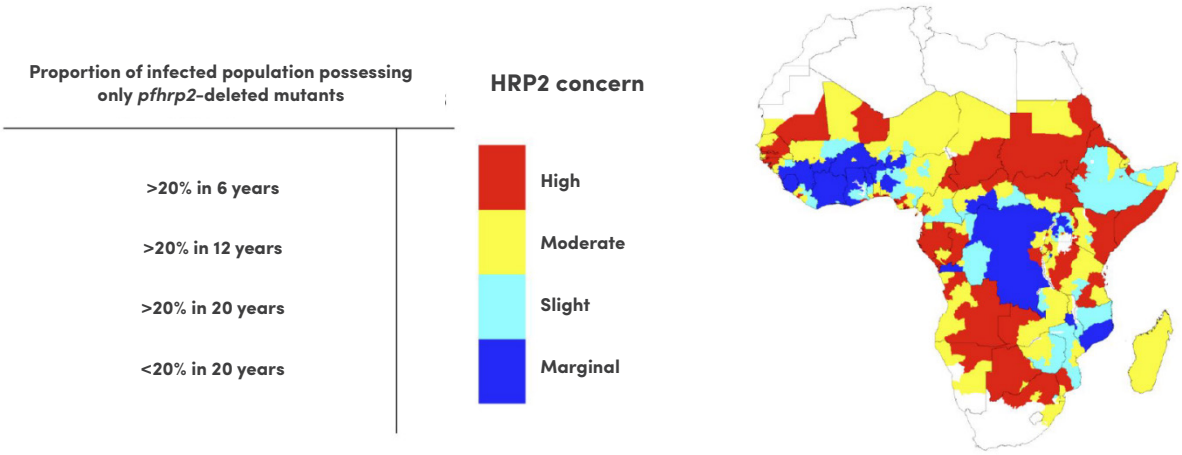


Note: The model assumes no change in prevalence and treatment coverage. The spread is therefore conservative.

Source: Oliver Watson, Imperial College London, unpublished data, 26 January 2023.

Fig. 14 shows the same map but categorized into four groups using discrete thresholds. Areas that have low malaria prevalence and high adherence to RDT results for treatment show the greatest risk for spread of *pfhrp2* deletions. However, low prevalence areas are at risk, as there is an increased chance that infections will be due to only one strain of malaria.

Fig. 14. Fitted model showing the risk of *pfhrp2*-deleted mutants over time by category



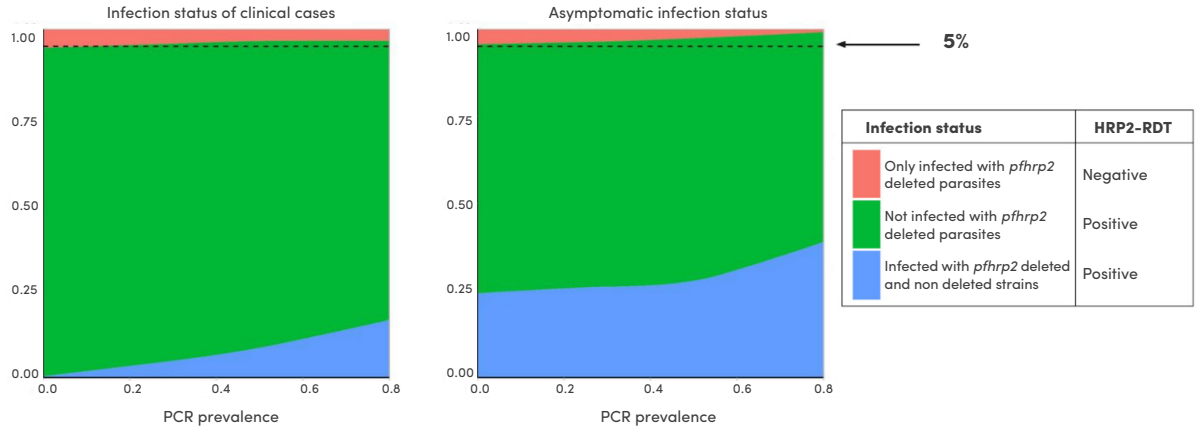
Source: Oliver Watson, Imperial College London, unpublished data, 26 January 2023.

2.8.2 Limitations of this approach

Metric that is reported: Several assumptions had to be introduced in the original model. First, the results presented tracked the proportion of all individuals (both asymptomatic and symptomatic infections) that were infected with only *pfhrp2*-deleted parasites. However, the probability that individuals will be infected with only *pfhrp2*-deleted parasites (i.e. an HRP2-based RDT will be less likely to yield a positive result) depends both on their infection status and transmission intensity.

Therefore, a model was constructed in which 5% of all parasites had *pfhrp2* deletions. Fig. 15 shows the results of this model. The proportion of individuals that are infected with only *pfhrp2*-deleted parasites is higher for clinical cases than for asymptomatic cases and decreases with increasing transmission intensity.

Fig. 15. Model showing proportion of *pfhrp2*-deleted and non-deleted strains, assuming 5% of parasites have *pfhrp2* deletions



Source: Oliver Watson, Imperial College London, unpublished data, 26 January 2023.

Simulation scenarios chosen: It was previously assumed that each country already had a 6% prevalence of *pfhrp2* deletions, which was a simplifying assumption based on 2016 data from the Democratic Republic of the Congo. Since 2016, significantly more data points have been generated that could be used to predict the continued spread of *pfhrp2/3* gene deletions.

Parameters selected for each region: The main drivers of selection for *pfhrp2* deletions are transmission intensity and treatment coverage. Although several other factors impact selection, in 2016, there was only limited information on:

- incidence of non-malarial fevers;
- proportion of clinical cases tested using microscopy;
- adherence to diagnostic test outcomes for treatment;
- size of the private drug market and the diagnostic practices of the private market;
- fitness costs associated with *pfhrp2/3* deletions;
- cross-reactivity of HRP2-based RDTs with HRP3;
- frequency of *pfhrp3* deletions; and
- whether *pfhrp2* deletions occur independently of *pfhrp3* deletions.

2.8.3 Updated plan for modelling the timeline for transitioning to alternative RDTs

Table 7 shows the updated parameters and data sources for modelling the timeline for transitioning to an alternative RDT.

Table 7. Updated parameters and data sources

Model parameter	Impact on speed of selection for *pfhrp2/3* gene deletions	Data source
Malaria prevalence	Lower malaria prevalence will increase selection by increasing the probability that individuals are infected with only *pfhrp2/3*-deleted parasites and are thus more likely to not be treated.	Malaria Atlas Project maps of slide positivity 2–10
Treatment-seeking rate	Increased treatment-seeking will increase the rate at which the selective advantage conferred by *pfhrp2/3* gene deletion is able to be exerted by evading diagnosis and treatment.	Demographic and Health Survey (DHS)/ Malaria Indicator Survey data used in generalized additive mixed models to predict care-seeking patterns
Size of the private drug market	Low use of malaria RDTs has been shown to exist in the private market in a number of locations *(10)*. If the use of RDTs is lower in the private market than in the public sector, then selection will decrease with an increasingly large private drug market.	As above, but two different generalized additive mixed models used: (i) estimating care-seeking from any (medical) source, (ii) estimating care-seeking in the public sector
Proportion of individuals seeking care who receive diagnostic test	Low use of any diagnostic test for guiding treatment decisions will reduce selective pressure for *pfhrp2/3* deletions.	DHS data (surveys in Africa asking if care-seeking febrile children received a finger/heel prick) and literature review

Model parameter	Impact on speed of selection for *pfhrp2/3* gene deletions	Data source
Non-adherence to RDT outcomes	Non-adherence to RDT outcomes (treating RDT-negative individuals) will decrease selection by negating the selective advantage conferred by *pfhrp2/3* deletions.	(i) DHS-data-based spatial model of the probability of people seeking care for fever receiving any antimalarial vs none; (ii) literature reviews of presumptive treatment
Microscopy-based diagnosis	The use of microscopy for malaria diagnosis will decrease selection by negating the selective advantage conferred by *pfhrp2/3* deletions.	(i) World Malaria Report's "proportion of cases confirmed by diagnostic" table; (ii) literature reviews
Cross-reactivity of HRP3 epitopes	Increasing cross-reactivity between HRP3 epitopes and HRP2-based RDTs will decrease selection for *pfhrp2* deletions.	Literature review of cross-reactivity testing of RDT types/brands
Fitness costs	Increased fitness costs due to *pfhrp2/3* gene deletions will decrease the speed of selection.	In vitro competition assay data *(11)*
RDT brands	Test sensitivity and cross-reactivity with HRP3 epitopes differ between brands of RDTs. In addition, a number of RDT brands detect non-HRP2 targets. Increasing cross-reactivity will decrease selection, while the use of non-HRP2-based RDTs will negate the selective advantage conferred by *pfhrp2/3* gene deletions.	Clinton Health Access Initiative (CHAI)/ United States President's Malaria Initiative data on volumes/proportions of RDT test types and brands used

Research into updating these parameters has been conducted by various teams, including the Malaria Atlas Project's CHAI commodities forecasting framework *(12)* by Tasmin Symons and researchers at Penn State University (Thu Tran). Although more data are available now than in 2016, considerable uncertainty still exists around key model parameters. In future, this uncertainty will be presented more transparently and incorporated when modelling administrative regions, for example, by providing the worst and best case scenarios based on uncertainty ranges.

2.8.4 Frequency of *pfhrp2* deletions without *pfhrp3* deletions

The United States Centers for Disease Control and Prevention published a study in 2021 showing that there was high cross-reactivity of HRP3 epitopes on HRP2-detecting RDTs *(13)*. As a result, an RDT can still give a positive result even in the absence of HRP2. Therefore, it is necessary to consider both types of deletion when making decisions about switching RDTs. Due to these cross-reactions, there is a need to understand how frequently *pfhrp2* deletions arise independently of *pfhrp3* deletions.

It has been hypothesized that *pfhrp2* deletions and *pfhrp3* deletions arise through different selection mechanisms and that a different evolutionary process underpins the two types of deletion. To estimate the relationship between these types of deletions, the WHO Malaria Threats Map was used to model the proportion of all *pfhrp2*-deleted parasites that were also found to have *pfhrp3* deletions. This was done by focusing on studies in which the same samples were tested for both *pfhrp2* and *pfhrp3* deletions.

Using beta-binomial distribution to model the proportion, it was estimated that 66% of samples with *pfhrp2* deletions also have *pfhrp3* deletions, although there was a large

range in this proportion across studies. This information will be incorporated into the modelling to account for *pfhrp2*-deleted samples still yielding positive HRP2-based RDT results due to cross-reactivity with HRP3.

2.8.5 Incorporating current data on the spatial distribution of *pfhrp2/3* deletions

Ideally, the data from the WHO Malaria Threats Map would be used to model the continued spread of *pfhrp2/3* gene deletions. However, survey quality and protocols vary greatly between studies, making this difficult, and negative survey results have not always been reported. Therefore, this approach may not be the most suitable.

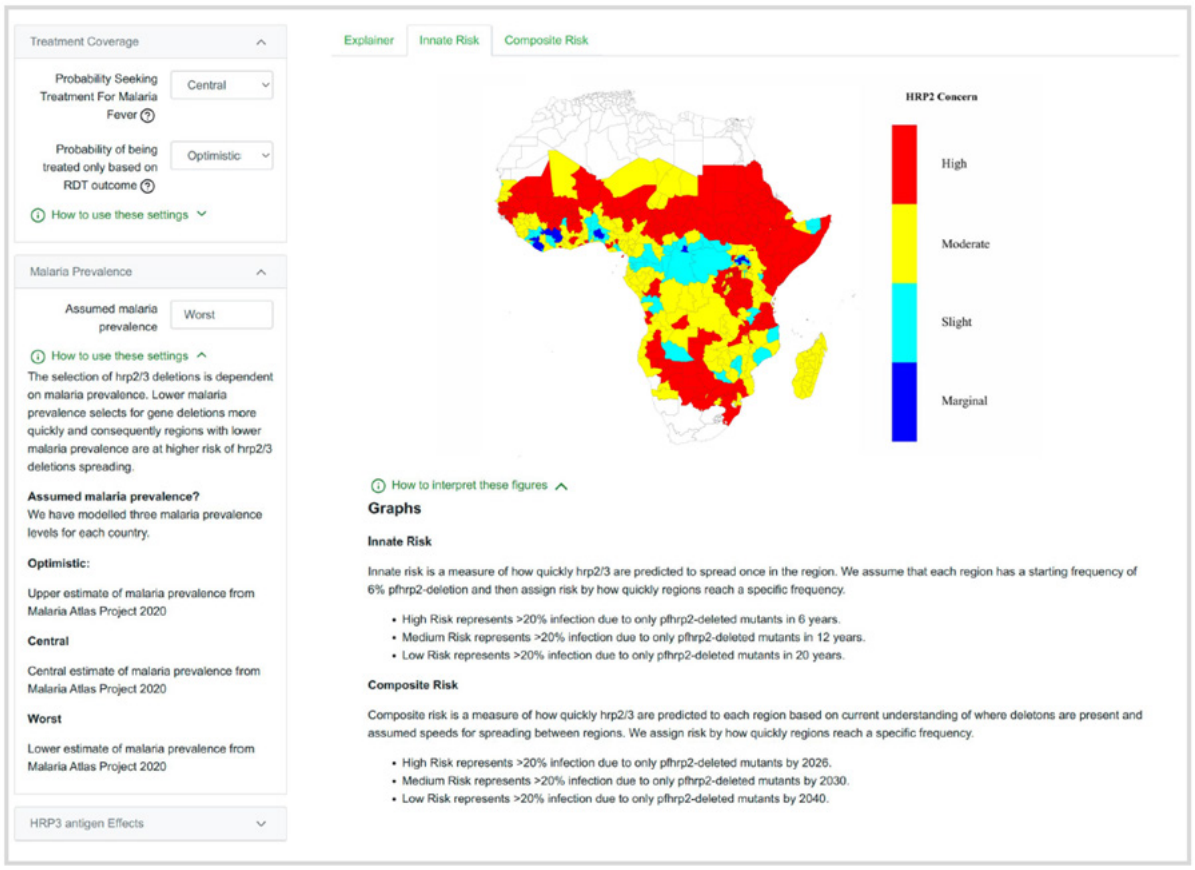
There are two approaches to modelling *pfhrp2/3* risk: innate risk score and composite risk score.

Innate risk score: This is the potential for *pfhrp2/3* deletions once present in a region to spread, based solely on the region's malaria transmission intensity and treatment-related data. To measure the innate risk, the same approach as used previously will be used, assuming a starting frequency of *pfhrp2/3* deletions and recording how quickly those deletions increase and subsequently cause clinical cases to be missed by HRP2-based RDTs.

Composite risk score: This will simulate the spread of infections caused by *pfhrp2/3*-deleted parasites out of the Horn of Africa – i.e. out of the clear "hot spot" of *pfhrp2/3* gene deletions in Africa that have been shown to be symptomatic and relevant. Different assumptions will be made for how quickly deletions will be introduced into regions based on (i) location adjacent to regions with *pfhrp2/3* deletions and (ii) the frequency of deletions in those regions. Once deletions are assumed to have been imported into a region, selection speeds from the innate risk score modelling will be used to estimate the subsequent time for different thresholds of concern with *pfhrp2/3* gene deletions to be reached.

The innate risk will be presented in an interactive public health tool (developed by Robert Zupko (Penn State Undergraduate Education)) (Fig. 16). Users will be able to explore different risk maps based on different assumed model parameters and look at the impact of parameter uncertainties. A risk score map based on best parameter estimates will be included in the updated response plan. This will help to focus future surveillance efforts and identify countries most likely to require non-HRP2 RDTs once deletions have been reported. Discussions are ongoing to decide whether to include composite risk scores in the updated response plan, as how quickly deletions pass from one country to another is much more difficult to determine and is more open to criticism.

Fig. 16. **Example of the interactive public health decision-making tool**



3. Discussion points and conclusions to support update to the WHO response plan to *pfhrp2* deletions

3.1 Threshold for changing RDTs

- One point that was repeatedly raised was the need to clearly articulate the difference between the population prevalence of *pfhrp2* gene deletions and the prevalence of false-negative RDT results caused by *pfhrp2* deletions, which are often confused. Due to a number of factors, including MOI and the cross-reactivity of HRP2 test lines with HRP3, these two values can be quite different. WHO recommendations for switching RDTs are based on the latter (the prevalence of false-negative RDT results caused by *pfhrp2* deletions), which is not always available, as it requires knowledge of both genotyping and RDT results. Therefore, there is interest in exploring whether a threshold of deletions can be determined at a population level. Hannah Slater's presentation showed that, based on their data set, a population threshold of 25% *pfhrp2* deletions could be used to inform the switch to an alternative RDT. Post-meeting analysis incorporating the effect of MOI revealed that the current WHO threshold of 5% false-negative RDT results caused by *pfhrp2* gene deletions corresponds to approximately a 15% level of *pfhrp2* deletions in a population (MOI=2). However, the model still does not account for the impact of cross-reaction of HRP2 with

HRP3, and therefore it is premature to draw conclusions. The model findings suggest that the current threshold is too low, but given that a switch will take time, it may be prudent to change before a higher population prevalence of 25% is reached. Therefore, a 15% population prevalence or 5% prevalence of false-negative RDT results caused by *pfhrp2* deletions may not be an unreasonable approach. It is necessary to incorporate additional data and factors such as HRP3 cross-reactivity into the modelling to support conclusions that could influence policy.

- There was a lot of discussion around the issue of MOI, including the difficulty of quantifying this value. If the issue of MOI distribution in clinical community-based infections were better understood, then some of these issues would disappear; however, the variation in MOI can be so great that it would be difficult to rely on MOI-based assumptions to draw connections between population prevalence and clinically relevant prevalence of *pfhrp2* deletions. However, this issue becomes irrelevant if the focus stays on the prevalence of false-negative RDT results caused by *pfhrp2* gene deletions, rather than on the prevalence of deletions in the parasite population.

- The review of the published literature comparing the performance of HRP2 and *Pf*-LDH test lines to microscopy and/or PCR in a range of clinical settings highlights the significant variability, likely reflecting differences in malaria epidemiology and RDT brands. However, on average, the difference (compared to microscopy) is close to 5%, which supports the current recommendation for switching to an alternative RDT. If countries have their own data/understanding of performance trade-offs, this could be used to inform RDT procurement decisions. Otherwise, based on this review, continued use of the 5% threshold is a reasonable measure to guide RDT selection based on HRP2 for detection of *P. falciparum*.

- Unpublished data were reviewed in the context of the recent commercialization and implementation of new *Pf*-LDH RDTs. Specifically, the data presented from Ethiopia and the United Republic of Tanzania suggest that the next-generation Rapigen *Pf*-LDH-based RDTs are comparably sensitive to HRP2-based RDTs in these clinical settings. This finding then raises the question as to why not simply scale up these new tests everywhere, regardless of *pfhrp2* deletion prevalence. There are, however, a number of factors that warrant consideration, such as the limited capacity of the manufacturer to scale up these RDTs to such an extent so quickly, the ability of countries or donors to pay for these new RDTs (which are currently significantly more expensive), and the dissemination of performance data and training materials for these RDTs so that countries feel comfortable making the switch. These factors mean that it will take time to fully introduce relevant new tests.

- Data from PATH used to support modelling and preliminary analysis from Senegal suggest that there is still a significant gap in analytical and clinical sensitivity between HRP2 and *Pf*-LDH RDTs. Therefore, the ideal configuration of future RDTs is to detect both HRP2 and *Pf*-LDH on the same test line. This would optimize sensitivity and make it easier for health workers to interpret the result (limiting re-training requirements).

3.2 Factors affecting determination of *pfhrp2/3* deletions

Participants who had experience analysing samples in the laboratory generally agreed that it is difficult to work with samples with low parasite densities, and it is challenging to ensure the accurate determination of the *pfhrp2/3* status of low-density samples. Discrepancies between methods and between laboratories have been observed. In this context, the 5% threshold allows for some error due to these laboratory challenges. However, while this is true, if the main interest were symptomatic people at health facilities, the low-density samples would not be so problematic. Measuring population-level gene deletions means that the technical accuracy is lower than just looking at clinically relevant infections, as there will be more low-density samples, which are prone to errors. In the published literature, investigators often only interrogate samples with > 100 parasites/µL for *pfhrp2* and *pfhrp3* deletions to increase confidence in deletion determinations.

3.3 Sample size for determining prevalence of *pfhrp2* deletions

In relation to the sample size calculations, there was recently a correction of an error in the original calculations (design effect not included), which led to significant increases in the sample size; however, the approach itself was also found to have power issues. Furthermore, the approach could be simplified by aiming to simply prove prevalence above the 5% threshold, instead of also trying to prove a prevalence that is less than 5%, as it is not so important to define this lower prevalence so accurately. The alternative approach proposed by Robert Verity was found to be convincing, but the group advised that the approach be reviewed by additional statisticians.

3.4 Risk-based transitioning to alternative RDTs

The modelling approach proposed by Oliver Watson for predicting the potential spread of *pfhrp2* deletions and planning the switch to non-HRP2 RDTs was generally appreciated. Although this modelling approach can serve as a guide for planning *pfhrp2* deletion surveys, it is still important to collect actual data on the rate of false-negative results with HRP2-based RDTs caused by *pfhrp2* deletions (reaching the 5% threshold) for making policy decisions regarding the switch to non-HRP2-based RDTs.

References

1. World malaria report 2022. Geneva: World Health Organization; 2022 (https://iris.who.int/handle/10665/365169, accessed 3 October 2023).

2. Thomson R, Parr JB, Cheng Q, Chenet S, Perkins M, Cunningham J. Prevalence of *Plasmodium falciparum* lacking histidine-rich proteins 2 and 3: a systematic review. Bull World Health Organ. 2020;98(8):558–68F. doi:10.2471/BLT.20.250621.

3. False-negative RDT results and implications of new reports of *P. falciparum* histidine-rich protein 2/3 gene deletions. Geneva: World Health Organization; 2017 (https://iris.who.int/handle/10665/258972, accessed 3 October 2023).

4. Response plan to *pfhrp2* gene deletions. Geneva: World Health Organization; 2019 (https://iris.who.int/handle/10665/325528, accessed 3 October 2023).

5. Mihreteab S, Anderson K, Pasay C, Smith D, Gatton ML, Cunningham J, et al. Epidemiology of mutant *Plasmodium falciparum* parasites lacking histidine-rich protein 2/3 genes in Eritrea 2 years after switching from HRP2-based RDTs. Sci Rep. 2021;11(1):21082. doi:10.1038/s41598-021-00714-8.

6. Gamboa D, Ho MF, Bendezu J, Torres K, Chiodini PL, Barnwell JW, et al. A large proportion of *P. falciparum* isolates in the Amazon region of Peru lack *pfhrp2* and *pfhrp3*: implications for malaria rapid diagnostic tests. PLoS One. 2010;5(1):e8091. doi:10.1371/journal.pone.0008091.

7. Watson OJ, Slater HC, Verity R, Parr JB, Mwandagalirwa MK, Tshefu A, et al. Modelling the drivers of the spread of *Plasmodium falciparum hrp2* gene deletions in sub-Saharan Africa. Elife. 2017;6:e25008. doi:10.7554/eLife.25008.

8. Malaria rapid diagnostic test performance: results of WHO product testing of malaria RDTs: round 6 (2014–2015). Geneva: World Health Organization; 2015 (https://iris.who.int/handle/10665/204118, accessed 3 October 2023).

9. Slater HC, Ding XC, Knudson S, Bridges DJ, Moonga H, Saad NJ, et al. Performance and utility of more highly sensitive malaria rapid diagnostic tests. BMC Infect Dis. 2022;22(1):121. doi:10.1186/s12879-021-07023-5.

10. Shelus V, Mumbere N, Masereka A, Masika B, Kiitha J, Nyangoma G, et al. "Testing for malaria does not cure any pain": a qualitative study exploring low use of malaria rapid diagnostic tests at drug shops in rural Uganda. PLoS Glob Public Health. 2022;2(12):e0001235. doi:10.1371/journal.pgph.0001235.

11. Nair S, Li X, Nkhoma SC, Anderson T. Fitness costs of pfhrp2 and pfhrp3 deletions underlying diagnostic evasion in malaria parasites. J Infect Dis. 2022;226(9):1637–45. doi:10.1093/infdis/jiac240.

12. Symons TL. CHAI commodities methodology. Boston: Clinton Health Access Initiative; 2022 (https://clintonhealth.app.box.com/s/3l7zni6ylp00a1oc6194re1s8nfdnzp0, accessed 3 October 2023).

13. Kong A, Wilson SA, Ah Y, Nace D, Rogier E, Aidoo M. HRP2 and HRP3 cross-reactivity and implications for HRP2-based RDT use in regions with *Plasmodium falciparum* hrp2 gene deletions. Malar J. 2021;20(1):207. doi:10.1186/s12936-021-03739-6.

Annex 1. Meeting agenda

Chairperson: Kumar V. Udhayakumar

Thursday 26 January 2023		
14:00–14:05	Opening of meeting	Andrea Bosman, Director, GMP a.i.
14:00–14:10	Meeting objectives and declaration of interests	Jane Cunningham
14:10–14:15	Introductions	All
14:15–14:20	Overview of key resources – response plan and survey protocol template	Jane Cunningham
14:20–15:10	Revisiting the 5% threshold	Michelle Gatton Rebecca Thomson Fitsum Girma Deus Ishengoma Gonzalo Domingo
15:40–15:50	Alternative approaches	Hannah Slater
15:50–16:20	Discussion	
16:20–16:35	Sample sizes based on thresholds and alternatives	Robert Verity
16:35–16:50	Discussion	
17:00–17:20	Risk based approach to transitioning to alternative RDTs	Oliver Watson
17:20–17:50	Discussion	
17:50–18:00	Next steps	Jane Cunningham
18:00	End of meeting	

Annex 2. List of participants

Expert panel

Alexander Asamoah
National Malaria Elimination Programme
Ghana Health Service
Accra
Ghana

Khalid Beshir
London School of Hygiene & Tropical Medicine
London
United Kingdom of Great Britain and Northern Ireland

Praveen K. Bharti
ICMR- National Institute of Malaria Research
New Delhi
India

Qing Cheng
Australian Defence Force Malaria and Infectious Disease Institute
Drug Resistance and Diagnostics Department
Brisbane
Australia

Anderson Chinorumba
Independent Consultant
Harare
Zimbabwe

Gonzalo Domingo
PATH
Seattle
United States of America

Dionicia Gamboa Vilela
Department of Cellular and Molecular Sciences
Universidad Peruana Cayetano Heredia
Lima
Peru

Michelle Gatton
School of Public Health and Social Work
Faculty of Health
Centre for Immunology and Infection Control
Queensland University of Technology
Brisbane
Australia

Deus Ishengoma
National Institute for Medical Research
Dar Es Salaam
United Republic of Tanzania

Cristian Koepfli
Department of Biological Sciences
Eck Institute for Global Health
University of Notre Dame
Notre Dame
United States of America

Selam Mihreteab
Ministry of Health
Asamara
Eritrea

Daouda Ndiaye
Cheikh Anta Diop University Dakar
Parasitology-Mycology Department
Faculty of Medicine and Pharmacy
Dakar
Senegal

Jonathan Parr
Division of Infectious Diseases
University of North Carolina
School of Medicine Department of Medicine
Chapel Hill
United States of America

Eric Rogier
Division of Parasitic Diseases and Malaria
Malaria Branch
The Centers for Disease Control and Prevention
Atlanta
United States of America

Hannah Slater
Malaria and NTDs and Malaria Diagnostics
PATH
Seattle
United States of America

Fitsum Girma Tadesse
Armauer Hansen Research Institute
Addis Ababa
Ethiopia

Rebecca Thomson (Rapporteur)
Independent Consultant
Glasgow
United Kingdom

Venkatachalam (Kumar) Udhayakumar
Independent Consultant
Atlanta
United States of America

Angela Patricia Guerra Vega
Grupo de Parasitología
Dirección de Redes en Salud Pública
Instituto Nacional de Salud
Bogota
Colombia

Robert Verity
Faculty of Medicine, School of Public Health
Imperial College
London
United Kingdom

Theodoor Visser
Clinton Health Access Initiative
Boston
United States of America

Oliver Watson
Faculty of Medicine
School of Public Health
Imperial College
London
United Kingdom

WHO Secretariat

Andrea Bosman
Director a.i.
Global Malaria Programme

Jane Cunningham
Technical Officer
Diagnostics, Medicines and Resistance Unit
Global Malaria Programme

Roberto Montoya
Regional Malaria Adviser
WHO Regional Office for the Americas

Abdisalan Noor
Unit Head, Strategic Information for Response Unit
Global Malaria Programme

Anne-Laure Page
Scientist
WHO Prequalification of IVDs Programme

Maria de la Paz Ade y Torrent
Advisor, Malaria Diagnostics and Supply Management
Department of Communicable Diseases and Environmental Determinants of Health
WHO Regional Office for the Americas/ Pan American Health Organization

Risintha Premaratne
Regional Malaria Adviser
WHO Regional Office for the South East Asia

For further information please contact:

Global Malaria Programme
World Health Organization
20 avenue Appia
1211 Geneva 27
Switzerland
Email: GMPinfo@who.int

9789240097919